ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ ІСТ-2021

ПРАЦІ
X МІЖНАРОДНОЇ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

13 – 19 вересня 2021 року
Харків - Одеса, Україна

Ministry of Education and Science of Ukraine
National Academy of Sciences of Ukraine
Polish Academy of Science, Branch in Lublin
Representative office „Polish Academy of Science” in Kyiv
Kharkiv National University of Radio Electronics
Science and Technological University Akademia Gorniczo-Hutnicza
named after St. Staszic in Krakow
Vasyl Stefanyk Precarpathian National University
Admiral Makarov National University of Shipbuilding in Mykolayiv
Zaporozhye National University
Academy of Applied Radioelektronics Sciences of Ukraine, Russia and Belarus
Ukrainian Oil and Gas Academy
Ukrainian Federation of Informatics
Belarus State University of Informatics and Radioelectronic
Belarus National Technical University

INFORMATION SYSTEMS AND TECHNOLOGIES IST-2021

Proceedings
of the 10-th International Scientific and Technical Conference

September 13-19, 2021
Kharkiv - Odesa, Ukraine

Kharkiv - Odesa 2021

Міністерство освіти та науки України
Національна академія наук України
Люблінський відділ Польської Академії Наук
Представництво „Польська академія наук” у Києві
Харківський національний університет радіоелектроніки
AGH науково-технологічний університет ім. Ст. Сташіца в Кракові
Прикарпатський національний університет ім. В. Стефаника
Національний університет кораблебудування ім. адмірала Макарова
Запорізький національний університет
Академія Наук Прикладної Радіоелектроніки
Українська нафтогазова академія
Українська Федерація Інформатики
Білоруський державний університет інформатики та радіоелектроніки
Білоруський національний технічний університет

Інформаційні системи та технології ІСТ-2021

П Р А Ц І
10-ї Міжнародної науково-технічної конференції

13-19 вересня 2021 р.
Харків - Одеса, Україна

Харків – Одеса 2021

УДК: 004.9

Наукові редактори: *А.Д. Тевяшев - доктор технічних наук, професор (Харківський національний університет радіоелектроніки);*
Л.Б. Петришин - доктор технічних наук, професор (AGH University of Science and Technology; Прикарпатський національний університет ім.В.Стефаника);
В.В. Безкоровайний - доктор технічних наук, професор (Харківський національний університет радіоелектроніки);
В.Г. Кобзєв – кандидат технічних наук, старший науковий співробітник (Харківський національний університет радіоелектроніки)

Матеріали статей рецензовано та опубліковано в авторській редакції.

Рецензенти: *Л.Б. Петришин – UST AGH, Польща; Б.О. Железко – БНТУ, Білорусь; Н.Г. Аксак, В.В. Безкоровайний, Є.В. Бодянський, А.Л. Єрохін, Л.О. Кіріченко, В.Г. Кобзєв, С.В. Машталір, К.Е Петров, Т.В. Полозова, О.Г. Руденко, І.В. Свид, А.Д. Тевяшев, В.О. Філатов – ХНУРЕ, Україна; О.В. Бісікало, Ю.О. Гунченко – ОНУ, Україна; С.О. Іванченко – НТУУ КПІ, Україна; Н.Ф. Казакова, О.О. Фразе-Фразенко – ОДЕКУ, Україна; В.В. Карасюк – НЮУ, Україна; В.А. Лужецький – ВНТУ, Україна; Л.І. Нефьодов – ХНАДУ, Україна; В.М. Синєглазов – НАУ, Україна; С.В. Чопоров – ЗНУ, Україна; Н.В. Шаронова – НТУ КПІ, Україна; І.В. Шевченко – КрНУ, Україна; О.О. Шумейко – ДГТУ, Україна; С.В. Яковлев – НАКУ ХАІ, Україна, Evgeni Schumm - Hochschule Hamm-Lippstadt, Німеччина*

Інформаційні системи та технології: праці 10-ї Міжнародної науково-технічної конференції, Харків-Одеса, 13-19 вересня 2021 року / наук. ред. А.Д. Тевяшев, Л.Б. Петришин, В.В. Безкоровайний, В.Г. Кобзєв. – Х.: ХНУРЕ, 2021. – 357 с.

ISBN 978-617-7519-59-0

Збірка містить матеріали статей Міжнародної науково-технічної конференції з проблем сучасних інформаційних систем та технологій.
Праці представляють інтерес для фахівців, науковців і аспірантів, діяльність яких пов'язана з розробкою та впровадженням сучасних інформаційних систем і технологій.

Visualization Modeling of Processes in Distributed Management Systems

Lubomyr Petryshyn [1], Mykhailo Petryshyn [2]

[1] *AGH University of Science and Technology, al. Mickiewicza 30, Cracow, 30-059, Poland*
[2] *Vasyl Stefanyk Precarpathian National University, 57 Shevchenko str., Ivano-Frankivsk, 76018, Ukraine*

Abstract
Simulation of the processes of sectoral cooperation management in distributed information systems allows to reduce the means of introduction and operation of such complex systems. The proposed method of visualization of information models reflects graphically the constituent processes and simplifies the understanding at the stage of analysis and design between the customer and the system developer. The basics of visualization modeling and simplified example of development of models of two-sectoral parking management system are presented.

Keywords
Visualization modeling, information processes, distributed systems, management

1. Introduction

Management of complex systems requires the use of information technology that reflects the state of the system and allows real-time management. Visualization of management processes allows to remove the psychological barrier and avoid mutual misunderstandings between the customer and the developer of information systems, as well as to reduce the means of development, implementation and operation of such systems.

The aim of the study is to present visual methods for modeling the management of complex systems in a sectoral cooperation, as well as to develop a simplified example of a two-sector parking network management system. The novelty of the work lies in the introduction of graphical modeling methods that provide visualization of control processes and simplify the understanding of their flow.

The practical significance lies in the ability to reflect the structure and course of management processes, to avoid misunderstandings in setting tasks and meeting customer requirements, as well as reducing the cost of development and operation of systems.

Fundamentals of visualization modeling are published in [1 - 3]. Below we analyze a simplified example of modeling an information management system for two selected sectors.

Parkspace company specializes in the management of urban space in order to create parking spaces, fulfills external orders at the request of cities or private institutions. The dynamic development of the company is ensured by the creation of P&R parking lots in large cities and numerous modifications of existing parking lots in urban centers, which led to the need of development an IT system that could improve the company's performance.

The presented example of the information system project will help to optimize the process of data flow in the company and will provide effective company computerization while reducing operating costs [4 - 14]. Figure 1 shows the hierarchical organizational structure of Parkspace company.

2. Analysis of company organizational departments

The company chairman coordinates the activities of the company and leads the work of all departments. His representative at business meetings is responsible for accepting orders and the current direction of the company's development.

EMAIL: l.b.petryshyn@gmail.com (A. 1);
m.l.petryshyn@gmail.com (A. 2)
ORCID: 0000-0003-4168-3891 (A. 1);
0000-0001-6319-3768 (A. 2)

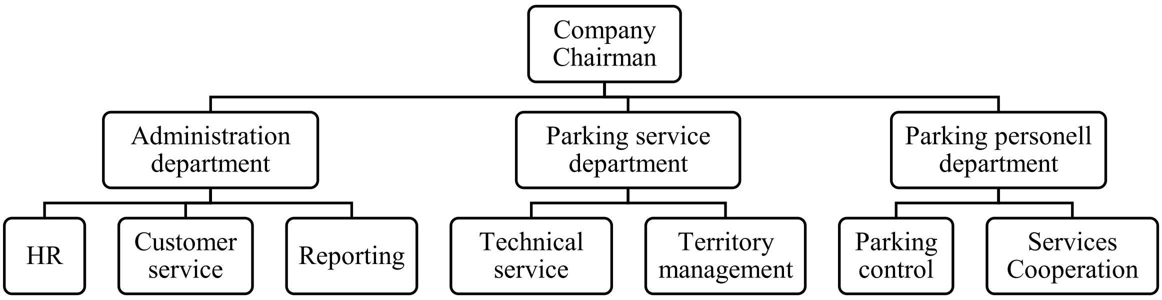


Figure 1: Hierarchical organizational structure of Parkspace Company

Administration department:

- human resources (HR) department is responsible for the current management and recruitment; is responsible for the training and development of the workforce; coordinates the company's salary and implements incentive programs,
- purchasing and maintenance department is responsible for preparing the company's proposals for tenders for the development of land for parking spaces, communicates and negotiates with the clients, reviews existing agreements and manages marketing research to find new customers,
- reporting and settlement department deals with the calculation of salaries and bills in the company, as well as the preparation of analyzes for the client, which represent the profit from the operation of car parks.

Parking service department:

- maintenance department repairs and installs parking meters, as well as provides technical services,
- territorial administration department prepares parking lots, determines parking places, markings of parking lines, cutting down trees, preparation of possible trade premises and adjustment of parking in accordance with the standard developed in the contract.

Parking Service Personnel Department:

- parking control unit coordinates the employees involved in the sale and receipts verification,
- department for cooperation with special services ensures the imposition of fines if necessary, or the evacuation of improperly parked vehicles by notifying the relevant services, transmits reports and statistics of activities to the department of reports and calculations.

Parkspace's management information system allows for the exchange of information between departments, sharing and storage of data and, as a result, the optimization of cooperation between departments.

3. Analysis of system processes

Table 1 presents the analysis results of system operations of the process for marking parking spaces, instead of Table 2 system operations of the process of replacing the parking meter are analyzed.

Table 1

System operations for marking parking spaces

Documents	Title	Responsible
D1.O2	Establishing contact with the client	Customer service
D2.O3	Opening an internal client account	Reporting and settlements department
D3.O2	Deployment of the client's project in the cloud	Customer service
D4.O5	Implementation of technical corrections in the project	Department of Territorial Administration
D5.O5	Marking the place and sending documentation	Department of Territorial Administration

Documents	Title	Responsible
D6.O2	Approval of performed services	Customer service
D7.O2	Send feedback to the O3 department	Customer service
D8.O3	Calculation of the order with the client	Reporting and calculations department
D9.O3	Payment of wages	Reporting and calculations department

Table 2
System operations to replace the parking meter

Documents	Title	Responsible
D1.O6	Registration in the notification system of a non-functioning parking meter	Parking department
D2.O4	The answer to the application is a decision to replace the parking meter	Technical service department
D3.O5	Installation of a new parking meter in the system	Department of Territorial Administration
D4.O5	Preparation of estimates	Department of Territorial Administration
D5.O3	Acceptance of cost estimates	Reporting and settlements department
D6.O5	Installation of a parking meter and sending information to the O4 department	Department of Territorial Administration
D7.O4	Testing of the parking meter to department O6	Technical service department
D8.O6	Parking test	Parking department
D9.O4	Sending the work report to the O3 department	Technical service department
D10.O3	Payment of expenses according to the estimate	Reporting and settlements department

4. Visual models

4.1. Matrix model

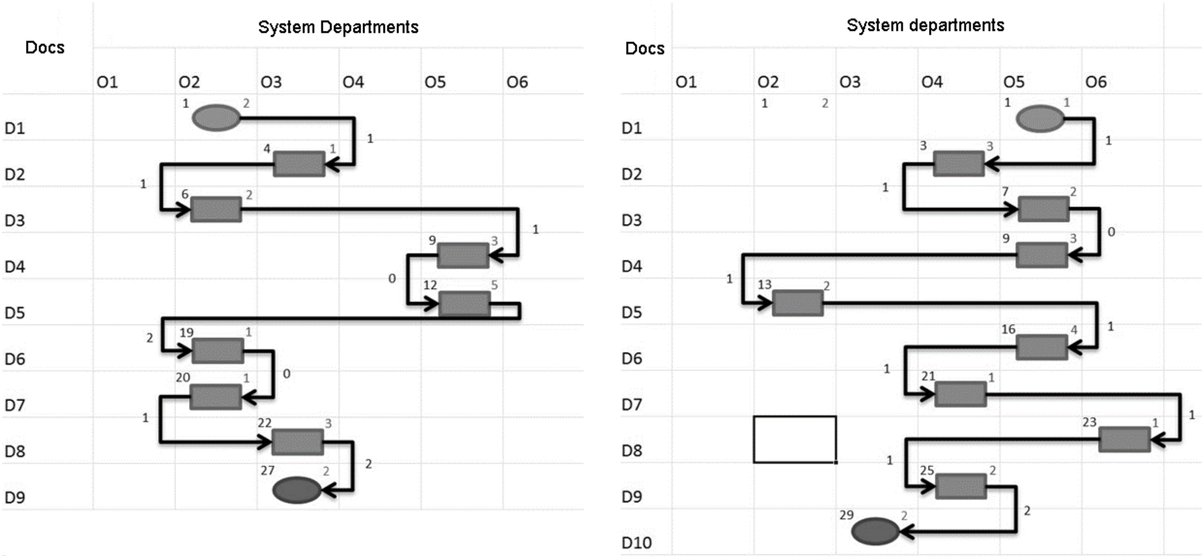
The matrix model allows to visualize the flow of system operations as part of enterprise processes performed in the relevant departments as a function of time. In Figure 2 shown matrix models of processes for: a) marking of parking places and b) replacement of parking meters.

4.2. The model of combined time graph

To estimate the total computational load of the control information system allows the model of the combined time graph of system operations of system processes, which for the example shown in Figure 3. Estimation of the total computational load of the control information system can be performed by a model of the combined time graph of the system operations, which for the example shown in Figure 3.

4.3. Block diagram of the algorithm

Based on the model of the combined time graph, a block diagram of the algorithm for executing the system control program is built, which with the use of object-oriented programming allows for rapid software implementation. In Figure 4 shows a block diagram of the algorithm for performing system operations. To avoid ambiguity in the notation of system operations for each of the two processes, the elements of process 1 are marked as "I", instead of the elements of process 2 - "II".



a) b)

Figure 2: Matrix models of processes: a) markings of parking places, b) replacement of the parking meter

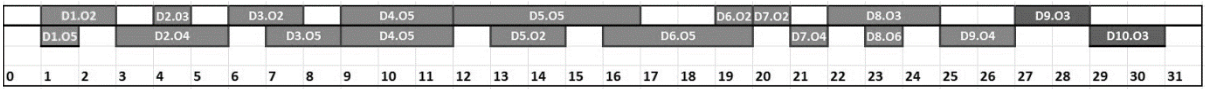


Figure 3: Model of combined time graph of the system operations

5. Processes simulation of the developed system

Simulation analysis for the first process of the developed system was performed in the software environment BP Simulator [15]. Due to the implementation of the project at the Polish company "Parkspace", the simulation models are displayed in the original language. The separate system operations, times of duration of their performance and executive divisions are defined (Figure 5). The simulation allowed to estimate the costs of the system and show the average execution time of each of the processes.

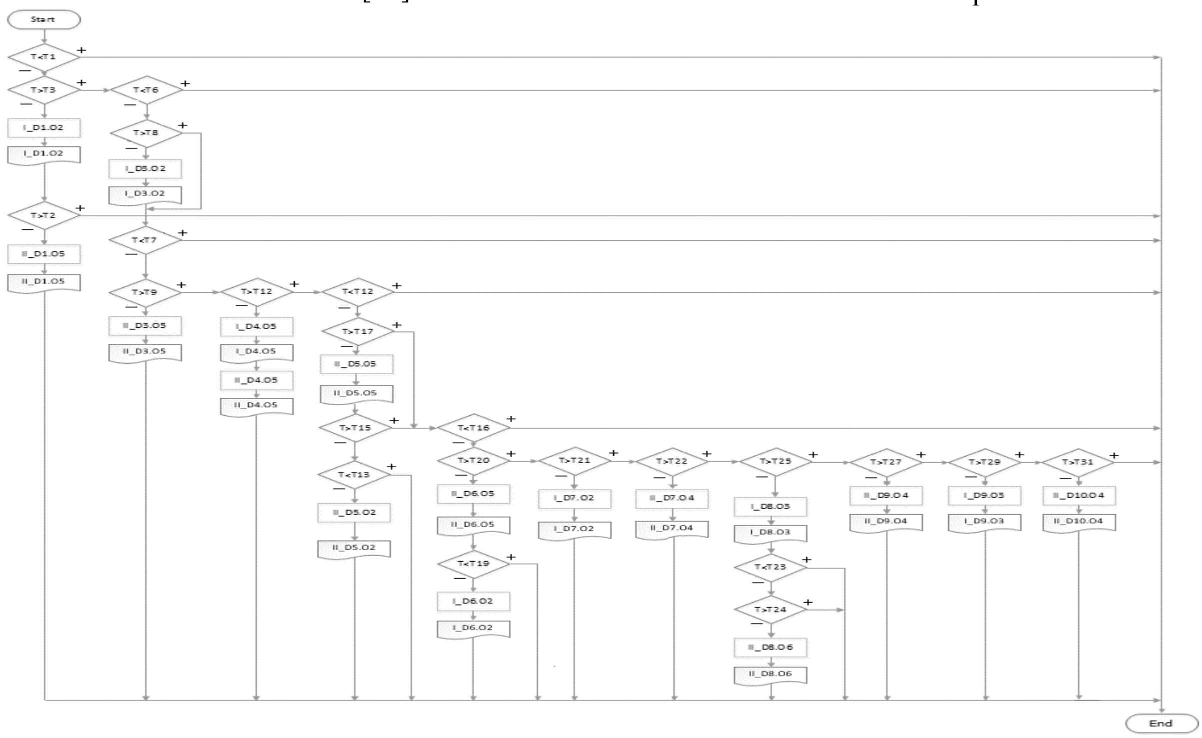


Figure 4: Block diagram of the algorithm for performing system operations

The report of simulation of functioning of the developed system is given in Figure 6.

6. Conclusions

Based on the proposed visualization methods of system operations of complex systems management in the conditions of intersectoral cooperation, the basics of modeling are laid, which allowed to display the status and enable information system management in real time. Visualization of management processes allowed to remove the psychological barrier and avoid mutual misunderstandings between the customer and the developer of information systems, as well as reduce the means of development, implementation and operation of such systems.

The management information system of a company that provides parking services is analyzed. The implementation of the developed system allowed to improve the information exchange between individual units and to move to an electronic system of registration of applications and proposals. Due to the implementation of changes, the company's competitiveness has increased, the localization of funds has been improved, and the functionality of the system has been adapted to market standards.

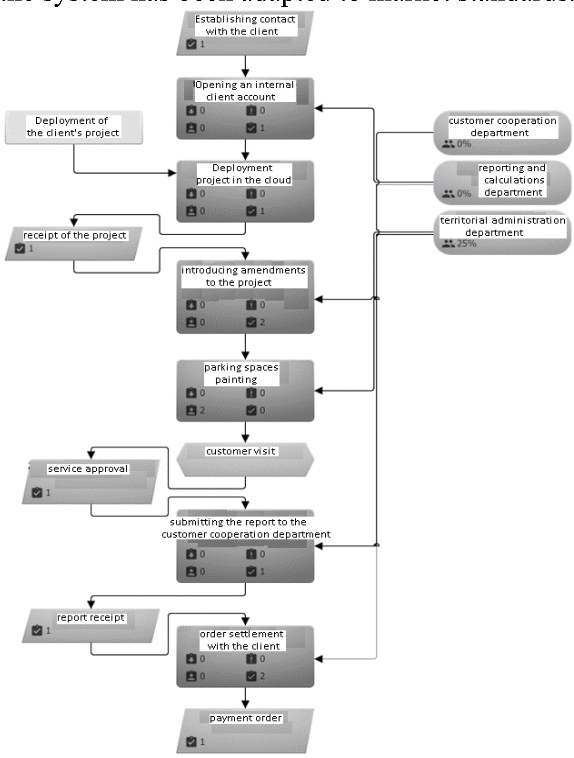


Figure 5: Simulation analysis of the first process of the developed system using the BP Simulator software environment

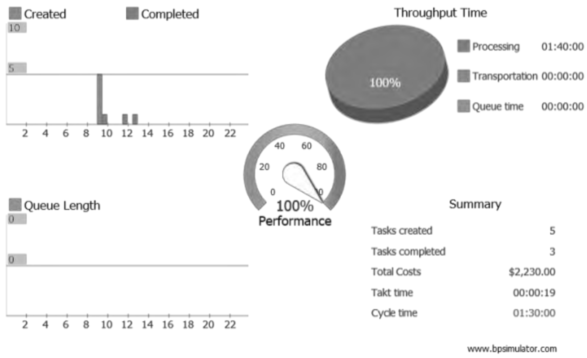


Figure 6: Simulation report of the developed system functioning

7. References

[1] L. Petryshyn, Y. Nykolaychuk, Analiticheskoe modelirovanie informacionnyh sistem avtomatizirovannogo upravlenia — Anatytical modelling of infosystems of automated management / Lyubomyr Petryshyn // W: Zarządzanie organizacjami w gospodarce rynkowej = Organisation management in free market economy / red. nauk. Wiesław Waszkielewicz. — Kraków : AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, 2007. — ISBN 978-83-7464-153-1. — S. 268–275.

[2] L. Petryshyn, M. Petryshyn, W. Cieslik, Processes visualization modelling in distributed management systems, in: Proceedings of the 1st International Conference on Intellectual Systems and Information Technologies, ISIT 2019, Odessa, Ukraine, August 19-24, 2019 / ed. by Nadiia Kazakova, [et al.]. — [Odesa : CEUR], cop. 2019. — (CEUR Workshop Proceedings ; ISSN 1613-0073 ; vol. 2683). — S. [53–57].

[3] L. Petryshyn, M. Petryshyn, A. Choczynska, Management processes informative modeling at the project development stage, in: Proceedings of the 1st International workshop IT Project Management, ITPM 2020 Slavsko, Lviv region, Ukraine, February 18–20, 2020, Vol. 1 / ed. by Sergey Bushuyev [et al.], — Lviv : [CEUR Workshop], cop. 2020. — (CEUR Workshop Proceedings ; ISSN 1613-0073 ; vol. 2565). — S. 254–265.

[4] IEEE 830-1998 Recommended Practice for Software Requirements Specifications.

[5] IEEE 1233-1996 Guide for Developing of System Requirements Specifications.
[6] IEEE 1362-1998 Guide for Information Technology – System Definition – Concept of Operations (ConOps).
[7] 29148-2011 - ISO/IEC/IEEE International Standard - Systems and software engineering - Life cycle processes -Requirements engineering.
[8] ISO/IEC TR 19759:2015 Software Engineering — Guide to the software engineering body of knowledge (SWEBOK).
[9] ISO 15288 System Life Cycle Process.
[10] ISO/IEC 15504 Software Process Improvement and Capability Determination.
[11] IEEE 1498-1995 - EIA/IEEE Interim Standard for Information Technology - Software Life Cycle Processes - Software Development: Acquirer-Supplier Agreement.
[12] IEEE 1042-1987 (Reaff 1993), IEEE Guide to Software Configuration Management.
[13] IEEE 1002-1987 (Reaff 1992), IEEE Standard Taxonomy for Software Engineering Standards (ANSI).
[14] IEEE 1045-1992, IEEE Standard for Software Productivity Metrics (ANSI).
[15] BP Simulator [Online]. 2021. URL: https://www.bpsimulator.com/run/.

Neutron radiography methods for object imaging

S.I. Prokhorets [1], M.A. Khazhmuradov [2]

[1,2] *National Science Center "Kharkiv Institute of Physics and Technology" 1, Akademichna st., Kharkiv, 61108, Ukraine*

Abstract

Different methods of image obtaining by neutron radiography are considered and their advantages and disadvantages are analyzed. Obtaining the neutron-radiography images using the projection transmission method are simulated. It is shown that slow neutrons spectrum from the neutron-radiography facility formation system allow to obtain the fairly clear image of the investigated object. The possibility of using fast neutrons to obtain a neutron image and information about the elemental composition of the object of study is investigated. The simulation of neutron imaging by fast neutron resonance radiography was performed.

Keywords

Neutron, photon, neutron radiography, neutron image, mathematical modeling, Monte Carlo method

1. Introduction

In neutron radiography (NR) method [1] the neutron beam is attenuated after passing through the investigated object and is registered with positional-sensitive detector. Neutron radiography is used for nondestructive control of materials and products using irradiation with collimated neutron beam and registration of flux density distribution. Different materials have different attenuation coefficients, so neutron beam after passing through the object can be interpreted as a signal that contains information about the composition and structure of the object of interest [1,2].

In general, the NR-method has sufficient sensitivity for all materials with which neutrons intensively interact. Especially effective is its use for studying of the equipment, which uses materials that are very different from each other at neutron attenuation coefficients. Specifically, NR-method advantage compared to X-ray and gamma radiography is that neutrons interact intensively with materials composed of elements with low atomic numbers, and easily penetrate materials from heavy elements. It enables contrasting imaging of areas, composed of light elements, including corrosion, fluctuations in density of ceramics, cavities with materials containing hydrogen in metals. The use of electronic imaging techniques in NR-method makes it possible to study the dynamics of fluid flows in engines, refrigerants in heat exchangers, hydraulic fluids, etc. Tracing with substances that absorb neutrons can be used to test the performance of complex mechanisms and devices with liquids and gas flows.

Another important advantage of NR-method is possibility of its using in intense gamma-fields for the studying the radioactive materials.

2. Methods of neutron-radiography images obtaining

In the channel of neutron flux forming and behind the object it is always present gamma-radiation coming from the neutron source and from (n, γ)-reactions in the materials of the object and surrounding structures. Therefore, the position-sensitive detector, the use of which is necessary for the registration of radiation spatial distribution behind the controlled object, have to meet the number of demands. The principal requirement is low efficiency of registration of background radiation. It is also one of the main requirements for nuclear particles registration at the electron accelerator, which in non-destructive control can be used as an environmentally friendly source of neutrons.

EMAIL: sprokhorets@kipt.kharkov.ua (A.1);
khazhm@kipt.kharkov.ua (A.2)

In addition, to reduce the impact of background radiation on the detector different methods of neutron-radiography imaging can be used: direct exposure method, transfer method, transmitted image projection or their combination [1].

When using the transfer method the image of the controlled object is obtained in two steps. In the first step NR-images is registered using intermediate detector with low sensitivity to background radiation. As these detectors, foils and plates made of materials that are activated under the influence of neutrons and maintain this activity for some time are used. Such detector is placed behind the object and expose in neutron field to obtain necessary activity. After exposure it is removed from the radiation field and transferred to the main detector that records the activity of the first detector. The duration of the second step of image obtaining is determined by half-life and activity of nuclei in the material of the intermediate detector and characteristics of the main detector such as efficiency of γ- and- β- radiation registration, spatial resolution and time of exposure.

The method of image transfer has its advantages and disadvantages. Its main advantage over other methods is its low sensitivity to gamma background. However, to obtain high activity of the intermediate detector is possible only in the presence of relatively high neutron density behind the controlled object, so even in modern neutron-radiography facilities with high neutron fluxes (~ 10^7 neutrons/cm^2·s and above) transfer method isn't used because of its complexity [2]. Finally, the transfer method prevents the conversion of neutron radiography facility into facility for neutron tomography (3D image).

When the direct exposure method is used the detector is placed in the radiation field behind the controlled object and is exposed for the time necessary to obtain NR-image. This time is determined by a number of factors, the most significant one are parameters of the detector and neutron flux density at the location of the detector.

It is known that the neutron registration methods are based on "transformation" of the neutron into charged particle followed by the registration of the latter. The first stage of this process is the nuclear reactions. As a result of such reactions charged particles appear almost instantly or during long period after removal of the detector from the neutron field. The first type of such reactions are neutron scattering on protons with recoil nuclei, reactions ^{10}B(n, γ), ^{3}He(n, p), interaction of neutrons with materials that are not activated in the process.

In the photographic methods of registration ionization radiation the X-ray films and photographic plates - glasses with photoemulsion layer are used.

In the process of obtaining NR-image by direct exposure method both neutrons and gamma-rays interact with the film. Based on the film chemical composition radiation interaction occurs due to several physical processes. For slow neutrons such processes are activation of silver and reaction ^{14}N(n,p)^{14}C at nitrogen nuclei in gelatin and for fast neutrons it is scattering at hydrogen nuclei. Background gamma-radiation interacts with the film due to photoelectric absorption, scattering by electrons (Compton-effect) and the formation of electron-positron pairs.

3. Modeling results

In [3,4] it is given the simulation results of neutron yields from targets of different materials – tungsten, tantalum, lead, gold and neutron flux at the output of the formation system of NR-facility. Among these materials tungsten is the most suitable one for the neutron target because of its characteristics. The simulation results of neutron flux at the output of the formation from polyethylene [3,4] are used for further imaging of objects by neutron radiography method.

Figure 1: Assembly used for modeling

Assemblies, made of graphite rods of various lengths and diameters (located in the direction of the beam of neutrons) are used as an objects for modeling. The sample of such assembly is shown at Figure 1

MCNPX software are used for modeling of radiographic image obtaining. This software implements method of transmitted image projection [5]. The scheme of this method is shown at Figure 2. Grid in this method is similar to X-ray film.

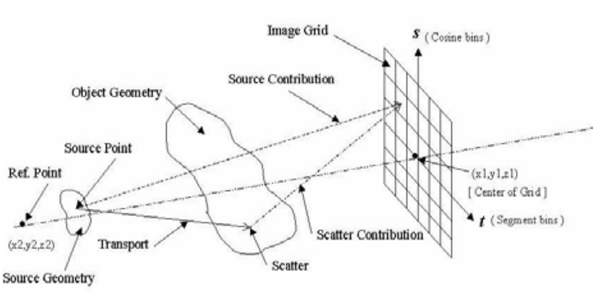


Figure 2: Scheme of transmitted image projection method

Figure 3 shows the results of modeling of assembly made of graphite base and 1 cm thick graphite rods from 0.3 cm to 4 cm in length and from 0.3 cm to 1.2 cmin diameter irradiated with neutrons from the output of the NR-facility formation system made frompolyethylene.

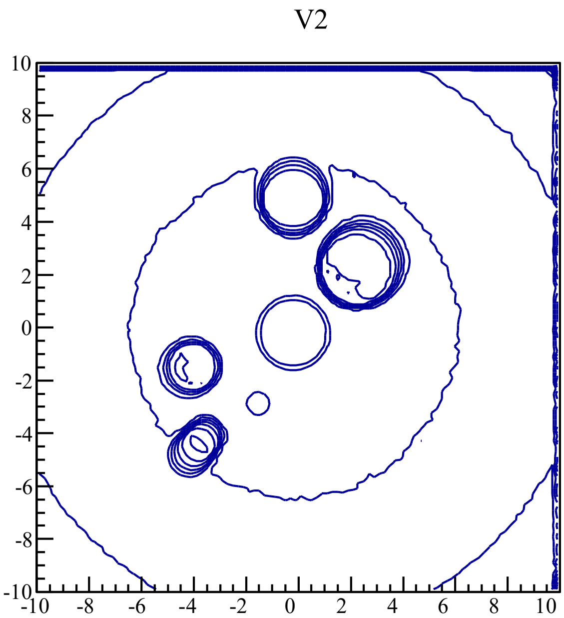


Figure 3: Results of modeling of assembly irradiation with neutrons at the NR-facility output.

The rod with 0.5 cm thickness is seen worst comparing to other rods, that is image contrast is low for small thicknesses, but for other rods fairly clear picture are obtained.

4. Using fast neutrons for neutron radiography

One of the proposed methods to measure the two-dimensional distribution of elements with low atomic number Z in the studied samples is resonance radiography at fast neutrons [6]. The basic idea of fast neutron resonance radiography is uses the variation in cross-section with energy and with isotopic composition. Fast neutrons can pass through massive samples and their interaction cross sections for the various elements are not very different from each other (as compared with cross-sections for X-rays or thermal neutrons). This kind of radiography can be used in airport safety systems and non-destructive control where it is necessary to have high sensitivity to light elements distribution in protective material with a large Z [6]. The method is based on the energy dependences of neutron cross sections for elements that are items of interest.

With a monoenergetic neutron source, it is possible to map one element simultaneously using a resonant peak / valley energy region for one element, while the cross sections for other elements are flat in the same energy range. One radiographic image is obtained at resonance energy, and the other - out of resonance. The difference between the two images gives information about the distribution of the corresponding element. After analyzing a number of radiographic images for different neutron energies it is possible to obtain spatial distribution of the elements of interest either by digital subtraction of images if the distribution of single elements is wanted or by unfolding methods if mapping of the distribution of several elements is of interest [7]. Neutron energies are normally chosen to cover the resonance cross-sections of one or more elements of interest. Fast neutrons fluxes with high intensity can be obtained at the accelerators with high currents using different gas targets.

Due to the high neutron yield scintillators with CCD-readout are normally used for image registration. Low time resolution is not needed because the images for different energies are read one by one. In this case thick targets can be also used to obtain neutron fluxes with almost "white" energy distribution in a range, for example from 1 MeV to 10 MeV.

The results of modeling of image of the sample assembly while irradiation by neutrons with 7

MeV and 4 MeV energies are shown at Figure 4 and Figure5.

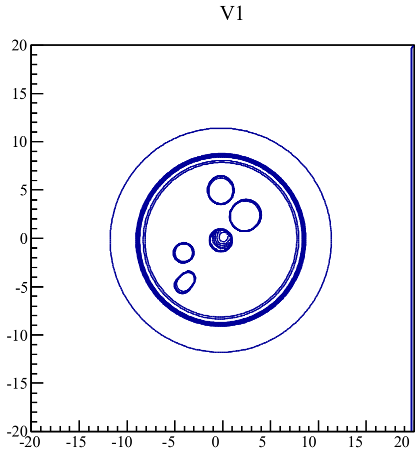


Figure 4: Modeling of image of the sample assembly for neutron energy 7 MeV

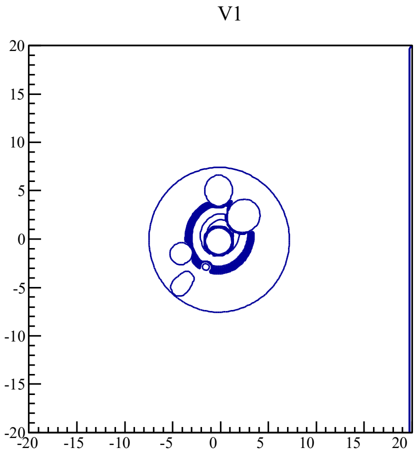


Figure 5: Modeling of image of the sample assembly for neutron energy 4 MeV

Obviously, neutron irradiation with energy 4 MeV of graphite sample assembly makes it possible to obtain an image of sufficient contrast for further recognition of different rods, including and those with small thickness.

Spheres with a diameter of 5 cm were considered for modeling. The composition of objects was considered as follows: hydrogen (H) – 1.0 g, carbon (C) – 7.1 g, nitrogen (N) – 7.1 g, oxygen (O) – 12.3 g; This can be considered as a simulation model of an explosive. Input neutron energies for modeling were 2.37 MeV, 5.36 MeV (carbon resonance energy), 2.23 MeV (nitrogen resonance energy), 3.43 MeV (oxygen resonance energy) and 0.01 MeV (oxygen valley energy O).

At Figure 6(a,b,c) the results of modeling of the NR- images obtaining for location of a sphere with a diameter of 5 cm in the soil at a depth of 1 cm are shown

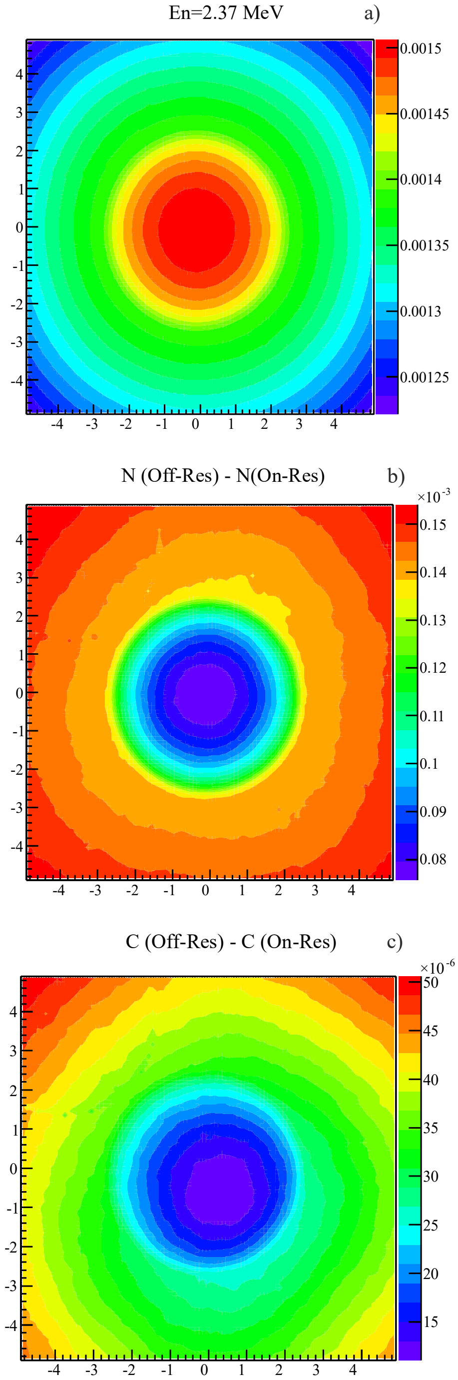


Figure 6: The results of modeling of the NR-images obtaining

Analysis of the images and their differences obtained in the simulation shows that the fast neutron resonance radiography method allows us to obtain information about the relative distribution of the elements of interest. For example, looking at the results in the soil, it can be concluded that nitrogen amount is about the same as carbon.

5. Conclusions

Different methods of NR-images obtaining are considered and their advantages and disadvantages are analyzed. Obtaining of neutron-radiography images using projection transfer method is simulated. It is shown that slow neutron spectrum obtained previously at the output of the formation system of the neutron-radiography facility allow a fairly clear images for the considered graphite sample. In addition, positive results are obtained for the possibility of using of fast neutrons for neutron-radiography imaging.

6. References

[1] N.D. Tyufyakov, A.S. Shtan, Fundamentals of neutron radiography, Atomizdat, Moscow, 1975 (in Russian).

[2] Yu.N. Burmistrenko, Photonuclear analysis of the composition of matter, Energoatomizdat, Moscow, 1986 (in Russian)

[3] S.I. Prokhorets, M.A. Khazhmuradov, Optimization of the setup parameters for neutron radiography, Automated Control Systems and Automation Instruments 126(2004) 4-9 (in Russian)

[4] I.M. Prokhorets, S.I. Prokhorets, M.A. Khazhmuradov, Formulation of criterion functional and set of constraints in the problems of physical settings designing, Problems of atomic science and technology 5(2004) 108-111.

[5] MCNP 2.4.0. RSICC computer code collection. Monte-Carlo N-Particle Transport Code System for multiparticle and high energy applications. CCC-715, 2002

[6] V. Dangendorf, G. Laczko, C. Kersten et al., Fast Neutron Resonance radiography in a pulsed neutron beam, in: Neutron Radiography. Proceedings of the Seventh World Conference, Rome, Italy, 2002, pp. 383-398.

[7] R.C. Lanza, G. Chen, "Fast Neutron Resonance Radiography for Elemental Imaging: Theory and Applications in: Neutron Radiography. Proceedings of the Seventh World Conference, Rome, Italy, 2002, pp. 383-398

Method of “atomic” spectra for speaker identification

Victor Solovyov [1], Oleg Rybalskiy [2], Vadim Zhuravel [3], Aleksandr Shablya [4], Yevhen Tymko [5]

[1] *“Silentium Systems”, Hornby Str 950 – 777, BC V6Z 1S4, Vancouver, Canada,*
[2] *National Academy of Internal Affairs, Solom’yanska Are. 1, Kyiv, 03035, Ukraine*
[3] *Kyiv scientifically-research expertly-criminalistics center Ministries of Internal Affairs of Ukraine, Vladimirska Str. 15, Kyiv, 01001, Ukraine*
[4] *Odessa Research Institute of Forensic Expertise of the Ministry of Justice of Ukraine, Uspenska Str. 83/85, Odesa, 65011, Ukraine*
[5] *Kyiv Scientific Research Institute of Forensic Expertise of the Ministry of Justice of Ukraine, Smolenska Str. 6, Kyiv, Ukraine*

Abstract
A method of accounting for a variety of factors affecting the parameters of the characteristics of a speaker's voice is considering, which provides the fundamental possibility of indirectly accounting for their practically unlimited number. The method is basing on the extraction of “atomic” structures from speech signals, which depend on the totality of the main factors influencing the speaker identification process. With this method, all significant factors affecting the characteristics of the voice will be indirectly taken into account at the level of these structures “Atomic” speech structures are understood as the spectra of any fragments of any vowel sounds allocated in a time window of 20 ms. “Atomic” structures are selecting automatically. Subsequent decisions made on the combinatorial set of a huge number of these “atomic” structures. The method provides a rational accounting for the multifactorial influence of various parameters. The decision on the identity of the voices of the announcers recorded on different phonograms is carried out on the basis combinatorics of “atomic” spectra of vowel sounds in both phonograms. The method has shown high efficiency in the examination of phonograms of short duration.

Keywords
probability, time window, vowel sound, speaker, identification, error curves, spectrum, intersection point, phonogram, expertise, efficiency.

1. Introduction

In the last decade, speaker identification systems have shown very high efficiency in test trials. When testing the most advanced systems on specialized databases, their minimum efficiency is only a few percent [1–6]. Note that the efficiency of such systems estimate by the value of the error probability at the point of intersection of the curves of errors of the first and second kind. As applied to forensic terminology, an error of the first type indicates the probability that the objects under investigation are incorrectly identified as different (probability of a false rejection), and an error of the second type – the likelihood that different objects are incorrectly identified as one (probability of a false omission). At the point of intersection of the curves, these probabilities are equal to each other (ER) [7].

However, it is known that testing the same speaker identification systems on different test data-bases often gives very different results [4–6]. This is probably due to the fact that many specific test databases are dominated by different factors, each of which has its own, different from the others, influence on the results of speaker identification by voice characteristics. This explains the fact that many practicing experts are very skeptical about the possibility of automatic speaker identification [4–6].

Many factors are known that affect the variability of the characteristics of a speaker's

EMAIL: edemsvi@gmail.com (A. 1); rov_1946@ukr.net (A. 2); fonoscopia@ukr.net (A. 3); alik.shablya@gmail.com (A. 4); e.tymko@kndise.gov.ua (A. 5)
ORCID: 0000-0003-3541-4740 (A. 1); 0000-0002-1440-8344 (A. 2); 0000-0002-2777-2102 (A. 4)

voice [1–6]. These are, for example, the language of the phonogram and dialects, the context of speech, the emotional state of the speaker, the duration of the phonogram, and many others. Under these conditions, the creation of test databases for testing and evaluating the effectiveness of speaker identification systems, which would cover all the main factors, is highly doubtful. The point is not that the size of such bases will be huge – at present this issue does not pose any problems. The difficulty lies in the need to quantitatively formalize a number of specific factors that affect the characteristics of the speaker's voice, which were mentioned above (for example, the context of speech, dialect, emotional state of the speaker, and others). Such a large set of parameters cannot but influence the position of the point of intersection of the error curves. Taking into account the influence of many different factors on the characteristics of the voice leads us from plotting the errors of the first and second kind on the plane, to plotting them as surfaces in the N-dimensional space. This statement of the problem may seem purely mathematical. But it will be shown below that it is this factor that is decisive in the discrepancy between the effectiveness of the speaker's identification systems and the needs of the examination practice.

Consider the projection of the N-dimensional space of surfaces of errors of the first and second kind onto the three-dimensional space, presented on fig. 1. Further, all graphs and examples implemented on the modules of the system of identification and verification of speakers “Avatar” [8–11].

This surface is built on the basis of a large database of speakers. The language of the phonograms is Ukrainian. On the OX axis − the parameter of the proximity of the characteristics of the voice in the “Avatar” system − SabsAI. This is the sum of the absolute values of the absolute difference of the total spectra for sounds [A] and [I]. On the OY axis − the duration of the phonograms T in seconds. On the OZ axis − the probabilities of errors of the first and second kind. Naturally, the points of intersection of the error surfaces lie on the spatial line and depend on the two adopted parameters. The important point in this illustration is the values of the probabilities at the intersection points. These values are much less than for 2D error plots. Which is natural, in view of taking into account two factors. In the multidimensional space of factors (when constructing and calculating multidimensional error surfaces), these values will decrease with an increase in the number of factors taken into account.

To illustrate this situation, consider the same graph, but normalized to the error probabilities for only one coordinate (OX) in Fig. 2.

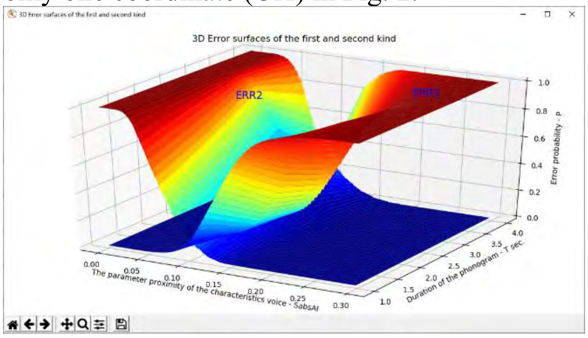


Figure 1: *3D* surface of errors of the first and second kind

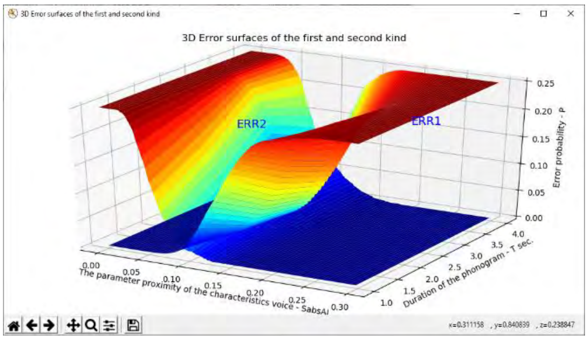


Figure 2: *3D* surface of errors of the first and second kind

On this graph, the probabilities at the points of intersection of three-dimensional surfaces are several times higher in terms of errors. For a more complete illustration, fig. 3 shows a two-dimensional cut of the surfaces of errors of the first and second kind, which corresponds to the usual classical representation of the error graphs.

Such three-dimensional surfaces can be built for almost any pair of parameters of voice characteristics (with the availability of appropriate software tools).

The presence of a large number of factors in the test database that directly or indirectly affect the characteristics of the voice increases the accuracy of assessing the effectiveness of systems with an increase in the amount of data. If a particular identification system also takes these factors into account, then the likelihood of errors at the point of intersection of the error curves is reduced, and the efficiency of the system in test tests increases. This applies to any factors, except for one − the duration of the phonograms. But it is, as a rule, one of the most essential in the examination of speech phonograms, since in fact it sets a certain limit of efficiency for any system.

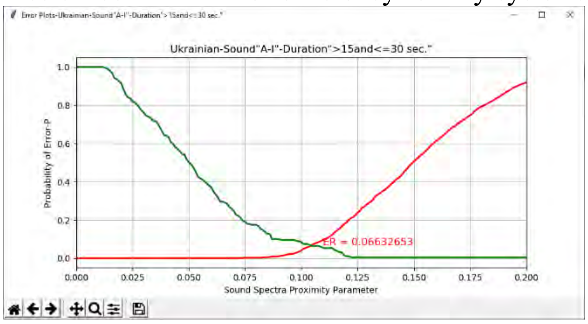


Figure 3: Two-dimensional plots of errors of the first and second kind

Thus, the fundamental issue with the existing methods of development and testing of speaker identification systems is, in general, an obvious position − the need for rational consideration of many significant factors that affect both the effectiveness of systems and the results of their testing. At the same time, the testing process, due to the practical impossibility of taking into account many real factors, in our opinion, does not sufficiently reflect the effectiveness of any speaker identification system.

The purpose of the work is to show the research method implemented in the developed systems, as much as possible, from our point of view, generalizing the possibilities of taking into account many factors that affect the parameters of the characteristics of the speaker's voice. This method creates the fundamental possibility of indirectly accounting for an almost unlimited number of factors.

2. Method of “atomic” spectra in problems of speaker identification

The classic research and development method for speaker identification systems usually involves the following sequence. Based on the well-known concepts of the factors and parameters affecting the characteristics of the speaker's voice, these parameters are distinguished from a large statistical material and quantitatively described. Next, the quantitative relationships between the characteristics of the voice and the selected parameters are determined. This method practically does not depend on either the research methodology or the technologies for constructing systems. The same scheme is using in developments based on neural network technologies. With this method, researchers and developers, on the basis of known ideas and their intuition, identify the main parameters that significantly affect the characteristics of the speaker's voice. Applying neural networks essentially doesn't change anything with this method. If a neural network is capable of identifying some new patterns, then this is carried out within a certain list of parameters that are fed to the input of the neural network for training. The main drawback of this method is the need to choose from subjective prerequisites a certain list of parameters for development and research. This set, ultimately, determines the effectiveness of the system being developed. Attempts to use the entire completeness of information contained, for example, in a sound wave with speech fragments, have not been known to us until now. And, despite the tendencies of using the entire set of initial data at the input of neural networks, there have been no examples of effective solutions obtained for this technological direction until now [4–6].

To resolve this contradiction, consider the following method, which we call the method of “atomic” structures. According to this method, structures are distinguishing from speech signals, which depend on the entire set of basic factors that affect the process of speaker identification. With this method, all significant factors affecting the characteristics of the voice will be indirectly taken into account at the level of these structures. Subsequent identification decisions based on the entire phonogram will be made on the basis of the combinatorial set of a huge number of these “atomic” structures.

By “atomic” structures of speech, we mean the spectra of any fragments of any vowel sounds allocated in a time window of 20 ms.

For the transformation, the Morlet wavelet is using as a basis

$$C_{mor}(t) = \frac{e^{j2\pi F_C(t)} e^{\frac{-t^2}{F_b}}}{\sqrt{\pi F_b}}, \qquad (1)$$

where t – is time; F_b – parameter of the wavelet width, F_C – is the center frequency (heterodyning frequency) of the wavelet when scanning a signal in a 20 ms window [12].

Based on the Morlet wavelet, we implement the spectral transformation

$$Y_{FC} = \sum_{i=1}^{N} A_i(t_i) \times C_{mor}(t_i, F_C), \qquad (2)$$

$$i = 1,2,\ldots,N$$

$$S_{FC} = \sqrt{\left(\left|Y_{FC}\right|\right)\frac{\left|Y_{FC}\right|}{N}}\ , \qquad (3)$$

where $A_i(t_i)$ – discrete samples of the audio signal in the time window of 20 ms, Y_{FC} – the result of complex transformation of the signal into the frequency domain, F_C – discrete frequency values with frequency scanning interval D_{FC} = 1 Hz, S_{FC} – normalized levels of spectral components, N – the number of averaging for each sample, t_i – discrete i-th time sample.

Consider redundant wavelet transforms with the number of samples in the time domain less than the number of samples in the frequency domain. Let us take, for example, an arbitrary 20 ms fragment of the speech signal of sound "A", with a sampling rate of 44100 Hz. Then the number of discrete samples falling on a segment with a duration of 20 ms will be N = 882. Then, for the applied non-orthogonal wavelet transform based on the Morlet wavelet with a scanning step in the frequency domain D_{FC} = 1 Hz, the maximum possible number of scanning steps in the frequency domain will be 22050 samples [12].

Equation (2) is formally a covariance function that connects the sample values of discrete amplitudes of an audio signal with harmonic functions of frequency within a time window. Equation (3) is considering as a normal spectrum with a frequency step of 1 Hz. At points F_C that are multiples of 50 Hz, the results of this transformation coincide with the results of the fast Fourier transform (FFT) (up to the Gaussian function used to smooth out the effect of the small size of the time window on the spectrum).

A feature of the presented method is the high accuracy of determining the position of local frequency maxima in the signal spectrum in a small time window, which is approximately 1 Hz. But it is precisely this accuracy that determines the accuracy of estimating the frequency of the fundamental tone and practically all other possible spectrum functions.

In classical spectral analysis, such a large number of spectrum frequency gradations is redundant. But this redundancy can significantly improve the accuracy of estimates for any research and development methodology. And, in particular, for solving problems based on deep learning neural networks [8–11].

Next, we consider the frequency spectra of vowel sounds at a time interval of 20 ms. It is known that the frequency spectra of almost any specific vowel sound, which is pronounced by one speaker at different times, differ from each other, both in the shape of the spectrum envelope and in the positions of the maxima at different time intervals (dynamics of the sound spectrum during pronunciation) [8 – 11]. Consider a model that characterizes the dynamics of the spectrum of one spoken vowel sound.

Let's select an arbitrary vowel sound from the soundtrack of a specific speaker. In the studies, this selection was made by an automatic software module in a special sound editor (this module is built into the Avatar system). This editor contains a special software module based on deep learning neural networks and provides automatic highlighting of vowels, regardless of language, speech context and speaker. In the international transcription standard, these are six vowels: [A], [E], [I], [I:], [O], [U]. This selection made from an average set of vowel sounds common to different language groups. It not base on a specific linguistic description of the phonemes included in vowel sounds and their combinations for specific language groups. Basically, these are the vowel sounds of Indo-European languages.

Based on the described method, a training base (Dataset) of a neural network was formed to identify the speaker at the level of "atomic" frequency spectra of fragments of vowel sounds with a duration of 20 ms. This database contained millions of fragments of frequency spectra for different speakers. Including hundreds of thousands of fragments of speech by the same speaker.

Note that the "Avatar" system uses the module of the phonemic machine. In a system with such a module, a sound identification technology is using, which differs in the interpretation of sounds from classical representations. In particular, signals not distinguished as sounds, physiologically perceived by the organs of hearing. This increases the efficiency of the speaker identification system. In the developed phonemic model machine, speech sounds represented as a certain set of several "atomic" components for each sound. The set of "atomic" constituents for each sound can partially overlap with the set of "atomic" constituents for other sounds. So, for example, many components of the frequency spectrum of sound structures [A] intersect with many structures of the frequency spectrum of sound [O] (there may be other intersections). Likewise, sound [I] with sound [I:]. When presented as averaged frequency spectra, these structures can be interpreted as the averaged spectra of the corresponding sounds uttered by a specific speaker. But these averaged spectra of “atomic” sounds do not fully correspond to physiologically perceived sound signals of speech. Thus, in the applied model of the phonemic machine, the structural components of speech sounds are distinguished.

During training, based on the proximity of the “atomic” spectra of fragments of vowel sounds, the network solved the problem of binary classification of the speaker (“It is he – It's not him”).

An important factor in the successful training of a neural network for binary identification of speakers (from the point of view of the similarity of the characteristics of their voices) is the multiplicity of models. Therefore, models developed for identifying a speaker by the proximity of spectral characteristics for each of the six vowel sounds, and combinations of these vowel sounds. In particular, the following combinations were used: [A] [E], [A] [I], [A] [I:], [A] [O], [A] [U], [E] [I] , [E] [I:], [E] [O], [E] [U], [I] [I:], [I] [O], [I] [U], [I:] [O], [I:] [U], [O] [U]. For each of the combinations of “atomic” spectra, separate models obtained in the learning process. The input of the neural network during training was the “atomic” spectra of vowel sounds and their combinations. The way out is the probability that two “atomic” spectra belong to one speaker.

An illustration of training a neural network for binary classification of speakers by a combination of sounds [A] [E] is shown in fig. 4.

The efficiency of speaker identification at the level of “atomic” spectral structures is relatively low. However, in this method, the speaker is identified by the totality of “atomic” structures of the spectra in two phonograms by a huge number of combinations of combinations of “atomic” spectra of vowel sounds. So, for example, on average, for a phonogram with a total duration of 10 s, the duration of vowel sounds is at least (3–4) s (taking into account pauses between words and consonant sounds in speech). Consequently, from the vowel sounds of a phonogram of such duration, it is possible to obtain about 200 intervals of 20 ms with “atomic” spectra. A significant factor in the number of “atomic” structures in the present study is the scanning of vowel sounds with a window of 20 ms. Usually, the duration of a vowel sound at an average speech rate is (30–60) ms [13, 14].

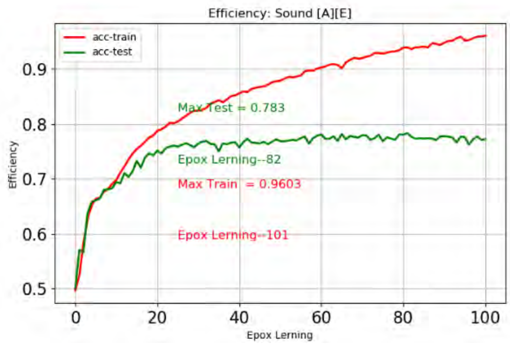


Figure 4: Illustration of the effectiveness of training a neural network for “atomic” structures

As the analysis of the dynamics of changes in the spectra of one vowel sound shows, the “atomic” spectra of sound in the interval of 20 ms are very variable for one sound, both in the position of the maximum and in the frequency structure of the entire spectrum. Under these conditions, in research and development, the vowel sound scanned with a 20 ms window with a scanning interval of no more than 5 ms.

Often for a phonogram with a duration of 10 s the number of different “atomic” spectra can reach 1000. And the number of different combinatorial combinations of “atomic” spectra in two phonograms with a duration of 10 s can reach the value N = 1,000,000. This is a huge statistics on different values of the probability P of speaker identification at the level of “atomic” spectra of vowel sounds (the output of a neural network in binary classification is the probability P “It is he – It's not him” of speaker identification). Further, decision-making on the entire set of combinations of “atomic” spectra of vowel sounds is carried out on the basis of the classical concepts of the theory of probability and mathematical statistics. The calculation of the probability distribution P over the entire set of values is carried out. The mathematical expectation is determined – P_{sr}, and the variance of the distribution – D. If $P_{s\,r} > 0.5$, the decision is made – “It is he”, if $P_{sr} < 0.5$ – ”It's not him”. The decision error is determining by the accuracy of calculating P_{sr} and is a function of the number of combinatorial combinations N and the type of the distribution function (for a Gaussian distribution, N and D are sufficient).

When developing a speaker identification system, the proposed method, based on the “atomic” frequency spectra of vowel sounds, does not require the selection of individual specific factors and parameters of voice characteristics. This is explaining by the fact that almost all factors that directly or indirectly affect the parameters of the speaker's voice characteristics are taken into account in the frequency spectra of “atomic” structures.

The experimental results show a very high efficiency of this method for analyzing phonograms of short duration, in particular, for phonograms with a speech duration of less than 1 s (if they contain vowel sounds). In fig. 5 shows an illustration of the possibility of speaker identification by a system based on the proposed method using phonograms less than 1 s long.

In fig. 5 shows an illustration based on the method of “atomic” structures of speaker identification by two sounds [A] from two different phonograms for the same speaker. The duration of the sounds is 63 and 81 ms.

From the point of view of the examination of speech phonograms, such identification (based on such short fragments) is “exotic”. But, within the framework of the stated methodology, this is a real opportunity, implemented in the current identification system.

At the same time, as the results of various tests show, even for single sounds isolated from various phonograms, it is not at all necessary that the decision error be very high. So, for example, in fig. 6 shows an illustration for two sounds [A] of the same speaker from recordings on different phonograms. But, the sounds are cut from the speech of the same speaker with a very close context. Therefore, the decision error, taking into account the proximity of many essential factors that determine the characteristics of the voice, can be very small.

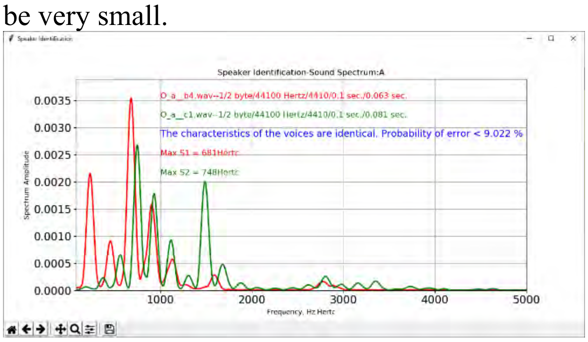


Figure 5: Illustration of the possibility of speaker identification by “atomic” spectra extracted from two short phonograms

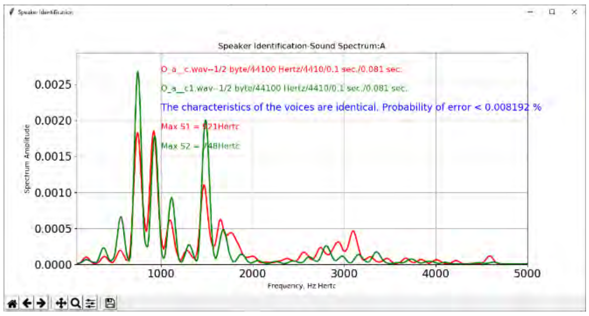


Figure 6: Illustration of spectra for speaker identification by the method of “atomic” structures, isolated from different phonograms, with a low probability of error

It is necessary to note a very important aspect of the research and development carried out. An increase in the number of speakers, speech variability, and the number of language groups for the Dataset database, starting from a certain moment, practically does not affect the “atomic” models (both for training arrays and for testing arrays). And, apparently, at this stage of research and development, it can be assumed that the presented method has a very high degree of generalization to various language groups. It practically does not depend on the variety of various factors and parameters of speech.

3. Conclusions

The proposed speaker identification method provides a rational consideration of many different factors affecting the characteristics of the voice. The method is based on the extraction of vowel sounds from voice signals of “atomic” structures in a time window of 20 ms. The spectra of these structures are influenced by all the main factors that characterize the individuality of the voice of a particular speaker. The decision on the identity of the voices of the announcers recorded on different phonograms is carried out on the basis of the “atomic” frequency spectra of combinations of vowel sounds in both phonograms. The method has shown high efficiency in the examination of phonograms of short duration.

4. References

[1] Lei, Y., Scheffer, N., Ferrer, L., and McLaren, M. (2014). A novel scheme for speaker recognition using a phonetically-aware deep neural network. In IEEE

International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.1695–1699.

[2] Kenny, P., Gupta, V., Stafylakis, T., Ouellet, P., and Alam, J. (2014). Deep neural networks for extracting baum-welch statistics for speaker recognition. Odyssey 2014 – The Speaker and Language Recognition Workshop.

[3] Kassin, S. M., Dror, I. E., and Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. Journal of Applied Research in Memory and Cognition, 2(1), 42–52

[4] Satyanand Sing. Forensic and Automatic Speaker Recognition System. International Journal of Electrical and Computer Engineering (IJECE) Vol. 8, No. 5, October 2018, pp. 2804–2811

[5] Amali Mary Bastina, N. Rama, "Biometric Identification and Authentication Providence using Fingerprint for Cloud Data Access," International Journal of Electrical and Computer Engineering , vol. 7(1), 2017, pp. 408-416.

[6] John H.L. Hansen and Taufiq Hasan, "Speaker Recognition by Machines and Humans." IEEE Signal Process. Mag., 2015, pp. 74–99.

[7] Rybalskiy O.V., Solovyov V.I., Cherniavskyi S.S., Zhuravel V.V., Zheleznyak V.K. A probabil-istic approach to making expert decisions on the analysis of complex objects. Bulletin of the National Academy of Sciences of Belarus. A Series of Physical and Technical Sciences. 2019, Vol. 64, No. 3, pp. 346–352. https://doi.org//10. 29235 / 15-8358-2019-64-3-346-352.

[8] Solovyov V.I., Rybalskiy O.V., Zhuravel V.V., Semenova N.V. Analyzing the Models of Speech Recognition on the Basis of Neural Networks of Deep Learning for Examination of Digital Phonograms. Cybernetics and Systems Analysis, 2021, Vol. 57, No. 1, pp. 133–138.https://doi.org/10.1007/s10559-021-00336-y

[9] Solovyov V. I.,Rybalskiy O. V., Zhuravel V. V., ZheleznyakV. K. Application of neuron networks of deep learning for exposures editing of digital phonograms. Proceedings of the National Academy of Sciences of Belarus. Physical-technical series, 2020, vol. 65, no. 3, pp. 383–389. https://doi.org/10.29235/1561-8358-2020-65-3-383-389

[10] Solovyov V.I., Rybalskiy O.V., Zhuravel V.V.Method of exposure of signs of the digital editing in phonograms with the use of neuron networks of the deep learning. Journal of Automation and Information Sciences, 2020, V..52, No. 1, pp. 22 – 28. https://doi.org//10.1615/JAutomatInfScien.v52.i1.30

[11] Solovyov V.I., Rybalskiy O.V., Zhuravel V.V.Verification of fundamental fitness of neuron networks of the deep educating for the construction of the system of exposure of editing of digital phonograms. Cybernetics and Systems Analysis, Vol. 56, No. 2, March, 2020, pp. 326–330. https://doi.org//10.1007/s10559-020-00249-2

[12] Mallat S. A wevlet tour of signal processing. AcademicPress, New York. – 1999, 671 p.

[13] J.L. Flanagan. Speech Analysis Synthesis and Perception. Springer, 1972, 397 p.

[14] G. Fant. Acoustic theory of speechproduction. The Hague, The Netherlands, Mouton, 1960, 284 p.

Evaluation the Quality of Measuring the Coordinates of Air Objects in the Synchronous Information Network of Surveillance Systems

Valerii Semenets [1], Iryna Svyd [2], Ivan Obod [3], Oleksandr Maltsev [4], Oleksandr Vorgul [5], Mariya Tkach [6] and Valeriia Chumak [7]

[1,2,3,4,5,6,7] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract

The paper shows that the hereditary transition from asynchronous to synchronous networks of airspace surveillance radar systems in which time is one of the coordinates of the detected air objects, allows to calculate the height of air objects due to long-range measurements with the desired quality.

Keywords

Synchronous information network, information network, surveillance system, radar system, airspace surveillance, accuracy, air object, friend or foe, height measurement, reception center

1. Introduction

In the leading countries of the world, for quite a long time, there have been unified national monitoring systems of the use airspace by both military and civil aviation [1, 2]. Obviously, this achieves the maximum efficiency of the use of airspace with relatively low material, technical and human costs [3].

One of the components of the airspace control system is a unified information network based on existing airspace surveillance systems. Significant attention is currently paid to the network construction of information resources [4-11].

In particular, the existing national unified airspace control systems, as a rule, are implemented on the basis of the network principle using separate information means [3].

The main tasks of these programs are to integrate existing radar surveillance systems of various departments into a common network and centralized management of this network by a higher authority. In this case, the information network of observation systems, as a rule, is implemented on an asynchronous principle [11-13] and a three-dimensional coordinate system is used [14-15].

The combined information of the considered information network of airspace observation systems is issued to consumers. However, this principle of organizing an information network impoverishes the information support of consumers and does not solve the problems of individual information facilities, in particular, secondary radar systems [16-22], as well as the joint functioning of primary and secondary radar systems [23-30], etc.

The transition to the synchronous principle of building an information network, in which a four-dimensional coordinate system is used, as shown in [31-36], will provide complete and reliable information support for consumers, as well as solve the problems of the functioning of individual information tools, in particular, systems for identifying air objects based on "friend or foe" [9, 22-25].

In addition, the use of a four-dimensional coordinate system in a synchronous information network expands the functionality of such a network, in particular, it becomes possible to implement cooperative reception of signals, measure the height of an aircraft by rangefinder measurements, etc.

The aim of the work is to assess the quality of measuring the height of an airborne object by rangefinder measurements in a synchronous

EMAIL: valery.semenets@nure.ua (A. 1); iryna.svyd@nure.ua (A. 2); ivan.obod@nure.ua (A. 3); aleksandr.maltsev@nure.ua (A. 4); oleksandr.vorgul@nure.ua (A. 5); mariia.zavorotna@nure.ua (A. 6); valeriia.chumak@nure.ua (A. 7)
ORCID: 0000-0001-8969-2143 (A. 1); 0000-0002-4635-6542 (A. 2); 0000-0002-9898-0937 (A. 3); 0000-0003-1520-9280 (A. 4); 0000-0002-7659-8796 (A. 5); 0000-0002-4248-7633 (A. 6); 0000-0002-2403-020X (A. 7)

information network of radar surveillance systems.

2. Assessment of the accuracy measuring the height of an airborne object in synchronous airspace observation systems

As is known from [8, 9], using measurements of the slant range to an airborne object at several points of the synchronous information network, it is possible to measure the flight altitude of the observed airborne object. It should be noted that such a task is faced by both primary and secondary radars. Indeed, as shown in [7], the synchronous information network of secondary location systems presupposes a transition to requestless systems and, therefore, requires measuring the height of the emitting object. In this regard, we will solve the problem of joint measurement of the height of an air object by systems of both primary and secondary radar.

Consider a synchronous information network that consists of n receiving points. For the primary system, this will be n ground receiving points of echo signals, one of which is emitting, and for the secondary system, these are n ground receiving points of response signals. Thus, the tasks of measuring the height of an airborne object by echo signals and response signals are identical. The accuracy of measuring the height of an airborne object depends on the accuracy of measuring the time of signal reception, the accuracy of synchronization of the time scales of the receiving points, as well as on the relative position of the receiving points and the airborne object, that is, the geometric factor.

Let's assume that an emission or an echo or a response signal occurs from an airborne object at time $T_k(t)$. Let's also assume that there are four ground receiving points.

Consequently, at each of the receiving points at the time $T_i(t)(i = 0, \ldots, 3)$ the signal emitted (reflected) by the airborne object is received. Considering the time scales formed at the receiving points of the synchronous information network to be highly stable, we can omit the dependence of time processes on t.

Thus, the time of arrival of the signal of the airborne object at each of the receiving points of the synchronous information network can be written as

$$T_i = T_k + r_i / c, \tag{1}$$

where c is the speed of light.

Subtracting the time of arrival at the base processing point (we consider it to be zero) from the time of the remaining receiving points, we obtain

$$R_i - R_0 = c(T_i - T_0) = r_i, \ i = 1,2,3. \tag{2}$$

Based on the geometry of the location of the receiving and emitting points, it is possible to write

$$R_0^2 = x^2 + y^2 + z^2, \tag{3}$$

$$R_i^2 = (x - x_i)^2 + (y - y_i)^2 + (z - z)^2. \tag{4}$$

From (4) it is easy to obtain

$$\begin{aligned} R_i^2 - R_0^2 = \\ = x_i^2 + y_i^2 + z_i^2 - 2(x_i x + y_i y + z_i z) \end{aligned} \tag{5}$$

Using expressions (5) and (4), we can write

$$\begin{aligned} R_i^2 - R_0^2 = (R_i - R_0)(R_i + R_0) = \\ = (r_i + 2R_0) r_i \end{aligned} \tag{6}$$

Substituting (6) into (5) and performing a permutation, we obtain

$$\begin{aligned} 2(x_i x + y_i y + z_i z + r_i R_0) = \\ = x_i^2 + y_i^2 + z_i^2 - r_i^2 \end{aligned} \tag{7}$$

Based on the set task, we need to evaluate the effect of errors in synchronization of time scales of reception points, i.e. T_i, for the quality of the height measurement, i.e. to the z coordinate. Differentiation of expression (7) allows writing

$$\begin{aligned} 2\left(x_i \frac{dx}{dT_j} + y_i \frac{dy}{dT_j} + z_i \frac{dz}{dT_j} + x_i \frac{dR_0}{dT_j} + \right. \\ \left. + R_0 \frac{dr_i}{dT_j}\right) = -2r_i \frac{dr_i}{dT_j} \end{aligned} \tag{8}$$

for all $i = 1,2,3; \ j = 0,1,2,3.$

Using the results of differentiation (8), and also proceeding from expression (4), we obtain

$$x\frac{dx}{dT_j} + y\frac{dy}{dT_j} + z\frac{dz}{dT_j} - R_0 \frac{dR_0}{dT_j} = 0. \tag{9}$$

In matrix form, the above expression can be written as the following relation

$$\vec{D}\vec{A} = \vec{R}, \tag{10}$$

where

$$\vec{D} = \begin{vmatrix} x & y & z & -R_0 \\ x_1 & y_1 & z_1 & r_1 \\ x_2 & y_2 & z_2 & r_2 \\ x_3 & y_3 & z_3 & r_3 \end{vmatrix}, \tag{11}$$

$$\vec{A} = \begin{vmatrix} \frac{dx}{dT_0} & \frac{dy}{dT_0} & \frac{dz}{dT_0} & -R_0 \\ \frac{dx}{dT_1} & y_1 & z_1 & r_1 \\ x_2 & y_2 & z_2 & r_2 \\ x_3 & y_3 & z_3 & r_3 \end{vmatrix}, \tag{12}$$

$$\vec{R} = \begin{vmatrix} x & y & z & -R_0 \\ x_1 & y_1 & z_1 & r_1 \\ x_2 & y_2 & z_2 & r_2 \\ x_3 & y_3 & z_3 & r_3 \end{vmatrix}. \tag{13}$$

It should be noted that the speed of light is omitted from expression (10) taking into account the fact that errors in the formation of time scales are given in terms of range.

Based on expression (10), we can write an expression for estimating the required matrix $\vec{A}$ in the following form

$$\vec{A} = \vec{D}^{-1}\vec{R} \tag{14}$$

Thus, for the selected location of the receiving points of the considered synchronous information network of observation systems and the position of the air object, the matrices $\vec{D}$ and $\vec{R}$ are known. This allows solving expression (14). As follows from (14), the third row of the estimated matrix $\vec{A}$ represents the sensitivity of measuring the height of an airborne object to errors in the synchronicity of the formation of time scales of receiving points. If all measured time intervals are equally sensitive to errors in the formation of a synchronous information network, then the sum of square errors is nothing more than the total value of the geometric factor [37]. Some results of calculating the sum of squared errors of range measurement in accordance with expression (14) are shown in Fig. 1 and 2 for different configurations of receiving points and the height of the air object. Fig. 1 is presented for the uniform location of three points at a radius of 45 km around the base receiving point, and Fig. 2 - for triangular. The flight altitude of the airborne object was 5 km.

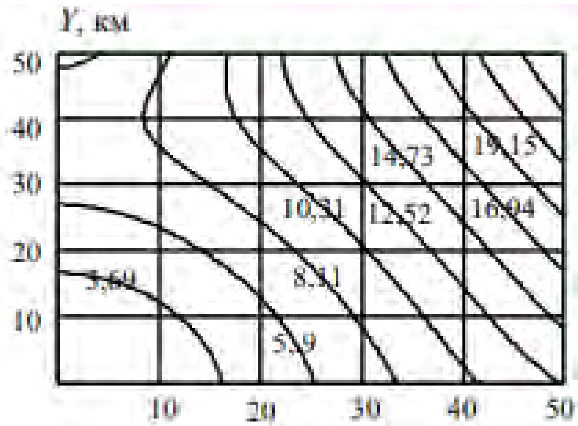


Figure 1: Evaluation the accuracy of measuring the height of an airborne object for the uniform location reception centers of three points at a radius

The presented calculations make it possible to estimate the required synchronization accuracy of the time scales of the receiving points of the synchronous information network of observation systems when calculating the height of an air object from range measurements carried out at the receiving points of the considered information network of radar observation systems.

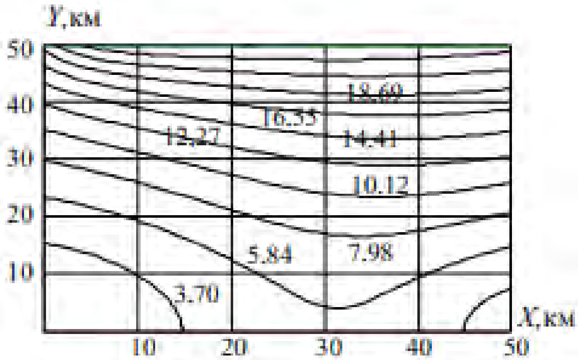


Figure 2: Evaluation the accuracy of measuring the height of an airborne object for reception centers the triangular location

3. Conclusions

The implementation of a synchronous information network of surveillance systems for primary and secondary radars, in which time is one of the coordinates, makes it possible to measure the height of the aircraft with the required accuracy by measuring the slant range at the receiving points of the considered information network of radar surveillance systems for the airspace with an appropriate number of receiving points.

4. References

[1] Y. Almathami, R. Ammar, Controlled Airspace infringements and warning system, 2016 IEEE International Symposium On Signal Processing And Information Technology (ISSPIT). (2016). doi:10.1109/isspit.2016.7886007.

[2] P. Massimini, J. Dieudonne, L. Monticone, D. Lamiani, E. Brestle, Insertion of controller-pilot data link communications into the National Airspace System: is it more efficient?, IEEE Aerospace And Electronic Systems Magazine. 15 (2000) 25-29. doi:10.1109/62.873472.

[3] J. Garcia, J. Molina, A. Berlanga, G. de Miguel, Data fusion alternatives for the integration of millimetre radar in airport surveillance systems, IEEE International

Radar Conference, 2005. (2005). doi:10.1109/radar.2005.1435935.

[4] H. Khudov, A. Fedorov, D. Holovniak, G. Misiyuk, Improving the Efficiency of Radar Control of Airspace with the Multilateration System Use, 2018 International Scientific-Practical Conference Problems Of Infocommunications. Science And Technology (PIC S&T). (2018). doi:10.1109/infocommst.2018.8632141.

[5] T. Otsuyama, J. Honda, A study of direction finding method for passive airport surveillance radar, 2014 International Symposium On Antennas And Propagation Conference Proceedings. (2014). doi:10.1109/isanp.2014.7026696.

[6] D. Pavlova, G. Zavolodko, I. Obod, I. Svyd, O. Maltsev, L. Saikivska, Comparative Analysis of Data Consolidation in Surveillance Networks, 2019 10Th International Conference On Dependable Systems, Services And Technologies (DESSERT). (2019). doi:10.1109/dessert.2019.8770008.

[7] A. Hume, C. Baker, Netted radar sensing, 2001 CIE International Conference On Radar Proceedings (Cat No.01TH8559). (n.d.). doi:10.1109/icr.2001.984634.

[8] T. Derham, K. Woodbridge, H. Griffiths, C. Baker, The design and development of an experimental netted radar system, 2003 Proceedings Of The International Conference On Radar (IEEE Cat. No.03EX695). (n.d.). doi:10.1109/radar.2003.1278755.

[9] I. Svyd, I. Obod, O. Maltsev, O. Strelnytskyi, O. Zubkov, G. Zavolodko, Method of Increasing the Identification Friend or Foe Systems Information Security, 2019 3Rd International Conference On Advanced Information And Communications Technologies (AICT). (2019). doi:10.1109/aiact.2019.8847853.

[10] C. Baker, A. Hume, Netted radar sensing, IEEE Aerospace And Electronic Systems Magazine. 18 (2003) 3-6. doi:10.1109/maes.2003.1183861.

[11] M. Weiss, Group sparsity techniques for data fusion of a passive MISO radar network, 2016 17Th International Radar Symposium (IRS). (2016). doi:10.1109/irs.2016.7497374.

[12] V. Chernyak, Fundamentals of Multisite Radar Systems, (2018). 475 p. doi:10.1201/9780203755228.

[13] H. Khudov, S. Kovalevskyi, A. Irkha, V. Lishchenko, O. Serdiuk, F. Zots, The Proposals for Synchronization Positions of MIMO Radar System on the Basis of Surveillance Radars, 2019 IEEE International Scientific-Practical Conference Problems Of Infocommunications, Science And Technology (PIC S&T). (2019). doi:10.1109/picst47496.2019.9061284.

[14] H. Griffiths, A. Farina, Multistatic and Networked Radar: Principles and Practice, 2021 IEEE Radar Conference (Radarconf21). (2021). doi:10.1109/radarconf2147009.2021.9455149.

[15] A. Maini, Handbook of defence electronics and optronics : fundamentals, technologies and systems, 1st ed., 2018.

[16] J. Franz, M. Weber, G. Hatke, L. Wood, Secondary surveillance phased array radar (SSPAR), 2014 Integrated Communications, Navigation And Surveillance Conference (ICNS) Conference Proceedings. (2014). doi:10.1109/icnsurv.2014.6820007.

[17] X. Du, K. Liao, X. Shen, Secondary Radar Signal Processing Based on Deep Residual Separable Neural Network, 2020 IEEE International Conference On Power, Intelligent Computing And Systems (ICPICS). (2020). doi:10.1109/icpics50287.2020.9202372.

[18] I. Svyd, I. Obod, O. Maltsev, A. Hlushchenko, Secondary Surveillance Radar Response Channel Information Security Improvement Method, 2020 IEEE 11Th International Conference On Dependable Systems, Services And Technologies (DESSERT). (2020). doi:10.1109/dessert50317.2020.9125018.

[19] M. Weber, J. Cho, H. Thomas, Command and Control for Multifunction Phased Array Radar, IEEE Transactions On Geoscience And Remote Sensing. 55 (2017) 5899-5912. doi:10.1109/tgrs.2017.2716935.

[20] G. Jiang, Y. Fan, H. Yuan, Assessing the Capacity of Air Traffic Control Secondary Surveillance Radar System, 2019 Cross Strait Quad-Regional Radio Science And Wireless Technology Conference (CSQRWC). (2019). doi:10.1109/csqrwc.2019.8799146.

[21] G. Galati, S. Bartolini, L. Mene, Analysis of SSR signals by super resolution algorithms, Proceedings Of The Fourth IEEE International Symposium On Signal

Processing And Information Technology, 2004. (n.d.). doi:10.1109/isspit.2004.1433713.

[22] I. Svyd, I. Obod, O. Maltsev, T. Tkachova, G. Zavolodko, Optimal Request Signals Detection in Cooperative Surveillance Systems, 2019 IEEE 2Nd Ukraine Conference On Electrical And Computer Engineering (UKRCON). (2019). doi:10.1109/ukrcon.2019.8879840.

[23] I. Svyd, I. Obod, O. Maltsev, G. Maistrenko, G. Zavolodko, D. Pavlova, Fusion of Airspace Surveillance Systems Data, 2019 3Rd International Conference On Advanced Information And Communications Technologies (AICT). (2019). doi:10.1109/aiact.2019.8847916.

[24] Y. Ma, S. Chen, C. Xing, X. Bu, L. Hanzo, Decomposition Optimization Algorithms for Distributed Radar Systems, IEEE Transactions On Signal Processing. 64 (2016) 6443-6458. doi:10.1109/tsp.2016.2602801.

[25] I. Svyd, O. Maltsev, I. Obod, G. Zavolodko, Fusion Method of Primary Surveillance Radar Data and IFF systems Data, 2020 IEEE 11Th International Conference On Dependable Systems, Services And Technologies (DESSERT). (2020). doi:10.1109/dessert50317.2020.9125040.

[26] J. Xu, X. Dai, X. Xia, L. Wang, J. Yu, Y. Peng, Optimizations of Multisite Radar System with MIMO Radars for Target Detection, IEEE Transactions On Aerospace And Electronic Systems. 47 (2011) 2329-2343. doi:10.1109/taes.2011.6034636.

[27] I. Obod, I. Svyd, O. Maltsev, G. Zavolodko, D. Pavlova, G. Maistrenko, Fusion the Coordinate Data of Airborne Objects in the Networks of Surveillance Radar Observation Systems, Data-Centric Business And Applications. (2020) 731-746. doi:10.1007/978-3-030-43070-2_31.

[28] V. Semenets, I. Svyd, I. Obod, O. Maltsev, M. Tkach, Quality Assessment of Measuring the Coordinates of Airborne Objects with a Secondary Surveillance Radar, Data-Centric Business And Applications. (2021) 105-125. doi:10.1007/978-3-030-71892-3_5.

[29] W. Liu, J. Wei, M. Liang, Y. Cao, I. Hwang, Multi-Sensor Fusion and Fault Detection using Hybrid Estimation for Air Traffic Surveillance, IEEE Transactions On Aerospace And Electronic Systems. 49 (2013) 2323-2339. doi:10.1109/taes.2013.6621819.

[30] I. Svyd, I. Obod, O. Maltsev, G. Zavolodko, Optimizing Airborne Object Detection of Secondary Surveillance Radar in Intra-System Interference Conditions, 2021 IEEE 16Th International Conference On The Experience Of Designing And Application Of CAD Systems (CADSM). (2021). doi:10.1109/cadsm52681.2021.9385224.

[31] M. Cherniakov, Bistatic radar, John Willey, 2008.

[32] C. Coleman, R. Watson, H. Yardley, A practical bistatic passive radar system for use with DAB and DRM illuminators, 2008 IEEE Radar Conference. (2008). doi:10.1109/radar.2008.4721007.

[33] C. Teague, D. Barrick, P. Lilleboe, Geometries for streamflow measurement using a UHF RiverSonde, IGARSS 2003. 2003 IEEE International Geoscience And Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477). (2003). doi:10.1109/igarss.2003.1295491.

[34] C. Wang, B. Wen, Z. Ma, W. Yan, X. Huang, Measurement of River Surface Currents With UHF FMCW Radar Systems, Journal Of Electromagnetic Waves And Applications. 21 (2007) 375-386. doi:10.1163/156939307779367350.

[35] L. Globa, M. Skulysh, O. Romanov, M. Nesterenko, Quality Control for Mobile Communication Management Services in Hybrid Environment, Lecture Notes In Electrical Engineering. (2019) 76-100. doi:10.1007/978-3-030-16770-7_4.

[36] I. Svyd, I. Obod, O. Maltsev, G. Zavolodko, Interference Immunity Assessment of a Secondary Surveillance Radar Synchronous Information Network, 2020 IEEE International Conference On Problems Of Infocommunications. Science And Technology (PIC S&T). (2020). doi:10.1109/picst51311.2020.9468021.

[37] O. Romanov, V. Mankivskyi, Optimal Traffic Distribution Based on the Sectoral Model of Loading Network Elements, 2019 IEEE International Scientific-Practical Conference Problems Of Infocommunications, Science And Technology (PIC S&T). (2019). doi:10.1109/picst47496.2019.9061296.

Accelerated learning of the neural network ADALINE in the presence of stationary correlated noise

Oleksandr Bezsonov[1], Oleg Rudenko[1], Natalia Serdiuk[1], Oleg Lebediev[1] and Valentyn Lebediev[1]

[1]Kharkiv National University of Radio Electronic, Kharkiv, 61166, Ukraine

Abstract

The problem of constructing a recurrent form of a multi-step ADALINA learning algorithm in the presence of stationary correlated interferences is considered. The basic relations describing the processes of accumulation of new and dumping of obsolete information are obtained. It is shown that the algorithms under consideration use the transformation of a random vector with correlated components into a random vector with uncorrelated components.

Keywords

Correlation, learning, recurrent algorithm, correlation matrix, centering, interference, inclusion of information, reset of information

1. Introduction

ADALINA (adaptive linear element) became the first linear neural network proposed by Widrow B. and Hoff M., which was an alternative to the perceptron [1]. Subsequently, this element and the algorithm of its training have found wide application in the tasks of identification, management, filtering, etc. The Widrow-Hoff learning algorithm is Kachmazh's algorithm for solving systems of linear algebraic equations. The properties of this algorithm in solving the identification problem are quite fully described in [2]. In [3], the regularized Kachmazh algorithm (Widrow-Hoff) was used to train ADALINA in the task of estimating nonstationary parameters and the corresponding estimates of convergence rate and accuracy were obtained.

It should be noted that the learning process of ADALINA can be significantly accelerated if instead of a one-step Widrow-Hoff (Kachmazh) algorithm to utilize a multi-step algorithm that uses a limited number of measurements, i.e. has limited memory, and is based on the current regression analysis algorithm [4-6].

Usually when constructing neural network models and their training algorithms, it is assumed that the noise in the measurements is not correlated. The presence of correlated noise complicates the evaluation procedure. Therefore, the correlation of noise is often neglected, while obtaining deliberately suboptimal results.

Recurrent estimation with correlated disturbances was considered in [7-9]. In [7] a recurrent form of MLS was proposed, in [8] a procedure of the stochastic approximation type was studied, the coefficients of which at each step are chosen to be optimal in the sense of minimizing the amount of a posteriori variance of estimates. In [9] it was shown that both of these methods have much in common and their experimental comparison was performed, which showed that in the case of a very strong correlation RMLS has greater computational stability.

The aim of this work is to obtain a recurrent form of a multi-step learning algorithm under conditions of stationary correlated noise.

2. Construction of a recurrent form of the current regression analysis algorithm

The problem of object identification described by the following equation is considered

$$Y_n = X_n c^* + \Xi_n, \tag{1}$$

EMAIL: oleksandr.bezsonov@nure.ua, (A.1); oleh.rudenko@nure.ua (A.2); nataliya.serdyuk@nure.ua (A.3); oleh.lebediev@nure.ua (A.4); lebedevvalen@gmail.com (A.5)

ORCID: 0000-0001-6104-4275 (A.1); 0000-0003-0859-2015 (A.2); 0000-0002-0107-4365 (A.3); 0000-0001-5998-0136 (A.4); 0000-0002-0095-7481 (A.5)

where $Y_n = (y_1, y_2, ... y_n)^T$ – vector of output signals; $X_n^T = (x_1, x_2, ..., x_n)^T = (X_{n-1}^T x_n)$ – matrix of input signals; $c^* = (c_1^*, c_2^*, .. c_N^*)^T$ – vector of parameters to be evaluated; $\Xi_n = (\xi_1, \xi_2, ..., \xi_n)^T$ – disturbance vector.

Covariance matrix D_n order n disturb ξ_{n+1} has the appearance

$$D_n = M\{\Xi_n \Xi_n^T\} = \begin{bmatrix} d_{1,1} & d_{1,2} \cdots & d_{1,n-1} & d_{1,n} \\ d_{2,1} & d_{2,2} \cdots & d_{2,n-1} & d_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{n,1} & d_{n,2 \ldots} & d_{n,n-1} & d_{n,n} \end{bmatrix} = \begin{bmatrix} D_{n-1} & d_{n-1} \\ d_{n-1}^T & d_{nn} \end{bmatrix}, \quad (2)$$

where $d_{ij} = M\{\xi_i \xi_j\}$;

$$d_{n-1}^T = (d_{n,1}, d_{n,2}, ..., d_{n,n-1}) = M\{\xi_n \Xi_{n-1}^T\}.$$

As is known [10], the application of evaluation

$$c_n = (X_n^T X_n)^{-1} X_n^T Y_n \quad (3)$$

to the model with correlated noise gives estimates whose variances will be underestimated.

Gaussian - Markov estimate (MLS), obtained by minimizing the quadratic functional, has the form

$$c_n = (X_n^T D_n^{-1} X_n)^{-1} X_n^T D_n^{-1} Y_n. \quad (4)$$

An algorithm of current regression analysis that has the form

$$c_{n|L} = (X_{n|L}^T X_{n|L})^{-1} X_{n|L}^T Y_{n|L}, \quad (5)$$

where

$$Y_{n|L} = \begin{pmatrix} Y_{n-1|L-1} \\ --- \\ y_n \end{pmatrix} = \begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L-1} \end{pmatrix} \quad (6)$$

– vector $L \times 1$;

$$X_{n|L} = \begin{pmatrix} X_{n-1|L-1} \\ --- \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{n-L+1}^T \\ --- \\ X_{n|L-1} \end{pmatrix} \quad (7)$$

– matrix $(L \times 1) \times N$;

was proposed in [4], and in [5] a modification of this algorithm is considered, which uses the mechanism of forgetting past information (smoothing). Here $L = const (L \geq N)$ – algorithm's memory.

Feature of algorithms with $L = const$ is that the matrices and observation vectors used in the construction of estimates at each step of the assessment are formed as follows: they include information about the newly arrived measurements and exclude information about the oldest. Depending on how these matrices and vectors are formed (new information is added first, and then obsolete information is excluded, or obsolete information is first excluded, and then new information is added), two forms of evaluation are possible.

Let's dwell on this in more detail.

By analogy with the Gaussian-Markov estimate (4), following estimate can be proposed

$$c_{n|L} = (X_{n|L}^T D_{n|L}^{-1} X_{n|L})^{-1} X_{n|L}^T D_{n|L}^{-1} Y_{n|L}, \quad (8)$$

where

$$D_{n|L} = \begin{bmatrix} D_{n-1,L-1+1} & \vdots & d_{n-1} \\ - & - & - \\ d_{n-1}^T & \vdots & d_{n,n} \end{bmatrix}, \quad (9)$$

$$d_{n-1}^T = (d_{n,n-L+1}, d_{n,n-L+2}, ..., d_{n,n-1}) = M\{\xi_n \Xi_{n-1|L-1}^T\}$$

Due to the fact that the matrix $Dn|L$ has a block representation, it can be written

$$D_{n|L}^{-1} = \begin{bmatrix} D_{n-1,L-1}^{-1} + \dfrac{D_{n-1,L-1}^{-1} d_{n-1} d_{n-1}^T D_{n-1,L}^{-1}}{\alpha_n} & \vdots & -\dfrac{D_{n-1,L-1}^{-1} d_{n-1}}{\alpha_n} \\ \cdots & \cdots & \cdots \\ -\dfrac{d_{n-1}^T D_{n-1,L-1}^{-1}}{\alpha_n} & \vdots & \dfrac{1}{\alpha_n} \end{bmatrix}, \quad (10)$$

where $\alpha_n = d_{nn} - d_{n-1}^T D_{n-1,L-1}^{-1} d_{n-1}$.

Similar relations are used in [7] to obtain a recurrent form of MLS.

Consider the case of stationary noise ξ.

Normalized covariance matrix for stationary noise with dispersion σ_ξ^2 looks like

$$R_{n|L} = D_{n|L} / \sigma_\xi^2 = \begin{bmatrix} 1 & \rho_1 \cdots & \rho_{n-L} & \rho_{n-L+1} \\ \rho_1 & 1 \ldots & \rho_{n-3} & \rho_{n-L} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{n-L+1} & \rho_{n-L} \cdots & \rho_1 & 1 \end{bmatrix}, \quad (11)$$

where ρ_i – the value of the normalized correlation function of the process ξ.

With this in mind, the estimate (8) takes the form

$$c_{n|L} = (X_{n|L}^T R_{n|L}^{-1} X_{n|L})^{-1} X_{n|L}^T R_{n|L}^{-1} Y_{n|L}. \quad (12)$$

Let's introduce $(n-m+1)$ – measurable vector

$$r_{m,n}^T = (\rho_n, \rho_{n-1}, ...\rho_m), \quad (m \le n).$$

Then the matrix $R_{n|L}$ can be represented in block form

$$R_{n|L} = D_{n|L} / \sigma_\xi^2 = \begin{bmatrix} R_{n-1|L-1} & r_{n-1,1} \\ r_{n-1,1}^T & 1 \end{bmatrix}, \tag{13}$$

where $r_{n-1,1}^T = (\rho_{n-1}, \rho_{n-2}, ...\rho_1)$, and the inverse matrix $R_{n|L}^{-1}$ will be calculated as follows:

$$R_{n|L}^{-1} = \\ = \begin{bmatrix} R_{n-1|L-1}^{-1} + \\ + \dfrac{R_{n-1|L-1}^{-1} r_{n-1,1} r_{n-1,1}^T R_{n-1|L-1}^{-1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} & -\dfrac{R_{n-1|L-1}^{-1} r_{n-1,1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} \\ -\dfrac{r_{n-1,1}^T R_{n-1}^{-1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} & \dfrac{1}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} \end{bmatrix}. \tag{14}$$

3. The procedure for adding new information

Suppose that on $(n-1)$-th tact an estimate is received

$$c_{n-1|L} = \\ = \left(X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L}\right)^{-1} X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} \tag{15}$$

The arrival of new information (adding a new dimension) leads to the calculation of the estimate, which by analogy with (15) can be written as follows:

$$c_{n|L+1} = \\ = (X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1})^{-1} X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} \tag{16}$$

where

$$Y_{n|L+1} = \begin{pmatrix} Y_{n-1|L} \\ --- \\ y_n \end{pmatrix} = \begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L} \end{pmatrix} \tag{17}$$

– vector $(L+1) \times 1$;

$$X_{n|L+1} = \begin{pmatrix} X_{n-1|L} \\ --- \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{n-L+1}^T \\ --- \\ X_{n|L} \end{pmatrix} \tag{18}$$

– matrix $(L+1) \times N$.

We introduce following notation

$$K_{n|L+1}^{-1} = \left(X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1}\right);$$
$$K_{n-1|L}^{-1} = \left(X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L}\right); \tag{19}$$
$$K_{n|L}^{-1} = \left(X_{n|L}^T R_{n|L}^{-1} X_{n|L}\right)$$

$$\alpha_{j,n} = \rho_j - r_{j+1,n+j-1} R_{n-1|L-1}^{-1} r_{1,n-1} \\ (j = 0,1,...) \tag{20}$$

and calculate $K_{n|L+1}^{-1}$

$$K_{n|L+1}^{-1} = X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L} + \\ + \frac{X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1} r_{n-1,1}^T R_{n-1|L}^{-1} X_{n-1|L}}{\alpha_{n,0}} - \\ - \frac{x_n r_{n-1,1}^T R_{n-1|L}^{-1} X_{n-1|L}}{\alpha_{n,0}} - \\ - \frac{X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1} x_n^T}{\alpha_{n,0}} + \frac{x_n x_n^T}{\alpha_{n,0}} = \\ = K_{n-1|L}^{-1} + x_n^* x_n^{*T}, \tag{21}$$

where $x_n^* = \dfrac{x_n - X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}$

Similarly, calculate

$$X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} = X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} + x_n^* y_n^*, \tag{22}$$

where $y_n^* = \dfrac{y_n - Y_{n-1|L} R_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}.$

Adding $x_n^* x_n^{*T} c_{n-1|L}$ to both parts (15)

$$K_{n-1|L}^{-1} c_{n-1|L} + x_n^* x_n^{*T} c_{n-1|L} = \\ = X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} + x_n^* x_n^{*T} c_{n-1|L}$$

and subtracting (15) from (16), taking into account the properties $K_{n-1|L}^{-1}$ and $X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1}$), we obtain

$$K_{n|L+1}^{-1}\left(c_{n|L+1} - c_{n-1|L}\right) = x_n^*\left(y_n^* - c_{n-1|L}^T x_n^*\right)$$

or

$$c_{n|L+1} = c_{n-1|L} + K_{n|L+1} x_n^* \left(y_n^* - c_{n-1|L}^T x_n^*\right), \tag{23}$$

where

$$K_{n|L+1} = K_{n-1|L} - \frac{K_{n-1|L} x_n^* x_n^{*T} K_{n-1|L}}{1 + x_n^{*T} K_{n-1|L} x_n^*}. \tag{24}$$

4. Procedure for discarding obsolete information

When discarding outdated information obtained on $n-L+1$-th step, we come from the assessment $c_{n|L+1}$ to evaluation $c_{n|L}$. To obtain the appropriate rules for the correction of the assessment, proceed as follows.

We use the block representation of the covariance matrix $R_{n|L+1}$

$$R_{n|L+1} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_{n-L+1} \\ \rho_1 & 1 & \rho_1 & \rho_{n-L+1,n} \\ \dots & \dots & \dots & \dots \\ \rho_{n-L+1} & \rho_{n-L} & \rho_{n-L-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & r_{n-1,1}^T \\ r_{n-1,1} & R_{n|L} \end{bmatrix}, \tag{25}$$

where $r_{n-1,1}^T = (\rho_{n-L+1}, \rho_{n-L}, \dots, \rho_{n-1})$, and the representation of the inverse matrix $R_{n|L+1}^{-1}$ as

$$R_{n|L+1}^{-1} = \begin{bmatrix} \dfrac{1}{\alpha_{n-L+1,0}} & -\dfrac{r_{n-L+1}^T R_{n,L}^{-1}}{\alpha_{n-L+1,0}} \\ -\dfrac{R_{n|L}^{-1} r_{n-L+1}}{\alpha_{n-L+1,0}} & R_{n,L}^{-1} + \dfrac{R_{n,L}^{-1} r_{n-L+1} r_{n-L+1}^T R_{n,L}^{-1}}{\alpha_{n-L+1,0}} \end{bmatrix}, \tag{26}$$

where $\alpha_{n-L+1,0} = 1 - r_{n-L+1,1}^T R_{n,L}^{-1} r_{n-L+1,1}$,

In this case

$$K_{n|L+1}^{-1} = \left(X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1}\right) = \frac{x_{n-L+1} x_{n-L+1}^T}{\alpha_{n-L}} + \frac{X_{n|L}^T R_{n|L}^{-1} d_{n-L+1} d_{n-L+1}^T R_{n|L}^{-1} X_{n|L}}{\alpha_{n-L}} - \frac{x_{n-L+1} d_{n-L+1}^T R_{n|L}^{-1} X_{n|L}}{\alpha_{n-L+1}} - \frac{X_{n|L}^T R_{n|L}^{-1} d_{n-L+1} x_{n-L+1}^T}{\alpha_{n-L+1}} + X_{n|L}^T R_{n|L}^{-1} X_{n|L} = K_{n|L}^{-1} + x_{n-L+1}^* x_{n-L+1}^{*T}, \tag{27}$$

where

$$x_{n-L+1}^* = \frac{x_{n-L+1} - X_{n|L}^T D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}.$$

We have the same

$$X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} = X_{n|L}^T R_{n|L}^{-1} Y_{n|L} + x_{n-L+1}^* y_{n-L+1}^*, \tag{28}$$

where $y_{n-L+1}^* = \dfrac{y_{n-L+1} - Y_{n|L} D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}$.

Subtract from both parts (16) $x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1}$. Then

$$\left(\left(X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1}\right)^{-1} - x_{n-L+1}^* x_{n-L+1}^{*T}\right) c_{n|L+1} = X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} - x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1}. \tag{29}$$

Taking into account that

$$\left(X_{n|L}^T R_{n|L}^{-1} X_{n|L}\right) c_{n|L} = X_{n|L}^T R_{n|L}^{-1} Y_{n|L}, \tag{30}$$

subtract from (30) the relation (16) taking into account the expressions for $K_{n|L}^{-1}$ and $X_{n|L}^T R_{n|L}^{-1} Y_{n|L}$ we obtain:

$$K_{n|L}^{-1}\left(c_{n|L} - c_{n|L+1}\right) = x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1} - x_{n-L+1}^* y_{n-L+1}^*,$$

where

$$c_{n|L} = c_{n|L+1} - K_{n|L} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n|L+1}^T x_{n-L+1}^*\right), \tag{31}$$

but

$$K_{n|L}^{-1} = K_{n|L+1}^{-1} - x_{n-L+1}^* x_{n-L+1}^{*T}.$$

Then

$$K_{n|L} = K_{n|L+1} + \frac{K_{n|L+1} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n|L+1}}{1 - x_{n-L+1}^{*T} K_{n|L+1} x_{n-L+1}^*}. \tag{32}$$

Thus, the algorithm will look like (the first two relations describe the inclusion of new information, and the next two – the rejection of obsolete)

$$c_{n|L+1} = c_{n-1|L} + K_{n|L+1} x_n^* \left(y_n^* - c_{n-1|L}^T x_n^*\right); \tag{33}$$

$$K_{n|L+1} = K_{n-1|L} - \frac{K_{n-1|L} x_n^* x_n^{*T} K_{n-1|L}}{1 + x_n^{*T} K_{n-1|L} x_n^*}. \tag{34}$$

$$c_{n|L} = c_{n|L+1} - K_{n|L} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n|L+1}^T x_{n-L+1}^*\right); \tag{35}$$

$$K_{n|L+1} = K_{n|L} + \frac{K_{n|L+1} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n|L+1}}{1 - x_{n-L+1}^{*T} K_{n|L+1} x_{n-L+1}^*}; \tag{36}$$

If at first the outdated information is rejected, and then new is included, the algorithm takes the form

$$c_{n-1|L-1} = c_{n-1|L} - K_{n-1|L-1} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n-1|L}^T x_{n-L+1}^*\right); \tag{37}$$

$$K_{n-1|L-1} = K_{n-1|L} + \frac{K_{n-1|L} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n-1|L}}{1 - x_{n-L+1}^{*T} K_{n-1|L} x_{n-L+1}^*}; \tag{38}$$

$$c_{n|L} = c_{n-1|L-1} + K_{n|L} x_n^* \left(y_n^* - c_{n-1|L-1}^T x_n^*\right); \tag{39}$$

$$K_{n|L} = K_{n-1|L-1} - \frac{K_{n-1|L-1} x_n^* x_n^{*T} K_{n-1|L-1}}{1 + x_n^{*T} K_{n-1|L-1} x_n^*}; \qquad (40)$$

Here

$$x_{n-L+1}^* = \frac{x_{n-L+1} - X_{n|L}^T D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \qquad (41)$$

$$x_n^* = \frac{x_n - X_{n-1|L}^T D_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}; \qquad (42)$$

$$y_{n-L+1}^* = \frac{y_{n-L+1} - Y_{n|L} D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \qquad (43)$$

$$y_n^* = \frac{y_n - Y_{n-1|L} D_{n-1|L}^{-1} d_{n-1}}{\sqrt{\alpha_n}}. \qquad (44)$$

5. Features of assessment for the case of stationary interference

In this case the output signal is converted taking into account the entered notations y_n^* can be written as follows

$$y_n^* = y_n - r_{1,n-1}^T R_{n-1|L-1}^{-1} Y_{n-1|L-1}. \qquad (45)$$

Calculating $M\left\{\overline{y_n^*} \, \overline{y_{n-i}^*}\right\}, i = 1, 2, \ldots,$ it is not difficult to show that all these mathematical expectations of the products are zero. Therefore, the centered component of the transformed process, which is determined by relation (45) is a discrete white noise, i.e. relation (44) describes the transformation of a random vector with correlated components into a random vector with uncorrelated components.

6. Conclusions

Thus, the problem of constructing a recurrent form of the current regression analysis algorithm is considered, which allows estimating unknown parameters in the presence of stationary correlated interferences.

The basic relations describing the processes of accumulation of new and dumping of obsolete information are obtained. It is shown that the considered algorithms use the transformation of a random vector with correlated components into a random vector with uncorrelated components.

7. References

[1] Widrow B. Adaptive switching circuits, / B. Widrow, M. Hoff // IRE WESCON Convention Record. Part 4. New York: Institute of Radio Engineers, 1960. – P.96–104.

[2] Либероль Б.Д. Исследование сходимости одношаговых адаптивных алгоритмов идентификации. / Б.Д. Либероль, О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2018. – 5. – С.19 – 32.

[3] Руденко О.Г. Регуляризованный алгоритм обучения адалины в задаче оценивания нестационарных параметров / О.Г. Руденко, А.А. Бессонов // Управляющие системы и машины. – 2019. –1. – С.22 – 30.

[4] Перельман И.И. Оперативная идентификация объектов управления./ И.И. Перельман // М.: Энергоиздат,1982. – 272 с.

[5] Руденко О.Г. Модифицированный алгоритм текущего регрессионного анализа в задачах идентификации и прогнозирования. / О.Г. Руденко, И.Д. Теренковский, А. Штефан, Г.А. Ода // Радиоэлектроника и информатика. – 1998. –№ 4(05). – С.58–61.

[6] Rudenko O.G. Adaline robust multistep training algorithm / O.G. Rudenko, O.O. Bezsonov // Control Systems and Computers. – 2020. – № 3. – P.15–27.

[7] Аведьян Э.Д. Рекуррентный метод наименьших квадратов при коррелированных помехах / Э.Д. Аведьян //Автоматика и телемеханика . –1975. –5. – С.67–75.

[8] Медведев Г.А. Рекуррентное оценивание при помощи коррелированных наблюдений / Г.А. Медведев // Автоматика и телемеханика . 1974. – 5. – С.110–116.

[9] Медведев Г.А. О рекуррентных оценках по коррелированным наблюдениям / Г.А. Медведев, Г.А. Хацкевич // Автоматика и телемеханика . – 1979. – 8. – С.69–75.

[10] Джонстон Дж. Эконометрические методы./ Дж. Джонстон – М.: Статистика,1980. – 444 с.

Analysis of duplication of publications of functional requirements in the form of mathematical expressions

M.V. Yevlanov [1], N.V. Vasyltcova [2]

[1, 2] *Kharkiv National University of Radio Electronics, Nauki ave. 14, Kharkiv, 61166, Ukraine*

Abstract

Functional requirements descriptions are the first descriptions of information systems that are formed during the life cycle of these systems. Therefore, all subsequent delays in the creation of an information system are in one way or another associated with errors made during the publication of functional requirements for the system. One of these delays is the correction of the information system, caused by the need to eliminate the detected duplication of functional requirements. Therefore, there is a constant need to eliminate duplication of individual functions of the information system even at the stage of formation and analysis of functional requirements for this system. The task is complicated by the fact that in a number of cases scenarios for fulfilling functional requirements are mathematical models or expressions. At the same time, the existing methods of analysis of functional requirements are not intended for the analysis of such publications.

In the course of the study, it was proposed to use an improved method for synthesizing descriptions of architecture options for the information system being created to search for duplicated functional requirements published in the form of mathematical expressions. This method allows to compare knowledge-based models of publications of functional requirements for an information system. To test the efficiency of the method, it was proposed to use the functional requirements for the task "Formation and maintenance of an individual plan for a scientific and pedagogical staff member of the department." These requirements contain a number of scenarios for calculating the desired parameters of an individual plan.

The studies carried out have shown that the usual publications of functional requirements in the form of mathematical expressions do not allow identifying situations of duplication of such requirements. The reason for this in the study is proposed to consider excessive detailing of mathematical expressions, which are publications of functional requirements. Recommendations have been developed and experimentally tested to eliminate redundant detailing of functional requirements publications.

Keywords

Functional requirement, requirements analysis, duplication.

1. Introduction

Requirements for an IT product are the first descriptions of this IT product, from which descriptions of all other components of this product are formed. At the same time, any IT product is presented as a system - a combination of interacting elements organized to achieve one or several set goals [1]. This view allows to divide the requirements for an IT product into two groups:

1. Functional requirements that answer the question: "What exactly should the system do?";
2. Non-functional requirements that answer the question: "How exactly should the system fulfill a specific functional requirement?"

Consequently, it will be possible to reduce the cost of creating an IT product if this IT product is presented as a system of functional requirements consistent with each other. This requires two main processes: the process for determining the needs and requirements of the interested party and the process for determining the system requirements [1]. Both of these processes have one thing in common: first, work is performed to collect individual requirements, and then work is performed to analyze the full set of requirements. At the same time, the main characteristics of such an analysis in [1] mean that the requirements are necessary, freely realizable,

EMAIL: maksym.ievlanov@nure.ua (A. 1); nataliia.vasyltsova@nure.ua (A.2)
ORCID: 0000-0002-6703-5166 (A. 1); 0000-0002-4043-487X (A. 2)

consistent, meaningful, complete, reflect the specifics, feasible, traceable, verifiable, acceptable and have boundaries.

However, the use of the considered characteristics of the analyzed requirements greatly complicates the automation of requirements analysis. The fact is that the overwhelming majority of these characteristics do not have a formal description. Consequently, the presence of these characteristics in the analyzed requirements cannot be confirmed or refuted by standard methods of data analysis, structured or unstructured texts. In particular, this difficulty is associated with such a characteristic of the requirements for an IT product as consistency. At the same time, the analysis of the full set of functional requirements for consistency is very important in the course of creating such IT products as information systems (IS). In accordance with the definition of the concept of "system" given in [1], the IS can be represented as a combination of interacting functions, and this interaction is carried out by transmitting and receiving structured data arrays. Identification of conflicting functional requirements for IS at the early stages of an IT project will reduce the cost of implementing an IT project for creating this IS. These costs primarily include the time spent on detecting conflicting functions of the IS, eliminating the found inconsistencies and redesigning the functional structure, database and software of the IS. Therefore, research in the field of automation of work on the analysis of functional requirements for IS and, in particular, work on the analysis of requirements for IS for consistency should be recognized as relevant.

2. Modern research in the field of analysis of system requirements

The basic understanding of the methods and methods of requirements analysis was formulated at the end of the XX - beginning of the XXI century. Examples of these methods are given in [2]. However, the effectiveness of these methods in practice is often insufficient. At the same time, the use of modern Agile methods to manage work with requirements does not significantly improve the efficiency of the application of requirements analysis methods [3].

One of the most promising areas of research in this area involves the development of methods for identifying and analyzing requirements based on their formal descriptions. At the same time, the most common way to formally describe the requirements for IP should be recognized as knowledge-based descriptions. The use of knowledge-based descriptions allows to pay significant attention to the improvement of existing models and methods of analysis of requirements [4]. So, the variant of comparative analysis of system elements based on their descriptions in the form of Use Case, Activity diagrams and data flow diagrams is considered in [5]. In [6], the transformation of stakeholder requirements publications into an executable system model based on Activity and State UML diagrams is investigated. In [7], a procedural model of the specification of functional requirements for software is described. In [8], a model of innovative service design is described, the use of which in the course of work with requirements presupposes a combination of formal knowledge-oriented methods and informal methods of the theory of inventive problem solving.

These and similar works determine the need for research aimed at improving the existing methods for analyzing the functional requirements for ISs for their compliance with the desired characteristics. It should be taken into account that the set of formulated requirements for IS is the main source of information for performing work on the synthesis of the description of the architecture of the system being created [1].

However, most functional requirements analysis methods are based on the assumption that the publications of these requirements are met in one of the following ways:

- As text in natural language;
- In the form of one or more visual models;
- As a software prototype.

At the same time, the question of the applicability of standard methods for analyzing the requirements for IS in the event that the scenario for meeting these requirements is described by mathematical expressions remains open. A similar situation is typical not only for IS management of technical or technical and economic objects, but also for IS management of organizations, the functional tasks of which require rather complex calculations. Therefore, this work aims to study the applicability of a method developed by one of the authors for identifying duplicate functional requirements published in the form of mathematical expressions.

3. Materials and methods of research

As a method for finding duplicated functional requirements, an improved method for synthesizing options for describing the architecture of the created IS is used. This method was described in [9, 10] and improved in [11]. The method consists of the following stages and steps.

Stage 1. Form the initial version of the description of the architecture of the created IS $Arch_{base}$.

Step 1.1. Determine the number of system-wide representations of functional requirements for IS at the knowledge level $n = \left|K_i^{fIS}\right|$.

Generate a set of descriptions of IT services IT_{acm} by performing the operation $IT_{acm_j} = K_i^{fIS}$.

Step 1.3. For the set $\{K_i^{fIS}\}$ form a matrix $Arch_{base}$ of the following form:

$$Arch_{base} = \\ = \begin{bmatrix} t_{11}(IT_{acm_1}) & \dots & t_{1n}(IT_{acm_1}) \\ \dots & \dots & \dots \\ t_{n1}(IT_{acm_n}) & \dots & t_{nn}(IT_{acm_n}) \end{bmatrix}, \quad (1)$$

where $t_{ij}(IT_{acm_j})$ - the number of occurrences of system-wide representations of functional requirements for IS at the knowledge level in the description of the IT service (function) of the created IS, *i,j=1,…,n.* End execution of the method step.

The values $t_{ij}(IT_{acm_j})$ during the formation of the matrix $Arch_{base}$ are determined as follows:

- $t_{ij}(IT_{acm_j}) = 1$ if for the proposed option of including K_i^{fIS} in the cluster IT_{acm_j} the value $Profit(IT_{acm_j}, r)$ not calculated;
- if $t_{ij}(IT_{acm_j}) = 1$, then $t_{ji}(IT_{acm_j}) = 0$.

Stage 2. Set the value of the repulsion coefficient r and calculate the value of the profit function $Profit(IT_{acm}, r)$

$$Profit = \frac{\sum_{j=1}^{k} \frac{S(IT_{acm_j})\left|IT_{acm_j}\right|}{W(IT_{acm_j})^r}}{\sum_{j=1}^{k}\left|IT_{acm_j}\right|} \quad (2)$$

where $S(IT_{acm_j})$ – the number of elements in a knowledge-oriented description of an IT service IT_{acm_j}; $W(IT_{acm_j})$ – the number of unique elements in a knowledge-oriented description of an IT service IT_{acm_j}; r – the repulsion coefficient, the value of which depends on the type of IS architecture; $\left|IT_{acm_j}\right|$ – the number of system-wide representations of functional requirements for IS at the knowledge level in the description of an IT service IT_{acm_j}; k – the number of IT services (functions) of the IS.

Stage 3. Conduct a synthesis of optimal and / or acceptable options for describing the architecture of the created IS.

Step 3.1. Take $Profit_{max} = Profit(IT_{acm}, r)$, $i = 1, j = i + 1$.

Step 3.2. If $t_{ij}(IT_{acm_j}) = 1$, then exclude K_i^{fIS} from IT_{acm_i} and include K_i^{fIS} in IT_{acm_j}. Otherwise, go to Step 3.8.

Step 3.3. For the obtained version of the architecture description, calculate the value $Profit(IT_{acm}, r)$ by expression (2).

Step 3.4. If $Profit(IT_{acm}, r) > Profit_{max}$, then take $t_{jj}(IT_{acm_j}) = t_{jj}(IT_{acm_j}) + t_{ii}(IT_{acm_i})$, $t_{ij}(IT_{acm_i}) = 0$ and $t_{ji}(IT_{acm_i}) = 0$ for $j = 1, \dots, n$, $K_j^{fIS} = K_j^{fIS} \cup K_i^{fIS}$ and go to Step 3.1.

Step 3.5. If $Profit(IT_{acm}, r) \in [Profit_{max} -\varepsilon; Profit_{max}]$, then take $t_{ij}(IT_{acm_i}) = 1$. Here and in the future, it is recommended to determine the value $\varepsilon = Profit_{max} \times 0{,}1$ [9, 10].

Step 3.6. If $Profit(IT_{acm}, r) < Profit_{max} - \varepsilon$, then take $t_{ij}(IT_{acm_i}) = 0$.

Step 3.7. Take $j = j + 1$. If $j \leq n$, then go to Step 3.2.

Step 3.8. Take $i = i + 1$, $j = i + 1$. If $i < n$, then go to Step 3.2. Otherwise, end the execution of the method step.

Stage 4. Exclude from consideration all variants of the description of the architecture of the created IS IT_{acm}, fixed at Stage 3, for which the condition $Profit(IT_{acm}, r) \in [Profit_{max} - \varepsilon; Profit_{max}]$ is not met. Complete the application of the method.

Step 4.1. Generate a version of the description of the architecture of the created IS, including those IT_{acm_i}, in which $t_{ii}(IT_{acm_i}) = 1$.

Step 4.2. For each $t_{ij}(IT_{acm_i}) = 1$ of the matrix $Arch_{base}$ form a variant of the description of the architecture of the created IS, taking at the time of formation $K_i^{fIS} = K_i^{fIS} \cup K_j^{fIS}$, $IT_{acm_j} = \emptyset$ and $t_{jj}(IT_{acm_i}) = 0$. End method application.

The result of this method is many options for describing the architecture of the created IS. In this set, each version of the description of the architecture will contain only unique functional requirements or functional requirements that duplicate each other to the minimum allowable degree. Thus, this method allows to identify and eliminate functional requirements that overlap each other completely or to a large extent.

To describe the functional requirements in the form of a mathematical model, the model of the problem "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" was used. This model was considered in [12] and has the following form

$$\alpha_e C_e = \sum_{a=1}^{k}\sum_{b=1}^{l}\sum_{c=1}^{m} t_{abc} + \sum_{d=1}^{p_1} t_d + \sum_{f=1}^{p_2} t_f + \sum_{g=1}^{r}\frac{a_g t_g}{q_g} + \sum_{i=1}^{s} t_i, \qquad (3)$$

where α_e – is part of teacher's rate; C_e – is quantity of hours allocated for one teacher's rate for the planned academic year; t_{abc} – is quantity of hours of academic work that is planned to groups of students and to academic discipline; t_d – is quantity of hours of methodological work that is planned to successful completion of the academic work; t_f – is quantity of hours of methodological work that is planned to improve quality of academic work; t_g – is quantity of hours of types scientific work that is planned for the academic year; a_g – is planned number of results of the type of scientific work; q_g – is planned number of co-authors for the result of the type of scientific work; t_i – is quantity of hours of organizational and educational work that is planned for the academic year.

Equation (2) must be met subject to the conditions [12]:

$$\begin{cases} 0 < \alpha_e \le 1; \\ t_{abc}, t_d, t_f, t_g, t_i, a_g > 0; \\ q_g \ge 1; \\ t_{abc}, t_d, t_f, t_g, t_i, a_g, q_g \in Z. \end{cases}, \qquad (4)$$

The result of the analysis of the functions of the task "Formation and maintenance of an individual plan of the scientific and pedagogical employee of the department" in [12] are the following precedents:

- Forming document section "position and stake";
- Quantification of hours of education work planned for the academic year;
- Forming document section “education work”;
- Quantification of hours of methodological work that is planned to successful completion of the academic work;
- Quantification of hours of methodological work that is planned to improve quality of academic work;
- Forming document section “educate-methodical work”;
- Quantification of hours of scientific work planned for the academic year;
- forming document section “scientific work”;
- Quantification of hours of organizational and educational work planned for the academic year;
- Forming document section “organizational and educational work”;
- Forming document section “final distribution of time by type of work for the academic year”.

Some of these precedents describe the functions that form individual sections of the document "Individual plan of the scientific and pedagogical staff member of the department." However, another part of these use cases describes the functions for calculating the number of hours for various types of employee work. It is difficult to estimate the degree of duplication for such functions in the usual way, since their descriptions are elements of the model (3). Therefore, recommendations should be developed for the use of an improved method for synthesizing options for describing the architecture of the created IS to identify duplicate functional requirements, which are presented in the form of mathematical expressions.

4. Statement of the main research results

Let’s first single out those precedents that describe the computational functions of the task "Formation and maintenance of an individual plan for a scientific and pedagogical staff member of the department." These use cases include:

1. Quantification of hours of education work planned for the academic year;
2. Quantification of hours of methodological work that is planned to successful completion of the academic work;

3. Quantification of hours of methodological work that is planned to improve quality of academic work;
4. Quantification of hours of scientific work planned for the academic year;
5. Quantification of hours of organizational and educational work planned for the academic year.

Now let's highlight the knowledge-based $K_i^{f_{IS}}$ for each of these use cases. To do this, let's first publish these use cases that characterize specific functional requirements. The publication of each of these requirements is a mathematical expression that determines the scenario for calculating the result of the corresponding functions (see Table 1).

Table 1
Functional requirements publications

Requirement designation	Requirement publication
f_{IS_1}	$\sum_{a=1}^{k}\sum_{b=1}^{l}\sum_{c=1}^{m} t_{abc}$
f_{IS_2}	$\sum_{d=1}^{p_1} t_d$
f_{IS_3}	$\sum_{f=1}^{p_2} t_f$
f_{IS_4}	$\sum_{g=1}^{r} \frac{a_g t_g}{q_g}$
f_{IS_5}	$\sum_{i=1}^{s} t_i$

It is proposed to represent the knowledge-oriented description for such functional requirements in the form of a set. The elements of such a set will be individual elements of a mathematical expression and action signs, and each of the elements of the set must have an independent meaning. Using this recommendation allows to present knowledge-oriented descriptions of publications of functional requirements from Table. 1 as follows (see Table 2).

Table 2
Knowledge-oriented functional requirements descriptions

Description designation	Description content
$K_1^{f_{IS}}$	$(\mathrm{a}, =, 1, \Sigma, \mathrm{k}, \mathrm{b}, =, 1, \Sigma, \mathrm{l}, \mathrm{c}, =, 1, \Sigma, \mathrm{m}, t_{abc})$
$K_2^{f_{IS}}$	$(d, =, 1, \Sigma, p_1, t_d)$
$K_3^{f_{IS}}$	$(f, =, 1, \Sigma, p_2, t_f)$
$K_4^{f_{IS}}$	$(g, =, 1, \Sigma, r, a_g, t_g, /, q_g)$
$K_5^{f_{IS}}$	$(i, =, 1, \Sigma, s, t_i)$

Given in Table 2 representations of knowledge-oriented descriptions of the considered requirements can be used in the course of applying an improved method for synthesizing options for describing the architecture of the created IS.

Let's consider the main results of the implementation of the stages of the improved method for synthesizing options for describing the architecture of the created IS. The result of the first stage will be the formed matrix (1), which in our case will have the form

$$Arch_{base} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (5)$$

The result of the second stage will be the value of the payoff function $Profit(IT_{acm}, r)$ for the variant of describing the architecture of the problem as a set of distinguished functional requirements. The initial data and the results of calculations of the second stage are given in Table 3.

Table 3
Initial data and calculation results of Stage 2 of the improved method

Description designation	$S(IT_{acm_j})$ value	$W(IT_{acm_j})$ value
$K_1^{f_{IS}}$	16	10
$K_2^{f_{IS}}$	6	6
$K_3^{f_{IS}}$	6	6
$K_4^{f_{IS}}$	9	9
$K_5^{f_{IS}}$	6	6
Calculation of the value $Profit(IT_{acm}, r)$		
$\lvert IT_{acm_j} \rvert = 5$	$r = 2{,}5$ $(service - oriented\ architecture)$	$Profit(IT_{acm}, r) = 0{,}058352$

During the third stage, possible situations of duplication of knowledge-oriented descriptions of the selected functional requirements are identified. In our case, the result of the third stage will be the corrected matrix (5), which has the following form

$$Arch_{base} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (6)$$

The form of matrix (6) allows to assert that there are no duplication situations among the identified precedents. Therefore, the result of performing Stage 4 of the improved method will be only one version of the architecture description. This option will include descriptions of each of the highlighted use cases.

To check the result obtained, let's search for duplicated functional requirements for their publications in the form of text in natural language. As such publications, let's use the names of the highlighted precedents. Then knowledge-based descriptions of these requirements will be formed by processing the publications by Porter's stemmer and then removing the stop words. The results of the formation of knowledge-oriented descriptions of functional requirements are shown in Table 4.

Table 4
Knowledge-oriented natural language descriptions of functional requirements publications

Description designation	Description content
K_1^{fIS}	Quantif hour educ work plan academ year
K_2^{fIS}	Quantif hour methodolog work plan success complet academ work
K_3^{fIS}	Quantif hour methodolog work plan improv qualiti academ work
K_4^{fIS}	Quantif hour scientif work plan academ year
K_5^{fIS}	Quantif hour organiz educ work plan academ year

For these knowledge-based descriptions, the result of the first stage of the improved method will also have the form of matrix (5). The result of the second stage of the improved method is shown in Table 5.

As a result of the third stage of the improved method, the corrected matrix (5) was obtained, which has the following form:

$$Arch_{base} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}. \qquad (7)$$

Table 5
Initial data and calculation results of Stage 2 of the improved method

Description designation	$S(IT_{acm_j})$ value	$W(IT_{acm_j})$ value
K_1^{fIS}	7	7
K_2^{fIS}	9	8
K_3^{fIS}	9	8
K_4^{fIS}	7	7
K_5^{fIS}	8	8
Calculation of the value $Profit(IT_{acm}, r)$		
$\left\|IT_{acm_j}\right\| = 5$	$r = 2{,}5$ $(service - oriented\ architecture)$	$Profit(IT_{acm}, r) = 0{,}05032$

The form of the matrix (7) allows to assert that among the selected precedents there are situations of duplication. Therefore, as a result of the implementation of Stage 4 of the improved method, two variants of the description of the architecture were formed. The first option will include descriptions of the following three use cases:

1. Quantification of hours of methodological work that is planned to successful completion of the academic work;
2. Quantification of hours of methodological work that is planned to improve quality of academic work;
3. Quantification of hours of education work planned for the academic year; Quantification of hours of scientific work planned for the academic year; Quantification of hours of organizational and educational work planned for the academic year.

The last of the generated precedents of the first variant of the description of the architecture includes the first, fourth and fifth of the previously identified precedents of computational functions.

The second option will include descriptions of the following two use cases:

1. Quantification of hours of methodological work that is planned to successful completion of the academic work; Quantification of hours of methodological work that is planned to improve quality of academic work;
2. Quantification of hours of education work planned for the academic year; Quantification of hours of scientific work

planned for the academic year; Quantification of hours of organizational and educational work planned for the academic year.

The first of the precedents of the second version of the description of the architecture includes the second and third of the previously highlighted use cases of computational functions.

Thus, the facts of the presence of duplicated functional requirements for the computational functions of the analyzed problem are confirmed.

It is proposed to explain the discovered discrepancy in the assessment of the duplication of the functional requirements of computational functions by excessive detailing of mathematical expressions, which are used in the form of publications of the analyzed requirements. The essence of the manifestation of excessive detail lies in the fact that each specific mathematical expression contains unique elements, which in reality are special cases of the same element of a more general nature.

To overcome the effect of excessive detail, it is proposed in the course of forming knowledge-oriented descriptions $K_i^{f_{IS}}$ from Table 2 replace the special cases of individual elements with the general elements given in Table 6.

Table 6
Descriptions of common elements of mathematical publications functional requirements

Private elements	Common element	Common element semantics
$(\mathrm{c}, d, f.g.i)$	$index_1$	Element denoting the index of the first operation
$(\mathrm{m}, p_1, p_2.r.s)$	top_1	Element denoting the upper limit of the operation range
(t_d, t_f, t_g, t_i)	t_{index_1}	Element denoting the variable used in the operation

Taking into account the entered in the table. 6 common elements, knowledge-oriented descriptions of functional requirements for the computational functions of the analyzed problem will take the following form (see Table 7).

Table 7
Adjusted knowledge-based functional requirements descriptions

Description designation	Description content
$K_1^{f_{IS}}$	$(\mathrm{a}, =, 1, \Sigma, \mathrm{k}, \mathrm{b}, =, 1, \Sigma, \mathrm{l}, index_1, =, 1, \Sigma, top_1, t_{abc})$
$K_2^{f_{IS}}$	$(index_1, =, 1, \Sigma, top_1, t_{index_1})$
$K_3^{f_{IS}}$	$(index_1, =, 1, \Sigma, top_1, t_{index_1})$
$K_4^{f_{IS}}$	$(index_1, =, 1, \Sigma, top_1, a_g, t_{index_1}, /, q_g)$
$K_5^{f_{IS}}$	$(index_1, =, 1, \Sigma, top_1, t_{index_1})$

The results of the implementation of the first and second stages of the improved method for the initial data shown in Table 7, will coincide with the results discussed above for the initial data shown in Table 2, and are given in the form of matrix (5) and Table 3.

The result of the third stage of the improved method for the initial data shown in Table 7, there will be a corrected matrix (5). In this case, this matrix will take the following form:

$$Arch_{base} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}. \quad (8)$$

The form of the matrix (8) allows to assert that among the selected precedents there are situations of duplication. Therefore, as a result of the implementation of Stage 4 of the improved method, two variants of the description of the architecture were formed. The first option will include descriptions of the following three use cases:

1. Quantification of hours of education work planned for the academic year;
2. Quantification of hours of scientific work planned for the academic year;
3. Quantification of hours of methodological work that is planned to successful completion of the academic work; Quantification of hours of methodological work that is planned to improve quality of academic work; Quantification of hours of organizational and educational work planned for the academic year.

The last of the generated precedents of the first variant of the description of the architecture includes the second, third and fifth of the previously identified precedents of computational functions.

The second option will include descriptions of the following two use cases:

1. Quantification of hours of education work planned for the academic year; Quantification of hours of scientific work planned for the academic year;
2. Quantification of hours of methodological work that is planned to successful completion of the academic work; Quantification of hours of methodological work that is planned to improve quality of academic work; Quantification of hours of organizational and educational work planned for the academic year.

The first of the precedents of the second variant of the description of the architecture includes the first and fourth of the previously identified use cases of computational functions.

Thus, the possibility of detecting duplicate functional requirements based on their publication in the form of mathematical expressions, provided that the redundant detailing of these expressions is eliminated, is confirmed.

5. Conclusions

The study of the possibility of searching for duplicated functional requirements published in the form of mathematical expressions has been carried out. It is shown that the use of mathematical models and expressions describing in detail the subject area for publications of functional requirements does not allow identifying duplication situations. It is proposed to consider the main reason for this result to be excessive detailing of mathematical expressions that act as a publication of a functional requirement.

Recommendations for the elimination of redundant detail are proposed and an experimental verification of these recommendations is carried out. The results obtained make it possible to assert that it is possible to detect duplicate functional requirements published in the form of mathematical expressions, provided that the excessive detailing of these expressions is eliminated. However, this conclusion needs additional verification on various types of requirements and mathematical expressions.

6. References

[1] GOST R 57193–2016. System and program engineering. System life cycle processes (ISO/IEC/IEEE 15288:2015, NEQ). (2016). Introduced 31–10–2016. M.: Standartinform, 98.

[2] M. Maguire, N. Bevan (2002). User requirements analysis. A review of supporting methods. Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, 25-30 August 2002, 133-148. doi: 10.1007/978-0-387-35610-5_9

[3] H.F. Martins, A.C. de Oliveira, E.D. Canedo etc. (2019). Design thinking: Challenges for software requirements elicitation. Information, 10, 12, 371. doi: https://doi.org/10.3390/info10120371

[4] T. Ambreen, N. Ikram, M. Usman, M. Niazi (2018). Empirical research in requirements engineering: trends and opportunities. Requirements engineering, 23, 1, 63-95. doi: 10.1007/s00766-016-0258-2

[5] Kaiya, K. Adachi, Y. Chubachi (2019) Requirements Exploration by Comparing and Combining Models of Different Information Systems. In: M. Virvou, F. Kumeno, K. Oikonomou (eds) Knowledge-Based Software Engineering: 2018. JCKBSE 2018. Smart Innovation, Systems and Technologies, vol 108. Pp. 64-74. doi: 10.1007/978-3-319-97679-2_7.

[6] S.-K. Kim, T. Myers, M.-F. Wendland, P.A. Lindsay (2012) Execution of natural language requirements using state machines synthesised from Behavior Trees. Journal of Systems and Software. 85, 11, 2652-2664. doi: 10.1016/j.jss.2012.06.013.

[7] V. Cornejo-Aparisio, S. Flores-Silva, N. Bedregal-Alpaca, D. Tupacyupanqui-Jaen. Procedural model for the specification of functional requirements in software construction [Modelo procedimental para la especificación de requisitos funcionales en la construcción de software]. RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao, 2020, E26, 571-586.

[8] C.-H. Lee, C.-H. Chen, A.J.C. Trappey (2019). A structural service innovation approach for designing smart product service systems: Case study of smart beauty service. Advanced Engineering Informatics, 40, 154-167. doi: https://doi.org/10.1016/j.eswa.2019.113062.

[9] M.V. Yevlanov, N.V. Vasiltcova, I.Yu. Panforova. “Models and methods of synthesis of the description of rational architecture of information system”. Bulletin of the Scientific University "Lviv Polytechnic". Information Systems and Networks Series (2015). No. 829: 135–152 .
[10] M.V. Evlanov (2016). Development of the model and method of selecting the description of rational architecture of information system. Eastern-European Journal of Enterprise Technologies, 1(2(79), 4–12. https://doi.org/10.15587/1729-4061.2016.60583.
[11] M.V. Yevlanov. “Improved method of synthesis of variants of description of architecture of the created information system”. ACS and automation devices (2018). No 175: 32–41.
[12] Vasyltsova N., Panforova I., Kuzma Ye. Formation of Function Use Cases Based on Its Mathematical Model // Fourth International Scientific and Technical Conference “COMPUTER AND INFORMATION SYSTEMS AND TECHNOLOGIES”. Kharkiv: DISA PLUS LLC, 2020: 59-60.

Determining of Factors Affecting the Azure Functions Response Time

Kravets Natalia[1]

[1] *Kharkiv National University of Radio Electronics*, Nauky Ave. 14, Kharkiv, 61166, Ukraine

Abstract

Function as a Service (FaaS) is a popular template for creating serverless applications. The concept of serverless computing implies that the cloud service provider is responsible for creating and managing the infrastructure, and the developer can focus on creating the business logic of the application. The functions are well suited for processing a stream of various events in real time. However, FaaS is not a universal solution. To make a decision about choosing a particular serverless technology for the implementation of their project, the developer must evaluate its effectiveness based on knowledge of the software model implemented by the cloud provider and factors that can increase or reduce the response time of the function. This article describes the features of implementing stateless and durable functions in Azure, taking into account the deployment plan. As well as a list of factors that directly affect the response time, which can help a developer who has knowledge about the characteristics of the input event flow of his future application to choose the optimal technology for implementing the project.

Keywords

Serverless computing, FaaS, Azure functions, response time, durable functions, parallel computing

1. Introduction

Serverless computing is a new approach that allows developers to focus on implementing business logic, while infrastructure management is carried out automatically. A cloud provider that provides serverless services uses servers, automatically scales and administers the necessary infrastructure. A company that uses the services of a serverless provider actually pays for the resources used. This technology is used for

- Event stream processing [1]
- Performing scheduled tasks [2]
- Automatic scaling of websites, applications and APIs[1,2]
- Processing of data coming from IoT sensors[3]
- Processing of transferred files[2]
- Real-time data processing. [1]

Azure serverless applications can be created based on various templates:

- Function-as-a-Service (FaaS)
- Container clusters managed by Kubernetes
- Serverless workflows
- Serverless API gateway
- Serverless environment for applications.

A function as a service is an independent small block of code that interacts with an array of cloud services to solve only one task. Building an application based only on functions is a non-trivial task, Azure offers special services for orchestration. Azure functions are not a universal solution. A consequence of multi-tenancy with dynamic resource sharing is the inability to guarantee a fixed response time when executing a function.

The response time includes, in addition to the actual "pure" execution time of the function and the time spent on communication, various delays resulting from a compromise between costs (optimization of equipment loading) and the results obtained (execution of service level agreements, SLA).

A lot of works are devoted to the study of the efficiency of using the function [1,3,4,5,11,12], the purpose of this article is to determine the factors that affect the response time of Azure functions.

2. Types of functions in Azure

Functions are often used to implement applications that process requests or respond to a

EMAIL: natalia.kravets@nure.ua (A. 1)
ORCID: 0000-0002-6753-3333 (A. 1)

specific event, applications built on the basis of an event-oriented architecture.

Using FaaS significantly reduces the service startup time, in addition, code management and scaling of the service is performed automatically (Figure 1). If the traffic grows, new instances of the function are created to process it, if the traffic decreases, the number of instances of the function decreases. If the application or hardware fails, the function is automatically restarted on another machine.

Figure 1: FaaS

An Azure function consists of code and a function configuration function.json, which defines the events that need to be tracked, and how to pass input data to the function and return data from it. In the function function.json defines the trigger, bindings, and other configuration parameters of the function. A function can have only one trigger, it defines the events that need to be tracked. Using bindings, resources are declaratively connected to the function for passing parameters. Bindings are not a mandatory element, they can be input and/or output. A function can have one or more input and output bindings.

Two types of functions are implemented in Azure: stateless and stateful (stable Functions, Durable Functions). The first type can be used to implement simple applications in which individual requests can be sent to different instances of the service.

The second type allows you to define workflows and entities with state tracking, create orchestrator functions. Automatic state management is implemented using control points.

2.1. Stateless Azure functions

Functions are independent blocks that do not have their own memory. Function instances are initialized by discrete events, can be executed in parallel, and have a limited execution time. Thus, FaaS is not suitable for applications that require long-term data processing or processing a large amount of data in RAM, but rather corresponds to an event-oriented model of building applications.

Azure guarantees that the function will be executed at least once, but the developer must solve the problem of idempotence independently.

For applications with high traffic, the problem may be the processing of an incorrectly formed event. In Azure, queues are used to ensure the reliability of event processing. When processing a message from the queue, the function may block an incorrect event. Locks are not used in event streams (for example, in Azure event hubs). For these services, the priority is to provide high bandwidth with support for multiple consumer groups and playback capabilities.

Stateless Azure functions can be effectively used for real-time event processing. At the same time, the most important characteristics for them will be the response time to the request (response time) and availability.

2.2. Durable Azure functions

Durable functions are an extension of Azure functions and allow you to implement the orchestration of the execution of a stateful function. A durable function consists of various Azure functions that can play different roles.

Currently, there are four types of durable Azure functions[6]:

1. Activity
2. Orchestrator
3. Entity
4. Client.

The orchestrator functions are used to coordinate the execution of other stable functions in the function application and have the following characteristics:

- define the workflow of a function using procedural code

- can call other stable functions synchronously and asynchronously
- are stable and reliable, the progress is automatically recorded when the function is waiting for or receiving a result
- can be long, the total lifetime of an instance of the orchestrator function can be measured in seconds, days, months, or never end.

As application templates in which the use of durable functions will be effective, Azure offers:

1. Function chaining
2. Fan-out/fan-in
3. Async HTTP APIs
4. Monitoring
5. Human interaction
6. Aggregator (stateful entities).

The template schemes are presented in Figure 2. These templates are remakes of well-known parallel programming templates: Event Pipeline, Fork/Join, asynchronous data processing, spin mutex, etc. [7]

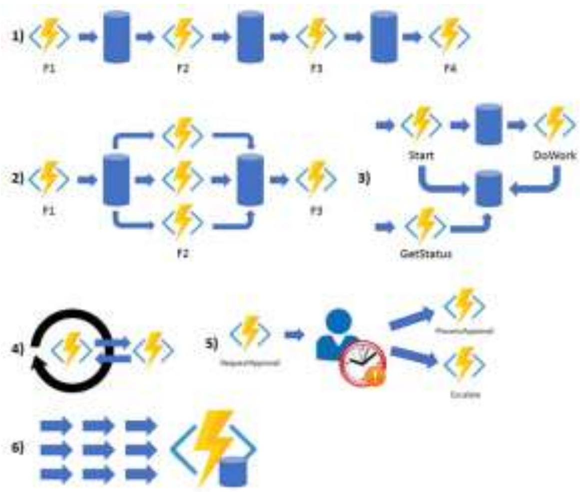


Figure 2: Application patterns for Durable Functions

In addition to the factors characteristic of stateless functions, the response time of durable functions will also be significantly affected by overhead costs due to the use of the presented templates.

2.3. Azure functions hosting plans

Currently, there are three hosting plans for Azure functions[2]:

- Consumption plan
- Premium plan
- Dedicated (App Service) plan.

The hosting plan defines the following parameters:

1. Scaling characteristics
2. Billing characteristics
3. Runtime Availability
4. Waiting time limits
5. Cold start behavior
6. Ability to integrate with networks
7. Restrictions on the number and parameters of function applications.

Items 1 and 5 directly affect the response time. During the scaling process, the number of function instances increases or decreases. Wherever the function is hosted in a virtual machine or in a container, in Azure there is the possibility of suspending the use of the infrastructure with a decrease in the amount of payment and the possibility of a complete stop (for virtual machines, this is a Stopped VM/ Deallocated VM) in this case, no payment is charged.

For the Consumption plan, scaling to zero is allowed (until the infrastructure is completely stopped) when idle, as a result, some requests will have a longer response time, additional time will be required to start the infrastructure from scratch. This problem is called "cold start" Consumption plan has some optimizations to reduce the cold start time, but this does not solve the problem completely.

For Premium plan, to avoid a cold start, unlimited instances of functions are created that do not scale to zero, only the function will work slower at the very first launch of the application.

When using the Dedicated (App Service) plan, the functions are performed within the framework of the application service plan and according to its tariffs. This solution is suitable if the script execution may exceed the time limits of the previous two placement plans, and Durable Functions cannot be used. As a result, the response time and the payment amount will be more predictable. This solution can be considered as a variant of the implementation of systems with strict requirements for response time (hard real-time systems).

3. Azure functions response time

Azure functions response time includes:

1. The time of the actual processing of the request (service time)
2. Communication time (transmission of information over the network)

3. Message delays in the queue
4. The time of delays caused by the features of the implementation of functions.

The factors affecting the duration of periods 1 and 2 are the same as for similar applications that do not use serverless computing — this is the complexity and implementation of the function, the amount of data being processed, and network characteristics. It should be mentioned the reasons for suspending the execution of processes in various runtime environments:

- The garbage collector (GC), which is part of the runtime environment of both declarative and many imperative programming languages, periodically suspends all running threads.
- The execution of the program can be stopped and restored at any time on the devices of end users.
- A virtual machine can suspend its work (suspend, suspend the execution of all processes and save the contents of memory to the hard disk) and restore execution (resume, restore the contents of memory and continue execution).
- A running thread can be suspended at any time of execution when switching the operating system context to another execution thread or switching the hypervisor to another virtual machine.
- The process of paging, swapping to disk, may suspend the application.
- When performing synchronous read/write from the disk, the work of the stream may be suspended.
- In the UNIX operating system, it is possible to suspend the execution of processes using the SIGSTOP signal. This signal stops the process from receiving CPU cycles, you can restore the process from the place where it was stopped via the SIGCONT signal. [8, 9, 10]

For period 3, the delay may be caused by a failure of one of the function instances or an incorrect event.

The factors that increase the response time due to the features of the implementation of functions (point 4) should include a cold start [11, 12], scaling features, overhead during synchronization and state preservation for durable functions. Problems can also occur when executing built-in triggers and bindings of Azure functions, calls to the Azure Base Services API, calls to RESTFUL endpoints, calls to client libraries, packages, or third-party APIs. As solutions, it is recommended to use standard monitoring tools, try/catch, implement idempotence and a repeat policy.

Thus, the currently used technology for implementing Azure functions does not allow us to guarantee a stable and / or fixed response time, except, perhaps, the Dedicated (App Service) plan. For several attempts to perform the same function, the response time will be different, so you should consider it as a distribution of values.

3.1. Azure functions response time for consumption plan

To estimate the response time for an empty Azure function that responds to an HTTP request created in the Consumption plan, 50 function launches were performed at intervals of 10s, 20s, 30s, 60s, 120s, 180s, 240s, 300s, 420s, 600s. The time of the first execution of the function was more than 2 times longer than the time of the 2nd start, which is explained by the problem of cold start.

When the pause between function calls increased from 3 minutes or more, the first call took longer than the median of the sample by ≈60%. This correlates with the standard idle time of Azure Load Balancer, which in turn caused the waiting time for functions whose triggers respond to an HTTP request to be limited to 230c[13,14,15]. For other types of events in the Consumption plan, the default waiting time is 5 minutes and the maximum is 10 minutes.

4. Conclusions

Serverless technologies are a new stage in the development of cloud services. The Azure FaaS service is a promising technology. Both the stateless Azure Functions and the Durable Azure Functions can be effectively used for real-time event processing [16,17,18,19]. However, the application response time requirements should not be strict.

There is a set of factors that increase the response time, depending on the features of the Azure Functions implementation: cold start, scaling features, overhead when synchronizing and saving state for durable functions, function instance failure, incorrect events.

The implementation of some part of the application in the form of functions can be very effective when taking into account the above features. For example, using a Consumption plan is most cost-effective if the stream of processed events will have a stable frequency of at least a certain value.

As a further development of this study, we can consider the construction of a model for estimating the response time, taking into account the influence factors described above. Next, it is planned to use this model to optimize resource consumption while maintaining the response time at a certain level for highly loaded applications that process the event flow in real time.

5. References

[1] Baldini, Ioana, et al. "Serverless computing: Current trends and open problems." Research advances in cloud computing. Springer, Singapore, 2017. 1-20.

[2] Rahul Sawhney, Beginning Azure Functions: Building Scalable and Serverless Apps, Springer Science+Business Media New York, 2019. doi: 10.1007/978-1-4842-4444-9.

[3] Benedetti, Priscilla, et al. "Experimental Analysis of the Application of Serverless Computing to IoT Platforms." *Sensors* 21.3 (2021): 928.

[4] Lloyd, Wes, et al. "Serverless computing: An investigation of factors influencing microservice performance." 2018 IEEE International Conference on Cloud Engineering (IC2E). IEEE, 2018.

[5] Manner, Johannes, et al. "Cold start influencing factors in function as a service." 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion). IEEE, 2018.

[6] Sawhney, Rahul. "Azure Durable Functions." Beginning Azure Functions. Apress, Berkeley, CA, 2019. 87-121.

[7] Czarnul, Pawel. Parallel programming for modern high performance computing systems. CRC Press, 2018.

[8] Krogh, Einar. "An introduction to Windows operating system." (2020).

[9] И. Пикуль, Принципы работы Garbage collection, 2018. URL: https://dou.ua/lenta/articles/principles-of-garbage-collection/.

[10] Room, Chat. "Linux Signals." system 16.52 (2019): 58.

[11] Manner, Johannes, et al. "Cold start influencing factors in function as a service." 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion). IEEE, 2018.

[12] Vahidinia, Parichehr, Bahar Farahani, and Fereidoon Shams Aliee. "Cold Start in Serverless Computing: Current Trends and Mitigation Strategies." 2020 International Conference on Omni-layer Intelligent Systems (COINS). IEEE, 2020.

[13] Soh, Julian, et al. "Developing and Deploying Azure-based Applications." Microsoft Azure. Apress, Berkeley, CA, 2020. 463-492.

[14] Azure Functions hosting options, 2020. URL: https://docs.microsoft.com/en-us/azure/azure-functions/functions-scale.

[15] Hole, Kjell Jørgen. "Tutorial on systems with antifragility to downtime." Computing (2021): 1-21.

[16] Basak, Anindita, et al. Stream Analytics with Microsoft Azure: Real-time data processing for quick insights using Azure Stream Analytics. Packt Publishing Ltd, 2017.

[17] Suresh, Visalakshmi. Scalable and responsive real time event processing using cloud computing. Diss. Newcastle University, 2017.

[18] Familiar, Bob, and Jeff Barnes. "Real-time processing using azure stream analytics." Business in Real-Time Using Azure IoT and Cortana Intelligence Suite. Apress, Berkeley, CA, 2017. 169-226.

[19] Kurniawan, Agus, and Wely Lau. Practical Azure Functions. Apress, 2019.

Air object recognition by the normalized contour descriptors

Valentyn Yesilevskyi, Andrew Tevyashev, Anton Koliadin

National University of Radio Electronics, Nauky ave.,14, Kharkiv, 61166, Ukraine

Abstract

This paper consider research into the methods for recognizing the type of an air object on a digital image acquired from video monitoring system. A method has been proposed that is applied on a feature vector built on the basis of a Fourier transform for the sequence of coordinates of its two-dimensional contour. This makes it easier to solve the classification problem owing to a more compact arrangement of the multidimensional feature vectors for similar air objects. The architecture of an air situation video monitoring system has been suggested, which includes an image preprocessing module and a module of neural network. Preprocessing makes it possible to identify an object's contour and build a sequence of normalized descriptors, which are partially independent of the spatial position of the object and the contour processing technique. Proposal method is easy in realization and do not require significant computational resources due take into consideration the specificity of recognizing objects in 3 dimensional. This research has shown that the reported results make it easier to train a neural network and reduce the hardware requirements for solve the task of air situation video monitoring.

Keywords

Air object recognition, contour analysis, Fourier descriptors, neural network

1. Introduction

Solving tasks of detection and recognition of air objects fast at high quality is of great importance for civilian applications in air traffic control, airport air situation monitoring, as well as military activities. Currently, the list of types of aircraft objects has been significantly expanded through the use of unmanned aerial vehicles (UAVs), quadcopters, cruise missiles, and helicopters. Optical video surveillance systems are increasingly being used to detect aerial objects. Video monitoring reduce the dimensions of detection systems and as well avoids problems related to the masking of characteristics of air objects in the radar detection area (stealth technology, small size, etc.). Video technologies require to solve tasks into the field of digital video processing and automated object recognition.

Solving them is typically based on the use of deep convolutional networks. Such a solution is too universal and not taking into consideration that in the area of recognition of the type of air objects, the contours of the object yield enough information to solve the problem. But the issue is that the resulting numerical characteristics of the contour should be invariant relative to the geometric distortion (displacement, orientation, scale) of the object's image. For air objects, this is especially relevant due to the three degrees of freedom in determining the position and is an important area to research.

One of the most effective and commonly used approaches to numerical description of the geometric shape of a flat object's contour is the application of a Fourier transform procedure. This approach is particularly interesting because it generates a unique one-dimensional identification sequence of the standard size, termed as Fourier descriptors, for all examined objects.

The deep convolutional networks that are used in laboratory image recognition experiments are too heavy both at the training stage and at the operational stage. For actual application, including mobile devices, methods that require fewer computing resources are in demand. This is confirmed by works [1–2] that address the use of object recognition methods based on feature

EMAIL: valentyn.yesilevskyi@nure.ua (A. 1); andrew.teviashev@nure.ua (A. 2); anton.koliadin@nure.ua (A. 3)
ORCID: 0000-0002-5935-1505 (A. 1); 0000-0001-5261-9874 (A. 2); 0000-0001-5552-5080 (A. 3)

vectors, including Fourier descriptors, acquired from a contour analysis. However, known solutions involving these methods either do not use information about phase from vectors complex vectors or do not parse the contours of objects into classes corresponding to three different projections. This suggests the relevance of our research into recognition air objects in an arbitrary spatial position, based on the Fourier descriptors of the contour.

2. Problem statement

Study [3] gives a detailed overview of the significant advances in the use of deep convolutional neural networks to recognize arbitrary images. However, as shown in papers [4, 5], deep learning methods require a large number of training sets and significant computer time to train a network.

It is shown in [6] that the use of transfer learning may become an option to overcome the heaviness of deep network training from scratch. This method assumes that a pretrained universal network additionally learns from specific images of air objects. The application of transfer learning makes it possible to bypass the limitations associated with the training time. However, some questions remain open. First, there is the issue of the redundancy of a universal approach based on deep neural networks when it comes to the recognition of types of air objects. Second, there is an issue related to the division of object classes, each of which is represented by too different (due to spatial orientation) images of objects on a flat image.

It should also be taken into consideration that explaining the recognition process in deep networks is based on highlighting the hierarchy of attributes in the image. This means that the network learning process involves excessive training the network, including for generating the contours' attributes, although this can be solved more effectively by classical methods of image processing.

Describing an image of an object by its contour is sufficient for the task of recognizing the types of air objects; it utilizes much less information than when analyzed using deep neural networks, which allows for a series of advantages. Studies [7, 8] report various methods for the mathematical notation of a contour; work [9] shows the use of these methods to recognize the types of air objects.

When recognizing 3-dimensional objects by their 2-dimensional image, there is an issue related to deriving a numerical descriptor invariant relative to the orientation of the object and its size in the image. The most interesting for description the contour of an air object is the use of Fourier descriptors. This method based on a discrete Fourier transform to the function that describes the contour of the object. As shown in papers [7, 8], this method is based on the representation of a closed flat curve by a sequence of points whose coordinates are considered complex numbers. Similar to how the spectrum of an audio signal, derived from a its Fourier transform, identifies this signal, the Fourier descriptors identify the closed contour of the flat shape.

Fourier descriptors partially resolve the issue of invariance in relation to the rescaling and rotation in the image plane. However, direct use of all the information derived from a Fourier transform is not possible as the phase component of the descriptors depends on the choice of the starting point of the contour. This issue requires separate research.

The task of recognizing aircraft types based on feature vectors derived from the Fourier descriptors has been solved since the 1980s by using a variety of classification methods. Work [10] applies methods of correlational analysis, distance assessment, support vector machine; paper [11] explored the use of neural networks to solve the task of categorizing objects by their descriptors of a Fourier contour.

A review of the results of the above studies reveals the following shortcomings in existing approaches to solving a task of air object recognition. Deep convolutional network methods are excessively heavy and general-purpose. The methods of network training do not take into consideration the possibility to simplify learning when dividing the training sample into subclasses corresponding to three different projections.

All this suggests that it is appropriate to construct an air object recognition method that would combine the advantages of a contour analysis based on Fourier descriptors and the specific training method of neural networks in order to overcome the above shortcomings

3. Studying the dependence of changes in the descriptors of the contour of air objects on the angle of rotation

This study is assumed that the image of an air object can be represented as an ordered sequence of $z(k) = (x_k, y_k)$, $k = 0, \ldots, N-1$ points that describe the contour.

The application of a Fourier discrete transform to this sequence generates a unique one-dimensional identification sequence of standard size values called Fourier descriptors, which possesses a series of interesting properties that are studied in detail below.

3.1. Exploring the properties of Fourier descriptors

A one-dimensional Fourier transform of the function $f(t)$:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \mathrm{d}t \qquad (1)$$

makes it possible to derive a continuous spectrum of this function in the frequency domain normally used to analyze time signals. For the case of the discrete sequence N of points $z(k)$ of the examined signal (1) takes the form of a discrete Fourier transform and leads to the calculation of the discrete spectrum:

$$F(n) = \sum_{k}^{N-1} z(k) e^{-i\frac{2\pi nk}{N}}, \qquad (2)$$

where $n = -N, \ldots, 0, 1, \ldots, N-1$.

It is important that this transform allows for a reverse operation – an inverse discrete Fourier transform:

$$z(k) = \frac{1}{N} \sum_{n}^{N-1} F(n) e^{i\frac{2\pi nk}{N}}, \qquad (3)$$

that restores the original sequence.

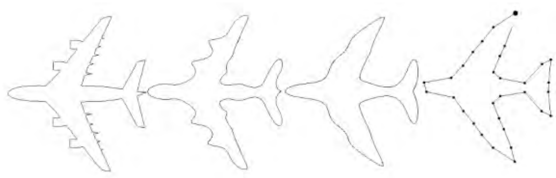
The sequence of Fourier descriptors possesses a series of important properties that explain their use in describing the contours of objects in two-dimensional images: invariant under shift, rotation, and scaling, as well as orderliness.

However, Fourier descriptors are sensitive to the choice of the starting point in Fourier transform processing. This effect makes it difficult to directly use the phase component of the descriptors to identify the contour.

There are also problems with arbitrary affine transformation. Such a transformation is one of the simplest, which roughly describes the distortion of the shape of a flat image of a contour when a 3-dimensional object is at an arbitrary angle relative to the plane of the camera.

The one of important useful feature in terms of contour identification is the orderliness of the descriptors by the degree of their importance to the image. This makes it possible to discard (zero) the high-frequency components without losing visual recognizability. In the sequence of the Fourier descriptors for a flat shape, the information about the shape of the contour is delivered by the first few elements of the sequence.

An inverse Fourier transform makes it possible to assess the degree of recognition of the shape at zeroed high-frequency components. The image contour shown in Figure 1, *a* contains about 2,000 points. Accordingly, a Fourier discrete transform would result in the same number of descriptors. The inverse transform would result in an accurate contour restoration.



a *b* *c* *d*

Figure 1: Illustration of the image contour filtering effect involving a Fourier transform

The zeroing of the descriptors above some of the filter's chosen boundary frequency results in a decrease in the details in the image at inverse transform while the object's recognition is retained. Figure 1, *b, c* shows the contour restored after zeroing all but 64 and 32 low-frequency descriptors, respectively. Figure 1, *d* shows an image obtained for the same original contour by discarding all descriptors except 32 for the rapid derivation of only 32 points of the restored image with linear interpolation between them.

3.2. Choosing a feature vector

In a general case, the Fourier descriptors $F(n)$ can accept different values for different images of the flat contour even of the same object. Their magnitudes depend on the scale r, the rotation angle φ, and the choice of the contour starting point k_0, as considered above.

Assuming that some reference value of the descriptors $F^*(n)$ has been selected, the calculated $F(n)$ sequence) for the recognized object can be represented as follows:

$$F(n) = re^{i\varphi}e^{-i\frac{2\pi n k_0}{N}} \cdot F^*(n). \qquad (4)$$

One can convert the Fourier descriptors to a form that lacks the influence of these factors. Consider the normalized descriptors according to [12]:

$$N(n) = \frac{F(1+n)F(1-n)}{F^2(1)}. \qquad (5)$$

It is then possible to show that:

$$N(n) =$$

$$= \frac{F^*(1+n)re^{i\varphi}e^{-i\frac{2\pi(1+n)k_0}{N}} F^*(1-n)re^{i\varphi}e^{-i\frac{2\pi(1-n)k_0}{N}}}{\left(re^{i\varphi}e^{-i\frac{2\pi k_0}{N}} F^*(1)\right)} = \qquad (6)$$

$$= \frac{F^*(1+n)F^*(1-n)}{\left(F^*(1)\right)^2} = N^*(n),$$

where $N^*(n)$ are the normalized coefficients, corresponding to the reference set of the Fourier descriptors $F^*(n)$.

It is easy to see that the normalized descriptors $N(n)$ (5), unlike Fourier descriptors, are independent of the above factors and can serve as an adequate sequence to describe the reference image of the contour.

3.3. Computational experiment

We have investigated the dependence of Fourier descriptors (2) and the normalized descriptors (5) on the angles of spatial position for four different types of aerial objects: aircraft, unmanned aerial vehicle, helicopter, quadcopter. In accordance with the filtering properties of a Fourier transform shown in Figure 1, it was determined that the first 32 descriptors would suffice to represent the shape of the air objects being examined to solve a classification problem.

One can see (Figure 2) that even similar profiles of the aircraft and helicopter have different «patterns» in terms of the amplitude component for their Fourier descriptors. The phase component, as discussed above, is not representative because of the uncertainty of choosing the starting point of the contour. However, this difference is especially noticeable for normalized descriptors where both the amplitude and the phase component are important.

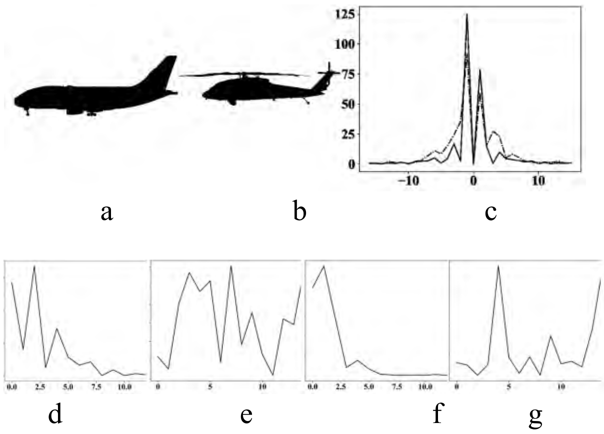


Figure 2. Difference in the shape of the objects represented by their Fourier descriptors and normalized descriptors: a – a binary image of the A380 aircraft profile; b – the same for the Apache helicopter; c – amplitude characteristics of Fourier descriptors (solid line – for A380, dotted line – for Apache); d, e – the amplitude and phase characteristics of the normalized descriptors for A380; f, g – the same for the normalized descriptors for Apache

When recognizing the type of any three-dimensional object in a two-dimensional image, there is an issue of the mutual location of the object and the video camera. Regardless of the classification method the proximity of the view of the contour of the real object and some pattern remains important. This is a particularly important issue for aerial objects that have the freedom to rotate around any of the 3 axes in threedimensional space.

There are a series of simplifications. Specific terms are typically used for aircraft rotation angles (roll, yaw, pitch). One can assume that the video plane is parallel to the XY plane of the system of coordinates of the aircraft Figure 3). At the same time, the aircraft can be simplified to present as a model in the form of flat surfaces (Figure 3, *a*). Then it is obvious that the roll of the aircraft would result in an easy computed change in the contour of the wing projection in the image.

In this case, for each point of this contour relative to the contour at a zero roll, the y-th coordinate would change proportionally to the cosine of the roll angle.

Figure 3 Change in the aircraft projection when its position changes

One can more accurately describe the change in the coordinates of a rotating object. Since of interest is a change in the x and y coordinates, one can record:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\beta & 0 \\ -\sin\alpha\cos\beta & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad (7)$$

where x' and y' are the coordinates of the new position of the arbitrary point of the object at coordinates x and y.

It is easy to show that due to the linearity of a Fourier transform, the proposed normalized descriptors (5) do not depend not only on such a change of angles but, in general, on an arbitrary affine transformation.

Studying the dependence of Fourier descriptors (2) and normalized descriptors (5) on the spatial angles for the examined air objects has shown that when the angle of the roll increases from 0 to 90 degrees, the side image smoothly transitions into the image from below (Figure 4), while the Fourier descriptors and the normalized descriptors are smoothly transformed into the horizontal projection descriptors.

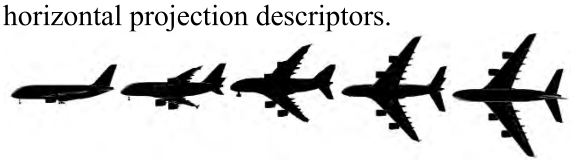


Figure 4 Change in the shape of the A380 aircraft's projection when the angle of the roll changes from 0° to 90°

However, in practice, the representation of an aircraft in the form of flat surfaces is too rough; the descriptors of different frequencies perform differently during this transformation. Figure 5 gives the results of a change in the amplitude of the first few normalized descriptors. Similar studies have been conducted for other types of aircraft (helicopter, UAV, and quadcopter).

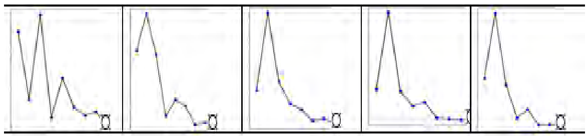


Figure 5 Change in the amplitude of the normalized contour descriptors when the angle of the A380 aircraft roll changes from 0° to 90°

It is generally accepted that the sequence of Fourier descriptors is invariant not only to the change in scale and displacement and only the roll and pitch influence the change in the two-dimensional projection of the three-dimensional silhouette of an aircraft.

Our analysis revealed that in the case when the shooting point is not directly under the plane, even yaw is not a pure image rotation, which can be compensated for when processing the descriptors. Thus, even in the trivial case, if one studies the contour of a cube or parallelepiped, at a yaw rotation the quadrangle contour can change to hexagonal.

The result of studying the contours of 4 types of objects: passenger plane, helicopter, unmanned aerial vehicle, quadcopter has established that the normalized descriptors $N(n)$ (5) even for low frequencies are sensitive to changes in the position of the object in space. That necessitates the use of separately normalized descriptors for each of the three orthogonal projections of the object as class references.

Therefore, three orthogonal projections as separate classes are used when solving a classification task and building a training set in this study, although this increases the number of classes to 11 for the 4 selected types of objects (the quadcopter has two projections that match).

4. Discussion of results of studying

Our results of the convergence of the training procedure indicate that the proposed approach to the selection of normalized contour descriptors as the vectors of object attributes in order to solve the task of recognizing the types of air objects using a neural network makes it possible to build an effective recognition subsystem for the complex of moving air object detection

The application of the proposed method is associated with a series of limitations in terms of its practical use. For example, solving the task implies, first, that an air object is large enough to detect a contour that can be used to calculate 15 normalized descriptors. Second, it is assumed that the image of the object is not distorted by fog or partially closed by clouds. The removal of these restrictions could be partially implemented by the pre-processing of the image and is the subject of further research.

5. Conclusions

A computational experiment involving model images was performed to investigate the dependence of change in Fourier descriptors and the normalized contour descriptors of images of 4 types of air objects on the angle of the object's rotation relative to 3 axes. It is shown that the use of normalized descriptors has a series of advantages over the Fourier descriptors. It has been established that to simplify the recognition task, one needs to parse the training set for each type of object into 3 classes corresponding to 3 orthogonal projections. This makes it easier to solve the classification problem owing to a more compact arrangement of multidimensional feature vectors for shapes with similar images.

Algorithmic maintenance for an air situation video monitoring system has been developed. The task of recognizing the type of an air object is tackled by a pre-trained neural network. The validation detection accuracy is as high as 99 %. This confirms that the proposed method of building a system for recognizing the types of air objects could simplify the requirements for the implementation of hardware while improving the accuracy when solving a task of air situation recognition.

6. References

[1] Strotov V., Babyan P., Smirnov S. (2017). Aerial object recognition algorithm based on contour descriptor. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. DOI: 10.5194/isprs-archives-XLII-2-W4-91-2017.

[2] Luciano da F. Costa, Roberto Marcondes Cesar Jr. (2009). Shape Classification and Analysis: Theory and Practice, Second Edition. 10.1201/9781315222325.

[3] Neha Sharma, Vibhor Jain, Anju Mishra, An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science, Volume 132, 2018, Pages 377 – 384. DOI: 10.1016/j.procs.2018.05.198

[4] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. Commun. ACM 2017, 60, 84–90. DOI: 10.1145/3065386.

[5]. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9. DOI: 10.1109/CVPR.2015.7298594, 10.1109/NAECON.2016.7856803.

[6] Valentyn Yesilevskyi, Andrew Teviashev, Anton Koliadin. Transfer learning in aircraft classification. IST-2019. the 8-th International Scientific and Technical Conference September 9 – 14, 2019 Kobleve-Kharkiv, Ukraine p.132 – 135. https://openarchive.nure.ua/bitstream/document/11942/1/3_IST.pdf.

[7]. Chaki Jyotismita, Dey Nilanjan. (2019). A Beginner's Guide to Image Shape Feature Extraction Techniques. 10.1201/9780429287794.

[8] Mingqiang Yang, Kidiyo Kpalma, Joseph Ronsin. Shape-based invariant features extraction for object recognition, Springer, 60 p. 2012. hal-00652422 DOI: 10.1007/978-3-642-24693-7_9.

[9] Hai-Jun Rong, Ya-Xin Jia, Guang-She Zhao, Aircraft recognition using modular extreme learning machine, Neurocomputing, Volume 128, 2014, Pages 166 – 174. ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2012.12.064.

[10] Makarov M.A., Berestneva O.G., Andreev S. Yu. (2014). Solving the problem of describing and classifying the contours of moving objects on video. Bulletin of the Tomsk Polytechnic University. Engineering of Georesources, 325 (5), 77 – 83.

[11] Nguyen, T. T. (2010). Algorithmic and software for shape recognition using Fourier descriptors and a neural network. Bulletin of the Tomsk Polytechnic University. Georesource Engineering, 317 (5).

[12] Granlund, G.H. (1972). Fourier Preprocessing for Hand Print Character Recognition. Computers, IEEE Transactions on. 21. 195 – 201. 10.1109/ TC.1972.5008926.

Topology Analysis of Hybrid Convolutional Neural Networks

Viktor Sineglazov [1], Olena Chumachenko [2] and Illia Boryndo [3]

[1] National Aviation University, ave. Lubomir Husar, 1, Kyiv, 03058, Ukraine
[2] National Technical University of Ukraine, ave. Peremogy 37, Kyiv, 03056, Ukraine
[3] National Aviation University, ave. Lubomir Husar, 1, Kyiv, 03058, Ukraine

Abstract

It is determined the perspective class of convolutional neural networks – hybrid convolutional neural networks. It's proved that these neural networks permit to supply less mean square error under less complexity. It is proposed a necessary composition of hybrid neural networks. It is shown that these networks must include beside traditional components (convolutional layers, pooling layers, feed-forward layers, additional layers: batch normalization layer, 1x1 convolutional layer, dropout layer, residual block). The useful properties of additional layers (blocks) have been determined and researched. It is considered that modern topologies of hybrid neural network as a combination of squeeze and excitation neural network, residual neural network, densely connected neural network, block attention module neural network, AlexNet, ResNeXt, etc. It is considered an example of hybrid convolutional neural network that based on using the densely connected convolutional blocks, residual layers, basic convolutional layers and additional supportive layers as batch normalization layer, flatten layer, attention merge layer, etc. Proposed architecture is the pair of two convolutional neural networks used for the parallel processing of input data and shows the positive results of using such type of neural networks systems. As the training dataset we're using CIFAR10. As a training algorithm it is used back propagation and adam algorithms. Achieved average accuracy of our system is approximately 93.1%.

Keywords

Hybrid Neural Networks, Convolutional Neural Networks, Neural Network Performance, Dynamic Neural Network Structure, Image representations

1. Introduction

Today, in modern world fully filled with "big data" tasks for the independent processing of graphic data there is a problem of lack of accuracy, performance [18] and other issues in the selection of special criteria. The urgency of this problem over time is only increasing due to the proliferation of the problem of digital identification.

In order to solve this problem the convolutional neural networks should be used. But even in this case the number of specific and complex tasks requires more flexible and targeted approach. It's the moment when hybrid convolutional neural networks take the place [1]. In this paper we will describe and present the research results on the usage and topology features of hybrid convolutional neural networks (HCCN). It's includes architectural global tasks that should be solved which contains not only learning samples and big data problems, but also a performance ones [2]. Most of CNN architectures are highly restricted due to its performance issues, low learning rates and at the same time requires a big number of high quality training material.

The main criteria of this research and overall using such systems by itself are to increase both accuracy and learning process performance [10], suppress such issues using global architectural approach.

2. CNN and HCNN topology

Before we'll dive in the details of hybrid CNN systems it's necessary to take a look onto CNNs itself.

Nowadays convolutional neural networks are

EMAIL: svm@nau.edu.ua (A. 1); elirankvik@gmail.com (A. 2); ibo.mistle@gmail.com (A. 3)
ORCID: 0000-0002-3297-9060 (A. 1); 0000-0002-6889-5221 (A. 2); 0000-0001-5375-6272 (A. 3)

the basic tool for graphical data processing and feature extracting. It's a notable deep architecture of deep learning. Due to its specific structure these neural networks can automatically extract the features or representation characteristics by processing the number of prepared graphical input data. Basically, one part of CNN is extracting features and another is processing them and classifying due to initial task requirements. [3]

In the same way hybrid convolutional neural network essentially is the combination of two or more different convolutional neural networks that should be configured and structured to work in pair (parallel or serial) and solve specific task or range of tasks. [4]

The idea of HCNN architecture is to implement the single-responsibility principle that makes each component of our systems (e.g. CNNs, classification/recognition algorithms, input/output data processors) to perform only one specific task. Therefore, we can decompose complex specific tasks into required dataflow steps that should be applied [5]. Each step will be processed with its specific element.

Having these simplified tasks its makes easier to train responsible neural networks and increases its accuracy and performance. As the example following networks can be combined: so called densely connected convolutional neural network in pair with squeeze and excitation convolutional neural network [6] based onto ResNeXt. It has a good potential due to global information holding at SE-CNN structure and DenseNet performance results. While combining networks the number of parameters should be considered [7]:

- input data parameters such as initial scale, resolution, number of channels and recognition task type;
- the output of first CNN should be acceptable by the second one, initial target information should be saved;
- the structures of both CNNs should be flexible and able to include supportive layers such as normalization layers, residual blocks, dropout layers, 1x1 convolution layers, etc.

In such approach any type of CNN can be paired. It's based mostly on the specifics of tasks and input data parameters (e.g. image resolution, scale, color channels, number of training samples, etc.).

3. Supportive structural blocks of CNNs

When the problem of low learning process performance takes place, there come out a number of solutions. One of them in to populate the architectural structure of current system with the supportive blocks. The main ones are:

- batch normalization layer;
- 1x1 convolution layer;
- dropout layer;
- residual block.

Normalization layer [20] is a layer that allows every layer of the network to do learning more independently [9]. It is used to normalize the output of the previous layers. The activations scale the input layer in normalization. Using normalization learning becomes efficient also it can be used as regularization to avoid overfitting of the model. The layer is added to the sequential model to standardize the input or the outputs. It can be used at several points in between the layers of the model. It is often placed just after defining the sequential model and after the convolution and pooling layers.

A 1x1 convolution layer is simply the convolution layer with filter size of 1x1. This convolutional layer should be applied while the pooling process of channel-wise type is required. This pooling operation also can be named as feature map pooling or a projection layer. Following approach takes place while we need to reduce processed data dimensionality, decreasing the number of feature maps whilst retaining their salient features [12, 19]. It can also be used directly to create a one-to-one projection of the feature maps to pool features across channels or to increase the number of feature maps, such as after traditional pooling layers.

Dropouts [17] are the regularization technique that is used to prevent overfitting in the model. The functional implementation of dropout layers is to randomly shuffle connections between some percentages of neurons in the network structure. When the neurons are switched off the incoming and outgoing connection to those neurons is also switched off. This is done to enhance the learning of the model. Dropouts are usually advised not to use after the convolution layers, they are mostly used after the dense layers of the network. It is always good to only switch off the neurons to 50%. If we switched off more than 50% then there can be chances when the model leaning would be poor and the predictions will not be good.

Residual blocks are skip-connection blocks that learn residual functions with reference to the layer inputs, instead of learning unreferenced functions. They were introduced as part of the ResNet architecture. The intuition is that it is easier to optimize the residual mapping than to optimize the original, unreferenced mapping. Having skip connections allows the network to more easily learn identity-like mappings.

4. Creating HCCN using structural modules

As we discussed before, now we can use both already existed CNN architectures and originally designed ones, prepopulate them with the supportive structural blocks if it's needed and apply it for our image recognition task [14].

As an example on Figure 1 is shown proposed and simplified HCNN architecture with the following basic parameters:

- batch size: 512;
- activation function: ReLU;
- learning rate: 0.001;
- growth coefficient: 16;
- input: 32x32x3.

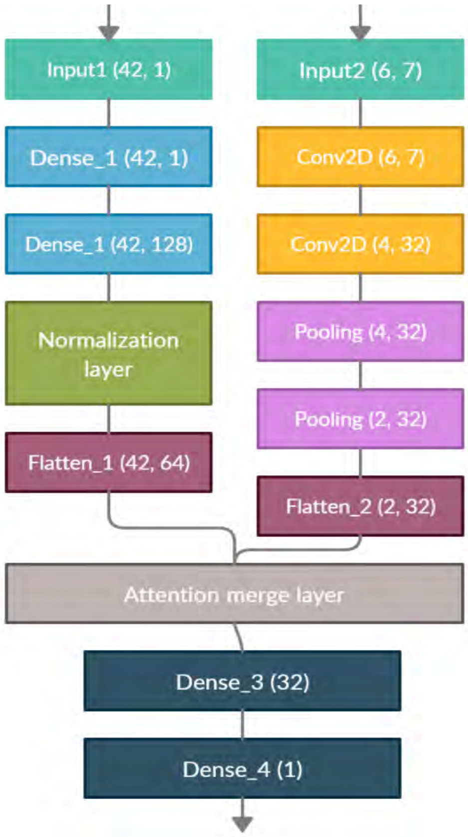


Figure 1: Proposed structure of hybrid convolutional neural network of parallel design

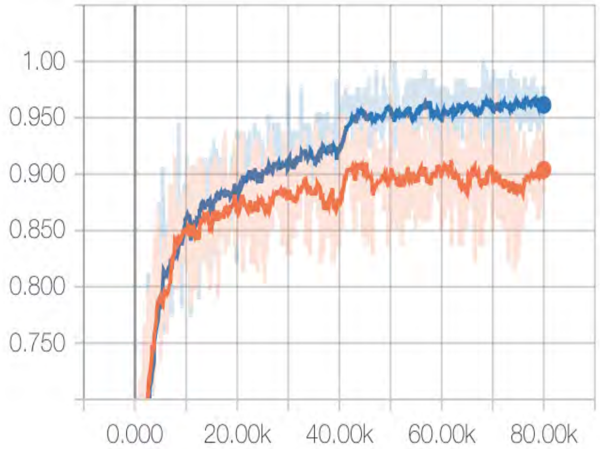


Figure 2: Training process accuracy graph for proposed HCNN structure

For the training [13] and testing purposes we will use the CIFAR-10 image dataset which contains images that fits for strict target recognition and classification processes. Raw resolution of this dataset is 32x32px and contains images of eight types. Initial weights of layered system are randomly generated before learning process was initialized.

As it's shown on Figure 2, there's the final result of such system pre-trained with CIFAR-10. Average accuracy is approximately 93.12711% and average loss is on 0.103 values but max values are going even lower.

5. Conclusion

It is considered the problem of hybrid convolution neural networks (HCNN) topology analysis. It is shown that is necessary to pay attention on some layers (blocks) possess some useful properties, which permit to increase the problem solution accuracy and decrease the complexity of HCNN. As such blocks, it was proposed to consider the following: batch size: 512, 1x1 convolution layer, dropout layer, residual block.

In the addition we can consider that it's necessary to look not only into supportive blocks but the generation algorithms [16] of hybrid convolutional neural network structure should be considered as well. It can be a genetic algorithm [13, 14, 15] or other structural synthesis algorithms as well.

The results based on proposed simplified HCNN are tense to 94% of accuracy as well as the error loss gentrification goes down up to 0.093 value.

6. References

[1] Iveta Mrazova, Marek Kukacka. "Hybrid convolutional neural networks" IEEE International Conference on Industrial Informatics 10.1109/INDIN.2008.4618146.

[2] Chaitanya Nagpal and Shiv Ram Dubey. “A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing”

[3] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014, pp. 675–678.

[4] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014.

[5] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[6] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu. “Squeeze-and-Excitation Networks”.

[7] C. Cao, X. Liu, Y. Yang, Y. Yu, J. Wang, Z. Wang, Y. Huang, L. Wang, C. Huang, W. Xu, D. Ramanan, and T. S. Huang, “Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks,” in ICCV, 2015.

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,” in ICCV, 2015.

[9] A. Krizhevsky and G. Hinton, “Learning multiple layers of features from tiny images,” Citeseer, Tech. Rep., 2009.

[10] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, “Places: A 10 million image database for scene recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.

[11] L. Shen, Z. Lin, and Q. Huang, “Relay backpropagation for effective learning of deep convolutional neural networks,” in ECCV, 2016.

[12] D. Mahajan, R. Girshick, V. Ramanathan, K. He, M. Paluri, Y. Li, A. Bharambe, and L. van der Maaten, “Exploring the limits of weakly supervised pretraining,” in ECCV, 2018.

[13] Yanan Sun, Bing Xue, Mengjie Zhang, Gary G. Yen, Jiancheng Lv, “Automatically Designing CNN Architectures Using Genetic Algorithm for Image Classification”, IEEE Transactions on Cybernetics, DOI:10.1109/TCYB.2020.2983860

[14] B. L. Miller, D. E. Goldberg et al., “Genetic algorithms, tournament selection, and the effects of noise,” Complex Systems, vol. 9, no. 3, pp. 193–212, 1995.

[15] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone, Genetic programming: an introduction. Morgan Kaufmann San Francisco, 1998, vol. 1.

[16] H. L. Liu, F. Gu, Y.-m. Cheung, S. Xie, and J. Zhang, “On solving WCDMA network planning using iterative power control scheme and evolutionary multiobjective algorithm [application notes],” IEEE Computational Intelligence Magazine, vol. 9, no. 1, pp. 44–52, 2014.

[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” The Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.

[18] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, “Automatic differentiation in pytorch,” 2017. [Online]. Available:https://openreview.net/forum?id=BJJsrmfCZ

[19] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. FeiFei, “Imagenet large scale visual recognition challenge,” International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.

[20] S. Ioffe, “Batch renormalization: towards reducing minibatch dependence in match-normalized models,” in Advances in Neural Information Processing Systems. Curran Associates, Inc., 2017, pp. 1945–1953.

Application of modern information technologies for solving the problem of finding a rational compromise in the interaction of partners under the conditions of absence analytical dependences of their objective functions

Oleh Borovyk [1], Serhii Lienkov [2], Oleksandr Sieliukov [3], Alexander Pashkov [4], Svitlana Kuznichenko [5], Dmytro Borovyk [6] and Iryna Basaraba [7]

[1, 7] *The National Academy of the State Border Guard Service of Ukraine named after Bohdan Khmelnytsky, Shevchenko str., 46, Khmelnytskyi, 29000, Ukraine*
[2] *Military Institute of Taras Shevchenko National University of Kyiv, Lomonosova Str., 81, Kyiv, 03189, Ukraine*
[3] *State Enterprise "Scientific Center for Precision Engineering", Boryspilska Str. 9, Kyiv, 02099, Ukraine*
[4, 5] *Ministry of Defense of Ukraine, Povitroflostky Ave, 6, Kyiv, 03168, Ukraine*
[6] *Khmelnytskyi National University, Instytutska Str, 11, Khmelnytskiy, 29000, Ukraine*

Abstract

For a large class of systems analysis tasks, an important issue is the disclosure of uncertainties. This is due to the variety of goals, properties and characteristics of the studied objects. Today, the task of revealing the conflicts uncertainty in the tasks of choosing the goals of plans and plans in the process of partners interaction or competitors opposition or opponents remains relevant. There are methods in systems analysis that allow you to solve these problems in some cases. They are based on the application of mathematical analysis methods and probability theory. However, these methods are applicable only to problems in which the number of partners and arguments in the analytically given objective functions that determine the purpose of their activities coincide. Since in practice, as a rule, such a restriction is not met, it is important to find approaches to solving problems of disclosing the uncertainty of conflicts in the choice of goals and plans in the process of partners interaction or countering competitors or opponents, arbitrary number of partners and arguments of their target functions, but in the absence of analytical dependencies of these functions. The paper formalizes and describes the problem features of finding a rational compromise in the interaction of any number of partners and any number of arguments that determine their target functions, but in the absence of analytical dependencies of these functions, and defines the approach to its solution using modern information technology. The proposed approach is applicable in the case of the component elements independence of the decision of the interacting entities. It provides for the capabilities integration of the previously developed author's method for the disclosure of uncertainty in the problems of interaction and the software product for the reproduction of functional dependencies in the class of additive functions.

Keywords

Mathematical model, uncertainty, objective function, information technologies, algorithm, software.

EMAIL: bov_nadpsu@ukr.net (A. 1); lenkov_s@ukr.net (A. 2); selukov@3g.ua (A. 3); db07757@gmail.com (A. 4); irynabasaraba2017@ukr.net (A. 5)
ORCID: 0000-0003-3691-662X (A. 1); 0000-0001-7689-239X (A. 2); 0000-0001-7979-3434 (A. 3); 0000-0003-4715-0658 (A. 4); 0000-0002-3209-9119 (A. 5)

1. Introduction

For a large class of formalized system analysis problems, an important issue is the disclosure of uncertainties. The most common in practice are uncertainties of goals, situations, conflicts. The essence of these uncertainties can be assessed from the work [1]. Today, the task of developing methods and algorithms for revealing the uncertainty of conflicts in the tasks of choosing the goals of plans and plans in the process of partners interaction or opposition of competitors or opponents, as well as the development of information technology for their study.

Currently, there are methods in systems analysis that allow you to solve these problems in some cases. They are based on the application of mathematical analysis methods and probability theory [2-3]. However, these methods are applicable only to problems in which the number of partners and the arguments of the given objective functions that determine the purpose of their activities coincide. Since in practice, as a rule, such restrictions and conditions are not always met, the task of finding approaches to solving problems of disclosing the uncertainty of conflicts in the tasks of choosing the goals of plans and plans in the interaction of partners or countering competitors or opponents (hereinafter, tasks) disclosure of conflicts), which would provide the possibility of solving the problem for any number of partners and the arguments of their target functions, especially in the absence of analytical dependencies of the latter.

Thus, the problem that needs to be solved can be formulated as follows.

Number of partners interacts, each of which has its own goal, which is determined by an unknown objective function, which can theoretically be formed on the basis of known empirical data.

Partners in the process of active interaction can exchange information about their actions.

Find values of arguments of unknown objective functions at which objective functions would reach values that would satisfy all partners.

Given the condition of the problem under study, we can conclude that its solution involves, first, finding a solution in the presence of an arbitrary number of partners and the arguments of their objective functions in the presence of analytical dependencies of the latter, and secondly, the establishment of analytical dependencies of objective functions in the case of only a certain set of empirical data that can determine the arguments and values of objective functions.

Regarding the first problem, it should be noted that in this statement the problem is investigated in [4]. In this paper, we formalize the problem of disclosing uncertainty in the interaction of partners, in which the number of arguments of the objective functions is not necessarily equal to the number of partners. The analysis of the existing approach to the solution of the formulated problem in the absence and presence of situational uncertainty for two and any number of partners is also carried out. Based on the application of technical constraints, an approach to solving the problem is proposed and software and algorithmic support for its implementation is formed. This approach is based on the preliminary formation of the area of acceptable solutions (Pareto area) and the subsequent search for a rational solution in this area, can be used to solve the problem of disclosing the uncertainty of conflicts in the absence and presence of situational uncertainty. Software-algorithmic implementation of the author's approach, presented in [4], allows to automate individual stages of problem solving.

Regarding the second problem, it should be noted that the problem of reproducing functional dependences on an experimentally obtained discrete sample is one of the rather important applied problems. This explains the fact that the theoretical aspects of its solution were given attention in [1]. And the substantiation of algorithm, structure, functional features of the software application of reproduction of functional dependence on discrete sampling was devoted to work [5].

However, despite the presence of approaches to solving individual components of the problem under study, there are no approaches to solving the problem in the above statement. It should be noted that the mechanistic combination of approaches to solving individual problems may not be applicable in the case under study. This is due to the presence of certain features of the physical content of the problem.

The purpose of this work is to formalize and describe the features of the problem of finding a rational compromise in the interaction of any number of partners and any number of arguments that determine their objective functions, but in the absence of analytical dependencies of these functions.

2. Formalization of the studied problem

To achieve this goal, it is considered appropriate: first of all to formalize the research problem; to carry out the analysis of features of statement of a problem which are caused by its physical maintenance; determine approaches to its solution taking into account the peculiarities of the production.

Formalized description of the problem.

Entities may be involved in solving a problem $C_1, C_2, ..., C_n$.

Each of the subjects is assigned its own unique set of tasks, the solution of which is entrusted only to the relevant responsible subject, and a set of tasks, which can be joined by other subjects or combinations thereof, but which are a priority for the responsible subject.

The ability to solve a problem depends on the decision elements that are made by the subjects $C_1, C_2, ..., C_n$. Each of the subjects $C_i \left(i = \overline{1,n} \right)$ makes decision $X_i \left(i = \overline{1,n} \right)$ elements of which are indicators $x_{i1}, x_{i2}, ..., x_{iC_i} \left(i = \overline{1,n} \right)$.

Solving the problem Π is a priority for one of the subjects. Let's denote this subject by $C_в$.

Among the elements of the decision of each subject are those indicators that determine the effectiveness of the subject in the territorial and temporal limits of the urgency of the problem Π (hereinafter, within the problem Π). These elements of the solution apply to the whole range of tasks of the subject, which he solves within the problem Π.

The decisions made by each entity must ensure that their target functions are max $y_i \left(i = \overline{1,n} \right)$.

If the subject makes decisions in the interests of solving the priority task of the subject $C_в$, i.e. the problem Π, the value of its own objective function should not decrease below the threshold level $y_i^* \left(i = \overline{1,n} \right)$.

At the same time, there is another objective function $\bar{y}_в$ for the subject $C_в$, which determines the possibility of solving not all the problems of the subject $C_в$ within the problem Π, but only the problem Π itself. The minimum possible value of the efficiency of solving the problem Π by the subject $C_в$ is denoted by $\bar{y}_в^*$.

Then the solution of the problem Π involves finding such elements of the solution of all subjects $C_1, C_2, ..., C_n$ that would ensure the fulfillment of the criteria

$$y_i \to \max \left(i = \overline{1,n} \right), \qquad (1)$$

$$\bar{y}_в \to \max \qquad (2)$$

and satisfied the conditions

$$y_i \geq y_i^* \left(i = \overline{1,n} \right), \qquad (3)$$

$$\bar{y}_в \geq \bar{y}_в^*. \qquad (4)$$

3. Analysis of the features of the problem

The features of the problem are as follows.

1. Entities $C_1, C_2, ..., C_n$ are partners. They can adjust their decisions in the interests of another entity (one or more).
2. The number of indicators-elements of the decision of the subjects $C_1, C_2, ..., C_n$ may be different. That is, the numbers $1C_1, 2C_2, ..., nC_n$ are not necessarily equal.
3. Target functions $y_i \left(i = \overline{1,n} \right)$ subjects $C_1, C_2, ..., C_n$ are not represented as analytical dependencies.
4. The problem under study may be set in such a way that the decision $X_i \left(i = \overline{1,n} \right)$ may contain indicators $x_{i1}, x_{i2}, ..., x_{iC_i} \left(i = \overline{1,n} \right)$ that relate to different elements of the decision of one entity, or the problem can be set so that decisions

$X_j\left(j=\overline{1,m}\right)$ can contain a set of elements of decisions of all entities $C_1, C_2, ..., C_n$ on individual indicators.

An example of the problem under study may be as follows.

The following subjects $C_1, C_2, ..., C_n$ of the security and defense sector of Ukraine act as subjects:

C_1 - Armed Forces of Ukraine (AFU);

C_2 - State Border Guard Service of Ukraine (SBGSU);

C_3 - National Guard of Ukraine (NGU);

C_4 - Security Service of Ukraine (SSU).

The problem Π is the activity of the sabotage and reconnaissance group (SRG) in the border area.

The solution of problem Π is to localize the actions of the SRG.

The solution $X_i\left(i=\overline{1,n}\right)$ is the following set of indicators:

x_{i1} - quantitative characteristics (for example, the percentage) of the personnel of the active division of the entity $C_i\left(\mathrm{i}=\overline{1,4}\right)$, which operates within the problem Π and may be involved in its solution;

x_{i2} - quantitative characteristics of the entity's personnel $C_i\left(\mathrm{i}=\overline{1,4}\right)$, which may be additionally involved in solving the problem Π;

x_{i3} - quantitative characteristics (for example, the percentage) of technical means (weapons and military equipment) of the active unit of the subject $C_i\left(\mathrm{i}=\overline{1,4}\right)$, which are used within the problem Π and can be involved in its solution;

x_{i4} - quantitative characteristics of technical means (weapons and military equipment) of the subject $C_i\left(\mathrm{i}=\overline{1,4}\right)$, which may be additionally involved in solving the problem Π;

x_{i5} - the possible duration of the involvement of resources (personnel and technical means) of the active unit of the subject $C_i\left(\mathrm{i}=\overline{1,4}\right)$ to solve the problem Π;

x_{i6} - the possible duration of the involvement of additional resources (personnel and technical means) of the subject $C_i\left(\mathrm{i}=\overline{1,4}\right)$ to solve the problem Π.

Functions $y_i\left(i=\overline{1,4}\right)$ are the effectiveness of the functions of the AFU, SBGSU, NGU, SSU, respectively, within the problem Π.

4. The author's approach to solving the problem under study, taking into account the peculiarities of its formulation

The author's approach to solving the research problem in the absence of analytical representation of objective functions involves the implementation of two stages: first, the establishment of analytical dependences for objective functions $y_i\left(i=\overline{1,n}\right)$; secondly, directly the search for a rational compromise, i.e. solving problem (1) - (4).

The method of realization of the second stage in the presence of analytical representation of target functions $y_i\left(i=\overline{1,n}\right)$ is described in detail in work [4].

Thus, the main problem of solving the problem is the implementation of the first stage, ie the establishment of functional dependencies $y_i\left(i=\overline{1,n}\right)$.

It should be noted that the problem of reproducing functional dependences on an experimentally obtained discrete sample is one of the rather important applied tasks, both operations research and system analysis [1, 6-7]. Its complexity is due to the fact that at the initial stage of formation of the concept and design of the operation can be known only incomplete, heterogeneous input information, which can act as empirical data, expert assessments, a priori information about analogues and prototypes, some information about the purpose and quality indicators. standard restrictions and data, etc. It is these data that can determine the arguments of the desired functional dependencies that determine the target functions of the partners in the problem studied in this paper. However, under these conditions, the choice of analytical forms of objective functions, justification of their content and purpose is an informal procedure that can be performed only by the researcher. The peculiarity of the problem is that the required functions should not only be as close as possible to the empirical data on a certain criterion, but also have extreme properties. The specificity of the extreme properties is due to the limited interval of setting the original data and is that the perturbations at the limits of the interval significantly affect the extreme properties of the function. This feature is fundamental and causes a more complex structure of approximation functions than in interpolation problems [8-14]. Another important feature is the need to choose a rational compromise between conflicting requirements: maximizing the level of reliability of the procedure for identifying the desired pattern, which necessitates increasing the complexity of the class of approximation functions, and minimizing the complexity and complexity of the procedure for forming the required dependence. Due to the poor choice of functions, it may happen that the reproduced function will approximate certain output data for most of a given interval, but in general will poorly describe the true functional dependence. This explains the fact that to reproduce the functional dependencies $y_i\left(i=\overline{1,n}\right)$ should use the methods described in [15-17].

The use of software and algorithmic software for the implementation of these methods, which is given in [5], allows you to automate the process of finding functions $y_i\left(i=\overline{1,n}\right)$. However, it should be noted that the application of such an approach has some difficulties.

To assess them, we analyze the initial data for the reproduction of functional dependencies, which are given in table. 1.

If all the data in the table 1 method of reproduction of functional dependencies, which is algorithmized in [5], is applied directly. If the table. 1 there are no individual data or their whole blocks, it is considered appropriate to pre-fill the table. 1, taking into account the minimum possible $\min x_{ij}$, maximum possible $\max x_{ij}$, or most probable x_{ij} values $\bar{x}_{ij}$ $\left(i=\overline{1,n},\ j=\overline{1,C_i}\right)$, that characterize the elements of the decision of the subjects $C_1, C_2, \ldots, C_n$. Filling in the table. 1 data should be carried out taking into account the physical content of the elements of the decision, their physically possible values and the forecast of the possibility of implementation in a particular case. After filling in the data of table. 1 then you should use the software product developed in [5]. It should be noted that the proposed approach is based on the reproduction of functional dependence in the additive form. The required functions $y_i\left(i=\overline{1,n}\right)$ are represented as a superposition of functions from variables $X_i\left(i=\overline{1,n}\right)$.

This approach is quite possible when the components of the vectors $X_i\left(i=\overline{1,n}\right)$ are independent. If the components of vectors $X_i\left(i=\overline{1,n}\right)$ are dependent, and this is possible in a number of applied problems, then the application of this approach can lead to large deviations of the obtained dependences from the real multifactor regularities, because the effects of vector components $X_i\left(i=\overline{1,n}\right)$ on the properties of functions $y_i\left(i=\overline{1,n}\right)$ will not be taken into account.

Therefore, in this case, when forming the structure of the models should take into account not only the influence on the functions $y_i\left(i=\overline{1,n}\right)$ of a certain group of factors $X_i\left(i=\overline{1,n}\right)$, but also the mutual influence of the components of different vectors $X_i\left(i=\overline{1,n}\right)$. This explains the need for the formation of a hierarchical multilevel system of models in the class of multiplicative functions.

Table 1

Initial data for reproducing analytical dependencies $y_i\left(i=\overline{1,n}\right)$

Empirical data set number (experiment)	X_1				…	X_n				y_1	y_2	…	$y_в$	$\bar{y}_в$	…	y_n
	x_{11}	x_{12}	…	x_{1C_1}	…	x_{n1}	x_{n2}	…	x_{nC_n}							
1																
2																
3																
…																
q_0-1																
q_0																
$\min x_{ij}$ $\left(i=\overline{1,n},\ j=\overline{1,C_i}\right)$	The minimum possible values x_{ij} for each column from a physical point of view									The minimum allowable limit level of values $y_i^*\left(i=\overline{1,n}\right)$						
$\max x_{ij}$ $\left(i=\overline{1,n},\ j=\overline{1,C_i}\right)$	The maximum possible values x_{ij} for each column from a physical point of view									y_1^*	y_2^*	…	$y_в^*$	$\bar{y}_в^*$	…	y_n^*
$\bar{x}_{ij}$ $\left(i=\overline{1,n},\ j=\overline{1,C_i}\right)$	The most probable values x_{ij} for each column															

5. Conclusions

Based on the results of the study, it can be concluded that the proposed approach to solving the problem of finding a rational compromise with the interaction of any number of partners and any number of arguments that determine their target functions, but in the absence of analytical dependencies of these functions component of the decision elements $X_i\left(i=\overline{1,n}\right)$ of the subjects $C_1,C_2,...,C_n$. It provides a combination of capabilities of the previously developed author's method and software product for the reproduction of functional dependencies in the class of additive functions.

Prospects for further research include the development of information technology, which would provide full automation of the proposed approach, development of software applications to solve the problem in a multiplicative form and testing the approach to a specific application problem.

6. Aknowledgements

The work was performed within the framework of joint research of the Department of General Scientific and Engineering Disciplines and the Department of Telecommunication and

Information Systems of the National Academy of the State Border Guard Service of Ukraine.

7. References

[1] M. Z. Zghurovskyi, N. D. Pankratova, Fundamentals of systems analysis. – K.: Publishing group VNV, 2007. – 544 p.

[2] O. V. Borovyk, L. V. Borovyk, L. O. Matokhniuk, Higher Mathematics: Differential calculus of a function of one and many variables. Part 2: textbook. – Khmelnytskyi: publishing house NABGSU, 2016. – 640 p.

[3] V. E. Hmurman, Theory of Probability and Mathematical Statistics. – M.: High school, 1972. – 304 p.

[4] O. V. Borovyk, D. O. Borovyk, Adequate knowledge of the scientific-methodical apparatus for the development of non-significance in the tasks of interaction // Library of Science Practitioners of Military Institute of Kyiv Taras Shevchenko National University. Issue № 55. – K: MIKNU, 2021. – P. 20-40.

[5] O. V. Borovyk, L. M. Traskovetska, Substantiation of the structure and functional features of the software application of the Chebyshev approximation problem as a component of the problem of reproduction of functional regularities by discrete sampling // Collection of scientific works № 57. Series: Military and technical sciences. – Khmelnytskyi: Publishing house. NABGSU, 2012. – P. 86-91.

[6] E. S. Venttsel, Operations research. – M.: Soviet radio, 1972. – 552 p.

[7] V. A. Abchuk, F. A. Matveichuk, L. P. Tomashevskyi, Handbook of operations research / Under the general editorship of F.A. Matveichuk – M.: Voenizdat, 1979. – 368 p.

[8] Greene, H. William, (2002). Econometric analysis (5th ed.). New Jersey: Prentice Hall.

[9] Least squares method // Higher mathematics in examples and problems: Textbook. 2nd edition. - Kyiv: Center for Educational Literature / Klepko V.Iu., Holets V.L.. – 2009. – 594 p.

[10] N. S. Bobrovnycha, Ye. H. Borysevych, Econometrics: textbook / N. S. Bobrovnycha, Ye. H. Borysevych – Odessa: O. S. Popov ONAC, 2010. - 180 p.

[11] G. G. Judge, and ets. The Theory and Practice of Econometrics edition / Judge, G. G. New York: John Wiley, 1985. – P. 3-120.

[12] Computational mathematics. Interpolation and approximation of tabular data: a textbook / L. V. Krylyk, I. V. Bohach, M. O. Prokopova. – Vinnytsia: VNTU, 2013. – 111 p.

[13] Continuous and smooth minimax spline approximation / P. S. Malachivskyi, V. V. Skopetskyi ; [resp. ed. I.V. Sergienko]; NAS of Ukraine, V. M. Hlushkov Institute of Cybernetics, Center of Mat. modeling Inst of applied. probl. Mechanics and mathematics named after Ya. S. Pidstryhach. - K.: Scient. opinion, 2013. – 271 p.

[14] B. P. Dovhyi, A. V. Loveikyn, E. S. Vakal, Yu. E. Vakal, Spline functions and their use. Textbook of Taras Shevchenko Kyiv National University. – K: KNU, 2016.

[15] A. O. Kalenchuk-Porkhanova, L. P. Vakal, Construction of the best uniform approximations of functions of many variables // Computer tools, networks and systems. − 2007. − № 6. P. 141−148.

[16] K. Lantsosh, Practical methods of applied analysis. – M.: Physics and math. lit., 1961. – 524 p.

[17] E. Ya. Remez, Fundamentals of numerical methods of Chebyshev approximation. − K.: Scient. opinion, 1969. - 623 p.

Characteristics of Multi-agent Systems

Dmytro Tolbatov [1], Serhii Holub [2]

[1] *Institute of Mathematical Machines & Systems Problems of the NAS of Ukraine, Academic Glushkov 42, Kyiv, 03187, Ukraine*
[2] *Cherkasy State Technological University, Shevchenko 460, Cherkasy, 18000, Ukraine*

Abstract
In today's world, multi-agent systems are widely used in various industries. They consist of various separate programs (agents) that interact with each other to solve problems. Multi-agent systems are very closely intertwined in the field of use with artificial intelligence, although there is no complete identity between them.

Keywords
Multi-agent system, Intelligent agent, GMDH

1. Introduction

Intelligent agent has some level of autonomy that allows it to perform specific, predictable, and repetitive tasks for users or applications.

It’s also termed as ‘intelligent’ because of its ability to learn during the process of performing tasks.

The two main functions of intelligent agents include perception and action. Perception is done through sensors while actions are initiated through actuators.

Intelligent agents consist of sub-agents that form a hierarchical structure. Lower-level tasks are performed by these sub-agents.

The higher-level agents and lower-level agents form a complete system that can solve difficult problems through intelligent behaviors or responses [3]. But often the systems are considered not separately from each other but in combination in a multi-agent system (MAS).

2. Multi-agent Systems

An agent is essentially an abstraction of a complex program that operates autonomously through the performance of certain tasks by the user. Each object in the agent's system has a detailed description of its behavior. The main characteristics of the agent are survivability (the code is constantly working and decides when to take action), autonomy (the agent makes decisions without human intervention), social behavior (they can be involved through other components) and reactivity (perceive and respond to the context in which they exist) [2].

Multiagent systems are made up of multiple interacting intelligent agents—computational entities to some degree autonomous and able to cooperate, compete, communicate, act flexibly, and exercise control over their behavior within the frame of their objectives. They are the enabling technology for a wide range of advanced applications relying on distributed and parallel processing of data, information, and knowledge relevant in domains ranging from industrial manufacturing to e-commerce to health care [4].

There are two reasons that led to the development of multi-agent systems. The first is the use of computer science. There are more and more different technical devices in the world, such as computers, servers, mobile phones, tablets, which in turn can be connected to the network, and those to the Internet. Previously, the number of connection points was less than the number of users, but today each of us can have several different devices. Computing resources are improving day by day, but the amount of data is growing even faster. All this contributes to the complexity of systems and their algorithms. To facilitate data handling, systems are divided into

EMAIL: dmytrotolbatov@gmail.com (A. 1); s.holub@chdtu.edu.ua (A. 2);
ORCID: 0000-0001-6418-2075 (A. 1); 0000-00002-5523-6120 (A. 2);

smaller subsystems. It is to solve such problems that there are multi-agent systems that study them at a high level of abstraction. The second reason is the development of society. Multi-agent systems play a significant role in the analysis of models of human interaction in different situations. People can independently predict the behavior of other people, conduct negotiations and discussions, resolve conflicts, form organizational structures. All this can also be analyzed and used by a smart system. This is done by an intelligent agent who can make decisions or execute assignments based on experience, nested data and environment. They can also be used to collect real-time information [1].

The agent must receive, store and process information about the current state of the subject area. An agent's view of the world around him, of other agents, and of himself can be specified in the form of an ontology. Definition and construction of ontology includes analysis of the subject area, selection of basic ontological elements (objects, their attributes, relations and processes), operations on these ontological elements. If agents operate on the same concepts, it will help solve many problems - from methods of communication between agents to ways to adapt to new conditions.

MAS has the following characteristics:

- each particular agent does not have enough information or ability to solve the problem, and thus he does not have a complete vision of the global task to be performed;
- there is no global control in the system, ie there are no agents controlling the whole system;
- no centralized data storage;
- agents, at least in part, independent;
- limited representation, ie none of the agents has an idea of the whole system or the system is too complex for knowledge of it to have practical application to the agent;
- calculations occur asynchronously.

Usually in multiagent systems software agents are investigated. However, MASs can also be robots, people, or teams of people. MASs can also contain mixed commands.

MAS can exhibit self-organization and complex behavior, even if each agent's behavioral strategy is simple enough. This is the basis of the so-called ant algorithms.

There is also an extended concept.

Agent-oriented systems (AOC) - hybrid systems that contain MAS and other information systems (expert, training, testing, decision support systems (DSS), distributed object programs, etc.).

The ideas of software agents in general and intelligent agents in particular are attractive because they allow people, as already mentioned, to delegate their powers. However, the development of MAS and truly intelligent agents requires special knowledge and is a complex resource-intensive task. After all, software agents are a new type of software that acts on behalf of the user. They are a powerful abstraction for visualizing and structuring the complex in the real world [5].

3. References

[1] Holub S.V. Bahatorivneve modeliuvania v tehnolohiiah monitorynhu otochuuchoho seredovysha. Monography. Cherkasy: Vyd. vid. Bo, 2007. – 220 c. ISBN 978-966-353-062-8

[2] Ivakhnenko A., Stekasko V.: Impulse immunity modeling. Kyiv: Scientific Opinion, pp 216. (1985)

[3] Mbaabu O. Intelligent Agents in Artificial Intelligence [Electronic resource] / O. Mbaabu. – Available from: https://www.section.io/engineering-education/intelligent-agents-in-ai/

[4] Weiss G. Intelligent Agents in Artificial Intelligence [Electronic resource] / G. Weiss. – Available from: http://www.the-mas-book.info/

[5] Multi-agent Systems [Electronic resource] / – Available from: https://stud.com.ua/163996/informatika/multiagentni_sistemi

Improving the process of carrier frequency estimation in modern satellite telecommunications

Valerii Kozlovskyi[1], Oleksandr Turovsky[2], Kateryna Nesterenko[3], Sergii Lazarenko[4],Yuriy Balanyuk[5]

[1]*National Aviation University, 1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058*
[2]*National Aviation University, 1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058*
[3]*National Aviation University, 1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058*
[4]*National Aviation University, 1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058*
[5]*National Aviation University, 1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058*

Abstract
Solving the problems of national security of the state, ensuring the functioning of its management system and maintaining sustainable economic development through effective communication requires the constant development of all types of telecommunications. Including modern efficient satellite communications. It is known that the efficiency of satellite telecommunications systems in general is directly affected by the efficiency of each component and device x of its structure.
The article directly considers the issue of increasing the efficiency of frequency estimation at the stage of synchronization of the input signal. Namely, the analysis of modern satellite telecommunication systems is carried out and the features that influence the efficiency of carrier frequency estimation and input signal synchronization are determined. The main discrepancies are the low energy of the input signal reception channel, significant frequency uncertainty of the signal and the presence of "adjacent channels" of the input signal reception. In general, multi-station satellite access systems with frequency division multiplexing are used, which operate in the mode of multiple access and when providing the channel on demand, as well as work in both continuous and packet modes of input signal reception. It is also shown that the demodulator of satellite telecommunications systems performs coherent processing of the input signal and uses noise-tolerant coding.
To solve the problem of estimating the carrier frequency of the input signal in coherent demodulators of modern satellite telecommunications, taking into account the above inconsistencies and features, the article proposes a methodological approach. The essence of this approach is to determine the potential variance limit of the obtained carrier frequencies and develop algorithms for estimating the carrier frequencies of the input signal separately for continuous and packet modes of input data based on fast Fourier transform.
Keywords: satellite telecommunication systems, satellite communication channel energy, frequency signal uncertainty, carrier frequency estimation, fast Fourier transform.

1. Introduction

Satellite communication is intensively used in the implementation of important national projects, including for the effective solution of national security problems, with the aim of socio-economic development of states and successful international cooperation.

Problems of signal transmission in modern satellite systems are determined by certain features of both the construction of the system itself and the problems of signal reception and transmission [1,2,3].

Direct input signal processing in modern satellite telecommunications includes coherent demodulation. And the development of algorithms for coherent signal demodulation involves the primary solution of the synchronization problem [4].

Complex envelope of the input signal of satellite telecommunications $z(t)$ - contains unknown values: carrier frequency offset, phase noise and time offset, respectively (ν, φ, τ).

The task of processing the input signal, which begins with its synchronization, is actually

EMAIL: vvkzeos@gmail.com (A.1); s19641011@ukr.net (A.2); nesterenko.kateryna.s@gmail.com (A.3); zzi.lazarenko@nau.edu.ua (A.4); alexanderturovsky@gmail.com (A.5);
ORCID: 0000-0002-8301-5501 (A.1); 0000-0002-4961-0876 (A.2); 0000-0001-7672-7386 (A.3); 0000-0003-3529-4806 (A.4); 0000-0003-3036-5804 (A.5)

reduced to the initial assessment of the true parameters of the specified signal ν, φ, τ, knowledge of the parameters of which is necessary for demodulation of the signal $z(t)$ [5,6].

The best results can be obtained by a joint assessment of unknown signal parameters.

However, in practice, such an assessment is possible only taking into account the values of the channel energy and the magnitude of

frequency uncertainty. Which, in turn, can lead to the use of a process in which the assessment of the shift of the carrier frequency of the signal received relative to the nominal value is carried out before other synchronization procedures are included, namely: phase synchronization and clock synchronization [3,5,7].

The complexity of the task of estimating the carrier frequency in the satellite channel is exacerbated by the possible presence of additional artificial and natural disturbances and radio space noise. Especially when these factors are close to the main signal type of modulation and singing measured data rate. Also crucial in the implementation of the algorithm for estimating the carrier frequency is the mode of transmission of useful data [3,6,7].

2. Analysis of research and publications

A number of works are devoted to the definition of the main features of modern satellite systems in the direction of increasing the efficiency of useful data transmission and directly to the issues of estimating the carrier frequency of the input signal.

In [2], the main direction of increasing the efficiency of modern satellite communication systems is justified by increasing the power of channel energy.

In [4], the author proposes to use low-orbit satellite systems, which somewhat devalues the shortcomings of geostationary systems. However, the issue of conducting a general analysis of the process of carrier frequency estimation in geostationary satellite systems is not presented in the paper.

In [8], the results of the study of multi-symbol differentiated detection of signals of different coding with phase shift in the presence of random frequency changes and additive white Gaussian noise are presented. It is determined that the distortion of the transmitted signals due to the attenuation of the amplitude and the introduction of a variable phase time to the information symbols can be caused by a change in frequency. To eliminate this distortion, a scheme of double differential phase shift coding PSK (DDPSK) and the following multi-character differentiated phase shift detection in the presence of frequency shift is proposed. But the article does not estimate the frequency uncertainty and does not take into account the possible modes of transmission of useful data.

The authors of [9] propose a joint assessment of the synchronization and shift of the carrier frequency and the detection of data using a filter of signals ranked by importance in the channels of additive white Gaussian noise. In this paper, we propose a weighted Bayesian Kramer Rao-boundary (WBCRB) for joint determination of time and carrier frequency shift, which takes into account the previous distribution of estimation parameters and is the exact lower limit for all considered signal-to-noise ratio (SNR). The issue of taking into account artificial and natural obstacles is not taken into account in this paper.

The work [10] combines the algorithm for estimating the frequency shift (FO) with the method of estimating phase noise. The estimation of phase noise is derived from the calculated coefficients of discrete-cosine transformation. A number of implementations of the proposed algorithm are analyzed. However, the modes of transmission of useful data and the impact of various obstacles on them are not taken into account.

Thus, a preliminary analysis of the process of functioning of modern satellite telecommunications and the development of a methodological approach to estimating the carrier frequency in them, taking into account the results of the above analysis is an urgent scientific problem to which this article is devoted.

3. Analysis of the process of functioning of satellite telecommunications systems

Analysis of the energy values of the satellite communication channel requires the implementation of coherent signal processing against the background of the use of noise-tolerant coding. The noise-tolerant coding system is one of the main parts of the satellite modem. The vast majority of systems use methods of high-precision coding with decoding by the Viterbo algorithm, as well as cascading codes and other methods [6,11,12,13,14]

Also widespread are turbo codes and codes with low density of parity checks [11].

Based on the low energy of the input signal of the satellite channel, satellite telecommunications systems mainly use phase modulation [1,2].

The most commonly used three types of phase modulation: FM-2, FM-4 and FM-8. Most ground stations of satellite communication systems operate in the bands 6 or 14 GHz for transmission and 4 or 11 GHz for reception [3,15,16]. The simplest satellite communication channel includes two satellite earth stations and a space communication station located on board the satellite. Note that there is at least one signal frequency conversion in this channel. This transformation is carried out in the on-board repeater of the communication satellite [1,2,15].

The presence of this transformation due to the internal instability of the reference generator-frequency converter of the on-board repeater of the communication satellite leads to accidental and non-stationary displacement of the carrier oscillation of the signal relative to the nominal value [15]. This causes the frequency uncertainty of the signal.

The entire frequency range in which the satellite repeater operates is divided into some bands (width 27 ... 36, 72 ... 120 MHz), in which the signal is amplified by a separate path - the trunk. The trunk, in turn, can transmit signals from many satellite earth stations. Thus, the satellite repeater can provide communication to a large number of subscribers. Therefore, the organization of access to virtually independent earth stations in the general communication system and the rapid establishment of connections between arbitrary stations and multi-station access are widely used in satellite telecommunications [5,15,16].

There are several different ways in which a large number of users can send information via satellite. Currently, two types of multi-station access are widely used in satellite telecommunications [12,13]:

– multi-station access with frequency division multiplexing;

– multi-station access with temporary division of channels.

In the first case, network subscribers separated by the carrier frequency, in the second subscribers have a common frequency resource, but separated in time of use of the common frequency path.

On the basis of these methods of channel division, as a rule, satellite telecommunications are built with the provision of a channel on demand. At the beginning of the communication session, in a certain way, the signals are switched at the respective satellite communication stations, the subscriber is allocated a frequency channel or a temporary position. In such communication networks it is possible to refuse to establish communication due to the busyness of all channels of the system.

The relative simplicity and low cost of equipment, as well as extensive experience in the development and operation of frequency division systems, gained during the development and operation of early communication systems, have led to the fact that the vast majority of satellite telecommunications are currently used multi-station access with frequency division multiplexing. Satellite telecommunications with channel on demand, operating with frequency division multiplexing, operate at relatively low information rates.

The relative simplicity and low cost of equipment, as well as extensive experience in the development and operation of frequency division systems, gained during the development and operation of early communication systems, have led to the fact that the vast majority of satellite telecommunications are currently used multi-station access with frequency division multiplexing. Satellite telecommunications with channel on demand, operating with frequency division multiplexing, operate at relatively low information rates.

As a result, in such systems it is possible to use relatively cheap satellite terminals with a small aperture class VSAT (very small aperture terminal) [12].

Narrowband channels are used in multi-station access with temporary division of channels systems. Therefore, the initial shift in the frequency of the carrier oscillation of the signal can be comparable to the bandwidth of the channel [15,16].

The structure of satellite telecommunications with frequency division multiplexing indicates another fundamental feature of the receiving channel - the presence of additional interfering effects, namely the so-called "neighboring channels" - signals with the same type of modulation and the same transmission rate as the main information channel. Since the uncertainty of the carrier oscillation frequency of the signal can be comparable to the data rate in the communication channel, a significant part of the spectrum of the "neighboring channel" can fall within the search range of the demodulator on the carrier frequency [16,17].

For satellite telecommunications with multi-station access with temporary division of channels, a control channel is required, the function of which includes the allocation of a free frequency resource to the subscriber. This channel is not symmetrical. The common channel from the central station to the peripheral stations of the network operates in continuous mode For the control information of each peripheral station of the network in this common channel is allocated its own temporary position. Signals from peripheral stations to the central station of the network are transmitted in packet mode at random times, occupying one fixed frequency channel. As a result, signals from many users cannot be separated in the receiver during simultaneous transmission [4,5].

Such channels use random access methods. When multiple network users try to transmit packets at the same time, the packets overlap over time (overlapping.) The overlap conflict must be resolved by using some channel protocols to retransmit the packets. In satellite telecommunications, the most widespread protocol is ALOHA [4,16].

This control channel is usually called a common alarm channel [4,15].

In order to analyze the carrier frequency estimation process, we determine the allowable variance of the carrier frequency estimation and analyze the known methods for estimating the carrier frequency of the phase-modulated signal (FM signal).

It should also be noted that in [1,2,15,16] the main attention is paid to the problem of information packet detection, and the algorithm for estimating the carrier frequency of the phase modulated signal is described insufficiently.

The coherence of signal processing in the demodulator causes extremely strict requirements for the accuracy of estimating the phase of the carrier oscillation in the corresponding loop of the demodulator with phase-automatic frequency tuning [17,18].

Because of this, the band bandwidth B_C should not normally exceed thousandths of the clock frequency of the received signal - about 10^{-3} 1 / T [17,18].

The values B_C are usually selected for reasons of compromise between the level of energy loss of the demodulator in a given range of signal-to-noise ratio per bit of information and the timing of the synchronization of the phase synchronization loop.

The smaller B_C the lower the level of energy loss of the demodulator, the longer the entry into synchronization. The value of the loop band is selected so that the additional energy loss of the demodulator does not exceed 0.1 dB.

Based on these considerations, the size of the loop is usually chosen in values $B_C \approx 3 \ {*}10^{-3}\, {}^{1}\!/\!_{T}$.

And for reliable entry of the auto tuning system into synchronization, the value of the variance of the estimation of the carrier frequency of the FM signal δ_C^2 should not exceed B_C.

That is, the value δ_C^2 should not be greater than 10^{-5} 1 / T 2 [19,20].

In the future, we assume that for the variance of the FM signal carrier frequency estimation at low signal-to-noise ratios per bit of information (from 0 to 12 dB) the following requirement must be met [5]:

$$\delta_C^2 T^2 \le 5 \ {*}10^{-6}. \qquad (1)$$

Given that the assessment of the carrier frequency in continuous mode and batch mode is carried out in fundamentally different ways, it is advisable to consider the evaluation methods in continuous and batch modes separately.

It is known that satellite information transmission systems work, including in modes with random access of signal packets [16].

That is, for them it is important to synchronize coherent phase demodulators operating in batch mode.

The synchronization methods for this case, described in some works, are intended for demodulators of communication systems that work with temporary channel allocation. The main disadvantage of these works is that we consider relatively small in relation to the bandwidth of the transmitted signal displacement of the carrier oscillation, while in real modern communication systems, these displacements can be correlated with the clock frequency of the received signal [5,18 , 19, 21].

To synchronize coherent phase demodulators operating in packet mode, we assume that to synchronize the packet demodulator on the carrier frequency at the beginning of the preamble is transmitted a segment of the harmonic signal.

The complex envelope of the received signal can be given as [5,19,21]:

$$z(t) = e^{j\left(2\pi\nu+\varphi\right)} + \nu(t), \qquad (2)$$

where ν – the carrier frequency offset relative to the nominal value;

$w\left(t\right)$ – complex additive Gaussian noise.

Note that coherent reception gives an advantage of noise immunity up to 1 dB for one-way phase-difference modulation when used in modern satellite communication systems of multiple modulation, when the number of signal positions reaches 512 signal variants, the gain can be 5 – 7 dB.

Thus, it is promising to consider for this case coherent methods of reception.

The task of estimating the carrier frequency of the received signal is reduced to the problem of estimating the frequency of the maximum in the spectrum of a fragment of a sinusoidal signal against the background of additive Gaussian noise, which follows from (2).

Currently, there are many methods for estimating the frequency of sinusoids. The work [23] is quite indicative.

It is known that the smallest estimation variance is determined by the Cramer – Rao boundary [5]. It seems that from all the variety of frequency estimation methods, one should choose those that provide an estimation variance that coincides with the minimum marginal estimation variance (MHD estimation) or close to it.

The analysis of effective estimates provided by different methods of estimating the frequency of a sinusoidal signal against the background of additive white noise provides a variance of the estimate that asymptomatically coincides with the Cramer – Rao boundary with increasing signal-to-noise ratio.

The estimate of the maximum similarity (MP - estimate) of the frequency of the sinusoidal signal is determined by the expression [5,19,21]:

$$\nu = \arg\left\{\max_{\nu}\left\{I(\nu)\right\}\right\}, \qquad (3)$$

where $I(\nu)=\left|\sum_{n} z(t_n) * e^{-j2\pi\nu t_n}\right|$

$t_n = nT_s$;

$n = 1,2...,N$;

t_S – the period following the calculation of the complex bending signal received;

NT_S – observation interval.

In fact $I(\nu)$, it is the Fourier transform module of the signal that is captured in the observation interval.

The value $I(\nu)$ is called the periodogram of the received signal.

Thus, finding the MP - estimation of the frequency of the sinusoidal signal is reduced to finding the maximum of the function $I(\nu)$ in the range of frequency uncertainty of the received signal. Given that the satellite communication signal is characterized by significant frequency uncertainty, finding the maximum from expression (3) requires significant computational problems.

Because of this, it seems appropriate to analyze the problems of implementing "fast" methods of finding the maximum.

Currently, the most widely used two approaches to finding the maximum of the periodogram of the signal.

The first approach is based on the use of auto-correlation analysis methods, the second - on the use of discrete Fourier transform algorithms.

The advantage of the first approach is a relatively simple calculation procedure. However, at low signal-to-noise ratios, methods based on a discrete Fourier transform are more efficient.

The procedure for finding the maximum of a periodogram based on a discrete Fourier transform includes [21,24]:

1. Calculation of fast Fourier transform (FFT) N_f and finding the maximum:

$$m_f = \arg\max_{k}\left\{|Ik|\right\},$$

where

$$I_k = \frac{1}{N_f}\sum_{n=0}^{N_f-1} z(t_n) * \exp\left(-\frac{j2\pi nk}{N_f}\right).$$

2. Finding the maximum $I(\nu)$ closest to km_f, determining $\overline{\nu}$ such that

$$\overline{\nu} = \arg\left\{\max_{\nu}\left\{I(\nu)\right\}\right\}$$

The above procedure is actually a two-step.

The first stage procedure is usually accepted as a rough evaluation procedure. Its accuracy is limited by FFT [5, 21,24].

To implement the second stage, you can use a number of methods, which are defined as methods of interpolation of calculations of the Fourier transform of the received signal. These methods are characterized by sufficient computational simplicity. However, the variance of the frequency estimation provided by interpolation methods at low signal-to-noise ratios loses the Kramer – Rao boundary.

Therefore, to implement the estimation algorithm with the highest efficiency as a procedure of the second stage, it is advisable to use the methods of finding the maximum, which are also implemented on the basis of Fourier transform [19, 21].

A common disadvantage of the known evaluation procedures based on the use of FFT is that their implementation requires the initial accumulation of data for the subsequent implementation of the calculation algorithm.

This provides an additional delay in the implementation of the evaluation procedure.

In batch mode, the additional delay is a significant disadvantage, as the limited length of the preamble of the packet imposes strict restrictions on the long duration of the processing procedure. Of interest are estimation methods based on recurrent computational procedures that combine the processing and accumulation of the received signal.

However, the variance of estimates that provide these methods, as shown in [19,25], significantly loses the Cramer – Rao boundary.

Currently, there are a number of methods for estimating the frequency of a sinusoidal signal, which is based on recurrent procedures. Such as Pisarenko method, MUSIK method, auto regression method [19,25].

Given the satellite channel (high frequency uncertainty, low energy, the presence of interfering actions in the form of Gaussian noise and interference such as "neighboring channel") and the need for coherent signal processing to date is a pressing problem of digital phase demodulator synchronization.

In the implementation of synchronization procedures in the satellite channel, the determining factor is the synchronization of the carrier frequency. The task of developing an algorithm for estimating the carrier frequency of the FM signal in coherent demodulators of modern satellite communication systems is of great importance.

The task of the study is essentially to develop an evaluation algorithm that takes into account the properties of the satellite communication channel, and to conduct a study of the effectiveness of the proposed solutions.

4. Formation of a methodical approach to improve the process of estimating the carrier frequency in modern satellite telecommunications

Thus, as the main methodological approach to improve the process of estimating the carrier frequency of satellite telecommunications is proposed:

1. Develop an algorithm for estimating the carrier frequency of the FM signal for coherent demodulators of satellite communication systems, which would provide optimal or close to optimal estimation.

This requires:

- to determine the potential boundaries of variances of the obtained estimates;
- to develop an algorithm for estimating the carrier frequencies of the FM signal, taking into account the specifics of the satellite communication channel;
- to develop implementation evaluation procedures based on FFT, while for the packet demodulator to offer such a procedure for estimating the carrier frequency of the received signal, which would provide, on the one hand, simplicity of hardware implementation, and on the other hand - optimal asymptotic estimation.

2. Investigate the effectiveness of estimates obtained using the proposed procedures for estimating the carrier frequency of the FM signal, by computer simulation.

This requires:

- to develop simulation models of signals and interference, taking into account the specifics of the satellite communication channel, including the presence of "neighboring channels";
- to construct dependences of the received variances of estimations on a signal / noise ratio on a bit of the information and to compare the received variances with potential borders;
- in the process of modeling to determine the minimum required length of the observation interval, which would provide the required evaluation efficiency (fulfillment of condition (1)) in the range of signal-to-noise ratio per bit of information (from 0 to 12 dB).

3. Show the feasibility of the proposed algorithms and their effectiveness in existing satellite telecommunications systems.

This requires:

- to show the hardware implementation of the developed procedures for estimating the carrier frequency of the FM signal on the basis of modern digital signal processors and thus show the feasibility of the developed procedures in real time;
- to conduct bench tests of the proposed evaluation algorithms;
- to conduct field tests of equipment using the proposed algorithms for estimating the carrier frequency of the FM signal in the existing satellite communication system.

It should be noted that the methodological approach defined in the work regarding the estimation of the carrier frequency of the input signal of satellite telecommunications can be applied in different volumes in various studies related to information processing.

This is evidenced by certain results published in [26,27,28,29,340,31,32, 33-37].

5. Conclusions

The considered modern satellite telecommunication systems are characterized by extremely low channel energy, significant frequency uncertainty of the signal and the presence of "neighboring channels".

To increase the efficiency of their application, coherent processing of the input signal in the demodulator and the use of powerful noise-tolerant coding is used, which involves the assessment of the carrier frequency with high requirements for its accuracy.

For satellite telecommunications systems with multi-station access, using the mode "frequency division multiplexing - providing a channel on demand" requires multiple access and operation, both in continuous and packet modes;

To estimate the carrier frequency of the FM signal of satellite telecommunications, it is necessary to develop an estimation algorithm for coherent demodulators, which would provide optimal or close to optimal estimates separately for the modes of continuous and packet data transmission.

6. References

[1] K.I. Kukk, Satellite communications: past, present, future. Moskow: Goryachaya liniya (2015) 262.

[2] Kryilov A.M. Satellite communication and broadcasting systems. State and development prospects. Moskow: Radio and communication. (2014) 182.

[3] O. Turovsky, V. Kozlovskyi, Y. Balanyuk, Y. Boiko, N. Lishchynovska, Consideration of limitations, which are formed by the input signal, on the phase error minimization process during carrier frequency tracking system of synchronization of radio technical device of communication. International Journal of Advanced Trends in Computer Science and Engineering. 9(5) (2020) 8922 – 8928. https://doi.org/10.30534/ijatcse/2020/290952020.

[4] A.S Dovgopoly, S.O. Ponomarenko, V.V. Tverdohlibov, O.O. Beloborodov, Improvement of satellite navigation systems of weapons and military equipment under conditions of intentional interference. Weapons and military equipment. 1(17) (2018) 67–71.

[5] O. Turovsky, O. Drobyk, A. Makarenko, O. Khokhlyuk, Estimates of the carrier frequency of the signal received by the satellite communication system in package mode. International Journal of Advanced Trends in Computer Science and Engineering. 9(3) (2020) 3223 – 3228. https://doi.org/10.30534/ijatcse/2020/115932020.

[6] Barabash Oleg, Laptiev Oleksandr, Tkachev Volodymyr, Maystrov Oleksii, Krasikov Oleksandr, Polovinkin Igor. The Indirect method of obtaining Estimates of the Parameters of Radio Signals of covert means of obtaining Information. International Journal of Emerging Trends in Engineering Research (IJETER). 8 (8) (2020) 4133 – 4139.

[7] O. Turovsky, Y. Khlaponin, M. H. Mohamed at al, Combined system of phase synchronization with increased order of astatism in frequency monitoring mode, Control, optimisation and analytical processing of social networks (COAPSN 2020): 2nd International CEUR Workshop Proceedings, Lviv, Ukraine, May 21, 2020. 2616(1) (2020) 53–62. http://ceur-ws.org/Vol-2616/paper5.pdf .

[8] A.M. Rabiei, N.C. Beaulieu, Frequency offset invariant multiple symbol differential detection of MPSK. IEEE Trans. Commun. 59(3) (2011) 652–657.

[9] A.A. Nasir, S Durrani, R.A. Kennedy, Particle filters for joint timing and carrier estimation: Improved resampling guidelines and weighted bayesian cramer–rao bounds. IEEE Trans. Commun. 60(5) (2012) 1407–1419.

[10] 10 J. Bhatti, N. Noels, M. Moeneclaey, Low–complexity frequency offset and phase noise estimation for burst–mode digital transmission in Proc. IEEE PIMRC. Toronto. (2011) 1662–1669.

[11] Oleg Barabash, Oleksandr Laptiev, Oksana Kovtun, Olga Leshchenko, Kseniia Dukhnovska, Anatoliy Biehun. The Method dynavic TF-IDF. International Journal of Emerging Trends in Engineering Research (IJETER), 8(9) (2020) 5713-5718. https://doi.org/10.30534/ijeter/2020/130892020

[12] Serhii Yevseiev, Roman Korolyov, Andrii Tkachov, Oleksandr Laptiev, Ivan Opirskyy, Olha Soloviova. Modification of the algorithm (OFM) S-box, which provides increasing crypto resistance in the post-quantum period. International Journal of Advanced Trends in Computer Science and Engineering

(IJATCSE). 9(5) (2020) 8725-8729.

[13] Oleksandr Laptiev, Oleh Stefurak, Igor Polovinkin, Oleg Barabash, Savchenko Vitalii, Olena Zelikovska. The method of improving the signal detection quality by accounting for interference. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (IEEE ATIT 2020) Conference Proceedings Kyiv, Ukraine, November 25-27. (2020) 172 –176.

[14] Oleksandr Laptiev, Vitalii Savchenko, Andrii Pravdyvyi, Ivan Ablazov, Rostyslav Lisnevskyi, Oleksandr Kolos, Viktor Hudyma. Method of Detecting Radio Signals using Means of Covert by Obtaining Information on the basis of Random Signals Model. International Journal of Communication Networks and Information Security (IJCNIS), 13(1) (2021) 48-54.

[15] INTELSAT. INTELSAT EARTH STANDARTS (ISS) Perfomance characteristics for broadband VSAT (BVSAT) digital carriers // Document IESS–313 (Rev.A) APPROVAL DATE: 08 August 2000.

[16] I.V. Gorbatiy Systems for Remote Sensing of the Earth from Space: Monograph / I.V. Gorbatiy. Lviv: SPOLOM, (2011) 612.

[17] O. Turovsky, «Clarification of the algorithm for estimating the frequency of the signal received by the satellite communication system in a continuous mode under the influence of «neighboring» channels», Advanced Information Systems. 4(3) (2020) 74-79. doi.org/10.20998/2522-9052.2020.3.09.

[18] O. Turovsky, V. Kozlovskyi, Y. Balaniuk, Y. Boiko "Minimization of phase error dispersion in closed type phase synchronization systems in carrier frequency tracking mode", 7th International Conference "Information Technology and Interactions (IT&I-2020, Satellite)», Kyiv, Ukraine, December 04, 2020, Taras Shevchenko National University of Kyiv. CEUR Workshop Proceedings. 2845 (2020) 382 – 390.

[19] O. Turovsky, L. Berkman, O. Drobyk, A. Zakharzhevskyi, O. Khakhlyuk, «Determination of the lower border of Cramer-Rao for evaluation of the carrier frequency of the radio technical communication channel signal», International Journal of Advanced Trends in Computer Science and Engineering, 9(4) (2020) 5838 – 5845. https://doi.org/10.30534/ijatcse/2020/243942020.

[20] Berkman, L., Tkachenko, O., Turovsky, O., Fokin, V., & Strelnikov, V. (2021). Designing a system to synchronize the input signal in a telecommunication network under the condition for reducing a transitional component of the phase error. Eastern-European Journal of Enterprise Technologies, 1(109) (2021) 66–76.. https://doi.org/10.15587/1729-4061.2021.225514.

[21] O. L. Turovsky, O. V. Drobik, Procedure for evaluation of the supporting frequency signal of the satellite communication system in continuous mode. Radio Electronics, Computer Science, Control. (2) (2021) 28–38. https://doi.org/10.15588/1607-3274-2021-2-3.

[22] L.N. Lee, A. Shenoy, Digital processor–based programmable BPSK(QPSK) offset–QPSK modems. Comsat Technical Review. 19(2) (1998) 57-64.

[23] A. L. Shestakov, A. S. Semenov, O. L. Ibryaeva, Estimation of the carrier frequency of a random sequence of pulses by the Prony method. Vestnik Yu.Ur.GU. 37(170) (2009) 106–115.

[24] O. L. Turovsky, Algorithm for estimating the carrier frequency of a phase-modulated signal of a satellite communication system when transmitting data in a continuous mode using the fast Fourier transform function. Zviazok, 2(144) (2020) 55-61. DOI:10.31673/2412-9070.2020.025560

[25] P. Kootsookos, E. Jacobsen, Accurate Frequency Estimators. IEEE Signal Processing Magazine. 24(3) (2007) 123-125.

[26] Oleksandr Laptiev, Savchenko Vitalii, Serhii Yevseiev, Halyna Haidur, Sergii Gakhov, Spartak Hohoniants. The new method for detecting signals of means of covert obtaining information. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (IEEE ATIT 2020) Conference Proceedings Kyiv, Ukraine, November 25-27. (2020) 176 –181.

[27] Valentyn Sobchuk, Volodymyr Pichkur, Oleg Barabash, Oleksandr Laptiev, Kovalchuk Igor, Amina Zidan. Algorithm of control of functionally stable manufacturing processes of enterprises. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (IEEE ATIT 2020) Conference Proceedings Kyiv, Ukraine, November 25-27. (2020) 206 –211.

[28] Vitalii Savchenko, Oleksandr Laptiev, Oleksandr Kolos, Rostyslav Lisnevskyi, Viktoriia Ivannikova, Ivan Ablazov. Hidden Transmitter Localization Accuracy Model Based on Multi-Position Range Measurement. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (IEEE ATIT 2020) Conference Proceedings Kyiv, Ukraine, November 25-27. (2020) 246 –251

[29] Oleg Barabash, Andrii Musienko, Spartak Hohoniants, Oleksandr Laptiev, Oleg Salash, Yevgen Rudenko, Alla Klochko. Comprehensive Methods of Evaluation of Efficiency of Distance Learning System Functioning. International Journal of Computer

Network and Information Security(IJCNIS), IJCNIS. 13(1) (2021) 16–28.

[30] A.O. Korchenko, V.O. Breslavskyi, S.P Yevseiev, N.K. Zhumangalieva, A.O. Zvarych, S.V. Kazmirchuk, O.A. Kurchenko, O.A. Laptiev, O. V. Severinov, S. S. Tkachuk. Development of a method for construction of linguistic standards for multicriterial evaluation of HONEYPOT efficiency. Eastern-European journal of enterprise technologies. 1(109) (2021) 14–23.

[31] Serhii Yevseiev, Oleksandr Laptiev, Sergii Lazarenko, Anna Korchenko, Iryna Manzhul. Modeling the protection of personal data from trust and the amount of information on social networks. Number 1 (2021), «EUREKA: Physics and Engineering». (2021) 24–31.

[32] Laptiev O., Savchenko V., Kotenko A., Akhramovych V., Samosyuk V., Shuklin G., Biehun A. Method of Determining Trust and Protection of Personal Data in Social Networks. International Journal of Communication Networks and Information Security (IJCNIS), 13(1) (2021) 15-21.

[33] O.Svynchuk, O. Barabash, J.Nikodem, R. Kochan, O. Laptiev. Image compression using fractal functions. Fractal and Fractional. 5(2) (2021) 1-14. https://doi.org/10.3390/fractalfract5020031 - 14 Apr 2021.

[34] Oleg Barabash, Oleksandr Laptiev, Valentyn Sobchuk, Ivanna Salanda, Yulia Melnychuk, Valerii Lishchyna. Comprehensive Methods of Evaluation of Distance Learning System Functioning. International Journal of Computer Network and Information Security (IJCNIS). 13(3) (2021) 62-71, Pub. Date: 2021-06-08 DOI: 10.5815/ijcnis.2021.03.06.

[35] Brumnik, R., Kovtun, V., Okhrimenko, A., Kavun, S. Techniques for performance improvement of integer multiplication in cryptographic applications (2014) Mathematical Problems in Engineering, 2014, art. no. 863617, . Cited 21 times. DOI: 10.1155/2014/863617.

[36] Gnatyuk, S., Kinzeryavyy, V., Kyrychenko, K., Yubuzova, K., Aleksander, M., Odarchenko, R. Secure Hash Function Constructing for Future Communication Systems and Networks (2020) Advances in Intelligent Systems and Computing, 902, pp. 561-569. Cited 21 times. DOI: 10.1007/978-3-030-12082-5_51.

[37] Gnatyuk, S., Kinzeryavyy, V., Iavich, M., Prysiazhnyi, D., Yubuzova, K. High-performance reliable block encryption algorithms secured against linear and differential cryptanalytic attacks (2018) CEUR Workshop Proceedings, 2104, pp. 657-668. Cited 20 times.

Intelligent system of energy-efficient microclimate control in greenhouses with energy consumption forecasting

Inna Yakymenko[1], Vitalii Lysenko[2], Valerii Koval[3], Ihor Bolbot[4]
[1,2,3,4]*National University of Life and Environmental Sciences of Ukraine, Heroyiv Oborony st., 15, Kyiv – 03041, Ukraine*

Abstract
Based on the analysis and control of energy flows in industrial greenhouses, which are complex biotechnological facilities operating under uncertain conditions, an intelligent system of energy-efficient microclimate control in greenhouses has been developed, based on the forecasting of energy consumption, including the zoning of payment for the use of electrical energy. The paper presents the results of analysis of energy flows and control of these flows in industrial greenhouses, which are complex biotechnical facilities operating in conditions of uncertainty, following on from the results of experimental and statistical analysis of relationships between external disturbances and energy costs that ensure compliance with the specified technology of growing plant products. It also presents the assessment results of the specificities of technological processes of growing vegetables in greenhouses. The authors conclude that heating and ventilation systems have the highest energy consumption and confirm the hypothesis of the existence of conditions of uncertainty in the operation of an industrial greenhouse. This creates conditions for the use of neural networks, which are able to operate effectively under conditions of uncertainty for energy consumption forecasting and proves the feasibility of developing an intelligent system of energy-efficient microclimate control in greenhouses based on energy consumption forecasting, taking into account electricity charges per zones.

Keywords
energy efficiency, resource efficiency, microclimate parameters, greenhouse, industrial greenhouse, intelligent control system.

1. Introduction

The account of all the variety of relationships between the parameters of growing vegetables in the greenhouse, their coordination and optimization require a high level of automation of technological processes. Automation of technological processes in greenhouses gives a significant effect: it increases productivity and improves working conditions, saves on natural gas and electricity, reduces the incidence of planting material and adult plants, increases yields and reduces the maturation time of plants and vegetables.

High energy prices result in a share of about 60 per cent in the cost structure of products grown in greenhouse. It is therefore important to develop new energy-efficient automated control algorithms that increase the overall profitability of production by minimizing energy costs [1].

Many factors, such as cloudiness, wind strength and direction, and other natural factors, poorly predicted by traditional methods, affect the energy consumption in greenhouses. The solution of these problems is possible through applying modern intelligent algorithms for processing information from the control object, and through using results to form appropriate control strategies to maximize profits from production output.

The issue of energy efficiency is considered by both domestic and foreign scientists, such as Agarkova O.M., Bolbot I.M., Dovgaliuk V.B., Dudnyk A.O., Zakharov N.G., Kuris Yu. V., Kurtener D.O., Levytska I.M., Lysenko V.P., Lee Yongwei, Pozin G.M., Pryschep L.G., Stroy A.F., Takakura, Terence Belvins, Tkachenko V.A. and other specialists. It was found that in order to

EMAIL: icheshun@gmail.com (A. 1); lysenko@nubip.edu.ua (A. 2); v.koval@nubip.edu.ua (A. 3); igor-bolbot@nubip.edu.ua (A. 4)
ORCID: 0000-0002-2316-2382 (A. 1); 0000-0002-5659-6806 (A. 2): 0000-0003-0911-2538 (A. 3); 0000-0002-5708-6007(A. 4)

achieve a better balance between energy consumption and product quality in greenhouses, it is necessary to improve existing microclimate control systems, through the use of modern intelligent algorithms for information processing and the formation of appropriate control strategies [2-7].

An important factor influencing the efficiency of vegetable growing in greenhouses is the widespread introduction of automated control and management systems of microclimate parameters, highly efficient and reliable systems and computer monitoring. One of the shortcomings of the existing systems is the increase in the temperature of the greenhouse due to inaccurate regulation. In fact, an increase in temperature by just 1°C results in an overexpenditure of additional 100 tons of conventional fuel in the greenhouse per year per hectare [8].

The operation of modern industrial greenhouses is characterized by transiency, by fuzzy processes and is exposed to random disturbances. At the same time, agronomic standards require high accuracy of temperature stabilization (+/1^{0}C), its timely change depending on the level of photosynthetic active irradiation, the phase of plant development and time of day. This creates conditions for increasing the requirements for the functioning of existing automation systems and technical improvement of hardware.

2. Development of an intelligent system for determining the parameters of the PI controller

The application of different methods of microclimate control, given in [9], allows to choose the best control algorithm.

Based on the fact that to maintain the control parameters at a given level, the settings of the PI controller (as one that is most common in industrial greenhouses) are variable under the same initial conditions, there arises a need to develop methods for finding controller settings that adapt to changing conditions and provide operation of the control system at a given level.

In order to improve the quality of regulation, the principle of genetic algorithm is used, the operation of which allows the settings to adapt to the consequences of changing working conditions. As the outputs of the model of the genetic algorithm we obtain the values of the settings of the PI controller. The optimal settings are: for the temperature control channel K_p=74,8, K_i=1,8; for the humidity control channel: K_p= 72,7, K_i=2,1.

To optimize the settings of the local control system, a fuzzy expert system was used to determine the parameters of the PI controller settings by deviation. The rules of the fuzzy expert system are based on ensuring a minimum area between the transition process and the time axis:

$$I = \int_{t_0}^{t}(x_{set} - x)dt \rightarrow \min_{k_p, k_i}, \qquad (1)$$

where x_{set}, x – respectively, the set value of the parameter and its value at the output of the object;

$k_p, k_i \in [k_{max}, k_{min}]$ – transmission factor and isodrome time of the PI controller.

Table 1 presents the quantitative comparison of the transient processes indicators.

Table 1

Results of the quality assessment of transition processes based on the application of various methods of setting the regulators (for a fixed state of the object)

	Method of setting	**Traditional method**	**Fuzzy conclusion**	**Genetic algorithm**
Temperature control	Kp	75.1	73.1	74.8
	Ki	3.1	3.3	1.8
	Re-regulation, %	33.4	11.1	0.3
	Regulation time, T, s	400	225	140
	Oscillation, N	2	-	-
Humidity control	Kp	64,3	71.2	72.7
	Ki	2.8	3.6	2.1
	Re-regulation, %	12.5	5	0
	Regulation time, T, s	270	315	62
	Oscillation, N	-	-	-

At the specified set values of temperature and humidity in the greenhouse: Θ_{in_set}=18°C and

φ_{in_set}=80% the settings of the PI controller for the temperature control channel at the output of the fuzzy expert system were: K_p= 73,1, K_i=3,3; for the humidity control channel: K_p= 71,2, K_i=3,6.

The traditional method of searching for the parameters of the PI controller is used as a reference model. The calculated values of the parameters of the PI controller settings under constant initial conditions were: transmission coefficient for the temperature control channel: K_p=75,2, and the isodrome time K_i=3,1; for the humidity control channel: K_p= 64,3, and the isodrome time K_i=2,8.

The best indicators are demonstrated by the use of a genetic algorithm, as in this case the least re-regulation occurs and there is no oscillation of the transition process. The application of this approach provides more accurate control under conditions of uncertainty, compared to traditional algorithms (both regulation and re-regulation time is reduced by two to three times).

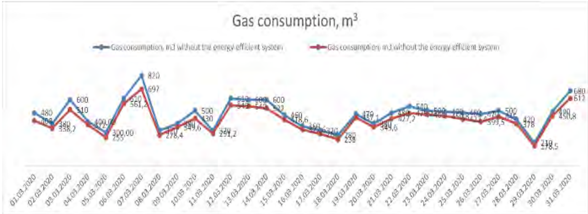
2.1. Development of a model for forecasting energy costs by means of applying neural network forecasting methods

The control at the upper level is achieved through applying forecast model that allows to forecast the cost of electricity and gas spent on maintaining a given technology for growing vegetables in greenhouses.

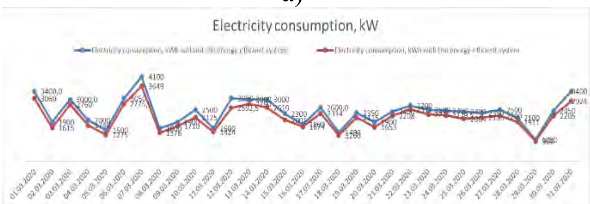
Taking into account the conditions of uncertainty generated by random disturbances and incomplete information about the biological component, the authors developed a neural network model for forecasting the energy consumption of natural gas and electricity in order to create an energy-efficient microclimate control system in greenhouses.

The outputs of the forecasting model are the values of natural gas and electricity consumption, which we use as inputs of the model of fuzzy control and genetic algorithm at the upper level of control. The input parameters of the neural network are: the value of the outside air temperature; solar radiation absorbed by the greenhouse and the level of CO_2 in the greenhouse.

The developed intelligent system of energy efficient control allows to forecast energy costs for the process of growing products in greenhouses taking into account current disturbances and to monitor the quality zones of input parameters that affect the energy consumption of growing vegetables. The comparison of the predicted values with the actual ones is shown in Figure 1.



a)



b)

Figure 1: Comparison of gas (a) and electricity (b) consumption in greenhouse vegetable production using the traditional approach and after the introduction of an intelligent automation system

2.2. Development of an algorithm for the operation of energy-efficient automatic control system with energy consumption forecasting

The methodology for the development of an intelligent energy-efficient system of automatic control of microclimate parameters in greenhouses is based on a combination of the following methods: intelligent setting of the PI controller parameters and energy consumption forecasting. The structure of such a system is shown in Fig. 2.

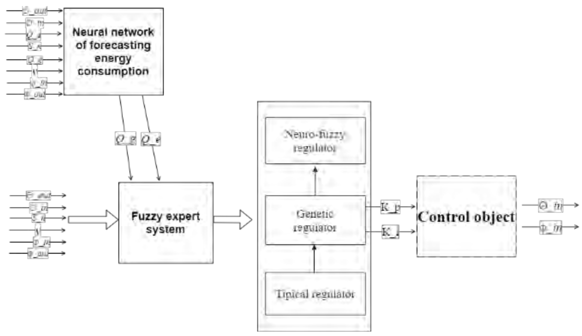


Figure 2: Structural and logical scheme of the system of energy efficient microclimate control

(where $\Theta_{in}, \Theta_{out}$ – air temperature inside and outside the greenhouse; S_{abs} – solar radiation absorbed by the greenhouse; Q_e, Q_g – consumption of gas and electricity in the greenhouse; ν – power of the ventilation system; $\varphi_{in}, \varphi_{out}$ – relative humidity inside and outside the greenhouse; K_p, K_i – parameters of the PI controller)

At the first stage of system development, an experimental study is conducted, on the basis of which a database of microclimate parameters is formed both inside and outside the greenhouse.

The next step is to create a fuzzy expert system for controlling groups of equipment for supplying heat to the greenhouse, for air ventilation and the formation of rules by experts based on the requirements for quality assurance of products grown. Further the processing of information on the projected energy costs takes place, based on which the equipment control strategy is chosen to minimize energy consumption by adjusting the settings of the regulator. The next step is to determine the parameters of the PI controller and conduct modeling for further synthesis of energy-efficient control system of microclimate parameters in the greenhouse. This allows to get an energy-efficient control system for growing vegetables in a greenhouse, the use of which reduces energy consumption by 10-15% compared to traditional technologies for controlling the parameters of the microclimate.

The algorithm for operation of the energy-efficient control system of microclimate parameters in the greenhouse is shown in Figure 3.

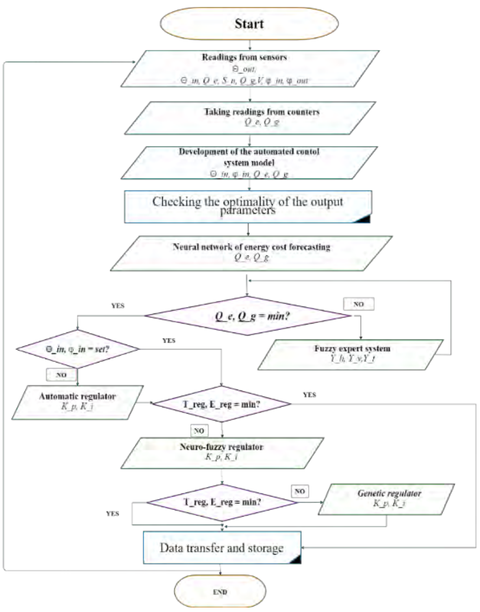


Figure 3: Algorithm for operating the energy-efficient control system of microclimate parameters in the greenhouse (where $\Theta_{in}, \Theta_{out}$ – air temperature inside and outside the greenhouse; S_{abs} – solar radiation absorbed by the greenhouse; Q_e, Q_g – consumption of gas and electricity in the greenhouse; ν – power of the ventilation system; $\varphi_{in}, \varphi_{out}$ – relative humidity inside and outside the greenhouse; Y_o, Y_v, Y_t – initial actin of controlling the activation of the heating system engines, ventilators and transoms; K_p, K_i – PI controller parameters; T_{reg} – regulation time, E_{reg} – regulation error)

To assess the economic efficiency we use indicators of minimization of energy resources and payback period of the implemented methods.

The theoretical studies allowed to calculate the capital costs, operating costs, profits from the implementation of the system and payback periods for the implemented systems. The results of the calculations are shown in Table 2.

Criterion for assessing economic efficiency:

$$\begin{cases} \text{Natural gas consumption} \rightarrow \boldsymbol{min} \\ \text{Electricity consumption} \rightarrow \boldsymbol{min} \\ \text{Profit of the enterprise} \rightarrow \boldsymbol{max}. \end{cases} \quad (2)$$

Table 2

The results of calculating the economic efficiency of the introduction of automation control system (ACS) of temperature-humidity regime

	Fuzzy expert system	**Genetic algorithm-based ACS**	**Neural network forecasting-based ACS**
Capital costs, thousand UAH	85.0	95.0	87.0
Operating costs, thousand UAH	165.1	168.3	165.74
Profit, thousand UAH	353.45	327.85	331.13
Payback period, days	**333.2**	**360.1**	**338.8**

3. Conclusions

Analysis of the quality of transient processes as a response to disturbances in automatic systems with a PI controller and an improved model of the object in terms of temperature and humidity has shown that the parameters of the controller settings are variable in nature. This proves the relevance of developing methods for finding the settings of the controller, which adapt to changing conditions of the automation system.

Based on the use of fuzzy logic methods and genetic algorithm, the authors developed and studied models for finding the optimal parameters of the PI controller settings that adapt to changing operating conditions of the automation system. According to the linear integrated quality indicator, the best quality of the transient process is demonstrated by the system, the PI controller of which is set up using a genetic algorithm for such indicators: K_p =74.8, T_i=1.8 (to comply with the parameters of the microclimate): Θ_{in_set}= 18°C and φ_{in_set}=80.

The present paper offers the developed system of intelligent automated system of energy efficient control of microclimate parameters in greenhouses and its algorithmic support, which includes: block of neural network forecasting of energy costs, block of decision support, which takes into account the prices of energy resources, block of optimizing the parameters of regulators based on the use of fuzzy logic and genetic algorithms. The introduction of such a system saves natural gas for heating and electricity – up to 10%.

4. References

[1] Intelligent Control System of Biotechnological Objects with Fuzzy Controller and Information Channel Filtration Unit // Alla Dudnyk, Vitaliy Lysenko, Natalia Zaets, Dmytro Komarchuk, Taras Lendel and Inna Yakymenko // 2018 International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, 2018. – P. 586-590.

[2] The Effect of Temperature and Mean Cumulative Daily Light Intensity on Fruiting Behavior of Greenhouse-grown Tomato: [Електрон. ресурс] / Sezgin Uzun // The University of Ondokuz Mayıs, Faculty of Agriculture, Department of Horticulture. 2007. – 132(4). – P. 459-466. – Режим доступу: до журн.: http://journal. ashspublications.org/content/132/4/459.full. pdf.

[3] Blevins T, Wojsznis WK, Nixon M. Advanced control foundation: tools, techniques and applications. International Society of Automation (ISA); 2013.

[4] Li Y, Ding Y, Li D, Miao Z. Automatic carbon dioxide enrichment strategies in the greenhouse: a review. Biosyst Eng 2018;171:101–19.

[5] Ding J-T, Yao H-Y, Zang Z-L, Huang M, Zhou S-J. Precise control and prediction of the greenhouse growth environment of Dendrobium candidum. Comput Electron Agric 2018;151:453–9.

[6] Tadashi Takakura Research exploring greenhouse environment control over the last 50 years// NARO Institute for Rural Engineering, Tsukuba, Ibaraki 305-8609, Japan)// Mathematics and Computers in Simulation 167 (2020) 19–31 // DOI: 10.25165/j.ijabe.20191205.5179

[7] Haibin Liu, Shengyu Fang, Xinqin Guo Research and Design of Intelligent Greenhouse Control System Based on AIoT Fusion Technology //To cite this article: Haibin Liu et al 2020 // IOP Conf. Ser.: Earth Environ. Sci. 474 032036 // doi:10.1088/1755-1315/474/3/032036

[8] Заєць Н. А. Експериментально-статистичне дослідження теплиці як об’єкта керування з метою підвищення ресурсоефективності виробництва/ Н.А.Заєць, А.О. Дудник, І.Ю. Якименко. // Енергетика і автоматика. – 2017. – № 4. – С. 200-211.

[9] Лисенко В.П., Якименко І. Ю., Гібридна система автоматизації для споруд захищеного ґрунту. Енергетика і автоматика: електрон. Наук. Фах. Вид. – 2020. – № 2. – С.37-50.

Fuzzy Modeling for Synthesis of Control Systems for Electro-Explosive Energy Conversion

Sergey Kozyrev

Department of Computer Technology and Information Security National University of Shipbuilding, Institute of Pulse Processes and Technologies NAS Ukraine, Mykolaiv, 54000, Ukraine

Abstract
The analysis of existing models of electro-explosive energy conversion in the channel of high-voltage discharge in a liquid was carried out. The application limits of existing models were determined. A fuzzy control model of electro-explosive energy conversion was developed using fuzzy approximation method based on the fuzzy sets theory, which describes the object in the whole space of states, taking into account its nonlinearity and stochasticity. An inverse model of a fuzzy controller was synthesized. Inverse model can be used in synthesis of control systems for electro-explosive energy conversion. Using this model provides adaptability of control systems in case of varying technological parameters and external influences.

Keywords
Fuzzy models, discharge-pulse technology, electro-explosive energy conversion, inverse model, control system

1. Introduction

Electro-explosive energy conversion is used in discharge-pulse technologies as a source of concentrated, metered impact in a given local volumes with high specific energy efficiency. It takes place in the channel of high-voltage discharge in a liquid. The discharge is conducted by pulse current generator (PCG), where the energy of the electric field of the charged capacitors' battery is transformed into a mechanical work which expands the discharge channel. The pressure impulse arises due to instantaneous expansion of the channel, caused by highly concentrated discharge energy. It is used as the main factor of technological operation on the processed objects in electro-discharge technologies. The main advantages of electro-explosive energy conversion are the possibility to achieve high specific energy efficiency and the controllability of the technological process. Leveraging the main advantage of the electro-discharge energy conversion (the controllability) requires its comprehensive study and formalization, i.e. the construction of the mathematical model of control.

Existing models are based on the study of electro- and hydrodynamic transients and are defined by a system of nonlinear differential equations [1]. But even with significant simplifications they are too complex and don't take into account many factors which influence the discharge process. Also, the discharge process has a stochastic nature due to the statistical patterns of elementary processes which take place at the stage of discharge channel formation. So existing models cannot be used for the control process. The analysis of the existing models, which are used in the systems of regulation and stabilization of individual parameters of the electro-explosive energy conversion process [2], shows that that they all are built using linearization of transfer functions of the object. So they define operations only around the nominal mode with certain assumptions about external perturbations. Those models only take into account the input perturbation of coordinate l (the length of the discharge gap, which is approximated by the linear function of time $\Delta l(t)$), and is considered additive, i.e. one that enters the input of the control object and is added to the current coordinate value.

EMAIL: skozyrev@gmail.com
ORCID: http://orcid.org/0000-0001-8995-131X

The assumptions made in development of existing control models of electro-explosive energy conversion are acceptable for a very narrow class of problems. In real-world conditions, there is a high number of unaccounted perturbations that significantly affect the discharge mode. Such perturbations include temperature rise, CO_2 absorption, decreasing of specific resistance ρ of the liquid where a high-voltage discharge occurs. The change rate of ρ depends on the intensity of processing modes. In the absence of regeneration systems, it is impossible to ignore changes in resistivity, so it is necessary to consider the additional coordinate ρ[n], which affects the operator of the control object and the statistical characteristics of the information coordinates. As for the assumption of linearity of perturbation in the coordinate l[n], it can be allowed only when processing surfaces with small height differences and around the point of the nominal mode. On the other hand, the adaptive control systems must provide the controllability of the process of electro-explosive energy conversion throughout the entire states space, under conditions of unpredictable external influences.

Therefore, the existing control models are not suitable for use in the synthesis of adaptive control systems.

The objective of this work is to develop an adequate control model of electro-explosive energy conversion under conditions of significant changes in environmental parameters and external influences, taking into account the significant nonlinearity and stochasticity of the control object. The model should be built on the basis of modern approximation methods using fuzzy logic. It has to improve the controllability and the accuracy of maintaining optimal modes in real-world conditions.

2. Synthesis of fuzzy control models of electro-explosive energy conversion

It's almost impossible to build an analytical control model of electro-explosive energy conversion for the whole space of states since the underlying physical processes in a discharge channel are way too complex, have a stochastic nature, are nonlinear and hard to formalize [1]. That's why the suggestion here is to use fuzzy-approximation methods, based on experimental data, in order to develop fuzzy control models [3].

For the purpose of building fuzzy control model over the entire space of states there were theoretical and experimental researches performed. The object of research was a discrete process with a period T of electro-explosive energy conversion in the discharge circuit of PCG. $\mathbf{Y}(nT)=\mathbf{AX}(nT)$, n – current step of a process. The parameters of discharge circuit can be considered as coordinates of input vector $\mathbf{X}(nT)$: U – charging voltage of accumulator, C – capacitance of capacitors' battery, L – inductivity of discharge circuit, $l(nT)$ – discharge gap length, $\rho[n]$ – specific resistance of the liquid. The output vector $\mathbf{Y}(nT, \tau)$ – the result of electro-explosive energy conversion in a discharge channel, which can be described by stochastic impulse functions: $i(\tau)$ – function of discharge current, $u(\tau)$ – function of a voltage in discharge gap, $p(\tau)$ – function of a pressure inside discharge channel, τ – duration of a discharge ($\tau<<T$). The functionals of the functions mentioned above, which quantitatively characterize them, can be considered as coordinates of output vector $\mathbf{Y}(nT, \tau)$: maximum values of discharge current $i_m[n]$ and pressure $p_m[n]$, discharge voltage $u_d[n]$. Experimental study of input vector's coordinates [2] has shown they are discrete stochastic functions with a Gaussian distribution in each point of factor space (according to fitting criterion χ^2 it is consistent with truth of significance value α=0,01) and therefore can be represented as: $y_m[n] = \mathrm{M}_y[n]+y^0$, where $\mathrm{M}_y[n]$ – expected value of output coordinate, y^0 – a spurious signal, i.e. stationary random value with a Gaussian distribution, whose stochastic characteristics are defined by processes of discharge channel formation and are dependent on object's position in the states space. The object's operator $\mathbf{A}$ puts a set of functional space of possible implementations of output functions $\mathbf{Y}$ in a correspondence to set of input states $\mathbf{X}$.

As for information coordinate, it is important to choose such output coordinate (or linear combination of corelated coordinates with correlation indexes of opposite sign) which provides best statistical efficiency. To determine such the most statistically effective information coordinate, a study of statistical characteristic and correlation ratios of output vector's coordinates was done. As a result, the linear combination $\Sigma[n] = i_m[n]+ku_d[n]/i_m[n]$ [2, 4] was chosen as the most statistically effective.

Fuzzy model of electro-explosive energy conversion is built as $\Sigma[n] = F(l[n], \rho[n])$. Since electro-explosive energy conversion is a stochastic object it's model has to be supplemented with relationship between the mean-square deviation σ_Σ of information coordinate and input coordinates: $\sigma_\Sigma = F(l[n], \rho[n])$. Building a fuzzy model requires a training data, which should consist of sets of groups of input signals ($l[n]$, $\rho[n]$) and corresponding expected values of output signal $\Sigma[n]$ along with its mean-square deviation σ_Σ. We form such signal groups by using the database of experimental data. In order to fill-in the experimental data for building fuzzy model of electro-explosive energy conversion the experimental research was conducted. The aim of the research was to study the relationship within the entire states space between output information coordinate $\Sigma[n]$ (along with its statistical characteristics) and the change of output vector $\mathbf{X}<l[n], \rho[n]>$ coordinates. A fractional factorial experiment was conducted, with simultaneous variation of all independent variables at all selected levels of values using experiment planning methods.

The experiment was performed at conditions of $U(t)$=const, $C(t)$=const, $L(t)$≈const, and with different combinations of factors $l[n]$, $\rho[n]$ and their values levels. The response variable of factorial experiment were the values of observed coordinates of output vector $\mathbf{Y}<i_m[n], u_{пр}[n]>$. Based on these coordinates' values the expectation of information coordinate $M(\Sigma[n])$ and its mean-deviation σ_Σ were calculated. Since the process of electro-explosive energy conversion has stochastic nature the number of samples was around 100-110 runs at each point of factor space. This ensured the confidence score of 0.95. The results of mathematical processing of experimental data are given in relative units in Tables 1, 2. The reference values are the amplitude of short-circuit discharge current I_{sc} and initial accumulator voltage U_0.

Table 1

Expectation of coordinate $M(\Sigma[n])$

l, m	ρ, Ohm m				
	6,0	7,5	10,0	15,0	20,0
0,025	0,56	0,57	0,58	0,60	0,64
0,050	0,40	0,42	0,44	0,56	0,60
0,075	0,26	0,28	0,28	0,49	0,53
0,100	0,13	0,16	0,18	0,40	0,42

Table 2

Mean-square deviation σΣ

l, m	ρ, Ohm m				
	6,0	7,5	10,0	15,0	20,0
0,025	0,028	0,024	0,023	0,015	0,010
0,050	0,037	0,036	0,036	0,014	0,012
0,075	0,042	0,040	0,039	0,015	0,013
0,100	0,074	0,069	0,046	0,017	0,016

The first step in constructing the fuzzy model is splitting the space of input and output signals into ranges which are going to serve as terms in order to build the membership function. The direct method was used for building membership function. This method doesn't require a precise determining of membership function, just the function type and its representative values (in our case – the values at experimental points according to experiment's plan, which are the pivot values for the ranges taken as terms). Let's accept the triangular shape of membership functions. One of the vertex of each membership function lays in the center of each range of coordinate's values. The value of the function $\mu^T(x) = 1$ corresponds to this central vertex. Two other vertexes lay in the centers of neighbor ranges. The corresponding function value for them is 0. The coordinates $\rho[n]$ and $l[n]$ are taken as the linguistic output variables. The number of terms (linguistic values, in our case – numeric ranges) for each variable is chosen equal to number of values levels per the plan of factorial experiment. In this particular case, the number of terms is equal to 5 and 4. Membership functions (MF, mf_i or $\mu^T(x)$) of coordinates of state vector $\rho[n]$, $l[n]$ are approximated with tent functions. The obtained membership functions are shown in Figure 1.

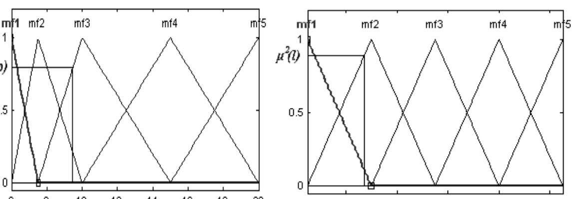


Figure 1: Membership functions $\rho[n]$, $l[n]$

The procedure of fuzzification of input signals (determining which range the current coordinates values of input vector belong to) is performed based on maximum value of membership function $\mu^T(x)$.

The rules base is composed on the basis of knowledge base (experimental data). The rules base in a form of fuzzy logical equations allows to connect membership function of output variable $\sum[n]$ and membership function of input vector's coordinates $\rho[n]$, $l[n]$. As a result, we get the linguistic value of output variable. Fuzzy simulation is done with a FuzzyTECH software, using Fuzzy Logic Toolbox package [3], which has easy to use interface for designing and debugging of fuzzy models. Graphic tools of Fuzzy Logic Toolbox provide means for interactive tracing of system behavior. The results of fuzzy approximation of relationship between input and output coordinates of object state vector $M\Sigma[n]=F(l[n], \rho[n])$ and relationship between statistical characteristics of information coordinate and object's position in factor space $\sigma_\Sigma = F(l[n], \rho[n])$ are presented in Figure 2, 3.

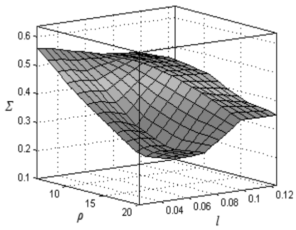


Figure 2:
$M\Sigma[n]=F(l[n], \rho[n])$

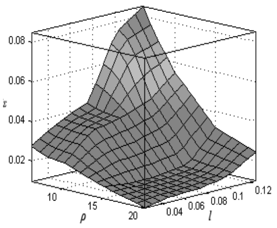


Figure 3:
$\sigma_\Sigma = F(l[n], \rho[n])$

The fuzzy model of electric explosion, built with fuzzy approximation methods, represents nonlinearity of object of control. To ensure the controllability in the entire range of input coordinates it is required to make changes into control law depending on the position of the object in the space of states. The values of statistical characteristics of information coordinates also depend on the position of the object within space of states, what requires to adapt the algorithm of estimation of information signal during the control process.

The results of modeling have proven that proximity operator between coordinates of input and output vectors of control object and statistical characteristics of information coordinates depend on the position within states space. Providing controllability of electro-explosive energy conversion within entire states space requires introducing of adaptation mechanisms in order to take into account nonstationary and stochasticity of the object. The mechanism of adaptation can be implemented using inverse model based systems, which generate reaction to perturbance by determining such a control action which guarantees the required values of output coordinates. The validity of the inverse model is ensured by the accuracy and integrity of specification of object properties and its statistical characteristics. Constructing the inverse model implies finding the reverse function of the control object, i.e. the function which allows to find such a value of control input which provides given value of output information coordinate and therefore provides the given operation mode of the electro-discharge energy conversion.

Building of inverse model electro-explosive energy conversion was done using fuzzy-approximation methods and experimental data. The rules base of fuzzy inverse model was formed based on numeric data [3]. The rules base includes only rules with highest truth degree. That is how the problems of inconsistent rules and rules multiplicity are solved. The block diagram of building the rules base for inverse model is shown in Figure 4.

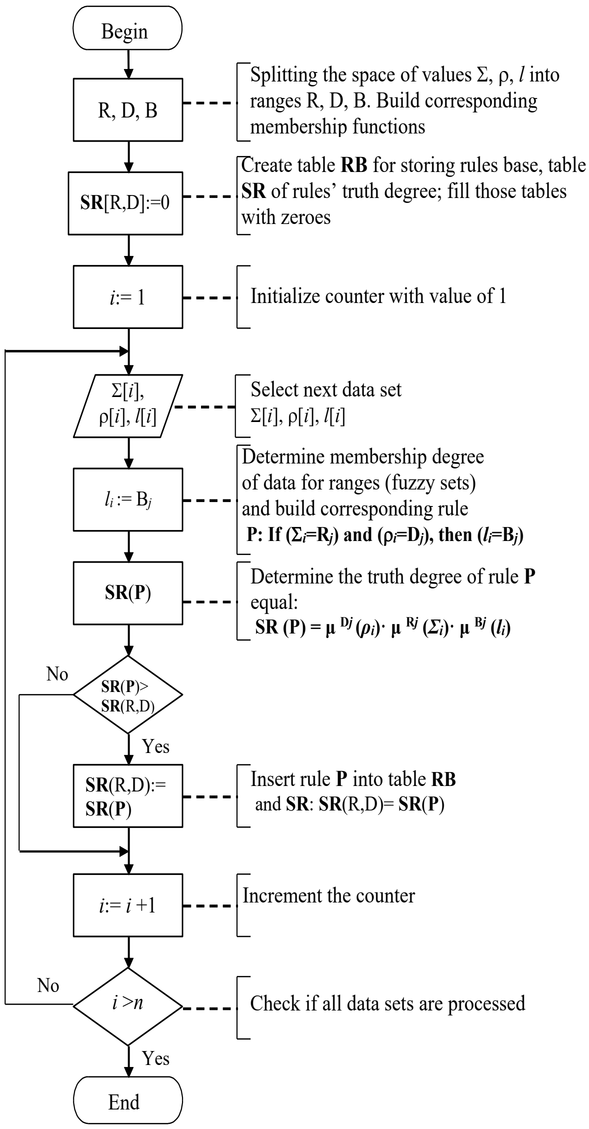


Figure 4: Block diagram of inverse model building algorithm.

For the output coordinate $\sum[n]$ and its mean-square deviation σ_Σ we perform the following procedures: splitting the space of signal values into ranges, which serve as the terms for constructing the membership function; building of tent-shaped membership function with the vertexes at expected values of output coordinate $\sum[n]$. The resulting membership function is shown in Figure 5.

After applying the above-mentioned algorithm to existing experimental data, we got consistent knowledge base, which is the basis for constructing of fuzzy inverse model $l[n]=F(\Sigma[n], \rho[n])$ in the MATLAB software using Fuzzy Logic Toolbox package (Figure 6).

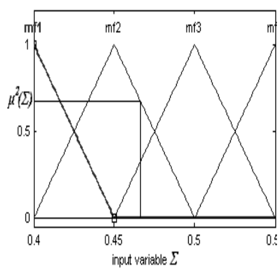


Figure 5: MF $\sum[n]$

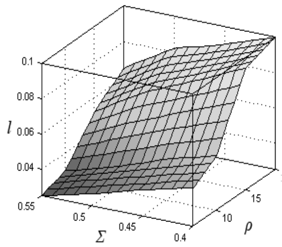


Figure 6: $l[n]=F(\Sigma[n], \rho[n])$

The fuzzy models, obtained by means of fuzzy approximation methods, can be used for synthesis of adaptive control systems for electro-explosive energy conversion in order to adjust the parameters of control system depending on the current state of control object. This allows to extend the controllability area, increase the established control precision and gives the control system the feature of adaptivity. The important feature of fuzzy regulators based control systems is their ability to provide required operating modes for the entire states space.

3. Conclusions

The fuzzy control model for electro-explosive energy conversion, based on fuzzy approximation methods, was developed. The model operates the expectation of information coordinate $M(\Sigma[n]) = F(l[n], \rho[n])$, describes the control object in the whole space of states, taking into account its nonlinearity. The experimental data was used for building a knowledge base during the synthesis of fuzzy model. In case of changes in technological parameters that base can be easily extended in order to provide adaptivity for new operating modes.

The fuzzy model of relationship between mean-square deviation of information coordinate σ_Σ of control object and its position in the space of states $\sigma_\Sigma = F(l[n], \rho[n])$ was developed. Using this model in a control system allows to increase the precision of information signal estimation and also allows to extend controllability area and improve overall control accuracy.

There was an inverse fuzzy model synthesized, which generates reaction to perturbations by determining a control action which guarantees required values of output coordinates. Using of inverse model, based on principles of fuzzy logic, for building control system provides adaptivity under conditions of wide range varying technological and environmental parameters. Such control system can generate adequate response to perturbations by determining relevant control action based on inverse model.

The use of developed fuzzy models for synthesis of control systems can provide effective control of electro-explosive energy conversion in the entire states space under conditions of indeterminacy of external perturbations and nonstationary environmental parameters, caused by changes in technological processes.

4. References

[1] Кривицкий Е.В. Динамика электровзрыва в жидкости. - Киев: Наукова думка, 1986. - 206 с.

[2] Вовк И.Т., Друмирецкий В.Б., Кривицкий Е.В., Овчинникова Л.Е. Управление электрогидроимпульсными процессами - Киев: Наук.думка, 1984. - 186 с.

[3] Леоненков А.В. Нечеткое моделирование в среде MATLAB и FuzzyTECH. – СПб.: БХВ - Петербург, 2003. – 736с.

[4] S. Kozyrev, N. Nazarova, D. Vinnichenko, Adaptive Filter of Input Information Signal for Discharge Pulse Installation Control System, Conference proceedings IEEE 2018 3rd International Conference on Intelligent Energy and Power Systems (IEPS), 2018, Kharkiv, Ukraine. - IEEE Catalog Number: CFP1805X-ART ISBN: 978-1-5386-9545-6. – P. 343-346.

Dynamic Configuration of Software Products Lines for Smart-Home Applications based on Recommender Systems Framework

Rustam Gamzayev[1], Mykola Tkachuk[1]

[1]*V.N. Karazin Kharkiv National University, Majdan Svobody 6, Kharkiv, 61077, Ukraine*

Abstract

The need to apply intelligent methods to configurate dynamic software product lines (DSPL) with respect to changes of user requirements is motivated, and as one way for this the usage of recommender systems (RS) is proposed. The briefly overview of the main components within the RS framework is outlined: a formal definition for RS, main methods for data processing in RS, and some technologies used to implement such systems. Based on these results the architecture of the RS prototype to support a configuring of software components in the "Smart-Home" (SH) applications is elaborated which uses the N-dimensional context-dependent tensor factorization method (CPTF) and the CARSkit open source software tool. The input data for the test-case in SH-domain are considered, and the first experiment results are obtained, where the CPTF method is compared with other similar ones used in RS, e.g.: UserKNN (Aggregation of K-nearest neighbors ratings), SVD++ (Advanced Singular Value Decomposition), MF (Matrix Factorization), and SPF (Semantic Pre-Filtering). This shows the higher accuracy of the CPTF method to generate the recommendations using the proposed RS prototype, and it allows to make positive conclusions, and formulate further steps to be done in this research.

Keywords

Software product line, configuration, smart-home, recommender system, tensor factorization

1. Introduction

For the complex solution of problems arising in such new and high-tech application domain as the Internet of Things system, in particular, in applications such as 'Smart-Home', automation of complex technological processes, etc., it is not enough to develop and maintain individual software solutions. In terms of resource consumption and time conservation, a more efficient way is the construction of their interconnected sets, which in modern software engineering are called software product lines (SPL) [1]. They have a shared set of common functions (components) that are customized according to the users' needs in a particular context of their operation and that provide variability of SPL. Dynamic SPL (DSPL) [2] are currently being actively developed and researched, in which the mechanisms of adaptation and configuration of variable functions and properties directly at the stage of execution and placement of their individual components play an important role.

The solution of these complex problems requires the development of new intelligent methods of creating DSPL, and in particular, a promising area of such research is the use of modern recommender systems (RS), which use knowledge-based methods and technologies for filtering consolidated information, which allows the choice of DSPL configurations in accordance with the environment.

The purpose of this study is the analysis the possibilities of applying methods and technologies for building RS to solve DSPL configuration problems, the development of software prototype architecture and its implementation, and the experimental study of feasibility and accuracy of the proposed approach in the subject area 'Smart-Home'.

EMAIL: rustam.gamzayev@karazin.ua (A. 1);
mykola.tkachuk@karazin.ua (A. 2);
ORCID: https://orcid.org/0000-0002-2713-5664 (A. 1);
https://orcid.org/0000-0003-0852-1081 (A. 2);

2. Related Work: Recommender Systems in Software Engineering

Nowadays there are not a large number of publications in this research area, but the growing interest in the topic of how to use RS exactly in the Software engineering (SE) domain should be noted.

For example, in [3], the solid analysis of RS usage in SE is provided, with the aim to improve system quality, make the development process more efficient and less expensive, and help developers to make better project decisions. In [4], the techniques and tools are presented for providing the open-source software developers with multiple recommendations to assist them in different phases of the development process, and for this purpose, a knowledge-aware RS was designed and evaluated on the real datasets. In [5], a similarity-based RS has been developed. It provides online recommendation support during the user's configuration process for a given SPL, based on the current feature selections (partial configuration) in run-time, and it helps users not only find the correct configurations but also to make such decisions more efficiently.

Even this brief overview shows the continuing need for research to answer the current question: how to use RS in SE efficiently, especially, with respect to SPL development and maintenance issues.

3. RS framework in a nutshell

The formal representation of the RS can be given following the definition given in [6]: let U be the set of users of the system (User - U), E be the set of all elements (Element - E) that fall under the category of their preferences, R ⊆E be the ranked list of subsets of such elements, and r is a specific element in the list R. The problem of providing a recommendation is to select r ∊ R so that it satisfies the requirements of the respective users u⊆U.

Let S be the metric of satisfaction (S) of an individual user u⊆U, which takes the value z, where z is some real number. Then if f () defines the recommendation function of r elements for u user, then the problem of forming a recommendation can be formulated as follows:

$$f(r, u) = S\ (z) \rightarrow max$$

Methods of constructing modern RS [7-9] mostly use the idea of collaborative filtering (CF), which, in turn, is divided into 2 approaches: a memory-based one and a model-based one. A number of studies suggest that for the use of RS to configure DSPL, it is advisable to choose the methods of model-based CF, which include: clusterization of many information objects based on certain criteria; matrix factorization (MF) as a way to decompose a certain matrix of the configuration of user interaction with an element of its content into the product of two rectangular matrices of smaller dimension, which allows the study of hidden factors for each user and each data element

Tensor factorization (TF) which expands the possibilities of the MF approach by decomposing the initial N-dimensional tensor at the rank k into N of individual matrices consisting of k columns, which represent the mapping of each individual dimension of the tensor into k factor dimensions of semantic space. It should be noted that TF is suitable for solving the problem of managing DSPL configurations, in which there is a problem of exponential growth of the configuration space [9].

Technologies and software for the implementation of RS should be considered on the examples of their open-source projects, a very detailed analysis of which can be found, for example, in [10]. In particular, such programming languages as Java, JavaScript, Python, the appropriate APIs, and data are effectively used for this purpose. Based on the reasons given in [10], the open-source project CARSKit (Context-aware Recommender System toolkit https://github.com/irecsys/CARSKit) was chosen for further research.

4. Design of the RS prototype for Smart-home applications configuring

The detailed analysis and modeling of the functional and software quality requirements for the target RS prototype was done in [10].

Basing on this one, and taking into account the special capabilities of the CARSKit tool (see, e.g., in [11]), the logical architecture of the RS prototype for testing the CF, MF and TF algorithms (see in Section 3) was developed. This solution at the high-conceptual level can be represented in Figure 1 in the form of a UML packages diagram [12].

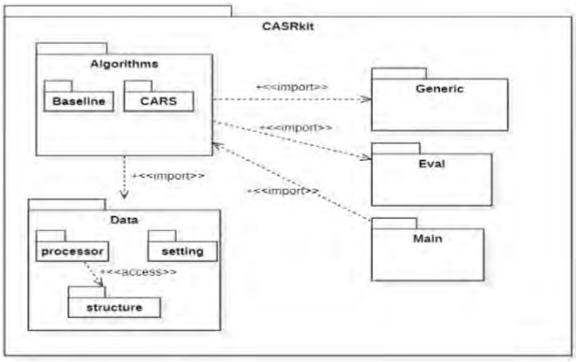


Figure 1: Diagram of the main functional packages of the RS prototype

The main *Algorithms* package is divided into two other packages: the *Baseline* package, which implements shared filtering algorithms (UserKNN, SVD ++, etc.) and ranking algorithms (RankSGD, RankALS, etc.), and the *CARS* package, which contains context-sensitive classes for configuring these basic algorithms. The *Data* auxiliary package includes other packages: *Structure*, *Processor*, and *Setting*, which, respectively, contain basic classes for implementing data structures (matrices and tensors), for basic data processing operations, and classes for configuring certain system parameters. The *Generic* package provides classes through the interfaces of which additional algorithms of recommendations can be implemented; the *Eval* package contains classes for calculating the metrics for evaluating the quality of the generated recommendations, and, finally, the *Main* package encapsulates the classes for loading and initializing the current session with the prototype RS.

Figure 2 shows the simplified deployment diagram [12] of the RS prototype components.

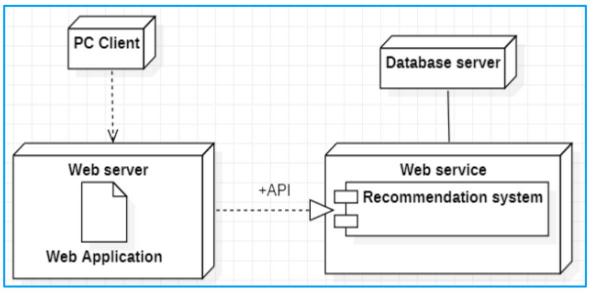


Figure 2: Diagram of RS components deployment

It includes the following hardware nodes and software components: the Client's computer (PC), the server-side Web application, and the recommendation system as a Web service that handles all retrospective data through some API, and the Database server as well.

5. Test-case input data, first experiment results, and their analysis

To prove the proposed approach, the test DSPL for the subject area ‘Smart-Home’ was considered and presented as a set of data on the configuration of certain functionalities of such a system [13], e.g.: temperature measurement, time determination, opening and closing doors, windows, etc. This set is a sparse tensor that describes these configurations in multidimensional information space: *‘user’* × *‘function’* × *‘weather’* × *‘temperature’* × *‘time’’*, where user contexts are relations that have the following attributes:

- Weather [‘sunny’ , ‘ rainy’];
- Temperature C (degrees Celsius) [NA, 10, 11, 12, 15, 18, 20, 22];
- Time (in hours) [7, 8, 11, 12, 13, 14, 19, 22].

A fragment of ‘raw data’ in relation to previous configurations is given in Figure 3, where the abbreviation NA (“Not Applicable”) means the lack of certain values:

```
userid,itemid,use,weather,temperatureC,time
1,1,0,sunny,20,12
1,2,1,sunny,20,12
1,3,0,sunny,20,12
1,4,1,sunny,20,12
1,5,1,sunny,20,12
1,6,-1,sunny,20,12
1,7,1,sunny,NA,12
1,1,0,sunny,NA,8
1,2,1,sunny,NA,8
```

Figure 3: ‘Raw’ data in the RS

Figure 4 shows a fragment of the file with the source data of the type *‘user’* × *‘configuration’*× *‘function’* × *‘rating’* (i.e., the value that indicates whether the user chooses this function). Each function with an asterisk (*) means that the obtained prediction for it coincides with the truth, i.e. the obtained value corresponds to the value for this function and the current configuration.

```
# userid: recommendations in (itemid, ranking score) pairs, where a correct recommendation is denoted by symbol *.
2, weather:rainy;temperaturec:22;time:14: (5*, 1.2725257), (4, 1.0836912), (1, 1.0204401), (3*, 1.0002029), (2*, 0.9998355)
3, weather:sunny;temperaturec:10;time:7: (1*, 0.95620173), (2, 0.8760538), (4, 0.80296594),
3, weather:sunny;temperaturec:18;time:11: (3, 1.0015707), (1, 0.95620173), (2*, 0.8760538), (4*, 0.80296594),
6, weather:rainy;temperaturec:22;time:12: (5*, 1.0016743), (1, 1.0009238), (3, 0.99720657), (2*, 0.9528067), (4*, 0.86005116)
4, weather:sunny;temperaturec:22;time:13: (3, 1.0731055), (1, 1.0702195), (2*, 1.0024319),
```

Figure 4: Result of ranking user-function pairs

To calculate the accuracy of the recommendations obtained, you need a set of metrics to assess the deviation between the projected and actual ratings. To do this, there are several metrics [14] available, of which the follows are chosen in this study: mean absolute error (MAE) and root mean squared error (RMSE):

$$MAE = \frac{1}{|Q|}\sum_{(u,i)\epsilon Q} |r_{ui} - \hat{r}_{ui}|,$$

$$\text{RMSE} = \sqrt{\frac{1}{|Q|}\sum_{(u,i)\epsilon Q}(r_{ui} - \hat{r}_{ui})^2}$$

where Q is a set of tests, |Q| is a cardinality of a set Q, r_{ui} represents the actual user estimates, and $\hat{r}_{ui}$ represents prediction ratings using the RS. The MAE metric is the simplest, but it does not take into account the direction of a single error (i.e., whether this error is positive or negative), and therefore RMSE is widely used in calculating the accuracy of RS prediction.

The accuracy of the N-dimensional tensor factorization approach [9] (or the Candecomp Parafac Tensor Factorization – CPTF, in terms of the CARSkit project [11]), was evaluated by the comparing its with the classical context-oriented approaches and approaches based on collaborative filtering [6,8], namely: UserKNN (Approach used to aggregate k-nearest neighbors) ratings), SVD ++ (Advanced Singular Value Decomposition), MF (Matrix Factorization), and SPF (Semantic Pre-Filtering).

Table 1 presents the values of the metrics MAE and RMSE for assessing the quality of forecasting, obtained by the calculations based on algorithms from Table. 1. The best values for the group are highlighted in bold.

Table 1
Values of metrics for assessing the quality of RS forecasting

	MAE	RMSE
CPTF	**0,6851**	1,0398
MF	0,7649	**1,0162**
UserKNN	0,946	1,2023
SVD++	0,9069	1,2318
SPF	1,1768	1,4756

Figure 5 below shows the same results in the form of graphs, where the abscissa axis shows the abbreviation of the 5 algorithms involved in the experiment, and the ordinate axis for each of them contains the values of the metrics MAE and RMSE respectively.

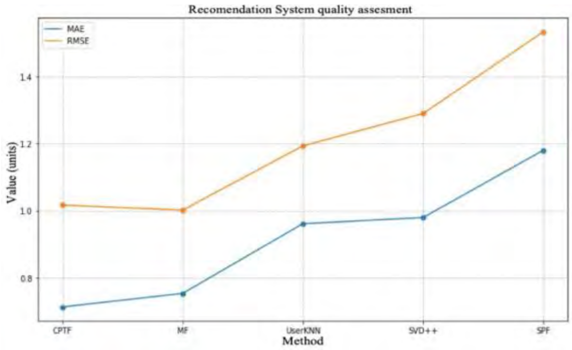


Figure 5: Graphs of MAE and RMSE metric values

Thus, the results of the first experiments indicate that the CPTF tensor factorization method is more accurate for determining dynamic configurations in the “Smart-Home” test DSPL than other methods of joint data filtering because it is a generalization of MF matrix factorization methods. The CPTF method generates recommendations based on the analysis of several dimensions of DSPL configurations, which allows them to dynamically process contextual information as for the needs and requirements of their users.

6. Conclusions and future work

We have analyzed the main possibilities of using RS to solve the problem of finding rational configurations of DSPL, and investigated in more detail the features of the N-dimensional context-dependent tensor factorization method for this purpose in ‘Smart-Home’ applications. This approach adapts the set of system functions to the current context and allows to process the appropriate information in multi-dimensional configuration spaces. The software implementation of this approach was carried out experimentally using the open-source RS project CARSkit [11].

Further work in this direction may include the following activities: 1) study the impact of the number of available configurations and context variables on the effectiveness of the proposed approach; 2) research into the possibility of using hybrid models of RS in this problem domain.

7. References

[1] Reinhartz-Berger, I., Sturm, A., Clark, T., Cohen, S., Bettin, J. (Eds.). Domain Engineering: Product Lines, Languages, and

Conceptual Models. Heidelberg, Springer, 2013. 420 p. (doi:10.1007/978-3-642-36654-3)

[2] Eleut rio J. A Comparative Study of Dynamic Software Product Line Solutions for Building Self-Adaptive Systems, 2017. 30 p.

[3] Marko Gasparic AndreaJanes. What Recommendation Systems for Software Engineering Recommend: A Systematic Literature Review In: Journal of Systems and Software Volume 113, March 2016, pp. 101-113(DOI: 10.1016/j.jss.2015.11.036)

[4] Phuong T. Nguyen, Juri Di Rocco, Davide Di Ruscio. Knowledge-aware Recommender System for Software Development // Knowledge-aware and Conversational Recommender Systems (KaRS) Workshop 2018, Vancouver, Canada.

[5] Yibo Wang, Lothar Hotz and Matthias Riebisch. Feature-based Recommendation for Product Configuration in the Software Product Lines. In: Proc. of 14th Workshop Model-Based Development of Embedded Systems (MBEES2018), 16-18 April, 2018, Dagstuhl, Germany. Fortiss, 2018, pp. 19-30.

[6] Shahab Saquib Sohail, Jamshed Siddiqui and Rashid Ali. Classifications of Recommender Systems: A review. In: Journal of Engineering Science and Technology Review 10 (4), (2017). - pp. 132-153 (doi:10.25103/jestr.104.18)

[7] Qin, J., Liu, B., Qian, J. A Novel Privacy-Preserved Recommender System Framework Based On Federated Learning // ICSIM 2021: The 4th International Conference on Software Engineering and Information Management, January 2021. pp. 82–88 (https://doi.org/10.1145/3451471.3451485)

[8] V. Lytvyn, V. Vysotska, V. Shatskykh et al. Design of A Recommendation System Based on Collaborative Filtering and Machine Learning Considering Personal Needs of The User // Eastern-European Journal of Enterprise Technologies. 2019, vol. 4/2 (100), pp. 6-28.

[9] Pereira J. et al. N-dimensional Tensor Factorization for Self-Configuration of Software Product Lines at Runtime // SPLC '2018: Proceedings of the 22nd International Systems and Software Product Line Conference. 2018, Vol. 1, pp. 87–97.

[10] Rustam Gamzayev, Mykola Tkachuk. Using Methods and Technologies of Recommendation Systems for Dynamic Software Product Lines Configuration // Bulletin of National Technical University "KhPI". Series: System Analysis, Control and Information Technologies, No. 1 (5) (2021). - pp. 91-97 (in Ukr.).

[11] Zheng, Y., Mobasher, B. CARSKit: A Java-Based Context-aware Recommendation Engine // ICDM’2015: Proceedings of the 15th IEEE International Conference on Data Mining, Atlantic City, NJ, USA. 2015, Vol. 1, pp.707-711.

[12] Official Web-site of OMG Consortium. URL: https://www.omg.org/spec/UML/2.0

[13] Pandit, D., Pattanaik, S. Software Engineering - oriented Approach to Iot - Applications: Need of the Day. In: Int. J of Recent Technology and Engineering (IJRTE), Vol.7 (6), 2019. - pp. 886-895.

[14] Mingang Chen, Pan Liuc. Performance Evaluation of Recommender Systems. In: Int J Performability Engineering, 2017, Vol. 13 (8), pp. 1246-1256. (doi: 10.23940/ijpe.17.08.p7.12461256)

Efficiency of Recognition of Pulse Signals at Training on Classified Samples

Volodymyr Orlov [1]

[1] *Military Academy (Odesa), Street Fountain road, 10, Odessa, 65009, Ukraine*

Abstract

Estimations of standards applied to formation decision statistics of recognizing system are analyzed. Simple analytical expressions for the analysis of efficiency of recognition from number of learning samples are received. The analysis of features of application of standards estimations for classes recognition algorithms in conditions of uncorrelated noise has allowed to reveal, that for the normalized standards efficiency of recognition depends on the following parameters signal -noise conditions: attitudes a signal / noise, standards correlation and the size of a standard vector.

Keywords

Recognition, learning samples, speed of convergence, correlation

1. Introduction

The multichannel construction of a system of recognition of radiation sources, forming decision statistics on entrance samples, is applied at designing systems of a passive location of pulse acoustic signals [1, 2].

As standards (reference signals channels) completely known signals are used, and in case of their absence the estimations of the standards received on taxonomic experimental samples are substituted [3].

Application of sample estimates of standards influences distribution laws of decision statistics and their difference from normal causes mathematical difficulties at the task solution of definition of sufficient number of learning samples, especially if criterion of efficiency is the probability of recognition error [3, 4].

In article the choice of signal parameters and the noise allowing to receive simple analytical expressions for the analysis of efficiency of recognition from number of learning samples is carried out. Because of big costs of receipt of experimental data and actual simulation of radiation sources it is necessary to determine models of signals and the noise, allowing to calculate volume of learning sample for maintenance of the set recognition error. In practice the radiation source generates equal signals, which are exposed to atmospheric distortions at propagation at a great distance. As time of correlation of such interferences is much less than intervals between acoustic pulses at carrying out of experiments, samples' fluctuations concerning average are independent.

2. Main part

2.1. The basic theoretical positions

Let's realize the analysis of a simple case of recognition of classes with equal covariance matrixes $C = C_i = C_j$ for known standards S_i and S_j, then we shall consider efficiency of use of their estimations $\hat{S}_i$ and $\hat{S}_j$. Vectors of X images are characterized by the following density of multivariate normal distribution

$$p(X,\omega) = (2\pi)^{-N/2}|C|^{-1/2} \cdot \exp\left[-\frac{1}{2}(X-S)^T C^{-1}(X-S)\right] \quad (1)$$

where S, X is the vectors of the size N of the standard and image's sample, T is the transpose sign.

EMAIL: vorlov@ukr.net
ORCID: 0000-0002-3065-3884

Suppose, all mistakes are equally dangerous and aprioristic probabilities of signals occurrence are equal, that allows for Gauss's signals to reduce the logarithm of the relation of likelihood to set of linear decision functions [4]

$$u_{ij}(X) = \ln p(X,\omega_i) - \ln p(X,\omega_j) = X^T C^{-1}(S_i - S_j) - 0{,}5 \cdot (S_i + S_j)^T C^{-1}(S_i - S_j) \quad (2)$$

and the condition of an X images belonging to a class w_i by criterion of a minimum of classification error probability is determined by decision statistics $u_{ij}(X) > \beta$, where threshold quantity $\beta = \ln p(\omega_j)/\ln p(\omega_i)$. Then $p(u_{ij} > \beta, \omega_j)$ represents probability of wrong image's classification belonging to a class ω_j, and $p(u_{ij} > \beta, \omega_j)$ - probability of wrong image's classification belonging to a class ω_i, and the logarithm of the likelihood ratio $u_{ij}(X)$ is described by normal distribution $\Phi(\overline{u}_{ij}, d)$ with average and a dispersion

$$\overline{u}_{ij} = (S_i - S_j)^T C^{-1}(S_i - S_j)/2 = r_{ij}/2, \quad (3)$$

where $u_{ij} \sim \Phi(r_{ij}/2, r_{ij})$.

In case of equal aprioristic probabilities of occurrence of classes $p(\omega_j) = p(\omega_i) = 0{,}5$ the logarithm of threshold quantity $\beta = 0$, $p(u_{ij} < \beta, \omega_i) = p(u_{ij} > \beta, \omega_j)$ and total error probability

$$p_{ij} = p(\omega_i)p(u_{ij} < \beta, \omega_i) + p(\omega_j)p(u_{ij} > \beta, \omega_j)$$

it is determined as

$$p(u_{ij} < \beta, \omega_i) = \int_{\beta}^{\infty} p(u_{ij}, \omega_j)du_{ij}$$
$$p_{ij} = \int_{0{,}5\sqrt{r_{ij}}}^{\infty} \sqrt{2\pi}\exp\{-y^2/2\}dy = erf(0{,}5\sqrt{r_{ij}}) \quad .(4)$$

where $erf(0{,}5\sqrt{r_{ij}})$ is the tabulated integral of probabilities [4].

In conditions of aprioristic uncertainty lets replace vectors of standards S_i with their maximum likelihood estimations on the k classified samples. Taking into account, that selective covariance matrix is a matrix of X-vectors deflections from the standard S, their estimations look like

$$\hat{S} = \sum_{m=1}^{k} X_{i,m};\ \hat{C}_i = \sum_{m=1}^{k}(X_{i,m} - S)(X_{i,m} - S)^T/k\ . \quad (5)$$

At use of estimations of standards $\hat{S}_i$, $\hat{S}_j$, in decisive statistics (2) it is necessary to average a frequency distribution on these parameters

$$p(u_{ij} > \beta, \omega_j) = \int_{\beta}^{\infty}\int\int p(u_{ij}, \hat{S}_i, \hat{S}_j)d\hat{S}_i d\hat{S}_j du_{ij}\ . \quad (6)$$

2.2. The analysis of efficiency of standards estimation

Owing to complexity of the distribution law (6) and the account of the set of signal parameters and the noise, a finding of mistake probabilitie's exact estimation is difficult. In this connection, it is necessary to choose model signal - noise conditions, allowing to define probabilities of a mistake from volume of learning sample. It is suppose, that covariance matrixes of standards' deviations are equal $C_i = C_j = I_0$, stationary and diagonal with an unit variance, where I_0-an unitary matrix, and all emitted signals have the identical energy equal $q = S_i^T S_i$ to the attitude signal / noise. Then taking into account of equality $p(\omega_i) = p(\omega_j)$ the decisive statistics (2) for i and j's classes is determined by expression

$$u_{ij}(X, \hat{S}_i, \hat{S}_j) = X^T(\hat{S}_i - \hat{S}_j) > 0, \quad (7)$$

and a standard estimation on the classified learning samples $X_{i,m} = S_i + Y_{i,m}$, where a vector uncorrelated noise $Y_{i,m} \sim \Phi(0,1)$, we shall present as

$$\hat{S}_i = k^{-1}\sum_{m=1}^{k}(S_i + Y_{i,m}) = S_i + k^{-1}\sum_{m=1}^{k} Y_{i,m} = S_i + Z_i\ , \quad (8)$$

where $Z \sim \Phi(0, k^{-1})$ is the normally distributed vectors. At presence of the sample containing a signal S_i on an entrance, the decisive statistics (7) looks like

$$u_{ij}(X, \omega_i) = (S_i + Y_i)^T(S_i - S_j + Z_i - Z_j)\ . \quad (9)$$

Having designated $Z = Z_i - Z_j$ distributed ~ $\Phi(0, 2/k)$ and having removed the brackets in (9), we shall receive

$$u_{ij}(X, \omega_i) = Y_i^T S_i - Y_i^T S_j + S_i^T Z + Y_i^T Z + S_i^T S_i - S_j^T S_j = v + t + a\ , \quad (10)$$

where $v = Y_i^T S_i - Y_i^T S_j + S_i^T Z$ has a normal distribution

$$v \sim \Phi(0, (S_i - S_j)^T(S_i - S_j) + 2S_i^T S_i/k);$$

$t = Y_i^T Z$ is the sum of the results from the multiplication of N random variables that have a normal distribution; $a = S_i^T S_i - S_i^T S_j$ is the average of decisive statistics. As t represents the sum of equally distributed random variables and the standard samples' sizes big enough N >50, distribution t according to central limiting Laplas's theorem, comes nearer to normal with parameters $t \sim \Phi(0{,}2N/k)$. Taking into account that v and t are uncorrelated $\overline{vt} = 0$ and normalization of the distribution law of decisive statistics (10), we shall determine its average a and a dispersion d

$$a = S_i^T S_i - S_i^T S_j = q - qr;$$
$$d = (S_i - S_j)^T (S_i - S_j) + 2S_i^T S_i / k + \quad (11)$$
$$2N/k = 2q(1-r) + 2(q+N)/k,$$

where $r = S_i^T S_j$ is a correlation coefficient between the normalized standards. The probability of the erroneous decision is defined by integral of probabilities

$$P_{ij} = 1 - \Phi(a/d);$$
$$a/d = (1-r)\sqrt{q/[2(1-r) + 2/k + 2N/kq]} \quad . (12)$$

According to (12) schedules of dependences of probability of a mistake from standards r correlation, attitudes signal / noise q (of the Figure 1) and numbers of learning samples k (of the Figure 2) for N =50 are constructed.

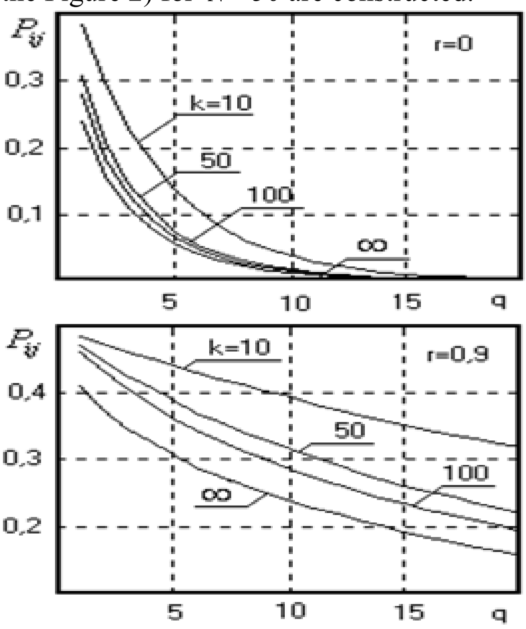


Figure 1: Dependences of probability of a mistake from attitudes signal / noise q: figure up – $r = 0$; figure at the bottom – $r = 0{,}9$

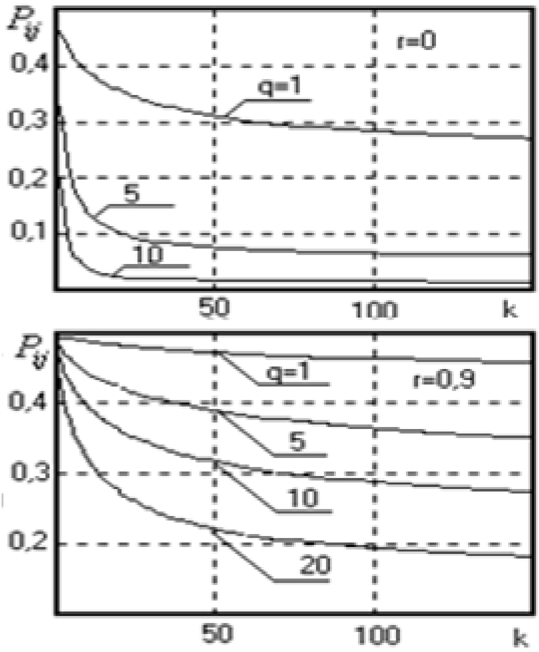


Figure 2: Dependences of probability of a mistake from numbers of learning samples k: figure up – $r = 0$; figure at the bottom – $r = 0{,}9$

The analysis of schedules has shown, that it is possible to reduce number of learning samples due to formation uncorrelated standards with the greater attitude signal / noise. With uncorrelated standards $r = 0$ and attitudes signal / noise $q > 10$ for achievement of a mistake $P_{ij} = 0{,}05$ the necessary number of learning samples is much less than size $k < N$ of a standards. Standards correlation worsens the efficiency of recognition and essentially slows down the speed of convergence to potential value of recognition mistake, causes additional increase in number of learning samples.

2.3. Modeling system on experimental data

Let us present a sample of input process as a vector $Y = X + S_i$ of size N additive mix of a signal S_i with a clutter X, and let us consider M classes $\omega_1 \ldots \omega_M$ of accepted signals, described in the density of normal distribution $p(X, \omega_i)$, $i = \overline{1, M}$ (1), where each density of distribution is completely determined by a vector of the standard S_i and covariance matrix $C_i = R = \overline{XX^T}$, $|R|$ is the determinant of covariance matrix of the clutters. The algorithm of detection - recognition consists of a choice maximal of solving statistics $d_i(Y)$ from signals exceeded a threshold level of detection [5]

$$d_i(X) = W_i^T Y\ ,\ \max d_i(Y) > c_i\ ,\ i = \overline{1, M}\ , \quad (13)$$

where c_i is threshold level of channel i; W is the weight vector of filter of size N in channel i.

Because threshold levels c_i are determined by probabilities of false alarm, depends on S_i and R, that take place computing difficulties for their definition, especial in conditions of adaptation in a real time scale.

Therefore, we are interested of construction of multichannel system with one threshold that is not calculated, but assigned. In [5] received, that a weight vector of solving statistics is

$$W_i = R^{-1} S_i / (S_i^T R^{-1} S_i)^{1/2}\ . \quad (14)$$

Under joint detection-recognition provides stabilization of clutters capacity on outputs of all channels

$$\sigma_i^2 = W_i^T R W_i = S_i^T R^{-1} R R^{-1} S_i / S_i^T R^{-1} S_i = 1 \quad (15)$$

Except of adaptive filters, filters with the constant weight factors adjusted on typical clutter's conditions or coordinated with standards are also widely applied in practice [6]. So, at ignoring of clutters correlation (in particular, because of absence of correlated clutters or complexities of processing) clutters covariance matrix take a diagonal form $R = \sigma^2 I_0$ (where $\sigma^2 = N^{-1} \overline{X^T X}$ is the clutter's power), and weight vector (14) becomes simpler and coordinated with a standard signal

$$W_i = S_i / \sigma (S_i^T S_i)^{1/2}\ . \quad (16)$$

In real conditions reception of signals is carried out in clutters conditions with unknown correlation properties. In connection with aprioristic uncertainty concerning to parameters of clutter, we'll replace a covariance matrix of clutters with its evaluation received by K samples of the size N of input process in absence of useful signals. If the number of samples is less than size of the filter $K < N$, that for exception of singular evaluation of matrix $\hat{R}$ it is applied regularly evaluation [5] with parameter of regularly $\alpha < 10^{-4} \sigma^2$

$$\hat{R} = \sum_{i=1}^{K} X_i X_i^T + \alpha I_0\ ;\quad \hat{\sigma}^2 = (NK)^{-1} \sum_{i=1}^{K} X_i^T X_i\ .$$

Let's carry out research of efficiency of algorithms of detection - recognition on a basis samples from the acoustic sensor received at influence of various shock mechanisms, located outside visibility which signals have pulse character, and quasi stationary wind clutters. For this purpose it was received 500 experimental samples of realizations containing over 1024 readout, removed with interval of time digitization about 1 ms. Under realization of adaptation procedure it is necessary to measure correlation properties of clutters on areas not occupied by a useful signal. Since in experimental data useful signals are located in the beginning of samples up to j <250 readout and their duration, comparable to the sizes of standards, do not exceed N <70, that the classified evaluation of clutters covariance function are provided on interval j =250-1024 readout of input process, carried enough in time from a useful signal.

Let's consider passing of signals in system under investigation. Input realization Y with elements y_j of a current moment of time j passes through everyone i from M filters on which output are formed a solving statistics $d_i(j)$ then it is made a choice of maximal value from them $d_{\max}(j)$

$$d_i(j) = \sum_{n=1}^{N} w_{i,n} y_{j-n}\ ;\ d_{\max}(j) = \max_{1<i<M} d_i(j)\ . \quad (17)$$

Output signal of detector-recognizer is number of the channel i >0, where decision $V(j) = i$ for $d_i(j) = d_{\max}(j)$. In this connection, at calculations of efficiency of recognition system on experimental data, it is necessary to choose a global maximum d on whole output realization $d = \max d_{\max}(j)$ of size j =1-1024, and then make a decision $V(j)$ to what class we need to attitude this sample. Let's consider by the example of illustrations the feature of input and output signals of the system adjusted to M =4 classes of signals. In a Figure 3 and Figure 4 are submitted epures: of input process containing a mix of a signal of 1-st class with a clutter (Fig. 3), from output of the scheme that choice a maximum (Fig. 4, a solid line).

On Figure. 4 we can see that signal from output of the filter of 1-st class exceeds false signals from outputs of other filters. These false peaks essentially exceeding a level of noise are caused by a mismatch of a useful signal on time and responses from others M -1 parallel filters. It is marked, that the average level of false peaks insignificantly differs by size from the filter to the filter, and their maximal interchannel level is defined of standards correlation.

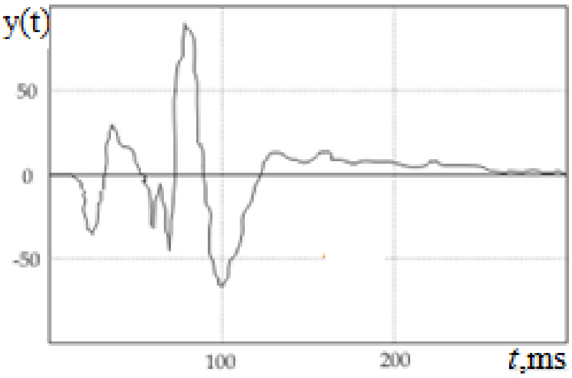


Figure 3: Dependence of the amplitude of the input signal on time

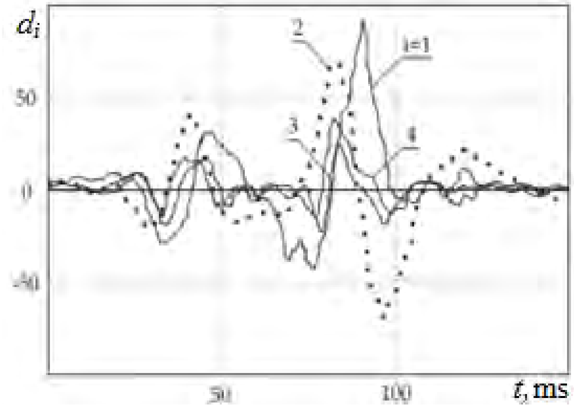


Figure 4: Dependence of the amplitude of the output signal on time

3. Conclusions

The design procedure of number of learning samples is resulted at use of maximum likelihood standards estimations for normally distributed classes learning samples, average values differing by vectors.

The analysis of features of application of standards estimations for classes recognition algorithms in conditions of uncorrelated noise has allowed to reveal, that for the normalized standards efficiency of recognition depends on the following signal-noise environment parameters: attitudes a signal / noise, standards correlation and the size of a standard vector.

The analysis of signals features in researched system has established, that signals after filtering have structure with many peaks, caused by mutually correlation functions of signals and filters which can essentially increase a number of false alarms under detection and complicate recognition of signals. Decrease of lateral peaks influence probably on the basis of detection algorithm choice and on a way of calculation weight factors of filters by covariance matrix of input process, which can contain a useful signal.

4. Acknowledgements

The results presented in the article are based on the work carried out in the Scientific Center of the Military Academy (Odesa).

5. References

[1] M. Hrabina and M. Sigmund, Gunshot recognition using low level features in the time domain, in, Proceedings of the 28th International Conference Radioelektronika, Prague, Czech Republic, 2018, pp. 1-5, doi: 10.1109/RADIOELEK.2018.8376372.

[2] S Raponi, I Ali, G Oligeri, Sound of guns: digital forensics of gun audio samples meets artificial intelligence, arXiv preprint arXiv:2004.07948, 2020 pp. 1-15. URL: https://arxiv.org/pdf/2004.07948.pdf.

[3] Fukunaga K. Introduction to statistical image recognition. Academic Press, New York, NY, 1972.

[4] Gonzalez R.C. Image recognition principles. Addision - Wesley Publishing Company London, 1974.

[5] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing, John Wiley & Sons, 2002.

[6] H. Sánchez-Hevia, D. Ayllón, R. Gil-Pita, M. Rosa-zurera, Gunshot Classification from Single-channel Audio Recordings using a Divide and Conquer Approach, in: Proceedings of the International Conference on Pattern Recognition Applications and Methods, volume 2, ICPRAM, University of Alcalá, 2015, Spain pp. 233-240. doi: 10.5220/000521830233024.

Machine Learning of Models for the Description of Positive Feedback on Social Networks

Nataliia Khymytsia [1], Maria Holub [2]

[1] *Lviv Polytechnic National University University, 12 Bandera street, Lviv, Ukraine*
[2] *Cherkasy State Technological University, Shevchenko Boulevard 460, Cherkasy, Ukraine*

Abstract
The emotional coloring of text messages is considered as one of the sources of information. Intellectual analysis of the emotionality of text messages is studied as a process of monitoring, which includes the organization of continuous observations, the formation of an array of numerical characteristics of the results of observations, and extraction from these arrays of information about the properties of objects. Information technology of intellectual monitoring is programmatically implemented in the form of a monitoring information system and used to detect the tone of the text content of one of the social networks.
The aim is to develop a method and means of using the public opinion of members of the social community in the process of examination of decisions, forming a collective mind. The existing method of text mining is being adapted to solve the new problem of classifying messages on social networks by tonality. The process of searching for texts of a given tonality, the processes of forming the initial description of the object of monitoring, the processes of constructing observation points, interpretation of modeling results have been improved. Thanks to the use of a deep level of decomposition of texts, it is possible to use the results of the analysis of short messages to identify the tone of the messages of the community as a whole. The technology of adaptive synthesis of classifier models, implemented in the monitoring information system, allows dividing texts with opposite tones into classes and use classifier models to assess the attitude of the social community to other given facts, topics, or events. The results of the test of the classifier models revealed that the tonality of all the texts that were selected from one of the thematic groups of the social community Facebook were classified correctly.

Keywords
Social Networks, Text Mining, Sentiment analysis, Classification, Monitoring.

1. Introduction

The use of social network content to study public opinion, in the process of examination of decisions [1], as a manifestation of the collective mind [2] and in other studies, requires automation of the process of identifying the emotional attitude of virtual community members to a message. Nowadays, the study of public opinion in the process of examination of management decisions is actively used [3]. Expressing positive and negative moods of the authors, the authors are happy or unhappy, evil or good, sad or happy, the event caused an urgent or non-urgent reaction, interest or disinterest of members of the social community [4].

This process is called Sentiment analysis [5] or determining the tone of text messages [6].

The classification of textual content in social networks by tone has its own characteristics - the content is formed from many short messages, the assessment of the tone is subjective, does not take into account sarcasm, the use of subtext, and more. The main purpose of this process is to build models-classifiers that can assign a text message to one of the classes: 1) positive assessment; 2) negative assessment.

To do this, the texts are described by numerical characteristics of information features, which will divide their individual properties and divide the texts into classes by key.

The main problems are the search for texts of opposite tones, the construction of the space of

EMAIL: nhymytsa@gmail.com (A. 1);
sashasokolovska92@gmail.com (A. 2)
ORCID: 0000-0003-4076-3830 (A. 1); 0000-0002-0553-8163 (A. 2)

informative features, the choice of teaching method of the classifier model, the interpretation of the classification results.

2. Related Works

To perform the task of determining the tone of the text, three main approaches are used:

1) on the basis of explicitly set by experts rules of parsing,

2) on the basis of the built rules of machine learning;

3) hybrid approach [7].

In each of the approaches, special attention is paid to the process of pre-processing of texts [8].

A detailed review of the use of methods of parsing the tone of text messages on Twitter is made in [9].

A method of constructing classifiers based on several models is popular [10].

It is already traditional to use machine learning methods to solve the problems of classification of text messages, in particular the classification of texts by the tone of formalized [11].

Features of this process are:

1) weak formalization of the processes of class formation (it is not entirely clear which texts are considered positive and which are negative; how to take into account sarcasm, irony, etc.);

2) a large number of short text messages.

Large messages are evaluated by content analysis by statistical processing of the number of "thematic" words [12]. For the analysis of short messages in social networks in [13] the probabilistic method of construction of the classifier model is used. The array of input data is formed at the expense of the characteristics of tags. It is claimed that the classifier, based on the Bayesian model, allows to correctly classify 77% of Twitter messages after training. The application of new methods of Text Mining is described [14].

The process of assessing the tone of the text by machine learning, ie the list of stages and their sequence, contains common stages [11]. Common stages in the technologies of text tone analysis are:

- “collection of a collection of documents, on the basis of which the machine classifier is studied;
- each document is decomposed into a vector of features (aspects) by which it will be studied;
- indicates the correct type of key for each document;
- the choice of classification algorithm and method for training the classifier is made;
- the obtained model is used to determine the tone of the documents of the new collection"[16].

3. Description of the problem and problem statement

The tone of social network messages on a given topic is a relevant source of information that can be used to perform monitoring tasks. To analyze the tone of text messages, it is necessary to solve the problem of their classification by constructing and using a classifier model [17]. The classifier must perform a complex task - assigning a new message to one of the classes of texts and measuring the tone of this text. To measure the tonality of the text, a scale should be used, which was used to evaluate texts in the formation of their classes.

Expertly set two classes of texts of opposite key:

$$K = \{k_1, k_2\}, \tag{1}$$

where k_1 – texts containing positive assessments; k_2 – texts, the authors of which negatively evaluate the topic given for discussion.
Texts with a predetermined key are selected to fill the classes:

$$T = \{t_1, t_2, ..., t_n\}, \tag{2}$$

where n – the number of text messages on a given topic, the tone of which could be estimated.

It is necessary to build a model-classifier f, which will display texts that were not used to build the classifier from the set $T^* = \{t_{n+1}, t_{n+2}, t_{n+p}\}$, $T^* \in T$ to the elements of the set of classes K (1):

$$f: T^* \to K. \tag{3}$$

To construct the classifier model f (3), the elements of the array T (2) are presented in the form of an array of numerical features

$$\begin{pmatrix} y_1 & x_{11} & x_{12} & \dots & x_{1l} \\ y_2 & x_{21} & x_{22} & \dots & x_{2l} \\ \dots & \dots & .. & \dots & \dots \\ y_m & x_{m1} & x_{m2} & \dots & x_{ml} \end{pmatrix} \tag{4}$$

where x_{ij} – value j-th sign i-th text message box, y_i – the value of the class attribute in i-th text box, m – the number of windows after text decomposition, l – the number of features of the text message for a given class.

The results of the classification are evaluated by the number of correctly classified texts N from the set T^*:

$$N \to max \qquad (5)$$

4. Research of the process of classification of tonal texts

The emotional attitude of the participants of the social network Facebook to a given topic was determined by classifying the textual content of the social network and on the basis of a combination of individual messages of authors belonging to one thematic group.

Classification of the attitude of representatives of the social network Facebook to a common theme, contains typical [17] stages:

- division of text messages by key;
- combination of texts of the same key in classes;
- decomposition of the text;
- formation of the space of classification features;
- construction of an array of input data (AID);
- construction of a decisive rule (model-classifier) for grouping texts;
- use of a classifier model to assess the tone of the content of the thematic group of the social community.

Figure 1 shows a diagram of the activity, which describes the process of analyzing the tone of text messages of the social network.

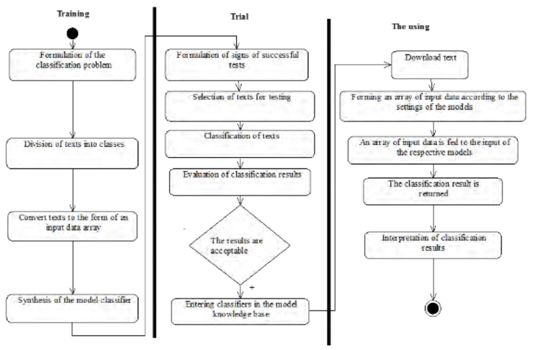


Figure 1:The process of analyzing the tone of text messages

The tone of the messages is determined in several stages: training, testing and use of the classifier model. When formulating the problem of classification of texts by key determine the list of classes, classification features, criteria for the success of the process.

Unlike content analysis, where tonality is assessed by the frequency of use of certain words, our method is based on the analysis of a combination of letters and letter combinations in texts of different tonality.

To fill classes with texts of opposite key, a hypothesis was put forward that the key of the text corresponds to the key of his keywords. According to this hypothesis, the search for texts of opposite tones was conducted among the messages of one content by keywords. Keywords were selected from the key dictionary [18].

Tables 1 and 2 contain examples of words that define the positive and negative tone of the texts.

Texts that contain words with a known tone were assigned the meaning of the tone of these words.

Because it is necessary to assess the tone of all content, classes for training classifiers were built by combining social network messages with the same tone in a single text, which was subjected to intellectual analysis [18].

To construct an array of input data, the decomposition of texts formed from the textual content, which was generated as a result of discussing a common topic, was used. To form the space of signs, a deep decomposition of the text to the level of signs was carried out [17].

To form observation points in the multidimensional feature space, the text was divided into sections (windows) of 50 characters, which is proportional to the volume of the social network message. A separate observation point described the numerical characteristics of a single window. The sequence of points in the table of the input array described the text formed from the content of one key of the thematic group of the social network Facebook. Thus, the input array contained points describing texts of opposite tones.

A model experiment was performed to test the hypothesis. The tonality of messages of the social network Facebook in the volume of 18075 characters was investigated. Classes of texts were formed from messages with a positive tone of 11063 characters and messages with a negative tone of 7012 characters. .

Table 1.
Keywords with a negative tone

№	Keyword	№	Keyword	№	Keyword	№	Keyword
1	suffering	36	devilish	71	inhuman	106	fusillade
2	diseasea	37	lifelong	72	hate	107	destroy
3	hostility	38	harassment	73	hateful	108	suicidal
4	weakness	39	to die	74	hatred	109	suicide
5	scantily	40	ass	75	untidy	110	satanic
6	hatred	41	stupidity	76	disrespectful	111	smelly
7	aggression	42	stupid	77	uncomely	112	deathly
8	satan	43	sneering	78	unwelcome	113	suicide bomber
9	apathetic	44	terribly	79	unhappy	114	sadly
10	incompetence	45	victim	80	misfortune	115	pimp
11	ruthlessly	46	ass	81	worthless	116	bitchy
12	unpunished	47	cruel	82	nobody	117	creature
13	cattle	48	brutality	83	paltry	118	terrorist
14	slaughterhouse	49	to perish	84	nauseating	119	terrorist
15	thug	50	threaten	85	burnt	120	tyrannical
16	to kill	51	abuse	86	cheesy	121	blunt
17	killer	52	malefactor	87	pederasty	122	severe
18	expulsion	53	crime	88	pedophile	123	motherfucker
19	die out	54	abuse	89	false	124	bastard
20	ridicule	55	treason	90	to spit	125	squalor
21	expendable	56	treason	91	walloping	126	defective
22	thug	57	idiocy	92	prostitute	127	bullshit
23	sitting	58	idiotic	93	impair	128	bribe-taker
24	despair	59	infection	94	foreign	129	clumsy
25	perfidious	60	backbreaking	95	punishment	130	dick
26	enemy	61	finished	96	pornography	131	hide
27	shameful	62	steal	97	intensification	132	ruffianly
28	shameful	63	collapse	98	ugly	133	bespatter
29	shame	64	bloody	99	ugly	134	shahid
30	slake	65	lindenly	100	frightful	135	rotten
31	heroin	66	loser	101	obscene	136	oppress
32	abomination	67	rage	102	oppression	137	degenerate
33	shittyly	68	abomination	103	regrettable	138	consumption
34	sodom	69	dead	104	humiliate	139	espionage
35	ashamed	70	drugs	105	humiliating	140	squintinly

Table 2.
Keywords with a positive tone

№	Keyword	№	Keyword	№	Keyword	№	Keyword
1	authentic	31	improvement	61	highly effective	91	persistent
2	authority	32	blissful	62	professional	92	persistence
3	authoritative	33	bliss	63	upscale	93	early
4	adequate	34	neighbor	64	highly productive	94	universal
5	adequately	35	near	65	highly professional	95	pleasure
6	neat	36	shine	66	high quality	96	delight

7	altruistic	37	shining	67	courageous	97	delight
8	appetizing	38	god	68	distinctive	98	grab
9	many	39	heroic	69	amaze	99	admire
10	multidimensional	40	divine	70	impressive	100	useful
11	multifaceted	41	godly	71	hostile	101	beloved
12	with many children	42	wrestler	72	adjustment	102	laureate
13	promising	43	weathered	73	omnipotent	103	moral
14	cheerful	44	polite	74	ecumenical	104	courage
15	desirable	45	politeness	75	all-powerful	105	super-profitable
16	colorful	46	solve	76	genial	106	incomparable
17	sinless	47	successful	77	heroically	107	exclusive
18	impeccable	48	successfully	78	hero	108	pomp
19	impeccably	49	improve	79	heroic	109	preferential
20	cheerful	50	grateful	80	worthy	110	useful
21	boon	51	fancy	81	effective	111	fastness
22	charity	52	lucrative	82	effectively	112	solemn
23	euphonious	53	profitably	83	unity	113	successfully
24	prosperous	54	benefit	84	desirable	114	charming
25	safely	55	win	85	compassionate	115	charm
26	noble	56	spectacular	86	pitiful	116	miraculous
27	prudent	57	liberating	87	humorous	117	remarkable
28	to bless	58	outpace	88	alive	118	perfectly
29	benevolent	58	exculpatory	89	feminine	119	happiness
30	wholesome	60	expressive	90	providing	120	happiness

The list of features was built according to the method described in [14].

To form an array of input data, the informativeness (frequency of use in the windows of texts of one class) of more than 12,000 features was studied. The informativeness of the signs was assessed by the formula:

$$P_i = \frac{\gamma_i}{\sum_{i=1}^{n} \gamma_i} 100\% \quad (6)$$

where P_i – indicator of informativeness i- signs (probability of using the sign in the window), γ_i – number of use cases i- signs in a separate window, n – the total number of characters in the window.

The list of informative features for each class of texts was individual. And the model-classifier was built for each class of texts separately. The limit of informative sufficiency (LIS) was determined expertly for the sign: $P_i^{min} = 3$. For the class of texts with a positive tone, 129 dictionary features were selected. After testing the model synthesis algorithms (MSA) for this array of input data, the monitoring information system (MIS) selected a multi-row algorithm GMDA [19].

AID contained 360 observation points, of which 50 points formed a test sequence and did not participate in the construction of the classifier model. When constructing the AID, the tone of each observation point of textual content that contained negative assessments, marked +2. The maximum tone of the positive content was set by the value –2. That is, the result "My" was given to a class that contained textual content with a negative tone.

It should be noted that the list of features of the AID for each class of texts was only partially compatible. Thus, an individual dictionary of features was formed for a separate class of texts. In the process of constructing the decision rule, the structure of models-classifiers was formed on the basis of different arrays of input data [14].

In the process of assessing the tone, the same text was described by two separate arrays of input data. They were intended for use by a given model-classifier. This provided better results for the classification of new texts.The result of the classifier test at each observation point is the tone value of this content window. Observation points were evaluated as "My" provided that the simulation result exceeded the value of 0.36. In the table, this result is marked with a sign «+». All other points were evaluated by the result "Alien" and marked with a sign in the table «–».

Table 3 presents the test results of the classifier model on textual content, the observation point of which was formed from the messages of different authors of the same key.

Table 3

Test results of the tonality classifier

№ observation	Content tone	The result of the classification	№ observation	Content tone	The result of the classification
1	N	+	7	N	+
2	N	+	8	N	+
3	N	–	9	N	+
4	N	+	10	N	+
5	N	+	11	N	+
6	N	–	12	N	+
13	N	-	19	N	+
14	N	+	20	N	+
15	N	+	21	N	+
16	N	-	22	N	+
17	N	+	23	N	+
18	N	+	24	N	+
25	N	+	31	P	+
26	P	–	32	P	+
27	P	+	33	P	+
28	P	+	34	P	–
29	P	+	35	P	–
30	P	+	36	P	–
37	P	–	43	P	+
38	P	–	44	P	+
39	P	–	45	P	–
40	P	+	46	P	+
41	P	-	47	P	–
42	P	+	48	P	+

5. Results

According to the classification, content with a negative tone contained points, 84% of which are classified with the result "My". Content with a positive tone is classified by the built model in 56% of the points describing these texts. Overall result - 70% of observation points of textual content that did not participate in the construction of the model are correctly classified.

The constructed model-classifier can be used to quantify the tone of the textual content of the social network. The tone of the content as a whole was quantified by averaging the values of the assessment of the tone of the points that describe this text. As a result of testing the classifier, the average tonality of content that contained negative ratings became 0.72. Thus, since the maximum value of the text tone is 2, the negativity of this content is set at 36%. From this we can assume that the social network, the textual content of which was studied, includes people with a high enough level of culture, who usually tolerantly express their negative attitude to the event.

6. Conclusions

In order to expand the capabilities of the monitoring information system, the process of assessing the tone of the text content of one of the thematic groups of the social network Facebook was studied. It is proved that the tonality of texts can be estimated by the tonality of its keywords. To search for texts of a given key, it is suggested to use a dictionary of key tones.

The limit of informative sufficiency of the space of classification features of text messages of social networks is exceeded under the condition of construction of these features on the basis of separate signs and description by a separate observation point of text areas, not less than 50 characters.Using a MIS model synthesizer, a classifier was built that provided the correct grouping of 70% of the AID observation points. Thus, an unmistakable qualitative assessment of the tone of textual content is provided. It is also proposed to quantify the tonality of the text using the signal at the output of the classifier model.

7. References

[1] Y. V. Kotel′nykov, M.V. Klekovkina, Vyznachennya vah otsinochnykh sliv na osnovi henetychnoho alhorytmu v zadachi analizu tonu tekstu, Prohramni produkty ta systemy (2013), № 4. 297–300.

[2] Dzheff Hau, Kraudsorsing. Kollektivnyj razum kak instrument razvitija biznesa, Moscov, M, 2012.

[3] Yaming Zhang, Majed Abbas,Wasim Iqbal, Analyzing sentiments and attitudes toward carbon taxation in Europe, USA, South Africa, Canada and Australia, Sustainable Production and Consumption: Elsevier (2021), Vol. 28., 241-253.

[4] G.R. Ramya, P. Bagavathi Sivakumar, An incremental learning temporal influence model for identifying topical influencers on Twitter dataset, Soc. Netw. Anal. Min. 11, 27 (2021). doi: 10.1007/s13278-021-00732-4.

[5] N.Vyrva, Sentiment Analysis in Social Media, Master's Thesis in Computer Science, Halden: Ostfold University College (2016).

[6] D.V. Shynhalov, E.V Meleshko, R.N. Mynaylenko, V.A. Reznychenko, Metody avtomatychnoho analizu tonal′nosti kontentu u sotsi-al′nykh merezhakh dlya vyyavlennya informatsiyno-psykholohichnykh vplyviv, Tekhnika v sil′s′kohospodars′komu vyrobnytstvi,haluzeve mashynobuduvannya, avtomatyzatsiya (2017), Vyp. 30. 196-202.

[7] W. Medhat, A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: a survey, Ain Shams Eng J., 2014, 5(4):1093–113.

[8] D-K. Nguyen-Nhat, H-T. Duong, One-Document Training for Vietnamese Sentiment Analysis. In: Tagarelli A, Tong H, editors. Computational Data and Social Networks, vol. 11917. Cham: Springer International Publishing, 2019, pp. 189–200.

[9] HT. Duong, TA. Nguyen-Thi, A review: preprocessing techniques and data augmentation for sentiment analysis, Comput Soc Netw 8,1 (2021). doi:10.1186/s40649-020-00080-x.

[10] K. Dashtipour, C. Ieracitano, F. Carlo Morabito, A. Raza, A. Hussain, An Ensemble Based Classification Approach for Persian Sentiment Analysis. In: Esposito A., Faundez-Zanuy M., Morabito F., Pasero E. (eds) Progresses in Artificial Intelligence and Neural Systems. Smart Innovation, Systems and Technologies, vol 184. Springer, Singapore. (2021). doi:10.1007/978-981-15-5093-5_20.

[11] F. Sebastiani, Machine learning in automated text categorization, ACM Computing Surveys, (2002), no. 34 (1), 1–47.

[12] Liu, Bing, Sentiment Analysis and Subjectivity, Handbook of Natural Language Processing (Second ed.). Editors: In Indurkhya, N.; Damerau, F. J. (2010).

[13] S. Bhatnagar, N. Choubey, Making sense of tweets using sentiment analysis on closely related topics. Soc. Netw. Anal. Min. 11, 44 (2021). doi:10.1007/s13278-021-00752-0.

[14] Holub Maria, Piven Oleg, Classification of texts in the technology multilevel information monitoring. Inzynier XXI wieku, Wydawnictwo naukowe Akademii Techniczno-Humanistycznej w Bielsku-Bialej. (2016) 119-122.

[15] Sebastiani F. Machine learning in automated text categorization. ACM Computing Surveys, 2002, no. 34 (1), pp. 1–47.

[16] Analiz_tonalnosti_tekstu.URL:https://uk.wikipedia.org/wiki/%D0%90%D0%BD%D0%B0%D0%BB%D1%96%D0%B7_%D1%82%D0%BE%D0%BD%D0%B0%D0%BB%D1%8C%D0%BD%D0%BE%D1%81%D1%82%D1%96_%D1%82%D0%B5%D0%BA%D1%81%D1%82%D1%83

[17] S. Holub, N. Khymytsia, M. Holub, S. Fedushko, The intelligent monitoring of messages on social networks. CEUR Workshop Proceedings, Vol 2616, Proceedings of the 2nd International Workshop on Control, Optimisation and Analytical Processing of Social Networks (COAPSN-2020), Lviv, Ukraine, 2020, pp. 308-317. URL:http://ceur-ws.org/Vol-2616/paper26.pdf

[18] Tone-dict-uk.URL:https://github.com/lang-uk/tone-dict-uk/blob/master/tone-dict-uk-manual.tsv

[19] A.G. Ivakhnenko, Inductive method of self-organization of models of complex systems, Kiev, Naukova Dumka, 1981. 296 p.

Making Predictions of the Stress-Strain State of a Cylindrical Shell under Extrenal Unform Pressure

Oksana Choporova[1], Serhii Choporov[2] and Andrey Lisnyak [3]

[1] *Zaporizhzhia National University, Zhukovsky str., 66, Zaporizhzhia, 69600, Ukraine*
[2] *Zaporizhzhia National University, Zhukovsky str., 66, Zaporizhzhia, 69600, Ukraine*
[3] *Zaporizhzhia National University, Zhukovsky str., 66, Zaporizhzhia, 69600, Ukraine*

Abstract

This article is devoted to the problem of prediction of the stress-strain state of a cylindrical shell using neural networks. The article shows how to train the neural network to calculate the von Mises stress of a cylindrical shell under external uniform pressure.

Machine learning, one of the six disciplines of Artificial Intelligence (AI) without which problems of having machines acting humanly could not be accomplished. Applications of ANNs to engineering structures arise in a variety of industries such as engineering, automotive, space structures, etc. ANNs allow to develop models e.g. for the stress-train state estimation of some type of solids. Thus, the development of machine learning methods for predicting the behavior of engineering structures is urgent.

The objective of this article is the development of a prediction model with a dataset of possible states of the cylindrical shell. The model should be able to predict the stress-strain state of a cylindrical shell of a homogeneous material. We also presume that uniform external pressure is applied on the shell. The model uses the following properties as an input: radius, length, thickness h, Young's modulus, Poisson's ratio, pressure. Edges of the shell may be clamped, simply supported, or free. Each node of the shell has 6 degrees of freedom: 3 translations and 3 rotations. The clamped edge implies that all nodes of the edge cannot be translated or rotated. The simply supported edge implies that for all nodes of the edge cannot be translated but rotation is allowed these nodes. The developed model allows to predict the von Mises stress with 10% of the mean absolute percentage error. The key advantage of an artificial neural network is the speed of prediction. The ANN predicts the von Mises stress almost instantaneously (milliseconds) comparing the finite element method.

Keywords

Machine Learning, Artificial Neural Network, Stress-Strain State, Shell, Prediction, Neurons, Data Generation Algorithm

1. Introduction

Artificial neural networks are information processing structures providing the (often unknown) connection between input and output data by artificially simulating the physiological structure and functioning of human brain structures.

The natural neural network consists of a very large number of nerve cells (about ten billion in humans), said neurons, and linked together in a complex network. The intelligent behavior is the result of extensive interaction between interconnected units. The input of a neuron is composed of the output signals of the neurons connected to it.

Artificial neural networks are composed of elementary computational units called neurons combined according to different architectures. For example, they can be arranged in layers (multi-layer network), or they may have a connection topology. Layered networks consist of input layer, made of n neurons (one for each network input); hidden layer, composed of one or more hidden (or intermediate) layers consisting of m neurons; output layer, consisting of p neurons (one for each network output).

EMAIL: o.choporova @gmail.com (A. 1); s.choporoff@znu.edu.ua (A. 2); a.lisnyak@znu.edu.ua (A. 3)
ORCID: 0000-0003-3167-7869 (A. 1); 0000-0001-5932-952X (A. 2); 0000-0001-9669-7858 (A. 3)

Neural networks are often used for statistical analysis and data modelling, in which their role is perceived as an alternative to standard nonlinear regression or cluster analysis techniques. Thus, they are typically used in problems that may be couched in terms of classification, or forecasting. Some examples include image and speech recognition, textual character recognition, and domains of human expertise such as medical diagnosis, geological survey for oil, and financial market indicator prediction.

ANNs have been applied to an increasing number of world problems of considerable complexity. Their most important advantage is in solving problems that are too complex for conventional technologies – problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found. In general, because of their abstraction from the biological brain, ANNs are well suited to problems that people are good at solving but computers are not. These problems include pattern recognition and forecasting (which requires the recognition of trends in data). However, unlike the human capability in pattern recognition, the ANN's capability is not affected by factors such as fatigue, working conditions, emotional state, and compensation.

Applications of ANNs to engineering structures arise in a variety of industries such as engineering, automotive, space structures, etc. ANNs allow to develop models e.g. for the stress-train state estimation of some type of solids. Thus, the development of machine learning methods for predicting the behavior of engineering structures is urgent.

In aviation engineering and space structures, the use of cylindrical solids with a small thickness is widespread. Such solids could be modeled by shells. Models based on ANNs should process the geometric and mechanical parameters of the solid, as well as boundary conditions.

2. Methods

2.1. Literature review

The rapid development of computer technologies led to the fact that a significant part of computing work was entrusted to computers. Due to the high speed of calculations, their low cost and sufficient accuracy for many applied problems, it became possible to use methods "heavy" in terms of computations and time expenditures for solving mathematical problems. Examples include enumeration methods, iterative, methods using large amounts of statistical data). However, along with the previously created solution methods and algorithms, new ones began to appear, the existence of which apart from the computer is difficult to imagine.

The increasing popularity of artificial neural networks leads to an increasing number of researches devoted to the development of ANN models for modeling various fields. Modelling of solids, the stress-train state is also possible domain of ANNs applications. The paper [4] describes the scheme of machine learning using for the stress-strain state analysis of a rectangular plate with a circular cut-out. The plate might be of arbitrary sizes and the cut-out might be of an arbitrary radius. Each side of the plate is supposed to be free, supported or fixed. Additional input parameters of the data set are following: size of plate's side, thickness of the plate, Young's modulus, Poison's coefficient, and pressure load. Initial parameters have been random generated. The training set is generated by the finite element method. The artificial neural network merges numerical and one-hot input layers. The developed regression model allows to predict von Mises stresses for a rectangular plate with a circular cut-out.

In [10], the proposed strategy demonstrates the effectiveness of machine learning to reduce experimental efforts for damage characterization in composites. In [6], a neural network based on the Kalman filter is employed to predict the collapse of a highway on a bridge processing temperature and oscillation data. In [7], a model based on the self-organizing map of Kohonen is developed to detect the fracture using vibration data. The article [15] deals with the use of a neural network technology for researching the wave damping effect of a solid foundation plate in the closed system "building – foundation – ground" under vibratory impact on the ground. For this purpose, the solutions of direct and inverse forecasting problems are studied on the basis of the developed practical method of step neuronet forecasting. The last has been successfully tested on the model problems of mathematics, the theory of elasticity and plasticity as well as static problems of structural mechanics and building structures. The article [16] formulates the efficiency suppositions of using neuronet approaches to construction engineering problems. It enumerates some results obtained by examples of the problems of the theory of elasticity and plasticity, structural mechanics and engineering structures in the field of control, optimization and forecasting. Machine learning and deep learning have become increasingly more widespread in computational mechanics and mechanical analysis [19–21]. For instance, in [22] authors developed a deep learning model to directly estimate the stress distributions of the aorta. In addition, in [23], they also used machine learning to estimate the zero-pressure geometry for two given pressurized geometries of the same patient at two different blood pressure levels for the human thoracic aorta. In [20] approximated the large deformations of a nonlinear, muscle actuated beam by using a deep-autoencoder to simulate soft tissue biomechanics. The article [24] demonstrated that machine learning can be used as a supplement to the finite element analysis of physical system modeling, and it can efficiently address some corresponding problems. Moreover, neural network algorithms have also been used to replace the conventional constitutive material model to improve finite element analysis [25]. Machine learning and neural networks also perform well in other civil and geotechnical engineering research areas, such as the prediction and classification of rockburst [26], [27].

Thus, the analysis of last researches and publications allows to conclude that problems of developing models based on neural networks for predictions stress-stain state are actual.

2.2. Problem statement

Typically, computer-aided engineering software uses numerical methods for the stress-strain state calculation. For example, the commonly used finite element method allows to calculate components of the stress-strain state with a good accuracy. Nevertheless, the finite element analysis could be a time-consuming process.

The design process implies carrying out some variety of experiments for the optimal construction development. In this process, artificial neural networks (ANN) could be used for fast estimation of the stress-strain state. ANN could be used as interactive assistant in the design process allowing fast examination engineering hypothesizes. In this case ANN should use geometric and mechanical parameters as an input and components of the stress-strain state as an output. The problem of prediction the parameters of the state of a cylindrical shell by its geometric and mechanical parameters could be classified as a regression problem.

The objective of this analysis is to make up a prediction model which should be able to predict the stress-strain state of a cylindrical shell of a homogeneous material. We also presume that uniform external pressure is applied on the shell. The shell has following properties: radius r, length l, uniform thickness h, Young's modulus E, Poisson's ratio ν. The shell is loaded transversely by a distributed load q per unit area (uniform pressure). Edges of the shell may be clamped, simply supported, or free. Each node of the shell has 6 degrees of freedom: 3 translations and 3 rotations. The clamped edge implies that all nodes of the edge cannot be translated or rotated. The simply supported edge implies that for all nodes of the edge cannot be translated but rotation is allowed these nodes.

3. Research

For machine learning, the initial set of values is divided into two parts: training and test. The training part contains the values of the input parameters. The test part is usually smaller than the training part, its values are used to check (in the case of regression analysis) the accuracy of training.

A dataset is generated using the finite element method. Parameters of a shell are randomly generated with following restrictions:

$l \in [0.1; 5]$ (meters);

$r \in [0.1; 5]$ (meters);

$h \in [\frac{1}{100}a; \frac{1}{20}a]$(meters), where $a = min(l, 2\pi r)$ "Generally, it is assumed that the thickness of shell is in the approximate range of $1/20 - 1/100$ of its span.";

$E \in [50000; 300000]$ (MPa);

$\vartheta \in [0; 0{,}45]$;

$q \in [0.01; 0.1]$ (MPa).

Boundary conditions are also randomly generated, boundary condition are categorical data. Possible boundary conditions are following: a free edge, a supported edge, a clamped (fixed) edge. Any combination of boundary conditions is possible excluding the situation of two free edges.

The data generation algorithm is shown below.

```
function generation(N)
begin
  D ← ∅
  ic ← 0
  while ic < N do
```

begin
$l \leftarrow \text{random}(0.1, 5)$
$r \leftarrow \text{random}(0.1, 5)$
$a \leftarrow \min(1, 2\pi r)$
$q \leftarrow -\text{random}(0.001, 0.1)$
$E \leftarrow \text{random}\ (50000, 300000)$
$\vartheta \leftarrow \text{random}(0, 0.45)$
$h \leftarrow \text{random}\left(\frac{a}{100}, \frac{a}{20}\right)$
$mesh \leftarrow CylinderQuads\left(m, \frac{l\,n}{\pi r}, r, l\right)$
foreach $C_0\ in\ [0; 2]$ **do**
begin
foreach $C_1\ in\ [0; 2]$ **do**
begin
if $C_1 + C_1 \geq 1$**then**
begin
$f \leftarrow \left(q \cdot \cos\left(\text{atan}\frac{z}{x}\right), 0, q \cdot \sin\left(\text{atan}\frac{z}{x}\right)\right)$
function left (x, y, z)
begin
if $C_0 = 2$ **and** $y = 0$ **then** **return** ALL
else if $C_0 = 1$ **and** $y = 0$ **then**
return FIRST|SECOND|THIRD
return FREE
end
function right (x, y, z)
begin
if $C_1 = 2$ **and** $y = l$ **then** **return** ALL
else if $C_1 = 1$ **and** $y = l$ **then**
return FIRST|SECOND|THIRD
return FREE
end
$D \leftarrow D \cup MindlinShell(mesh, h, E, \vartheta, left, right, f)$
$ic \leftarrow ic + 1$
end
end
end
end
return D

where MidlinShell calls a function of the finite element analysis of shells; random (a, b) is a function that generates a float value from the interval [a; b]; N is the number of experiments.

The model of an artificial neural network includes two input branches of neurons (see Fig. 1). Branches use dense input layers of neurons to separate an input for numerical and categorical data. Each branch could include few hidden dense layers of neurons. Then the branches are merged combining the output the hidden layers for numerical and categorical data processing. An arbitrary number of hidden dense layers could follow the layer for the merged output. Finally, the last hidden dense layer is connected with the top dense layer.

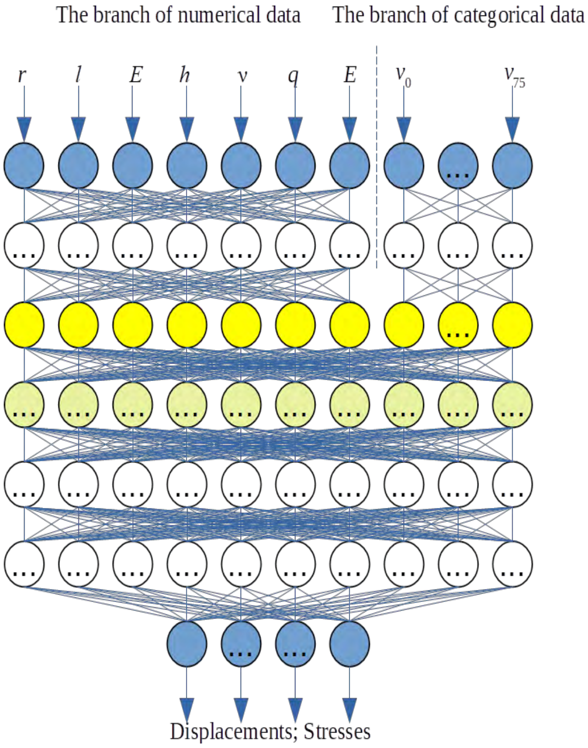


Figure 1: The layers of neurons in the model

The model is implemented using Keras and Python as following codes:

```
float_model = Sequential()
float_model.add(Dense(21,
input_dim=number_of_floats,
activation='tanh'))
onehot_model = Sequential()
onehot_model.add(Dense(190,
input_dim=number_of_categories,
activation='relu'))
combined_input =
concatenate([float_model.output,
onehot_model.output])
combined_layer = Dense(200,
activation='relu')(combined_input)
combined_layer = Dense(200,
activation='relu')(combined_layer)
combined_layer = Dense(200,
activation='relu')(combined_layer)
combined_layer =
Dense(len(Y.columns),
activation="linear")(combined_layer)
model =
Model(inputs=[float_model.input,
```

```
onehot_model.input],
outputs=combined_layer)
```

Where number_of_floats is the number of numerical input features and it is equals to 7; number_of_categories is the number of boundary conditions combinations which is equals to 8; the branch for numerical data is called float_model uses 21 neurons and the hyperbolic tangent activation; the branch for categorical data is called onehot_model and uses 190 neurons and Rectified Linear Unit activation; the combined branch concatenates outputs of the numerical branch and the categorical branch, it uses 3 hidden dense layer, each hidden layer has 200 neurons and Rectified Linear Unit activation; the output layer is a dense layer with the linear activation.

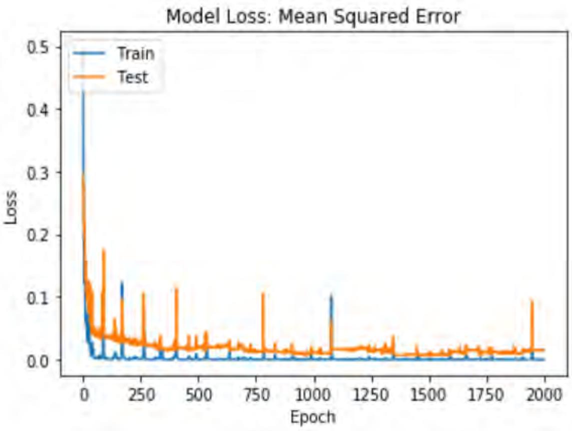


Figure 2: The mean squared error in the best model

The dataset is split into train and test subsets. Train is 80%, test 20%. Train and test are stratified by categorical data (boundary conditions). Training is performed in 2000 epochs. The Fig. 2 shows the mean squared error in the best model for training and testing. The Fig. 3 shows the mean absolute percentage error.

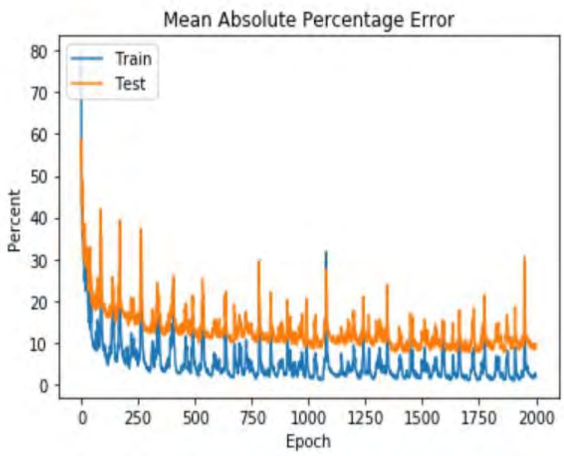


Figure 3: The absolute percentage error in the best model

4. Conclusions

The developed ANN allows to predict the maximum von Mises stress of a cylindrical shell loaded transversely by a distributed load. ANN includes two separate branches that are concatenated in the combined branch. The first branch uses numerical parameters of the model. The second branch processes one-hot encoded boundary conditions as categorical data. Such design allows to train the ANN with a good convergence. The mean absolute percentage error of ANN predictions is approximately 10%. Prospects for further studies are associated with genetic algorithms using in ANN optimization.

The key advantage of an artificial neural network is the speed of prediction. The ANN predicts the von Mises stress almost instantaneously (milliseconds) comparing the finite element method. The von Mises stress is typically used in the strength analysis. So, pretrained artificial neural networks might be used as an interactive assistant in CAE or CAD software.

Further researches are related to the development of artificial neural networks that will predict the stress-strain state according to the drawing or image of shell structures using machine vision and classification algorithms.

5. References

[1] Hertzmann A, Fleet D. Machine Learning and Data Mining, Lecture Notes CSC 411/D11, Computer Science Department, University of Toronto, Canada(2012).

[2] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5(4):115–133 (1943).

[3] Hern A. Google says machine learning is the future. Available at: www.theguardian.com/technology/2016/jun/28/all.

[4] Choporova O., Lisnyak A. “Using machine learning to predict the stress-strain state of a rectangular plate with a circular cut-out” CEUR Workshop Proceeding s. 2020. T. 2791. C. 1-6. URL:https://www.scopus.com/record/display.uri

?eid=2-s2.0-85098720535 &origin= resultslist

[5] Keras. URL: https://www.tensorflow.org/guide/keras.

[6] Abambres M., Marcy M., Doz G. “Potential of Neural Networks for Structural Damage Localization”, engrXiv. 2018. URL: Available: engrxiv.org/rghpf/ doi: 10.31224/osf.io/rghpf.

[7] C. Jin, S. Jang, X. Sun “Damage detection of a highway bridge under severe temperature changes using extended Kalman filter trained neural network” Journal of Civil Structural Health Monitoring. 2016. Vol. 6, Iss. 3, pp. 545–560.

[8] Maksymova O. M. “Razvitie s primenenie neurosetevih tehnologiy dlia zadach mehaniki i stroitelnih konstrukcij”, Vestnik IrGTU. 2013. № 8 (79). pp. 81–88.

[9] Oliynyk A. O., Subbotin S. O., Oliynyk O. O. “Evolyutsiyni obchyslennya ta prohramuvannya”. Zaporizhzhya: ZNTU. 2010. p. 324.

[10] Kozyn I. V. “Evolyutsiyni modeli v dyskretnoyi optymizatsiyi”. Zaporizhzhya: ZNU. 2019. p. 204.

[11] Maksymova O. M. “Neural network forecasting technologies for problems of the dynamics of building structures” XIV All-Russian Scientific and Technical Conference "Neuroinformatics-2012"; Moscow 2012.

[12] Maksymova O. M. “Neuronet technology development and application for solving mechanical and engineering structures problems” Vestnik IrGTU. 2013. № 8 (79). p. 82.

[13] M. Raissi, P. Perdikaris, and G. Em Karniadakis, ‘‘Multistep neural networks for data-driven discovery of nonlinear dynamical systems,’’ 2018, arXiv:1801.01236. [Online]. Available: http://arxiv.org/abs/1801.01236

[14] Peter Z. Szewczyk Neural Net-work Based Extrapolation Strategies in Structural Analysis and Design, 1999. p.238 – 255.

[15] M. Raissi, P. Perdikaris, and G. Em Karniadakis, ‘‘Multistep neural networks for data-driven discovery of nonlinear dynamical systems,’’ 2018, arXiv:1801.01236. [Online]. Available: http://arxiv.org/abs/1801.01236

[16] F. Roewer-Despres, N. Khan, and I. Stavness, ‘‘Towards finite element simulation using deep learning,’’ in Proc. 15th Int. Symp. Comput. Methods Biomechan. Biomed. Eng., 2018, pp. 1–20.

[17] J. Ghaboussi, ‘‘Neural networks in computational mechanics,’’ Arch. Comput. Methods Eng., vol. 3, no. 4, p. 435, 1996.

[18] L. Liang, M. Liu, C. Martin, and W. Sun, ‘‘A deep learning approach to estimate stress distribution: A fast and accurate surrogate of finite-element analysis,’’ J. Roy. Soc. Interface, vol. 15, no. 138, Jan. 2018, Art. no. 20170844.

[19] L. Liang, M. Liu, C. Martin, and W. Sun, ‘‘A machine learning approach as a surrogate of finite element analysis-based inverse method to estimate the zero-pressure geometry of human thoracic aorta,’’ Int. J. Numer. Methods Biomed. Eng., vol. 34, no. 8, p. e3103, May 2018.

[20] Onur Avci P. O., Abdeljaber A. O. “Self-Organizing Maps for Structural Damage Detection: A Novel Unsupervised Vibration-Based Algorithm”, Journal of Performance of Constructed Facilities. 2016. Vol. 30, Iss. 3. pp. 1–11.

[21] Tao S. “Deep Neural Network Ensembles”. Available: https://arxiv.org/abs/1904.05488

[22] Hany Sallam, Carlo S. Regazzoni, Ihab Talkhan, and Amir Atiya “Evolving neural networks ensembles”. IAPR Workshop on Cognitive Information Processing, pp. 142-147

[23] Webb A. M., Reynolds C., Iliescu D.-A., Reeve H., Lujan M., Brown G. “Joint Training of Neural Network Ensembles”. Available: https://arxiv.org/abs/1902.04422.

Method for Automatic Detection and Recognition of Aerial Objects

Volodymyr Zhyrnov [1], Svitlana Solonska [2] and Oleksandr Dokhov [3]

[1] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61014, Ukraine*
[2] *Kharkiv National Automobile and Highway University, st. Yaroslava Mudrova, 25, Kharkiv, 61000, Ukraine*
[3] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61014, Ukraine*

Abstract

A method for automatic detection and recognition of aerial objects is proposed. This method uses a system of unary and binary predicate attributes, such as the presence of a signal in the analyzed cell, transition of a signal from this cell to cells adjacent in azimuth or in range. Such a system of predicate attributes makes it possible to describe the semantic component of a symbolic model of radar signals. Further, predicate attributes of point and extended aerial objects are given. Based on this method, a system and software for automatic detection and recognition of aerial objects with processing of real records on surveillance radars have been developed. As a result of semantic analysis of pulse burst fluctuations in a time domain, distinctive features of stationarity or nonstationarity are determined which make it possible to identify moving or static air objects. The semantic components of recognition, which are similar to decision-making algorithms by an operator, have been investigated. The proposed approach is based on the algebra of finite predicate.

Keywords

Radar, aerial object, recognition, detection, predicate attribute, symbolic model

1. Introduction

The results of application of the method, software and experimental studies of the system for automatic detection and recognition of aerial objects (AO) with processing of real records in surveillance radars are presented.

The relevance of research deals with the development of algorithm for automatic information processing that provide effective detection and recognition of aerial objects due to the accumulation of both signal (energy) and semantic information in the analyzed cell and in its surroundings under clutter.

The existing techniques of detection and recognition of aerial objects [1, 2] don't take into account the semantic aspect of the processes that allow one to develop the effective systems for automatic radar signal processing. Therefore, it is proposed to use a symbolic model of radar signals which includes the image of the studied object and the semantic component of the process.

As a rule, radar information systems [1, 2] analyze inter-period changes in the signal environment, but they are limited to creating a geometric image. If an aerial object can be described as point and moving, then from a pack of received reflected signals an image extended along the azimuth is formed. If the object is extended (weather clutter, birds and local air inhomogeneities) then from the received signals an image of this object is formed. There are also intelligent information systems based on process knowledge model which are associated with the possibility of parallel perception of information for further decision-making on the analysis of semantic attributes [3-5].

The proposed symbolic model of process knowledge makes it possible to take into account the semantic component of processing radar signals which includes both the creation of the object image and developing the semantic

EMAIL: nauka123@ukr.net (A.1); solonskaya@ukr.net (A.2); dokhov.alex@gmail.com(A 3)
ORCID: http://orcid.org/0000-0002-2397-3126 (A.1); http://orcid.org/0000-0002-8841-7825 (A.2); XXXX-XXXX-XXXX-XXXX (A.3)

component. Therefore, we can combine detection and recognition operations and use them at the same time.

The method is based on the determination of system of unary and binary predicate attributes: the presence of a signal in the analyzed cell; signal drift from this cell to neighbouring cells in range or in azimuth. Such a system of predicate attributes allows one to describe the semantic and image components of the symbolic model of radar signals. Using predicate attributes obtained, it is possible to determine whether it is a point or an extended aerial object. In addition, distinctive features of stationarity or nonstationarity of a signal burst are determined which make it possible to identify whether is the moving aerial object or static. In accordance with the symbolic model of aerial object and taking into account its image and semantic components, the type of object is identified.

2. Predicate attributes of the symbolic model of aerial objects taking into account its image and semantic components

The predicate attributes of the symbolic model of aerial objects describe the semantic and image components of radar data. This is a system of unary and binary predicates and predicate attributes of symbolic models for point and extended aerial objects.

In this study, we use the experimental data (Fig. 1) obtained by recording the reflected signals of surveillance radar in the centimetre range (pulse duration 1 μs, probing frequency 365 Hz). The figure shows amplitude distribution of the radar images in a 129 × 129 matrix of real radar signal records. The proposed model includes procedures to formalize and analyze the geometric signal image of observable aerial objects [9-11] and the operation of creating a predicate model of the semantic component for recognizing the observed location objects.

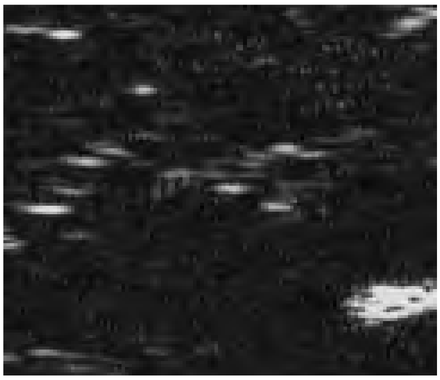


Fig. 1. Amplitude distribution of radar images in a 129 × 129 matrix of real radar signal records.

Let $M = \{q_{11}, q_{12}, ..., q_{ij}, ..., q_{mn}\}$ be a fixed set which is a rectangular matrix $\|A\|$ of dimension $M \times N$, consisting of elements $k = m \times n$ – the signal amplitudes in the processing elements of the radar field of view , and $B \subseteq M$ – some of its subsets whose signal amplitudes q_{ij} exceed the threshold values V_{ij}. A set of logical elements is made: if $q_{ij} \in B$, then $t_{ij} = 1$; if $q_{ij} \notin B$, then $t_{ij} = 0$, $i = \overline{1,m}$, $j = \overline{1,n}$.

Predicate $A(x)$ on a set M corresponding to a set of processing elements B that exceed a threshold value with the characteristic $(t_{11}, t_{12}, \cdots, t_{ij}, \cdots, t_{mn})$ is

$$A(x) = t_{11}x^{q_{11}} \vee ... \vee t_{mn}x^{q_{mn}} = \bigvee_{i=1,j=1}^{mn} t_{ij}x^{q_{ij}} \quad (1)$$

The expression $x^{q_{ij}}$ is a form of event recognition. When $x = q_{ij}$, $x^{q_{ij}} = 1$.

2.1. Formation of predicate attributes of the air object's symbolic model

The predicate attributes of the symbolic model of aerial objects in general form is a system of n unary and binary predicates Z_j

$$M = \{Z_j,\ j = 1...n\}. \quad (2)$$

Such a system of predicates allows one to describe the situation around the analyzed information cell and to formalize the process of making up a symbolic image of a signal $A(x)$ during several radar soundings. In the intelligent surveillance radar systems these can be:

- Z_{pij} – an unary predicate of presence or presence of a signal in the information cell a_{ij} (i, j - number of elements of radar coverage sector);
- Z_{dij} – a binary predicate of signal drift to the information cell adjacent in range;
- Z_{aij} – a binary predicate of signal drift to the information cell adjacent in azimuth.

$$Z_{pij} = 1 \text{ at } A_{ij} > 0 \tag{3}$$

$$Z_{dij} = 1 \text{ at } A_{i-1j} > 0 \wedge Z_{pij} = 1 \tag{4}$$

$$Z_{aij} = 1 \text{ at } Z_{pij} = 1 \wedge A_{ij-1} > 0, \tag{5}$$

Further, predicate attributes of the symbolic model of radar signals based on semantic and image components are obtained.

2.2. Formation of predicate attributes of the point air object’s symbolic model

Further, if there is a predicate attribute of adjacent in azimuth cell Z_{aij+l_1}, then in the next processing step the presence of a predicate attribute Z_{aij+l_2} in the information cell a_{ij+l_2} is checked.

$$Z_{aij+l_2} = (A_{ij} > 0 \wedge Z_{pij+l_2} = 1) = 1 \tag{6}$$

The values l_2 are found from equation (6) and then the following equation is obtained:

$$Z_{aij+l_2+1} = (A_{ij+l_2} > 0 \wedge Z_{pij+l_2+1} = 1) = 1 \tag{7}$$

At the n-step the predicate equation is:

$$Z_{aij+l_n} = (A_{ij+l_{n-1}} > 0 \wedge Z_{pij+l_n} = 1) = 1. \tag{8}$$

From the system of predicate equations (7) – (9) all the values $l_1 \cdots l_n$ and symbolic model for the point aerial object are defined:

$$Z_{mij} = \bigwedge_{l_1}^{l_n} Z_{ai,j+l_n} = Z_{ai,j+l_1} \wedge Z_{ai,j+l_1} \wedge \ldots$$
$$\ldots \wedge Z_{ai,j+l_{n-1}} \wedge Z_{ai,j+l_n} = 1. \tag{9}$$

The types of the point air object’s symbolic models are shown in Fig. 2.

2.3. Formation of predicate attributes of the extended air object's symbolic model

Further, if there is a predicate attribute of adjacent in range cell Z_{di+k_1j}, then in the next processing step the presence of a predicate attribute Z_{di+k_2j} in the information cell a_{i+k_2j} is checked.

$$Z_{di+k_2j} = (A_{ij} > 0 \wedge Z_{pi+k_2j} = 1) = 1 \tag{10}$$

Fig. 2. Types of the point air object’s symbolic models

The values k_2 are found from equation (10) and then the following equation is obtained:

$$Z_{di+k_2+1j} = (A_{i+k_2j} > 0 \wedge Z_{pi+k_2+1j} = 1) = 1 \tag{11}$$

At the n-step predicate equation is:

$$Z_{di+k_nj} = (A_{i+k_{n-1}j} > 0 \wedge Z_{pi+k_nj} = 1) = 1 \tag{12}$$

From the system of predicate equations (10) – (12) all the values $k_1 \cdots k_n$ are defined and a column of the symbolic model for extended air object is written as follows:

$$Z_{cdij} = \bigwedge_{k_1}^{k_n} Z_{di+k_n,j} = Z_{di+k_1,j} \wedge Z_{di+k_2,j} \wedge \cdots \wedge Z_{di+(k_{n-1}),j} \wedge Z_{di+k_n,j} = 1 \qquad (13)$$

If the signal arrives at the cell under study from the adjacent in range cell a_{i-1j} and adjacent in azimuth cell a_{ij-1} (binary predicates $Z_{d\,ij}$ and $Z_{a\,ij}$ are generated simultaneously), then predicate attribute Z_{bij} of the symbolic model of extended aerial object is formed. In this case, the presence of a column of the extended object's symbolic model Z_{cdij} in the information cell $a_{i+k_2,j+l_2}$ of the next radar sensing $j+l_2$ is checked

$$Z_{di+k_2,j+l_2} = (A_{ij+l_2} > 0 \wedge Z_{pi+k_2 j+l_2} = 1) = 1\,. \qquad (14)$$

From (14) we define the values k_2, l_2 and write the predicate equation:

$$Z_{di+k_2+1,j+l_2} = (A_{i+k_2,j+l_2} > 0 \wedge Z_{pi+k_2+1,j+l_2} = 1) = 1\,. \qquad (15)$$

From (13) – (15) we find $k_1 \cdots k_n$ and write down predicate for the next symbolic model column of the extended aerial object:

$$Z_{cdij+l_2} = \bigwedge_{k_1}^{k_n} Z_{di+k_n,j+l_2} = Z_{di+k_1,j+l_2} \wedge Z_{di+k_2,j+l_2} \wedge \cdots \wedge Z_{di+(k_{n-1}),j+l_2} \wedge Z_{di+k_n,j+l_2} = 1 \qquad (16)$$

Finally, from (13) - (16) we define all possible values of the predicate attribute Z_{cdij} for $j+l_1 \cdots j+l_n$ azimuthally directions and the predicate attribute Z_{bij} of the static extended object's symbolic model

$$Z_{bij} = \left(\bigwedge_{l_1,k_1}^{l_n,k_n} Z_{cdi+k,j+1} = 1\right) = \bigwedge_{l_1}^{l_n}\left(\bigwedge_{k_1}^{k_n} Z_{cdi+k,j+1} = 1\right) = 1\,. \qquad (17)$$

To evaluate an energy attribute of the symbolic model, the concept of accumulated energy of an extended air object as sum of signal amplitudes of information cells in the direction determined by vectors (k_n, l_n) according to the predicate equation (17) is introduced. Taking into account the distribution of amplitudes within (k_n, l_n) the energy attribute for extended air objects is determined:

$$I_{b2} = \sum_{k_1,l_1}^{k_n,l_n} q_{i+k_n,j+l_n} Z_{cdi+k,j+1}\,. \qquad (18)$$

The recognition of extended stationary objects is carried out in accordance with equations (10) – (12), (13) – (17) and by the energy attribute of the symbolic model (18).

2.4. Results of the study

To test the effectiveness of proposed technology, the experiments have carried out: 1) if the predicate attribute of the symbolic model for extended stationary air object is generated, then this object belongs to the angel-echo class; 2) if the predicate attribute of the symbolic model for the point aerial object is formed, then it belongs to the aircraft class. The probabilities of moving objects detection for the proposed method of processing in comparison with the classical ones for different the signal-to-noise ratios are shown in the Table 1.

Table 1
Probabilities of recognition of moving air objects at various signal-to-noise ratios.

Signal-to-noise ratio	Proposed method of processing	Classical method of processing	Difference
5	0,12	0,06	+0,06
10	0,34	0,27	+0,07
15	0,5	0,48	+0,02
20	0,66	0,658	+0,002
25	0,77	0,76	+0,01
30	0,82	0,82	0
35	0,86	0,86	0
40	0,94	0,89	+0,05

Experiments have shown that at small values of the signal-to-noise ratio, which do not exceed 10 dB, the probability of air objects recognizing is higher than when processing by traditional methods (Table 1) and at values is over 15 dB, the results approach the reference.

3. Conclusions

Thus, the method for automatic detection and recognition of aerial objects based on the determination of a system of predicate attributes that makes it possible to describe the semantic component of the symbolic model of radar signals has been developed. The procedures to form predicate attributes of the symbolic model for point and extended aerial objects and software for automatic detection and recognition of aerial objects with processing of real signal records of surveillance radars is produced. In accordance with the symbolic model of aerial object and taking into account its image and semantic components, the type of object is identified.

The advantages of the proposed method for automatic detection and recognition of aerial objects in comparison with the existing ones are:

➢ automatic detection and recognition of aerial objects taking into account a symbolic model of radar signals which includes the image of the studied object and the semantic component of the process;

➢ combining detection and recognition procedures and using them simultaneously that allow one to develop the effective system for automatic radar signal processing.

4. References

[1] Skolnik M. I. (eds) (2021) Radar Handbook, McGraw-Hill, New York

[2] Jian Li, Hummel R., Stoica P., Zelnio E.G. (eds) (2013) Radar Signal Processing and Its Applications. Springer.

[3] Russell S. (2019) Human compatible: Artificial intelligence and the problem of control, Penguin.

[4] Luger G. (eds) (2005). Artificial Intelligence: Structures and Strategies for Complex Problem-Solving. Williams.

[5] Bondarenko M. F., Rusakova, N. E., Shabanov-Kushnarenko, Ju. P. (2010) O mozgopodobnyh strukturah [About brainlike structures]. Bionika intellekta, Kharkiv, Ukraine, (2), pp.68-73. (In Russian).

[6] Solonska S., Zhyrnov V. (2021) Adaptive semantic analysis of radar data using fuzzy transform (Book Chapter). Springer, 2020, Lecture Notes on Data Engineering and Communications Technologies. Vol 48, pp. 157-179.

[7] Zhyrnov V., Solonskaya S. and Zima I. (2014). Application of wavelet transform for generation of radar virtual images. Telecommunications and Radio Engineering, 73(17), pp.1533-1539.

[8] Solonskaya S., Zhirnov V. (2018). Intelligent analysis of radar data based on fuzzy transforms. Telecommunications and Radio Engineering, 77(15), pp.1321-1329.

[9] Solonskaya S., Zhirnov V. (2018). Signal processing in the intelligence systems of detecting low-observable and low-doppler aerial targets. Telecommunications and Radio Engineering, 77(20), pp.1827-1835.

[10] Zhyrnov V., Solonska S. (2020). Process knowledge about observed objects in intellectual monitoring systems. Telecommunications and Radio Engineering, 79(18), pp.1599-1607.

[11] Zhyrnov V., Solonska S. (2020). Intelligent system for detection of low-visible air objects in surveillance radars. Telecommunications and Radio Engineering - 2020. Volume 79, Issue 17, pp. 1513-1519.

A set of logical rules for computer systems of decision support of the operation and recovery of weapons and military equipment during hostilities

Roman Litvinchuk 1, Andrii Levchenko 2

[1] *Military Academy (Odesa), Fontanskaya Road 10, Odesa, 65009*
[2] *Odesa Mechnikov National University, Dvorianska, 2, Odesa, 65026, Ukraine*

Abstract
The set of logical rules for software implementation of logic in the corresponding computer systems used for operation and recovery management is chosen in the work. The selection of logical rules is based on the criterion of increasing the speed of the algorithms for processing the source data. Using a set of selected logic rules allows you to implement a logic circuit using fast Warshell algorithms. To this end, the paper considers a number of properties of production-logical connections in the FS-system.
The proposed approach, in contrast to the existing ones, uses the advantages of algebraic systems in computer programs and provides the necessary quality and efficiency of vehicle management with a significant reduction in the requirements for computing resources, which is especially current for military information systems.

Keywords
Decision support, technical support, FS-system, logical rule, transitive circuit

1. Introduction

Formulation of the problem. With the development of new weapons and military equipment (WME) and, as a consequence, the forms and methods of combat operations, the requirements for the management of evacuation and recovery of WME have significantly changed, which contributed to improving and enhancing the effectiveness of rational distribution and use of forces and means to achieve victory in battle in a short time [1].

The greatest influence on the technical support system during the battle was exerted by the following factors:

timely and complete provision of troops with material means (primarily weapons, ammunition);

in the conditions of modern local conflicts and restoration of armaments and military equipment [2, 3].

These issues have become most relevant especially at this time during the anti-terrorist operation (ATO) and the Joint Forces Operation (JFO) in the eastern regions of Ukraine.

Consequently, the success of hostilities requires a quality management system for the evacuation and recovery of WME, which would allow for combat management of troops (forces) on a scale close to real time. Therefore, the state pays much attention to the development of the technical basis of the management system of the Armed Forces of Ukraine. An important area of its improvement is the creation and implementation of a decision support system (DSS) in the evacuation management and WME recovery system, which is based on theories of artificial intelligence.

Analysis of recent research and publications, which initiated the solution of the problem of creating an intellectual DSS allows us to identify significant activity in this direction [6, 10, 11]. And it also justifies the need for further research to improve the mathematical and software of information systems.

In this sense, against the background of the complexity of mathematical models of

EMAIL: Litvinsanr@gmail.com (A. 1);
katyaandreylev@gmail.com (A. 2).
ORCID: https://orcid.org/0000-0002-5681-691 (A. 1);
http://orcid.org/0000-0003-4423-8267 (A. 2)

multicriteria optimization of resource allocation management by tasks, it is advisable to justify a set of logical rules for DSS evacuation management and recovery of WME to ensure the required speed of source data processing algorithms.

Thus, the scientific task of developing new mathematical models of formal description of knowledge is relevant, namely, models of logical rules of inference in the automation of the DSS process of evacuation management and recovery of WME. For these models, along with the requirement to provide a rational solution, there is a requirement to reduce computational costs.

The purpose of the article. Substantiation of a new approach in the model of logical inference rule, which, unlike the existing ones, uses the advantages of algebraic systems in computer programs and provides the necessary quality and efficiency of WME evacuation and recovery management at logistics control points with significant reduction of computing resources.

The paper investigates the possibility in the process of constructing a logic circuit to identify the stage corresponding to the transitive circuit [12]. The selection of logical rules is based on the criterion of increasing the speed of the algorithms for processing the source data. The use of a set of selected logic rules allows you to implement a logic circuit using fast Warshell algorithms [13].

To this end, a number of properties of production-logical connections in the FS-system have been studied.

2. Main material

FS-system system is a system of algebra (algebraic structure, algebraic system), in which, together with operations on any pair of its elements $\cup$ («union») and $\cap$ («intersection») additionally set a binary relationship R with production-logical properties.

The binary relation with production-logical properties R is indicated $A \xrightarrow{R} B$, where A and B elements of the FS-system.

To substantiate the necessary set of logical rules for DSS management of WME evacuation and recovery, it is necessary to investigate the features of the implementation of logical relations to the corresponding problems of transitive relations, which is covered by a set of theorems.

Theorem 1. Let the R relation in the FS system $A_i \xrightarrow{R} B$, $i=1,...,n$. Then the logical connection is fair $\bigcup\limits_{i=1,...,n} A_i \xrightarrow{R} \bigcup\limits_{i=1,...,n} B$.

This theorem is proved in works of Makhortov S.D. [7]

Result 1. On the basis of Theorem 1, a generalized inference rule is introduced: an ordered pair $A, B \in F$ is logically connected by a relation $R(A \xrightarrow{R} B)$, if the following condition is satisfied:

There are those $A_i, B_i \in F$, $i = 1,...,n$ that $\bigcup\limits_{i=1,...,n} A_i = A,\ \bigcup\limits_{i=1,...,n} B_i = B$ where $A_i \xrightarrow{R} B_i$, $i = 1,...,n$.

Thus, if R - relation on the FS-system at the conclusion of any logical connection $A \xrightarrow{R} B$ it is enough to apply only the rule $u)$.

This means that during the construction of the algorithm for resource allocation in DSS management of WME evacuation and recovery, the number of branching processes is significantly reduced.

Theorem 2. Let be a binary relation R in an FS-system. Then, when deriving any logical connection, $A \xrightarrow{R} B$, all applications of the transitive rule can be excluded or transferred to the final stage of this process, as proved in the works of S.D. Makhortov. [7]

If it is true for $m \geq 0$, , then you need to prove this statement with equal recursion $m+1$. . Assuming induction, the first m steps of inference can be organized so that the transitive rule will be applied only at the end of the process. In this case, it all depends on which output rule is applied in the last $(m+1)$-m step. Let the connection $A \xrightarrow{R} B$ ultimately be derived from the condition $u)$. Each base ratio $A_i \xrightarrow{R} B_i$, $i = 1,...,n$ ratio has a recursion level $\leq m$ and, assuming induction, uses all its transitive connections only at the end of the output. In other words, for each $i = 1,...,n$ there is a chain of elements $A_i = C_i^0, C_i^1, ..., C_i^{n_i} = B_i$ such that the relations are satisfied, $C_i^{k-1} \xrightarrow{R} C_i^k$, $k = 1,...,n$, , in the derivation of which the transitive rule is not used. As a result, member n of value n_i are limited in the aggregate, , which means $n_i \le N$, $i = 1,...,n$. Due to this fact, all the above chains of elements are equal in length N and can be supplemented, if necessary, with a repeating element $C_i^{n_i} = B_i$ on the right. Then for chains $A_i = C_i^0, C_i^1,..., C_i^N = B_i$, $i = 1,...,n$, constructed elements $\tilde{C}_k = \bigcup_{i=1,...,n} C_i^k$, $k = 0,...,N$. It is likely that $\tilde{C}_0 = A$, $\tilde{C}_N = B$. In addition, since $C_i^{k-1} \xrightarrow{R} C_i^k$, $i = 1,...,n$, ,, the rule *u)* can be applied to these pairs, ie,. $\tilde{C}_{k-1} \xrightarrow{R} \tilde{C}_k$, $k = 1,...,N$.

Thus, all applications of the transitive rule in basic relations $A_i \xrightarrow{R} B_i$, $i = 1,...,n$ are replaced by its application at the final stage of derivation of construction of relations of elements $\tilde{C}_k$, $k = 0,...,N$.

A significant consequence of the proof of Theorem 2 is the understanding of the linearization of the resource allocation algorithm in DSS for evacuation management and WME recovery. That is, the algorithm can be implemented partially linearly with a reduction of comparison operations and reduction of branching processes.

Consider the model of logical reduction. It is known that in the software implementation of DSS it is advisable to avoid unjustified duplication of functional elements of code or data. When using logical systems, the question also arises about their minimization equivalent to algorithms. In mathematical logic, the minimal system of axioms is called the basis. Problems of existence of bases of rules for various logics were considered in [7]. The theory of artificial intelligence emphasizes the fact of additional possibilities of minimization in systems of production type.

Assuming that the set of logical rules for DSS of evacuation management and recovery of WME is an FS-system with a mathematical model of the algebra of the production system, the solution of the minimization problem gives a way to represent the binary relationship, when the computer memory stores only a unique part of information relations, and other information can be obtained from the general properties of logical relations. The unique part means the relationship selected according to certain criteria, $R_0 \subset R$, from which R follows as a logical circuit. The method of reduction is aimed at reducing from complex to simpler.

This allows to use as a source of knowledge for DSS of evacuation management and recovery of WME relational databases.

The logical reduction of the ratio in the FS-system is called the equivalent minimum ratio R_0 Minimality in this definition is understood in the sense of partially ordered large quantities. First, for this R the logical reduction R_0 may not be the only one. Secondly, after exclusion from R_0, at least one pair, a relationship is obtained that is not equivalent R.

Based on studies of other algebraic systems, it is expedient to draw a conclusion about the transfer of these results also to FS-systems [7]: namely, that the logical closure of this relation R is a transitive closure of some other relation $\tilde{R} \supseteq R$. This result is useful for the development of efficient algorithms for constructing logic closure and reduction. It makes it possible to reduce the study of the questions of finding the logical closure and reduction of order relations to the consideration of the corresponding properties of transitive relations.

For an arbitrary binary relation R in the FS-system it is considered the relation $\tilde{R}$, constructed on R consecutive performance of the following steps is considered:

1) add to R all pairs of the form (A,A), $A \in F_R$, and denote the new relationship R_1;

2) add to R_1 all pairs of the form (A,B), with elements of the form $A = \bigcup_1 A_i$, $B = \bigcup_1 B_i$, where all (A_i, B_i), $i = 1,...,n$, belong to R_1.

3) combine the obtained relation with the inclusion relation $\supset_R$.

From the previous theorems it follows that the relation $\underset{\approx}{R}$ is equivalent to R. This is justified as follows. Let R be a binary relation in the FS-system. Then if $A \xrightarrow{R} B$, then $(A,B) \in \tilde{R}$. Also let it take place $A \xrightarrow{R} B$. If the specified ratio occurred directly from the condition of production and logical properties, then immediately. $(A,B) \in \tilde{R}$.

It remains to consider the case of application of rule 3. All the necessary reflexive pairs can be prepared at the beginning of the withdrawal process. This type of pairs $A \supset_R B$ will not be required. Therefore, rule 3) can only complete this process.

If we compare the above 2 stages of inference with the sequence of constructing the relationship it turns out that the inference $A \xrightarrow{R} B$ corresponds to the construction of some subset $\tilde{R}$, which proves the inclusion $(A, B \in \tilde{R})$.

Theorem 3. For an arbitrary relation R in an FS-system, a logical closure $\xrightarrow{R}$ is a transitive closure $\tilde{R}^*$ of the corresponding relation $\tilde{R}$**.** A detailed rendering of this theorem is presented in the works of Makhortov S.D. [7].

Rationale. First, from the description of the construction process $\tilde{R}$ it is easy to see that if $R_1 \subseteq R_2$, then $\tilde{R}_1 \subseteq \tilde{R}_2$. A similar statement exists for the transitive closures of these relations, which means a proven fact. Next, for an arbitrary relationship, we consider the relationship R we consider the relationship $\underset{\approx}{R}$, built on this R by successive steps, , the inverse of the construction $\tilde{R}$, namely:

1) exclude from R the pair of forms that contain it $A \supset_R B$ and indicate a new relationship R_{-1}**;**

2) exclude from R_{-1} all pairs (A, B) with elements of the form $A = \bigcup\limits_1 A_i$, $B = \bigcup\limits_1 B_i$, where all (A_1, B_1) $(i = 1, \ldots, n)$ belong to R_{-1} to and do not coincide with (A, B);

3) exclude from the received relation all pairs of reflections.

Thus the ratio $\underset{\approx}{R}$ is equivalent to R.

Note that such an approach to constructing a transitive reduction ("remove all transitive pairs") would be erroneous. The reason is that in some situations (the presence of cycles) transitive reduction is not the only one, and the simultaneous removal of all existing transitive pairs can lead to loss of connections. However, since the FS system itself is acyclic, the removal of all the above "combined" pairs (A, B) leads to the same result, regardless of the removal order. For this reason, we can talk about the simultaneous removal of all such pairs.

The lemma arising from the implementation of Theorem 3 is essential for the case of substantiation of the set of logical rules of computer DSS control of operation and recovery of WME during combat operations.

Lemma 1. Let R the binary relation in the FS-system. In order for it R to be a logical reduction, it is necessary and sufficient that R does not contain any such pair (A, B), that is satisfied by the relation $A \xrightarrow{R \setminus \{(A,B)\}} B$.

This lemma is proved in works of Makhortov S.D. [7].

To prove the opposite, it is assumed that there is no pair $(A, B) \in R$, for which it is true $A \xrightarrow{R \setminus \{(A,B)\}} B$. It is necessary to prove that in this case R is a logical reduction. The inverse is also allowed - that there is a relation $R_0 \subset R$, equivalent R to what $(A, B) \in R / R_0$. Then because $(A, B) \in R$, by virtue of the equivalence of these relations is true $A \xrightarrow{R_0} B$. Since the relationship R_0 does not contain a pair (A, B), then $R_0 \subseteq R \setminus \{(A, B)\}$, and the logical connection $A \xrightarrow{R_0} B$ contradicts the assumption - there are no such pairs (A, B) in R. The resulting contradiction proves the necessary statement

The fact of proof of the lemma allows us to state that the connections implemented in relational databases do not violate the logic of resource allocation in DSS of operation management and recovery of WME in the case when the data of the database management system is used as source data for decision making.

The following theorem indicates a sufficient condition for the existence and method of constructing a logical reduction of this relation.

Theorem 4. Suppose that for a binary relation R, given in an FS system, a corresponding relation $\underset{\sim}{R}$ is constructed. Then, if for $\underset{\sim}{R}$ there is a transitive reduction R^0**,** then accordingly its relation $\underset{\approx}{R^0}$ is a logical reduction of the initial relation R**.**

This theorem is proved in works of Makhortov S.D. [7].

Let (A,B) – be an arbitrary pair of relations R^0. . It must be shown that a logical connection $A \xrightarrow{R\backslash\{(A,B)\}} B$ is impossible. It is assumed that this connection exists. Then the ratio $R\backslash\{(A,B)\}$ is equivalent $\underset{\approx}{R}^0$. Therefore, the application of condition $(A,B)\in R$ for the output $A \xrightarrow{\underset{\approx}{R}^0\backslash\{(A,B)\}} B$ is not possible, because the pair (A,B) is not contained in the set $\underset{\approx}{R}^0\backslash\{(A,B)\}$.

Any logical connection can be built in such a way that all applications of the transitive rule will be made only in the final stage of its derivation. This fact means that there is a chain of elements $A = C_0, C_1, ...C_N = B$ such that the relations $C_{k-1} \xrightarrow{R^0\backslash\{(A,B)\}} C_k$, $(k=1,...,N)$, are satisfied,, in deriving each of which the rule - there is an element $C \in F$ that $A \xrightarrow{R} C$ **i** $A \xrightarrow{R} B$ does not apply**.** Hence by Lemma 1 – $(C_{k-1}, C_k) \in \tilde{R}$**.** Thus when $N>1$ the pair (A,B) is transitive to $\tilde{R}$**.** The result is a contradiction to the initial assumption $(A,B)\in \underset{\approx}{R}^0$**.**

It remains to investigate the case $N=1$. In this situation, there is a logical inference $A \xrightarrow{\underset{\approx}{R}^0\backslash\{(A,B)\}} B$, which may include the application of rule 1) only in the final stage. The obtained pair (A,B) is described by step 2) the process of constructing the relationship $\tilde{R}$. Accordingly, when found $\underset{\approx}{R}^0$ (reverse process), it will be excluded.. Thus, the contradiction to the initial assumption $(A,B)\in \underset{\approx}{R}^0$. is confirmed.

And research potentially possible situations for the supposed logical conclusion $A \xrightarrow{\underset{\approx}{R}^0\backslash\{(A,B)\}} B$. As a result, it is established that in each case the existence of a connection $A \xrightarrow{\underset{\approx}{R}^0\backslash\{(A,B)\}} B$ contradicts the fact $(A,B)\in \underset{\approx}{R}^0$. Thus, the relation $\underset{\approx}{R}^0$ is a logical reduction.

Thus, all of the above made it possible to substantiate the study of the adequacy of the proposed model of the logical rule.

3. Conclusions

To confirm the proposed model, an experiment was conducted on the basis of automated tests with a knowledge base of DSS control and management of the communication system deployed at the rear control points of the operational unit (module for ensuring the functional stability of the communication management system).

Validation tasks were generated randomly with the ability to control the depth of output, the number of generated rules and objects. The described experiments allow to determine the number of requests to external sources of information depending on the amount of initial facts and rules. The obtained results are compared by the number of external requests with the procedures of normal inverse inference and inference based on the proposed model of the logical rule. The results of the experiment were processed in the package Statistica 13.3. Quantitative assessment of efficiency gains in relation to the average number of external requests indicates their reduction when justifying the option of dynamic routing and construction of the core of the reference network. On average, due to the application of the model of the logical rule, it is possible to state an increase in efficiency up to 12%.

The article for the first time shows a set of logical rules of computer DSS for the management of operation and recovery of WME during combat operations, as a logic circuit with the general properties of the transitive circuit. The use of the model will reduce the study of some important issues related to logical relations to the corresponding problems of transitive relations. In particular, the construction of a logic circuit or reduction can be carried out using the fast Worshell algorithm.

This approach, unlike the existing ones, uses the advantages of algebraic systems in computer programs and provides the necessary quality and efficiency of vehicle management with a significant reduction in computing resources, which is especially essential for military information systems.

3. References

[1] Development of tactics of the Land Forces in local wars and armed conflicts of the second half of the XX and the beginning of the XXI

centuries.. Kiev - : National University of Defense of Ukraine, 2007. - 328 p.
[2] Strategic Defense Bulletin of Ukraine for the period up to 2025.
[3] Military-political, strategic, operational and tactical content of local wars and armed conflicts. National University of Defense of Ukraine, 2001. 222 p.
[4] Mykus S.A. Functional stability of information and telecommunication system. Yu. V. Kravchenko, O.A. Leschenko. East European Scientific Journal. 2016. №.2(6), P.47-52.
[5] Mykus S.A.. Management of hardware and software resources in a computer system based on methods and models of artificial intelligence. Y.V Kravchenko, N.V. Rudenlo. Telecommunication and information technologies. 2014. № 4, P. 7-15.
[6] Savchenko V.A. Model of a multilevel real-time decision support system based on intelligent integration. 2011. №1(13), P. 106-112.
[7] Mahortov S.D. Mathematical Foundations of Artificial Intelligence: LP-Structures theory for building and researching production-type knowledge models. Edited by. A.V. Vasenina. Moscow. 2009. 304 p
[8] Burstein Frada, Holsapple W. Clyde (editors), (2008) Handbook on Decision Support Systems 1, Berlin: Springer-Verlag., p. 21.
[9] Turban Efraim, Aronson E. Jay, Liang Ting-Peng (2005) Decision Support Systems and Intelligent Systems, 7th edn., New Jersey: Pearson/ Prentice Hall, p. 292.
[10] Davis K. Paul, Kulick Jonathan, and Egner Michael, Implications of Modern Decision Science for Military Decision-Support Systems, Santa Monica: RAND Corporation, 2005.
[11] Laskey Blackmond Kathryn, D'Ambrosio Bruce, Levitt Tod S., and Mahoney Suzanne, “Limited Rationality in Action: Decision Support for Military Situation Assessment”, in Minds and Machines, 10, 1, February 2000.
[12] Studies in Logic and the Foundations of Mathematics: Theory of Relations, Edited by Roland Fraié, Volume 145, Pages 1-451 (2000) doi:10.1016/S0049-237X(00)80048-8.
[13] BrigitteJaumard, MichelMinoux, An efficient algorithm for the transitive closure and a linear worst-case complexity result for a class of sparse graphs. Information Processing Letters. Volume 22, Issue 4, 17 April 1986, Pages 163-169. doi:10.1016/0020-0190(86)90021-9.

Modeling regional level control system using fuzzy cognitive maps

Viktoriia Dyomina [1], Tetiana Bilova [2], and Iryna Pobizhenko [3]

[1] *V. V. Dokuchaev Kharkiv National Agrarian University, post office «Dokuchaevske - 2», Kharkiv region, Kharkiv district, 62483, Ukraine*
[2] *Kharkiv National University of Radio Electronics, Pr. Nauki, 14 Kharkiv, 61166, Ukraine*
[3] *Kharkiv National University of Radio Electronics, Pr. Nauki, 14 Kharkiv, 61166, Ukraine*

Abstract
The problem of constructing a numerical model for managing a region in situations of various kinds of uncertainties, when these uncertainties cannot be clearly formalized using the methods of probability theory and mathematical statistics, is considered. The use of fuzzy logic methods for describing data based on a variety of parameters characterizing a variety of systems and controls, i.e. based on the description within regional differentiations. Fuzzy cognitive support and decision-making systems for problem areas of regional governance are considered.

Keywords
fuzzy logic, describe data, numerical model, regional management.

1. Introduction

The current stage of socio-economic development of Ukraine is characterized by the presence of spatial deformations at the regional and intraregional levels. The problem of asymmetry of regional development trends is solved with the help of state regional policy. However, the question of intra-regional socio-economic imbalances remains insufficiently studied. At the same time, at the intra-regional level, differentiation is of great importance for a number of indicators of the economic and social development of administrative-territorial entities. This situation determines / creates the existence of fragmentation and fragmentation of the socio-economic space of the regions, the expansion of the areas of depression. This is due to the concentration of financial resources in regional development centers (particular in large cities), with the widening gap between urban and rural areas, with the deepening of the problems of territorial imbalance at the intra-regional level [1].

These problems are exacerbated during the implementation of the reform of local self-government and administrative and financial decentralization. Their basic principle is the development of territorial communities. New economic conditions compelled researchers and practitioners in our country to seek adequate forms, methods and tools of strategic management in the region. In particular, an attempt is made to apply the existing principles and methods of strategic management, implemented at the enterprise level in regional systems [2]. Of course, there is a direct analogy between the regional socio-economic system and the economic system (corporation). But the study of foreign theory and methods of strategic management and planning is necessary to develop a domestic concept of strategic management. This concept is necessary for adaptive methodological work at the level of regional studies, especially in conditions of describing reality, which allow the processing of large amounts of data in the conditions of growing uncertainty.

2. Literature review

New economic conditions forced researchers and practitioners to look for adequate forms, methods and tools of strategic management in the region.

In work [2], an attempt was made to apply in regional systems the existing principles and methods of strategic management, implemented

EMAIL: vvdemina17@gmail.com (A. 1); tetiana.bilova@nure.ua (A. 2); irina_pob@ukr.net (A. 3)
ORCID: 0000-0001-6467-5021 (A. 1); 0000-0002-1085-7361 (A. 2); 0000-0002-0723-1878 (A. 3)

at the enterprise level. To develop the concept of adaptive methodological work at the level of regional studies, especially in terms of describing reality, which allow processing large amounts of data in conditions of accrescent uncertainty, multivalued logic is used and the nature of uncertainty is considered [3, 4].

An extension of this approach can be the introduction of the concept of a typology of boundaries in terms of the categories that describe them [5]. If the boundaries are well distinguishable, then they can be described by fuzzy membership functions that underlie the theory of fuzzy sets of type 1. If the boundaries are poorly distinguishable, for example, expressed by some zones, then they are expressed by fuzzy sets of type 2 and higher. In fuzzy sets of type 2, membership values are specified by membership functions, i.e., the inaccuracy of the membership definition is taken into account. It can be linguistic or linguological uncertainty [6]. For example, it is applicable in the case of decision-making processes or the description of natural phenomena [7], social upheavals [8].

For fuzzy sets, the number of results is infinite, but limited to the range from zero to one. Each numerical coefficient in fuzzy logic can be associated with a linguistic variable.

Fuzzy sets create the basis for the development of a more flexible approach to the analysis of reasoning and the modeling of complex systems. Methods and models of fuzzy mathematics have found their application in pattern recognition, image analysis, expert systems, decision support systems and in many other areas.

Among the works in the field of fuzzy logic, one should note the publications of D. Dubois and H. Prade on the theory of fuzzy measure and measure of possibility; J. Bezdek on fuzzy clustering and pattern recognition; M. Sugeno on fuzzy inference and fuzzy integral; R. Yager on fuzzy logic. But, despite the large number of theoretical works, the applied value of fuzzy models has been questioned for a long time [9]. It was only in 1975 that the English engineer Ebrahim Mamdani first applied fuzzy sets in the actuated control system.

There is Bart Kosko's Fuzzy Approximation Theorem (FAT). According to it, any mathematical system can be approximated by a system based on fuzzy logic [17].

The complex apparatus of differential and integral calculus is used in management and identification. Instead, it is possible to fairly accurately reflect an arbitrary "input-output" relationship employing the theory of fuzzy sets [15]. The use of fuzzy linguistic models based on fuzzy logic of sets in control problems and decision planning is developed in works [10-13].

In works [18-25] Fuzzy Cognitive Maps (FCM) are used as a modeling tool. FCM are presented in the form of a graphical structure that reflects the causal relationships of influencing factors in the decision-making system. FCM have the ability to model complex chains of cause and effect relationships through weighted cause and effect relationships [5].

Therefore, in [18], a cognitive model was formed, based on the identification of the "action force" of links between the elements that determine the level of influence of factors of the micro- and macroenvironment of the multi-level transfer of innovations. In [19], FCM is used to form the competitive potential of agricultural territories in Kazakhstan. The study of connections in a complex system of agribusiness makes it possible to form a set of strategic goals and choose the best option for management impact, ensuring the formation of the agricultural export potential of the region.

In [20], FCMs with group consensus based on linguistic assessments are used to study the IoT industry in Taiwan. A learning algorithm is proposed for adjusting the weights of the interaction between factors at all stages. The proposed model requires revision in order to adapt it to other subject areas. In [21], a method for supporting the components of modeling processes in conditions of uncertainty is considered on the example of building FCM of the Nigerian rice agri-food system. FCM applies to scenario development and defining points of influence on the system.

In works [22-25], FCM is used to model procedures for managing innovation processes for scientific and industrial clusters [22], to support strategic planning [23], adaptation to climate change in arid and semi-arid regions [24], ecology evaluation and management [25].

Thus, the effectiveness of using existing methods for constructing models of complex fuzzy-target systems is due to the following capabilities: visually represented analyzed system, modeling in the absence of statistical information about the states of the system, as well as evaluating the effectiveness of management decisions.

Expansion of FCM with elements of linguistic description of various processes allows obtaining only fuzzy information about the transitions of the system from one state to another. And the use of fuzzy logic will make it possible to describe non-trivial complex systems evolving in time with elements of random behavior. The latter is especially important when constructing a numerical model for managing a region.

3. Development of a mathematical model

To build a model for managing the region, we will analyze the current state of the vector of criteria (concepts). The problem in the control loop is associated with the impossibility of a clear formalization of all concepts, as well as with the fact that the "user" partly hides the true goals and objectives solved by the control system. Thus, it is necessary to achieve the fuzzy goals set in this way, i.e. formalize the tasks of managing fuzzy target systems.

In the linguistic description of nontrivial processes, there is only fuzzy information about the transitions of the system from one state to another. Fuzzy logic allows one to describe non-trivial complex systems that develop over time, with elements of random behavior, which is especially important when building a numerical model for managing a region.

The problem can be solved using the methods of fuzzy logic. The concepts of the system under study are represented in the form of terms of fuzzy sets. However, the use of fuzzy logic procedures does not take into account individual features of regional innovation processes. First, there is a need to determine systemic indicators (consonance, dissonance, mutual influence, etc.), which are traditionally calculated. Secondly, the connections between concepts change over time, and not only the value, but also the nature of the influence can change.

To take into account the first feature, when modeling the relationship between the results of the implementation of strategies for the socio-economic development of regions and innovative strategies at the regional level, it is possible to take into account the uncertainty of system characteristics. These system characteristics differ in the following:

- relations of influence between concepts are represented by fuzzy sets, while the problem of taking into account the negative influence of concepts is solved by expanding the base set on the area of negative values for linguistic variables;
- system characteristics of a fuzzy cognitive map are determined based on a transitively closed matrix of mutual influences, the elements of which represent fuzzy sets defined on the areas of positive and negative values of the base sets;
- as a result of the analysis, the uncertainty of the used system characteristics is taken into account.

The main stages of building a mathematical model using linguistic methods that use a fuzzy description of the management process at the regional level are considered further.

Analyzing the set of parameters that characterize the region as a complex system and methods of its management. The types within regional differences can be classified according to the following criteria: the result of the impact on the socio-economic development of the territories of the region, the factors of occurrence, the nature of the impact, the level of coverage, the duration of these trends.

For this, it is best to use FCM, which qualitatively characterize the structures of complex systems, reflect the uncertainty and dynamics of system elements, and also reflect / recreate the relationships between them.

A complex spatial socio-economic system with fuzzy logic is represented by FCM in the form of a parametric functional oriented graph

$$R_n = \langle G, X, E, \Theta \rangle,$$

where G – cognitive structure;

$X|V \to \Theta$, $X = \{X_1, X_2. \dots, X_m\}$ – set of vertex parameters V, $V^i = \{V_1^i, V_1^i. \dots, V_n^i\} \in V$, $V^i \in X_i, i = \overline{1, m}$, which form the nodes of the graph are fuzzy sets and correspond to a set of concepts;

$E: (V_j^i, V_k^l) \to w_{jk}^i$ – function associating value (V_j^i, V_k^l), where w_{jk}^i – the weight of the directed edge from V_j^i to V_k^l, if $j \neq k$, and $w_{ij} = 0$, if $j = k$ u $i = l$. Stated differently, $E = (V \cdot V)$ – connection matrix. The values of the weights on the main diagonal of the matrix are equal to zero, since changes in knowledge about the concept cannot affect the concept itself;

$E = (V \cdot V) = (V_j^i) \longrightarrow G_j, G_j$ – a function that every concept V_j^i assigns the sequence of its activation degrees so that for each $t \epsilon V, G_j(t) \epsilon L$ – this is the degree of the concept activation V_j^i at the moment *t*. $G(0) \epsilon L^n$ is the original vector containing the initial values of all concepts. $G(t) \epsilon L^n$ – the final state vector of concepts at a certain iteration L;

$\Theta \to L$ – transformation function, which establishes a connection between $G(t+1)$ and $G(t)$ for all $t \geq 0$ upon condition $\forall i, j \in \boldsymbol{N}$,

$$G_j^i(t+1) = f\left(\sum_{\substack{j=1 \\ j \neq k}}^{n} w_{jk}^i \cdot G_j^i(t)\right).$$

The transform function is used to bring the weighted sum of the concepts states into a certain range, which is set at the level [0; 1].

As a transformation function, the logistic function of the form is used

$$G_j(x) = \frac{1}{1 + \mathrm{e}^{-v \cdot x}}.$$

Modeling with FCM is a combination of fuzzy logic and cognitive modeling. Figure 1 shows a fragment of the FCM.

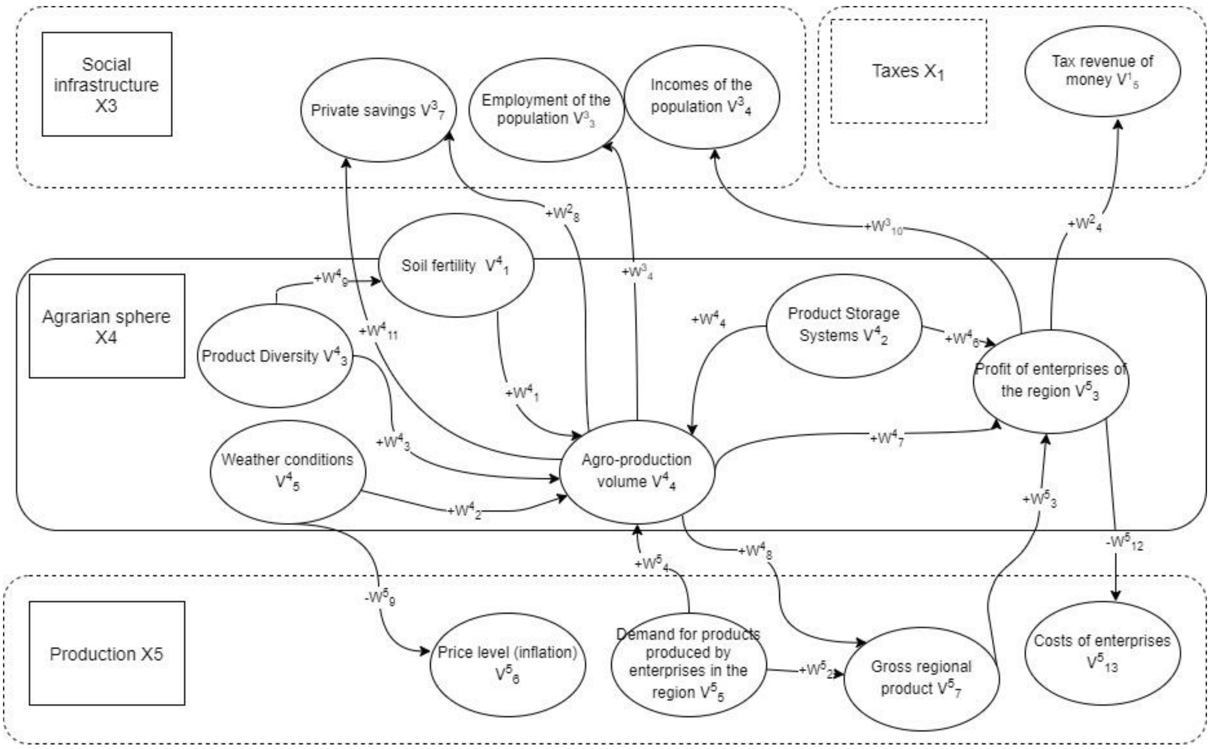


Figure 1: Fragment of the FCM, which reflects the socio-economic situation in the region at a fixed t using the example of agriculture

The FCM causal relationship can be represented as a connection matrix. On its basis, it is also possible to calculate the measure of the centrality of the concept and obtain information on direct and indirect causal relationships in the map.

Let us dwell on the problem of decision-making based on fuzzy sets, so a fuzzy system of regional management can be described using a mapping [3]

$$\Phi: F(X^m) \to F(X^m)$$

where $F(X^m)$ – set of fuzzy matches. Fuzzy mapping Φ characterizes transitions from state to state and can be set using a linguological model.

Let it be given $\Xi = \{\xi_1, \ldots, \xi_m\}$, $\Psi = \{\psi_1, \ldots, \psi_l\}$, ξ_j и ψ_k, $j \epsilon J = \{\overline{1,m}\}$, $k \epsilon K = \{\overline{1,l}\}$, correspond to the set of values of linguistic variables Λ_j, Z_k with corresponding basic sets (concepts) X_i, V_j^i. The dynamics of the system can be described by linguistic display

$$\Phi: \Lambda^m \times V^l \to \Lambda^m \qquad (1)$$

where

$$\Lambda^m \equiv \Lambda_1 \times \ldots \times \Lambda_m = {}_{j=1}^{m}\Lambda_j,$$

$$\Lambda^l \equiv \Lambda_1 \times \ldots \times \Lambda_l = {}_{k=1}^{l}\Lambda_k.$$

Linguistic mapping (1) can be specified using a fuzzy reasoning scheme of the type:

«if $\xi_1 = a_{11}$ and $\xi_2 = a_{12}$ and … and $\xi_m = a_{1m}$» and $\psi_1 = c_{11}$ and $\psi_2 = c_{12}$ and … and $\psi_j = c_{1j}$ then $\xi_1 = b_{11}$ and $\xi_2 = b_{12}$ and … and $\xi_m = b_{1m}$ else…. (2)

«if $\xi_1 = a_{n1}$ and $\xi_2 = a_{n2}$ and … and $\xi_m = a_{nm}$» and $\psi_1 = c_{n1}$ and $\psi_2 = c_{n2}$ and … and $\psi_j = c_{nj}$ then $\xi_1 = b_{n1}$ and $\xi_2 = b_{n2}$ and … and $\xi_m = b_{nm}$»,

where a_{ij}, b_{ij}, c_{ij} – corresponding values of linguistic variables $a_{ij}, b_{ij} \in \Lambda_j, c_{ik} \in V_k, i \in I = \{\overline{1,n}\}$. The values of linguistic variables correspond to fuzzy subsets defined on the corresponding base sets X_j, Z_k :

$$a_{ij} \Leftrightarrow a_{ij},$$

$$b_{ij} \Leftrightarrow b_{ij},\ a_{ij}, b_{ij} \in F(X_j);\ c_{ij} \Leftrightarrow c_{ij}\ \in F(Z_k).$$

So, according to the results of the impact on the socio-economic development of territories [1] the variable ψ_1 can take the value c_{11} = «positive (a decrease in disagreement is accompanied by an improvement in indicators in underdeveloped areas of the region)»; c_{21} = «negative (deterioration of indicators of socio-economic development of more developed areas, deterioration of interconnections between territories)»; c_{31}= «neutral (reduction or growth of differences does not significantly affect the level of socio-economic development of region)».

By the nature of the influence ψ_2 the variable can take the value c_{12} = «acceptable», c_{22} = «excessive».

A variable ψ_3 (object of study) can take the values c_{13} = «Inter-district»; c_{23} = «Urban area»; c_{33}= «Rural territory»; c_{43} = «Administrative center»; c_{53} = «Periphery»; c_{63}= «Industrial areas»; c_{73} = «Agricultural areas»; c_{83} = «Recreational areas»; c_{93}= «Centers of economic growth»; $c_{10,3}$ = «Depressed areas».

A variable ψ_4 (subject of research) can take on values c_{14} = «social»; c_{24} = «economic»; c_{34} = «environmental»; c_{44} = «informational» etc.

A variable ψ_5 (coverage level) can take the values c_{15} = «partial (local)»; c_{25} = «general».

A variable ψ_6 (causes) may take values c_{16} = = «objective (caused by objective differences in the availability of human, natural, economic, financial and other resources)»; c_{26} = «subjective (generated by the level of rationality of the use of natural resources, the effectiveness of management processes)».

A variable ψ_7 (duration) may take values c_{17} = «short term», c_{27} = «continuous».

To get the result in the form of a conclusion as an element of the linguistic variable "stable socio-economic growth", "threat to socio-economic development", etc. linguistic mapping (1) correspond fuzzy mapping

$$\Phi: F(X^m) \times \boldsymbol{F}(Z^l) \rightarrow \boldsymbol{F}(X^m) \qquad (3)$$

The fuzzy representation Φ is determined by the formula

$$\Phi: \bigcup_{i=1}^{n} ((\mathop{\times}_{j=1}^{m} a_{ij}) \times (\mathop{\times}_{k=1}^{l} c_{ik}) \Rightarrow (\mathop{\times}_{j=1}^{m} b_{ij})) \qquad (4)$$

where $\Rightarrow$ – denotes an implication operation, $\times$ – denotes the Cartesian product operation.

If $A, B \in F(X)$ – fuzzy subsets having membership functions $\mu_A(x)$, $\mu_B(x)$, $x \in X$, then the intersection operation takes the form [15]

$$A \cap B \Leftrightarrow \mu_{A \cap B} = \mu_A \mathrm{T} \mu_B.$$

Fuzzy mapping (3) is defined as

$$\Phi \Leftrightarrow \mu_\Phi(x, z, x') = \bigvee_{i \in I} \left(\left(\left(\Delta_{j \in J} \mu_{a_{ij}}(x_j) \mathrm{T} \mu_{b_{ij}}(x'_j) \right) \right) \mathrm{T} \left(\mu_{c_{ik}}(z_k) \right) \right) \qquad (5)$$

where $\Delta_{k \in K} \mu_{c_{ik}}(z_k) = \mu_{c_{i1}}(z_1)\mathrm{T}, \ldots, \mathrm{T}\mu_{c_{il}}(z_l)$.

Therefore, the linguistic mapping (3) - (5) is constructed in terms of the membership function, which can be specified using a fuzzy reasoning scheme.

4. Conclusions

The proposed approach makes it possible to identify problem areas of socio-economic activity of regions when describing data and create prerequisites for the formation of a regional development strategy. At the same time, interregional imbalances, the pace and trends of development of the regional economic system, as well as the need for regional initiatives should be taken into account.

The developed model of regional management is the basis for the formation of a regional development strategy to eliminate imbalances and reach a new economic and social level. The next stage of the study will be the concluding quantitative assessment of the qualitative characteristics of the level of development of social policy, which is an integral indicator of the effectiveness of social policy in the region using expert methods and fuzzy logic methodology. To obtain groups of development directions that are homogeneous in terms of qualitative and quantitative indicators, it is necessary to carry out clustering (determine the place of each identified development sector at time t).

5. References

[1] Ya. Tibinka, Konceptualni zasady regulyuvannya vnutrishno-regionalnoyi socialno-ekonomichnoyi dyferenciaciyi v umovax akty`vizaciyi miscevogo samovryaduvannya [Conceptual bases of regulation of intra-regional socio-economic differentiation in conditions of activation of local self-government]. Economic Forum, vol. 1, 2016, pp. 77-84.

[2] V. V. Oglikh, T. I. Efanova, Modelirovanie protsessov razvitiya regiona v kontekste strategii planirovaniya [Modeling of regional development processes in the context of planning strategy]. Mathematical methods, models and information technologies in economics, vol. 12, № 174, 2015, pp. 430-436.

[3] V. B. Silov, Prinyatie strategicheskih resheniy v nechetkoy obstanovke [Making strategic decisions in a fuzzy environment], INPRO-RES, Moskva, 1995.

[4] F. Wenstop, Deductive verbal models of organizations, Int. J. Man-Machine Studies, vol. 8, 1976, pp. 293-311.

[5] Many-Valued Logic, 2015. URL: https://plato.stanford.edu/entries/logic-manyvalued

[6] M. Black, Vagueness: An exercise in logical analysis, Philosophy of Science, vol. 4, 1937, pp. 427–455.

[7] L. A. Zadeh, Toward a theory of fuzzy information granulation and its centrality

in human reasoning and fuzzy logic. Fuzzy Sets and Systems, vol. 90, 1997, pp. 111-127.
[8] O. A. Osoba, B. Kosko, Fuzzy Cognitive Maps of Public Support for Insurgency and Terrorism, Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, Vol. 14, № 1, 2017, pp. 17–32.
[9] K yubileyu Lotfi Zade (podborka statei) [For the anniversary of Lotfi Zadeh (selection of articles)]. Novosti iskusstvennogo intellekta, vol. 2-3, 2001, pp. 3-27.
[10] M. Mizumoto, Fuzzy reasoning under new compositional rules of inference, Kybernetes.. vol. 14, № 2, 1985, pp. 107-117.
[11] B. W. Pitsworth, Axiomatic approach to implication for approximate reasoning with fuzzy logic, Fuzzy Sets and Systems, vol. 3, № 2, 1980, pp. 193-219.
[12] W. J. M. Kickert, E. H. Mamdani, Analysis of a fuzzy logic controller, Fuzzy Sets and Systems, vol. 1, № 1, 1978, pp. 29-44.
[13] R.M. Tong, Some properties of fuzzy feedback systems, IEEE Trans. Syst, Man, and Cybern, vol. 10, № 61, 1980, pp. 327-330.
[14] S. D. Shtovba, Vvedenie v teoriyu nechetkikh mnozhestv i nechetkuyu logiku [Introduction to the theory of fuzzy sets and fuzzy logic], 2019. URL: http://matlab.exponenta.ru/fuzzylogic/book1/0.php.
[15] M. E. Lebedeva, Nechetkaya logika v ekonomike – formirovanie novogo napravleniya [Fuzzy logic in economics – the formation of a new direction], Ideas and Ideals, vol. 11, № 1, Part 1, 2019, pp. 198-212.
[16] L. Zadeh, Fuzzy Sets, Information and Control, Vol. 8(3), June 1965, pp. 338-353.
[17] B. Kosko, Fuzzy thinking, Hyperion, 1993.
[18] O. Shkarupa, V. Boronos, D. Vlasenko, K. Fedchenko, Problems and Perspectives in Management, vol. 19, № 1, 19 February 2021, pp. 151-162. doi: 10.21511/ppm.19(1).2021.13
[19] M. Y. Anokhina, R. Abdrakhmanov, Y. E. Gridneva, M. Arrieta-López, N. R. Dzhalilova, A. Meza-Godoy, Formation of the competitive potential of the agricultural territories, Entrepreneurship and Sustainability, vol. 7, № 3, 2020, pp. 1921-1936. doi:https://doi.org/10.9770/ jesi.2020.7.3(32).
[20] Chen Chen-Tung, Chiu Yen-Ting, A study of dynamic fuzzy cognitive map model with group consensus based on linguistic variables, Technological Forecasting and Social Change, vol. 171, October 2021. URL: https://www.sciencedirect.com/science/article/pii/S0040162521003802. doi: 10.1016/j.techfore.2021.120948.
[21] G. I. Edwards, K. Kok, Building a Fuzzy Cognitive Map from stakeholder knowledge: An Episodic, asynchronous approach, Current Research in Environmental Sustainability, vol. 3, 2021. URL: https://www.sciencedirect.com/science/article/pii/S2666049021000293.
[22] A. E. Zaenchkovski, E. A. Kirillova, M. V. Golovinskaya, V. L. Sazonova, E.A. Borisova, Cognitive fuzzy-logic modeling tools to develop innovative process management procedures for scientific-industrial clusters Frontier Information, Technology and Systems Research in Cooperative Economics Studies in Systems, Decision and Control, vol. 316, 2021, pp. 209-221. doi:10.1007/978-3-030-57831-2_22.
[23] P. Hajek, O. Prochazka, P. Pachura, Fuzzy cognitive maps based on text analysis for supporting strategic planning, 5th International Conference on Research and Innovation in Information Systems, 2017. doi: 10.1109/ICRIIS.2017.8002479.
[24] P. K. Singh, H. Chudasama, Pathways for climate change adaptations in arid and semi-arid regions, Journal of cleaner production, 2021, p. 284. doi: 10.1016/j.jclepro.2020.124744.
[25] A. Mourhir, Scoping review of the potentials of fuzzy cognitive maps as a modeling approach for integrated environmental assessment and management, Environmental Modelling and Software, vol. 135, October 2020. doi: 10.1016/j.envsoft.2020.104891.

Using Neural Networks To Solve The Differential Equation

Natalia Marchenko[1], Ganna Sydorenko[2] and
Roman Rudenko[3]

[1] National Technical University "Kharkiv Polytechnic Institute", str. Kyrpychova 2, Kharkiv, 61002, Ukraine
[2] National Technical University "Kharkiv Polytechnic Institute", str. Kyrpychova 2, Kharkiv, 61002, Ukraine
[3] National Technical University "Kharkiv Polytechnic Institute", str. Kyrpychova 2, Kharkiv, 61002, Ukraine

Abstract

The article shows that an improved form of construction of the solution of differential equations to use the direct distribution of a multilayer neural networks has been found. The solution satisfies the initial conditions by construction, but has less influence on the solution error at a distance from the initial conditions compared to the original (or "older") form of construction solutions. The solution of differential equations with neural networks is a function given in an analytical form, which can also be differentiated or integrated analytically. It has been also found a method to modify the calculation of the loss function for cases where the solution process stops at a local minimum, which will cause a large dependence of the subsequent values of the functions on the accuracy of finding the previous values. In this article, the backpropagation method is used to train a neural network. This improves the accuracy of finding the solution to the differential equation. A comparative analysis of comparative methods for constructing solutions to differential equation in partial derivatives (the thermal conductivity equation) is carried out and it is shown that the choice of the form of construction of the solution depends on the differential equations to be solved and the formulation of the problem that requires solving these equations. Also among the results, it can be noted that the solution of differential equations using artificial neural networks may have an accuracy comparable to the classical numerical methods for solving differential equations, but it usually takes much longer to achieve similar results on small dimensions.

Keywords 1

Differential equations, neural networks, multilayer neural networks, thermal conductivity equations, finite-difference methods, numerical methods for solving differential equations.

1. Introduction

In the modern science the differential equations are the basis for building mathematical models that describe physical, natural, economic, social and other phenomena. It is known that analytical or numerical solutions are used to solve differential equations both ordinary and partial derivatives. Most cases of analytic solutions are the time consuming or impossible process. Therefore, numerical methods are used in ones are Runge-Kutta methods, linear multistep methods, methods of finite differences, volumes or finite elements, spectral and Galorkin method [1].

In the general case, differential equations in partial derivatives have an infinite number of solutions and in order to find a specific solution associated with the technical, physical, natural or economic process to the equation we have to put forward some additional conditions: initial or boundary. Also all finite methods are amenable to discretization, so as the domain, partial derivatives are replaced by difference-relevant ratios (no centered or centered) on relevant independent variables. But in this case a problem appears: during the solution of an equation there is a process of discretization of the equation on a given area of the solution. This process significantly impairs the obtaining of a more accurate solution and affects the limitation of the differentiation of the equation. To avoid this problem, you can use numerical methods that are based on the apparat of neural networks.

EMAIL: Natalia.Marchenko@khpi.edu.ua (A. 1);
Ganna.Sydorenko@khpi.edu.ua (A. 2);
roman.rudenko.a@gmail.com (A. 3)
ORCID: 0000-0001-9889-3713 (A. 1); 0000-0002-0761-2793 (A. 2); 0000-0002-9424-6639 (A. 3)

The objective of this article is to use a multi-ball neural network to solve partial differential equations (for example of the thermal conductivity equation). Its main advantage in comparison with classical numerical methods is the possibility of presenting the solution in analytical form where it is possible to take derivatives multiple times [3]. This ensures the stability of the neural network model in relation to errors in the data, namely inaccuracies in the problem of the coefficients of equations, boundary and initial conditions, perturbations of limits and errors of calculations. Also using the neural networks approaches allows us to parallelize calculations, organize calculation using a different set of neural networks including various types [4-7].

2. The formulation of the problem

Let us to consider the problem of solving differential equations in the form.

$$G\big(x, F(x), F'(x), F''(x)\big) = 0, \quad (1)$$

with initial conditions

$$\begin{cases} F(x_0) = a_0 \\ F'(x_1) = a_1, \\ F''(x_2) = a_2. \end{cases}$$

where x is the vector of variable values;

$F(x)$ – desired function;

x_0, x_1, x_2 – coordinates of initial conditions;

a_0, a_1, a_2 – their matching values.

The solution is represented as:

$$F^*(x) = N(x, p), \quad (2)$$

where N is the function of the neural network with parameters p and inputting values x, or in the form of:

$$F^*(x) = A(x) + Q(x)N(x, p), \quad (3)$$

where $A(x)$ is a function that satisfies the initial conditions in advance;

$Q(x)$ – function constructed in a way to be equal to zero at the points corresponding to the coordinates of the initial conditions.

An approximate solution F^* for case (2) can be obtained by minimizing of the expression:

$$\left(G\left(x, F^*(x), F^{*\prime}(x), F^{*\prime\prime}(x)\right)\right)^2 + \left(F^*(x_0) - a_0\right)^2 + \left(F^{*\prime}(x_1) - a_1\right)^2 + \left(F^{*\prime\prime}(x_2) - a_2\right)^2.$$

The task of building function $A(x)$ is reduced to the task of the function that takes a certain values in the given points, and can take any value at all other points. To find the function, for example, an interpolation polynomic of Lagrange can be used in this case that looks like:

$$L(x) = \sum_{i=1}^{n} y_i l_i(x), \quad (4)$$

where $l_i(x)$ - basic polynomials are determined by the formula:

$$l_i(x) = \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}.$$

The construction of an expression $Q(x)$ is more complicated, and usually comes down to creating a polynomial that has roots at points that coincide with the points of the initial conditions. In the case of setting the initial conditions for the derivatives of the desired function, the polynomial is created so that its derivatives of the appropriate order have roots at the points that define the initial conditions.

$$Q(x) = \prod_{i=1}^{s} (x - x_i)^{d_i + 1}, \quad (5)$$

where d_i – order of derivative at setting i-th initial condition.

For case (3), the minimization expression is simplified to the form:

$$\big(G(x, F^*(x), F^{*\prime}(x), F^{*\prime\prime}(x))\big)^2 -> \min.$$

However, this writing has drawbacks in accuracy of finding of the solution.

3. Solving differential equations using neural networks

The solution in the form (2) doesn\`t completely satisfy the initial conditions, but the solution in the form (3) despite of satisfying the initial conditions and may have a bigger error on the distance from the point that it sets. In addition to this, there is the problem of stopping the optimization at the local minimum.

For its solution, function loss can be modified so that it can give the greatest weight to points, which are closer to the original terms of keeping the amount of loss function for all the points:

$$loss^*(x) = loss(x) * e^{-kx/x_{max}}, \quad (6)$$

$$loss^{**}(x) = loss^*(x) * \frac{\sum_{i=1}^{p} loss(x_i)}{\sum_{i=1}^{p} loss^*(x_i)},$$

where $loss(x)$ is the loss function;

x – the argument of the desired function;

k – coefficient that regulates the degree of redistribution of error towards the initial conditions;

x_{max} – the maximum value of the argument of the desired function;

x_i – the point at which optimization is made;

p – the number of points at which optimization is performed [8].

Software for solving differential equations was developed by using Python and R programming language, a machine learning library with the support for learning of neural networks TensorFlow.

As multilayer neural network of direct propagation was chosen as the neural network structure. The generalized structure of multilayer feedforward neural network is shown in Figure 1.

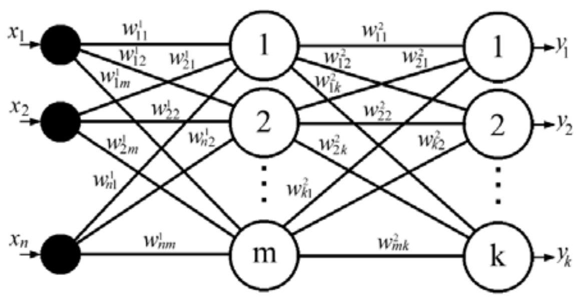


Figure 1: Multilayer neural network of direct propagation

Multilayer networks with full connection provide the ability to transmit information from each neuron of the previous layer to any neuron of the next. Most often these are multilayer perceptron. In the case of direct interlayer propagation, the information is transmitted unidirectional in the direction of increasing the layer number. Direct propagation within one layer is used in the case when a group of neurons of this layer is connected to the previous layer indirectly through selected neurons [8-10].

Thus, the undirectionality of connections leads to the construction of exclusively hierarchical structures in which information processing is distributed by levels. Each level of hierarchical information processing is responsible for its own layer of neurons. The output information of the higher level of the previous layer is the input for the neurons of the next layer, which provides a deeper level of processing [10].

Let the training sample $\Psi = \{(X^{(n)}, y^n)\}_{n=1}^N$, containing N pairs: $x^{(n)}, y^n$ respectively, the input and output vectors and the set of parameters $W = \{w, v\}$, which consists of the parameters of the neurons of the hidden layer w and the parameters of the output layer v. The method of inverse propagation [10] is to minimize the objective function:

$$E = \frac{1}{2N}\sum_{n=1}^{N}[y^{(n)} - \tilde{y}^{(n)}(x)]^2 -> \min.$$

The number of layers and the number of neurons in each layer are selected based on the structure of the problem and the complexity of the form of solutions after the experiments. The results of the application of the proposed method with a simple example ordinary differential equation given in [11].

The form of construction of expression $Q(x)$, as seen from numerical experiments, greatly affects the error of approximation of the desired solution in distances from a given point of initial conditions. To reduce this effect we can write the expression for $Q(x)$ in the form:

$$Q(x) = \prod_{i=1}^{S} th(x - x_i)^{d_i+1},$$

where $th(x) -$ is the hyperbolic tangent.

Thus, the function $Q(x)$ should satisfy the condition of zero equality at given points, but will have a much smaller deviation at the but will have a much smaller deviation at a distance from the points that give the initial conditions.

4. The main results for the thermal conductivity equation

One of the classical problems is the modeling of the thermal state in a flat plate, which is described by a differential equation in partial derivatives with Dirichle conditions

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = -\frac{Q}{k}, \quad (7)$$

$$\begin{cases} T(0, \mathrm{y}) = 0, \\ T(\mathrm{a}, \mathrm{y}) = 0, \\ T(x, 0) = 0, \\ T(\mathrm{x}, \mathrm{b}) = 0. \end{cases}$$

where T – temperature,

x, y – coordinates along the respective axes $x\epsilon[0,\ a]$, $y\epsilon[0,\ b]$;

Q – heat flow;

k – coefficient of internal thermal conductivity.

The analytical solution has the form [12]:

$$T(x, y) = \frac{Q}{2k}[x(a - x) - \frac{8a^2}{\pi^3}\sum_{n=0}^{\infty}\frac{\sin\left(\frac{(2n+1)\pi x}{a}\right)}{(2n+1)^3}\frac{ch\left(\frac{(2n+1)\pi(b-2y)}{2a}\right)}{ch\left(\frac{(2n+1)\pi b}{2a}\right)}].$$

The algorithm of neural network training process for the thermal conductivity equation is:

1. Create the model:

1.1. Input variables vector (x, y).

1.2. Add positional encoding:

$$x_i = \sin(\frac{x}{a} * n)\,, n = 2, 3, \dots, 33\,,$$

$$y_i = \sin(\frac{y}{b} * n), n = 2, 3, \dots, 33\,.$$

1.3. Create 3 layers of a neural network of the perceptron type with output dimensions: 512, 32, 32 and the tanh activation function: $\mathrm{out} = \tanh(\mathrm{W} * \mathrm{in} + b)$, where in – for the first layer – the input vector, and for the following layers – the output of the previous layer, W – the weight matrix, b – bias vector.

1.4. Add one output layer of a neural network with an output dimension of 1 without an activation function: $u = W * \mathrm{in} + b$, where in – output vector of the previous layer.

1.5. Add the transformation to satisfy the initial conditions:

$$ucond = \sin\left(\frac{x}{a}\pi\right)\sin\left(\frac{y}{b}\pi\right)(u+1)\sqrt{a * b}.$$

2. Create the optimizer Adam.

3. Training loop (4000 iterations):

3.1. 200 random points are selected and the training method is called. Selected points are uniformly distributed in a given range, while for each 4 of 47 points, one of the coordinates is moved to the border.

3.2. The optimizer calculates the loss function for these points and updates the weights (stochastic gradient descent algorithm Adam with back-propagation of error is used).

3.3. If over the last 40 iterations there was no significant decrease in the loss function, then learning rate is decreased by 2%.

The loss function of this model with the satisfaction of the initial conditions by the construction of the cost function is displayed:

```
def loss(u, x):
du_dx = tf.gradients(u, x)[0][:, 0]
du_dx2 = tf.gradients(du_dx,
x)[0][:, 0]
du_dy = tf.gradients(u, x)[0][:, 1]
du_dy2 = tf.gradients(du_dy,
x)[0][:, 1]
eq_right = - self.Q / self.k
eq_left = du_dx2 + du_dy2
loss_diff = tf.reduce_mean((eq_left
- eq_right) ** 2)
return loss_diff
```

The obtained optimization result in the basic form:

```
10000/10000 - 1s - loss: 5.7227e-5 -
rmse: 8.2836e-4 - val_loss: 5.2799e-
5 - val_ rmse: 8.8854e-4
```

The results for solving equation (7) with parameters $\mathrm{Q} = 10,\ \mathrm{k} = 5, \mathrm{a} = 5,\ \mathrm{b} = 8$ in General (2) are shown in Figure 2, and for solutions in the form (3) - in Figure 3.

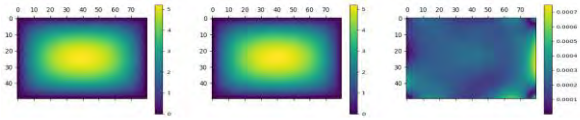


Figure 2: Solution of the equation in general (left), analytical solution (middle), solution error (right)

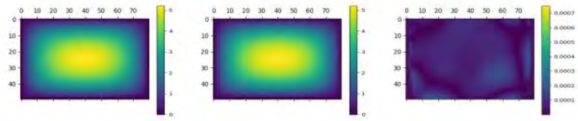


Figure 3: Solution of the equation in the form (3) (left), analytical solution (middle), solution error (right)

In the first case with mean-square error is 8.8854×10^{-4}, moreover, the error in a closer in half of the points to the boundary conditions 1.06 times less than the error in half of the points away distance from the boundary conditions. The solution in the basic form (2) had the largest errors in the field of setting boundary conditions.

When using the solution in the form (3) with the satisfaction of the initial conditions, the root mean square error is 2.3488×10^{-3} and at the same time the ratio between the error in half the points at a distance from the boundary conditions and near to them is slightly different from the previous case and is 1.07.

To compare the obtained results, we solved the equation of the example with the same parameters by the finite-difference method (grid method) with the grid 80×50. The Figure 4 shows the solution of the equation by the finite-difference method.

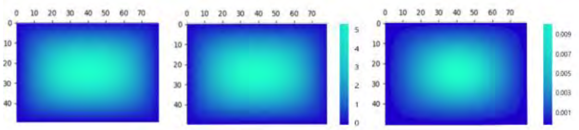


Figure 4: Solution of the equation by the finite-difference method in the form of (3) (left), analytical solution (middle), solution error (right)

In the grid method, the mean-square error is 9.7254×10^{-3}, and the largest errors are observed in the middle of the plate and it is undesirable at the edge. Therefore, we can conclude that the replacement of differential operators by difference analogues introduces errors in the numerical solution of the equation, because there are errors associated with steps in time and space.

Therefore, comparing the results of research, we can conclude that the method of grids works, but with greater errors and greater time and resources. Therefore, our study showed better performance of the proposed method. The neural network apparatus can also be used to solve systems of differential equations. Testing of the developed software was performed for systems efficiency mall order (up to 10) and showed satisfactory results in comparison with classical numerical methods.

5. Conclusions

In this article the method of solving of differential equations is given, which requires only a task of a loss function to optimize of the direction and does not require solving equations with highest derivative. The solution of differential equations using neural networks is a function, given in analytic form that can be differentiated or integrated analytically as well.

Also among the results, we can note that the solution of differential equations using artificial neural networks may have bigger accuracy comparable to classical numerical methods, but at the time, we can consume, that it usually takes longer time to achieve similar results in the tasks of small dimensions.

In the article, to find a solution to a partial differential equation, a multilayer feedforward neural network neural network was used for training with using the backpropagation method of the loss function. This improved the accuracy of finding the solution to the heat equation.

The main advantage of using neural networks in solving some differential equations is that the solution is performed analytically and can be found not only for individual values of parameters of equations, but also for all values of parameters in a limited range of values.

6. References

[1] E. Hayrer, GWanner, Solution of ordinary differential equations. Rigid and differential-algebraic problems, Myr, Moskow, 1999.

[2] A. S. Mazmanishvili, G. Yu. Sydorenko Methodical instructions for laboratory work “Solution of differential equations in partial derivatives”, Kharkiv, NTU "KhPI", 2015.

[3] Lagaris I. E. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations / Isaac Elias Lagaris, Aristidis Likas, Member, IEEE, and Dimitrios I. Fotiadis // IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 9, NO. 5, SEPTEMBER 1998.

[4] S. Mall, S. Chakraverty, Comparison of Artificial Neural Network Architecture in Solving Ordinary Differential Equations. Hindawi Publishing Corporation Advances in Artificial Neural Systems Volume 2013: 2-12.

[5] A. N. Vasilyev, D. A. Tarkhov, Neural network approach to problems of mathematical physics. Sankt-Petersburg: Nestor-Istoria, 2015.

[6] A. N. Kovalenko, A. A. Chernomorets. M. A. Petina. On the use of neural networks for the solution of partial differential equations. Scientific Bulletin of BelSU . Ser. Economics. Informatics, volume 9 (258), issue 42, pp. 103-110, 2017.

[7] Martin Magill, Faisal Z. Qureshi, Hendrick W. de Haan. Neural Networks Trained to Solve Differential Equations Learn General Representations, 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

[8] Maithra Raghu, Justin Gilmer, Jason Yosinski, and Jascha Sohl-Dickstein. SVCCA: Singular Vector Canonical Correlation Analysis for Deep Learning Dynamics and Interpretability. In Advances in Neural Information Processing Systems, pages 6078–6087, 2017.

[9] Neha Yadav, Anupam Yadav and Manoj Kumar. An introduction to neural network methods for differential equations. Springer, 2015.

[10] Modjtaba Baymani, Asghar Kerayechian, Sohrab Effati. Artificial Neural Networks Approach for Solving Stokes Problem, Applied Mathematics 01(04) (2010):288-292. DOI:10.4236/am.2010.14037.

[11] R. O. Rudenko, N. A. Marchenko. Development of mathematical and software for solving differential equations using neural networks // XIII International Scientific Conference undergraduates and graduate students" (19-22.11.2019 y.): Conference materials / ed. prof. Ye.I. Sokol. Kharkiv: NTU "KhPI", 2019.

[12] Yu. P. Boglaev Computational mathematics and programming. Moskow: Vysshaya shkola, 1990.

Application of free software FreeCAD for modelling and analysis of generated stresses of Jib crane beam

Viktor Strelbitskyi[1], Liubov Bovnegra[2] and Nataliia Punchenko[3]

[1] *Odessa National Maritime University, Odesa, Ukraine*
[2] *Odessa Polytechnic State University ,Odesa, Ukraine*
[3] *Odesa State Environmental University Odesa, Ukraine*

Abstract

The jib cranes typically used to load and unload many items in many industries. The paper aims to investigate the potential FreeCAD for the desing of elemehts of jib crane. The basic premise of the FEM is that a solution region can be analytically modelled or approximated by replacing it with an assemblage of discrete elements. The numerical analysis of mechanical stress was made to the jib crane beam which, which cross-section was the I-beam, are presented in this paper. The beam is modeled as a three dimensional design and the analysis of the construction is performed using finite element method.

The sample analysis results cover various load configurations with respect to selected beam construction components. As a result of the influence of the loading force position on the equivalent stress in the gantry crane beam was evaluated. Using finite element method in the analysis is appropriate and this method provides great convenience in the calculations.

It was found that the maximum values of the stresses induced in the beam do not exceed the strength of the material; therefore, the structure is safe..

Keywords

FreeCAD ,FEM, jib crane, load, stress.

1. Introduction

The jib cranes are currently one of the most common devices for transporting loads and can be applied in many industry areas. The jib cranes provide an economical way to lift materials.

Jib cranes are very useful for lifting and transferring heavy loads in circular work volume, the crane provides easy, safe and faster transfer of load from one place to another [1-4]. These cranes can help to improve material handling and work flow. Jib crane provides flexibility in design.

First used of gantry cranes for doing the lifting the work piece from raw material to finished product by using Greek persons.

The main component of crane is girder beam, usually made of steel, which transfers the load on its other elements therefore it is tested for strength [1-4]. To ensure the safety of the structure, it is necessary that the maximum stress value be less than the strength of the material [1-4].

The construction of a jib crane can fail due to various factors. However, the most common cause is plastic deformation due to excessive loading [1].

The maximum deflection of the cantilever beam is critical criterion for the jib cranes.

It should be noted that the design of modern mechanical constructions is a complex task that requires the use of suitable tools. Moreover, experimental studies are expensive and time consuming [1-5]. Therefore, finite element modeling (FEA) is a suitable tool for solving this problem [6-11].

Initially, finite element analysis was used in aerospace structural engineering.

The FEA is very useful especially for validating the strength and stability of the design under actual loading conditions. At FEA method helps in assessing the structural performance of the design under static, dynamic or cyclic loads.

The program’s predictions of the modelling of uniform beams and cantilevers under various

EMAIL: vict141174@gmail.com (A. 1); dlv5@ukr.net (A. 2); iioonn24@rambler.ru (A. 3)
ORCID: 0000-0001-7027-9498 (A. 1); 0000-0003-0429-2816 (A. 2); 0000-0003-1382-4490 (A. 3)

loading and restraint conditions are generally in close agreement with existing predictions and test results.

To solve a problem, the FEM subdivides a large system into smaller, simpler parts that are called finite elements that are joined at nodes. This is achieved by a particular space discretization in the space dimensions, which is implemented by the construction of a mesh of the object: the numerical domain for the solution, which has a finite number of points. It is in fact a particular numerical method for solving partial differential equations in two or three space variables, i.e. some boundary value problems. The finite element method formulation of a boundary value problem finally brings to a system of algebraic equations [8,9].

The method approximates the unknown function over the domain [3,4,6,7]. The finite element method, therefore, has three main stages: build the model; solve the model and display the results.

The subdivision of a whole domain into simpler parts has the following advantages [2,8,9]:

1) accurate representation of complex geometry;

2) inclusion of dissimilar material properties;

3) easy representation of the total solution;

4) capture of local effects.

FEA simulations provide a valuable resource as they remove multiple instances of creation and testing of hard prototypes for various high fidelity situations [6-10]. The modelling is helpful in improving the speed of product development.

The differential equation can be solved with sufficient accuracy using with the help of computer.

In this study, a gantry crane is modelled in 3D using FreeCAD computer software [10,11].

FreeCAD is a 3D CAD/CAE parametric modeling application, is free alternatives, developed in 2001. Nowadays FreeCAD is a multidisciplinary software application, due to long term active collaboration of developers and users.

FreeCAD contains many components to adjust dimensions or extract design details from 3D models to create high quality production ready drawings. Besides, it allows both GUI modeling and script-based modeling using Python [10,11].

The simulation itself is done by FEM module, which is implemented as an integration of Calculix solver. This is the first solver that was integrated in FreeCAD to work with the FEM analysis. The Calculix solver is able to do linear and non-linear calculations, for static, dynamic, and thermal problems. The input file that CalculiX uses can be prepared and edited before the solver is started. The results displayed by the FEM workbench are of course currently not enough to perform real-life decisions about structures dimensionning and materials [10,11].

Open source scientific hardware contributes to the development of Open Science by facilitating the comparison of the scientific experiments. The Open Source Hardware Statement of Principles is made publicly available so that anyone then can study, modify, distribute, make. [6,10,11].

However, FreeCAD can prove useful in a much wider area than its current purpose [6,9,10].

It should be noted that the calculated maximum stresses are compared with the allowable stresses of the beam material. The calculated deflection values were also compared with the allowable ones. Since, only if the strength condition is met, the structure is considered safe.

In addition, bending of the beam will damage the bogie bearings.

2. Samples and analysis method

2.1. Description of the problem

As previously mentioned, in this study the main beam of the 5 ton jib crane is examined. (Fig.1).

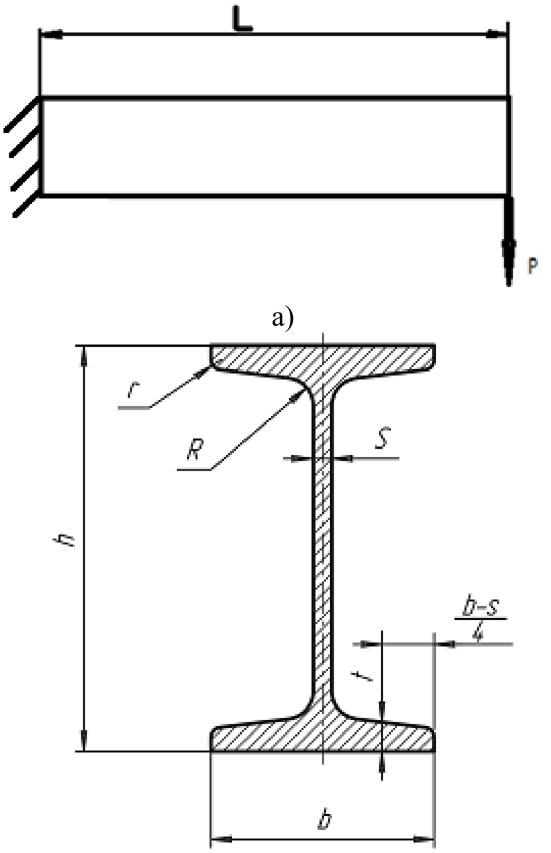
Figure 1: Jib crane

A boom of jib crane experiences a maximum load when the load is acting at the maximum span tends to bending.

The boom of jib crane which is made of I-section beam.

To analyze the magnitude of the load on the stress distribution in the jib crane, the design scheme was presented in the form of a static single-span beam (Fig. 2), with dimensions shown in Table 1.

A trolley moves along the bottom of the main beam, which has a hook-block, which is connected to the cables and can lift up to 5 tons (Fig.2). The length of the beam – 6 m.



a) b)

Figure 2: Considered the beam loaded (a) of the gantry crane with cross section of the I-beam (b)

Table 1

Dimensions beam

Dimension	mm
h	330
b	160
S	7,5
t	10,2
R	13
r	5
L	6000

The structure is loaded by load perpendicular to the bottom flange of the beam and directed downwards. The amount of load is determined by the total weight of the load and the bogie.

Since the beam is made of S steel, the corresponding material constants were used in the calculated calculations.

Usually, the load movement is neglected and a numerical analysis is performed for a certain position of load. Then the load inertial forces are ignored [2,6,7,9].

Meshing is done using automatic meshing method.

The analysis of the stresses of the gantry crane structure was made using FreeCAD which is implemented as an integration of Calculix solver.

Numerical modeling of the gantry crane girder makes it possible to assess the state of stresses, indicating the critical sections of the structure.

2.2 Results and discussion

The figure 3 represents the CAD model the jib crane. Corresponding boundary conditions which are applied to the finite element models are illustrated in Fig.4.

The Von Mises stress and displacement in the I- beam of the load F = 55 kN (a), F=45 kN (b), F=35 kN (c) and F=25 kN (d) are shown in Figures 5-6.

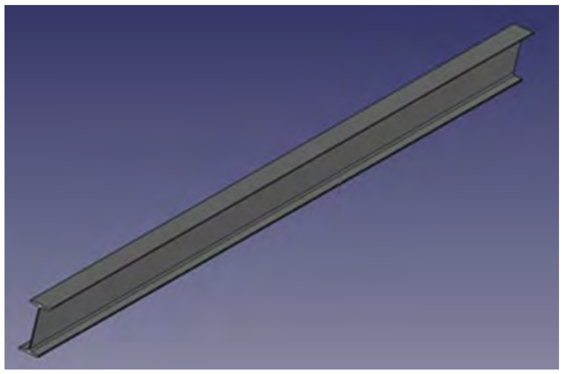


Figure 3: CAD Model of Cantilever I-Type beam

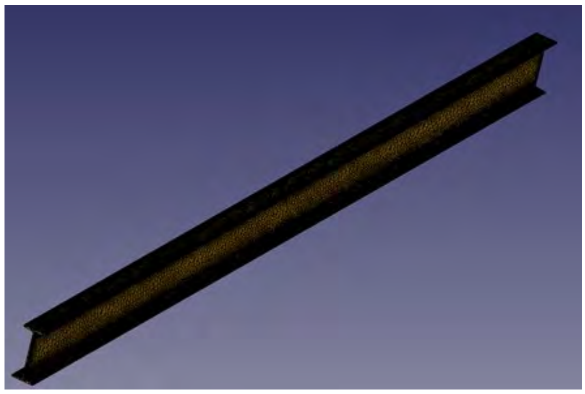
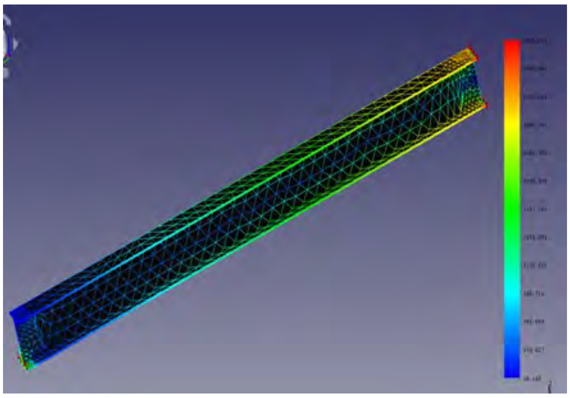
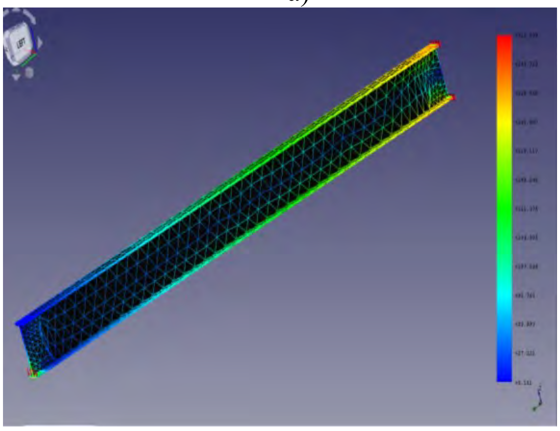


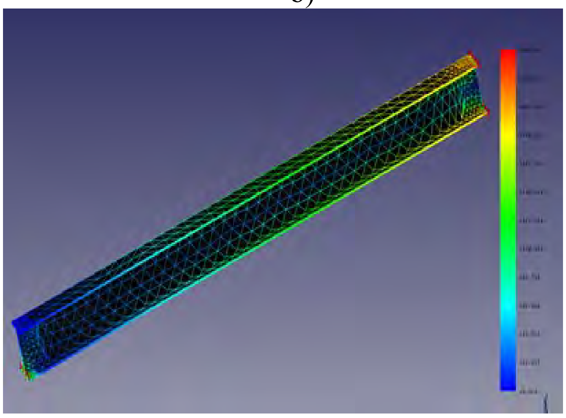
Figure 4: Fnite element models of Cantilever I-Type beam



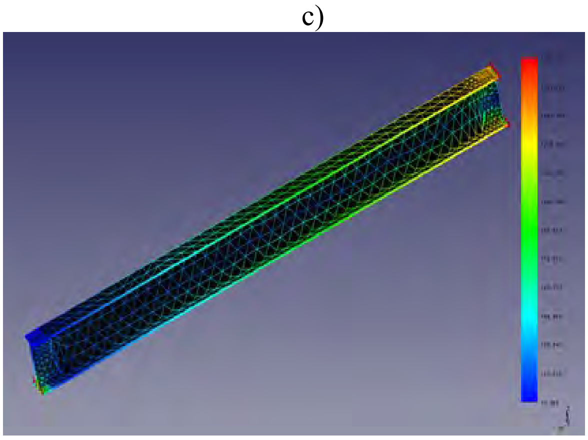
a)



b)

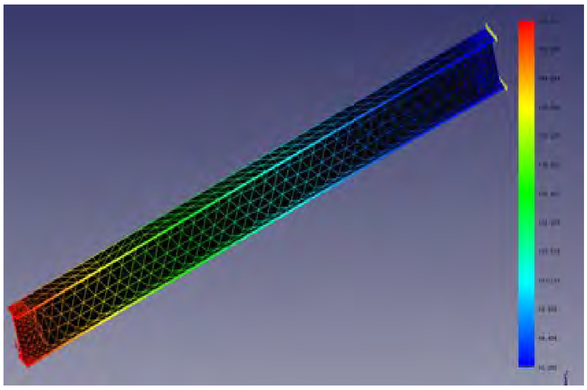


c)

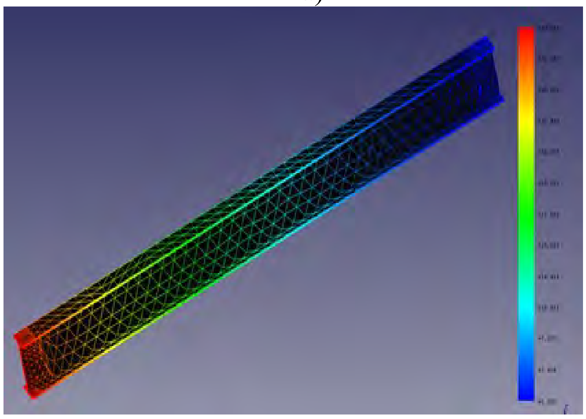


d)

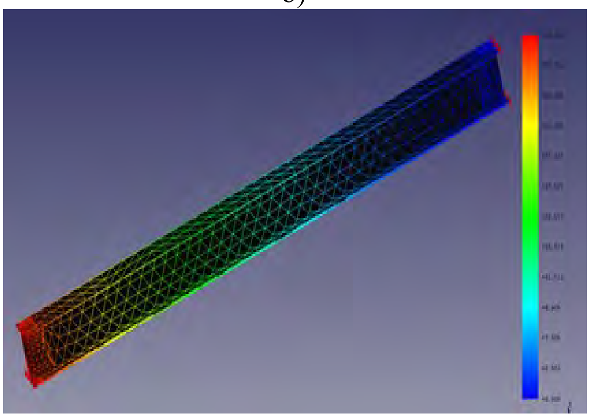
Figure 5: Von Mises stress in the I-beam of the load F = 55 kN (a), F=45 kN (b), F=35 kN (c) and F=25 kN (d)



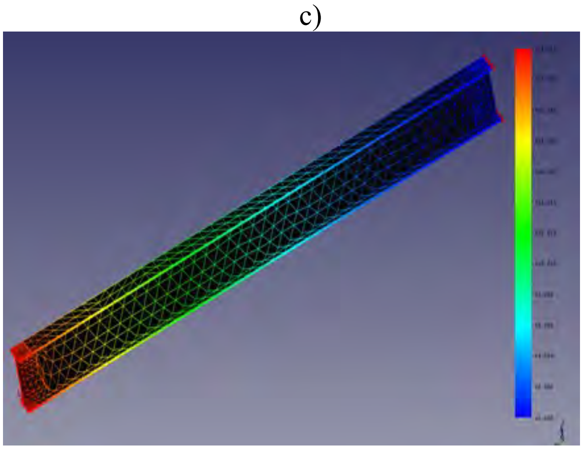
a)



b)



c)



d)

Figure 6: Displacement in the I-beam of the load F = 55 kN (a), F=45 kN (b), F=35 kN (c) and F=25 kN (d)

The results show that an increase in load for beam leads to an increase in voltage and displacement. In addition, a maximum of stresses is located at the points where the beam is attached.

The maximum von Mises equivalent stress in the cross-section of the I-beam ilength 394,3 N/mm^2 which does not exceed the permissible .

It should be noted that results of numerical simulations are consistent with the results found in literature [1,3].

The research performed allows for the evaluation of the stress state, pointing out the critical areas and values.

3. Conclusions

In this work the analysis of the strength analysis the jib crane I- beam are presented. With the development of computers, FEA has become a practical way to solve many problems. The analysis has been made on the FreeCAD, by using the finite element method, as an integration of Calculix solver.

It provides a convenient method for optimizing the structural material and examining the applicability of the structural material.

The influence of force value on the equivalent the von Mises equivalent stress (Fig.5) and deformation (Fig.6) in the beam was evaluated.

The use of the I cross-section in the jib crane beam design seems to be a better solution in terms of strength.

The maximum Von Mises Stress occurs at the junction of at the beam and column (rigid fastening), at the free end of the beam - highest deformation.

The maximum voltage value for the boom of the I-beam section is 394,3 N/mm^2, which does not exceed the permissible. As the computed stress values in the jib are smaller than the allowable stress of Material (Structural Steel) of the components, thus the jib crane is safe according to norms ISO/TR 25599:2005 [3].

4. Acknowledgements

We thank the anonymous reviewers for their important and valuable suggestions.

5. References

[1] Verschoof I.J. Cranes - design, practice, and maintenance, Professional Engineering Publishing Limited London and Bury St Edmunds, UK., 2002.

[2] Gerdemeli .I, Kurt .S and Tasdemir “Design and Analysis with Finite Element Method of Jib Crane”, Mechanical Engineering Istanbul Technical University –Turkey, 2012.

[3] ISO/TR 25599:2005, Cranes - Jib cranes - International Standards for design, manufacturing, use and maintenance requirements and recommendations.

[4] EN 13001-2:2011: Crane safety - General design - Part 2: Load actions.

[5] Yakimov, Alexey, Isak Karabegovic, Sergey Uminsky, Viktor Strelbitskyi, and Julia Shichireva. Stability of the Quality Parameters for the Surface Layer of Parts During Circular Grinding Operations. In Grabchenko’s International Conference on Advanced Manufacturing Processes, Springer, Cham, 2020, pp. 464-476. https://doi.org/10.1007/978-3-030-68014-5_46.

[6] Chaudhary S. A., Khan N.S., A review paper on structural analysis of cantilever beam of jib crane, International Journal of Engineering Research and General Science, (2015), Vol 3, No 3, pp. 612-615.

[7] Machado, F.; Malpica, N.; Borromeo, S. Parametric CAD modeling for open source scientific hardware: Comparing OpenSCAD and FreeCAD Python scripts. PLoS ONE (2019), 14.

[8] Reddy J. N. An Introduction to the Finite Element Method (Third ed.). McGraw-Hill, 2006.

[9] Tirupathi R. Chandraputla and Ashok, D. Belgundu” Introduction to Finite Elements in Engineering”, Pearson Education, 2002, 3rd Edition

[10] Machado F, Malpica N, Borromeo S (2019) Parametric CAD modeling for open source scientific hardware: Comparing OpenSCAD and FreeCAD Python scripts. PLoS ONE 14(12): e0225795. https://doi.org/10.1371/journal.pone.0225795.

[11] FreeCAD Documentation Available from: http://www.freecadweb.org/wiki/Main_Page

Neuro-fuzzy intelligent network for control of microclimate systems

Kharchenko R. Yu. [1], Kochetkov O. V. [2]

[1] National University "Odessa Maritime Academy", Didrikhson str.8, Odessa, 65029 Ukraine
[2] Odessa National Maritime University, Mechnikova str.8, Odessa, 65029 Ukraine

Abstract
Metods in the construction of a neuro-fuzzy network for control of parameters of microclimate systems based on expert data are considered. A method for intelligent adaptation of the control is proposed. This allows maintaining high accuracy of the specified parameters for different operating modes of the system and reducing the complexity of control. Software has been developed that implements the proposed network. Computer simulations have shown the ability of the network to self-learn based on expert experience.

Keywords
neurocontroller, adaptive PID - controller, parameter approximation, active identification, fuzzy logic, neural network.

1. Introduction

It is explained by the imperfectness of mathematical models of the objects represented both in an analytical (often too simplified) and experimental forms. It is difficult to obtain mathematical models of the control systems of functioning complex objects experimentally as well as of those operating in remote or automatic modes. These difficulties are connected with the influence of external and parametric disturbances, some of which are of non-stationary nature and cannot be controlled. Taking the above-mentioned into account, the developed APCS (automated process control systems) require readjustment accompanied with lower quality of the control process, which involves additional consumption of material resources.

As frequent works on setting up ACS (automated control systems) at the time of commissioning as well as during subsequent operation (when tasks and loads are changed) are required, organizational problems arise during their implementation. For example, the number of control system circuits to be adjusted at a modern heat power object may reach several tens, which makes it almost impossible to provide high-quality and rapid manual execution of the work by a limited number of the employees from the operational personnel [1-3]. The above-mentioned shows that application of intelligent control systems of complex multi-mode heat power objects (microclimate systems) has become an urgent and necessary measure of improving efficiency of the energy complex management in general. The goal of this paper is development and training of hybrid networks in order to determine the optimal values of PID - controller settings in cascade ACS of the temperature under changing load conditions of the object (adjusting mode) and in the presence of start-up mode.

2. Literature review

The study of scientific publications in the field of adaptive ACS of heat and power systems [4 - 7] led to the conclusion that the traditional methods of active identification and associated algorithms are widely used for calculating the optimal settings of PI and PID regulators by analyzing CFC (comprehensive frequency characteristics) of the objects or auto-oscillation mode. It should be noted that for the temperature ACS under consideration auto-oscillatory process

EMAIL: romannn30@gmail.com (R. Yu. Kharchenko); 0679016767@ukr.net (O. V. Kochetkov)
ORCID: 0000-0003-3051-7513 (R. Yu. Kharchenko); XXXX-0000-0001-8200-4897 (O. V. Kochetkov)

is invalid due to the process procedure requirements as the temperature deviation from the norm can lead to a premature wear of the equipment. Thus, there arises a scientific problem of finding optimal object identification methods in the cases of changing load and the algorithms for calculating the settings of PID regulators, taking into account the expert opinion, in order to ensure the expected transient process (overshoot G < 30%, the degree of attenuation Ψ = (0,75 - 0.95) with a minimum regulation time T_r.

At present, scientific approaches related to the use of intelligent systems are very popular in the theory of adaptive control [8 - 11]. These systems successfully implement the experience and knowledge of experts (fuzzy regulators) as well as possess the self-learning ability (neural controllers). Joint or combined use of these areas gave rise to the emergence of a new scientific trend of hybrid, or fuzzy neural networks (FNN) [12]. Consideration of this technology as applied to identification and adaptation of the superheated temperature ACS is an important current scientific problem.

3. Functioning algorithm for adaptive ACS parameters of microclimate systems

The authors proposed a structure of the hybrid adaptive control system (Fig. 1). In fig. 1 the following designations are introduced: K - object transfer coefficient, T - object time constant, τ - delay, n -order of the object, y - output parameter.

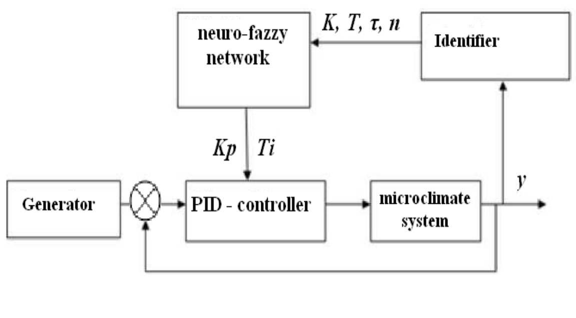


Figure 1: Structure of adaptive ACS

The structure of PID regulator adaptive setting is based on the application of frequency identification methods and finding optimal PID regulator settings by the hybrid network. Application of the frequency methods has made it possible to provide interference immunity of the algorithm and rational organization of the active experiment using the functioning system in terms of maintaining the stability region. Identification was conducted by applying two different-frequency sinusoidal signals from the generator to the system input, the frequencies belonging to a significant range [7]. Structure of the object transfer function consists from several inertial links with the lag of $W(s) = \frac{K}{(T(s)+1)^n} e^{\tau(s)}$, the values varying over time within a definite range depending on the load type or on the steam boiler operating mode. The identifier determines values of the object parameters and its order. These values are further used by the optimizer in the form of a neural fuzzy network operating according to the algorithm of Sugeno [2 - 3] for searching the optimal values of PI regulator settings (K_r, T_c). The hybrid network training should be conducted taking into account the opinions of expert fitters of ACS.

4. Object identification of the optimal PID-regulator parameters

In order to determine the values of the four object parameters *(K, T, τ,* n), It is proposed to use a harmonic oscillator that estimates CFC of the object at two different frequencies belonging to a significant range. The optimal values of PID regulator settings were calculated using MathCAD program.

It should be noted that PID regulator settings, calculated using the proposed method, required manual corrections since a number of the obtained fading transient processes did not meet the given criteria (the first deviation and the regulation time). To obtain a test sampling (a matrix of the neural fuzzy network training) of the optimal parameter values for PID regulator of the cascade ACS, a computer experiment was conducted using MatLab (Simulink) software with manual correction of the settings of K_r and T_c. Parameter values of the object transfer function were changed taking into account different modes of the steam boiler operation (nominal, economical, starting, stopping, regulation and peak modes) [8].

Experimental results are presented in table 1.

Table 1
Optimal PID regulator settings and parametric disturbance action on the object

K_{ob}	1	5	12	4,43	7,05	11,9
T_{ob}	3	1,5	1,5	2,48	1,5	1,5
τ	1	0,5	0,5	0,92	0,5	0,5
K_r	0,99	0,135	0,031	0,22	1,92	0,031
T_i	5,55	1,92	1,58	6,61	0,1	1,58

As practice has shown, it is possible to use both PID and PI regulator without significant loss of quality of object characteristics [5-8]. Therefore, in experiments, the differential component of regulation is sometimes excluded.

5. Development of the fuzzy neural system

Fuzzy neural network (FNN), or a hybrid network (HN), is a multi-layer neural network without feedback [13]. The inputs of such network (K, *T, τ)* are represented in the form of linguistic variables (small, average and high values). In Matlab (ANFIS) program the process of building an adaptive neural fuzzy inference system was conducted (Fig. 2) for approximating the dependence that represents a casual relationship between K, *T, τ* and K_r , T_c. Proceeding from recommendations [14, 15] and computer experiments in MatLab (Fuzzy Logic Toolbox) environment, the type of membership functions (trapezoidal and triangular), describing the input values, was chosen [9]. During training the number of cycles we used was equal to 40 and the selected method of instruction - back error propagation [10].

Thus, the hybrid network implemented mapping of PID regulator parameters according to the regulation object characteristics: $S^k = f(x^k) = f(x_1^k, x^k_2, ..., x_n^k)$, $k = 1,2,\ldots N$, in the presence of training set $((x^1, y^1), \ldots,(x^N, y^N))$ - table 1.

For simulating an unknown function *f* Sugeno algorithm was used with the knowledge base of the following type: n : IF x_1 is A_{i1} H x_2 is A_{i2} H x_1 is A_{in}, THEN $T_C = S_i$, $i=1$, 2, …, m, where Ay - fuzzy sets of a triangular form, describing the expert statements (small (S), average (A), big (B)), *S*, - output values of the regulator.

The degree of truth p of rule *i* is determined by the operation of conjunction. The fuzzy system output was determined by the center of gravity method [13].

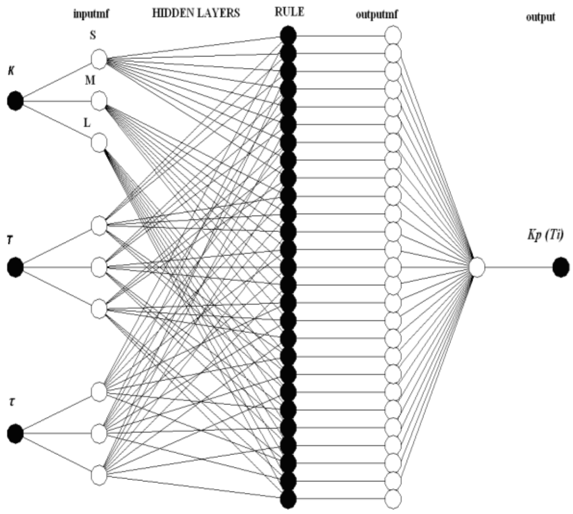


Figure 2: Structure of neuro-fuzzy network

6. Computer experiments to test adaptive and neuro-fuzzy systems

Using MatLab (Simulink) software, models of the cascade ACS were elaborated. They include PID (PI) regulators and third-order inertial objects with time delay (a model via the regulation channel) with a non-linear element (restriction on the control action).

A single jump was supplied to the system input. If the values of the transfer function parameters are changed in the cases of power unit transition to a peak or a start-up mode. Substitution of these values into Simulink software scheme has made it possible to obtain the following transient processes (Fig. 3).

From the analysis of the transient processes of Fig. 3 (1, 2) it could be concluded that under nominal or stable operating conditions of the object the hybrid and traditional PI regulators demonstrated the same quality indicators (regulation time T_r = 45s). However, under the influence of parametric disturbances (3, 4) the hybrid system had shorter regulation time (T_{ri} = 138 s) as compared with the traditional adaptive ACS (T_{r2} = 173 s), the hybrid ACS overshooting G^{hb} = 28% while that of a traditional ACS G^{tr} = 50%, the degree of fading for the hybrid ACS Ψ^{hr} = 0,91, for the traditional Ψ^{tr} = 0,68, i.e. the proposed ACS is optimal and energy-saving while the traditional cascade ACS requires additional adaptation.

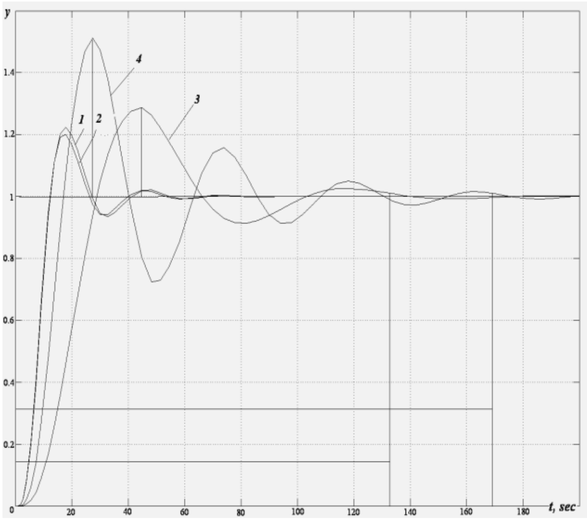


Figure 3: Transient processes over the regulation channel: 1 - for a traditional ACS and 2 - for the fuzzy neural ACS (under nominal load), 3 - for the fuzzy neural ACS and 4 - for a traditional ACS (under changing load).

The training of the NNC - neural network controller and the emulator was carried out on the model of an adaptive PID controller with manual adjustment of its parameters.

Computer modeling in MatLAB of neural network control with an emulator and a PID-controller showed that the trained neural network (NN) compensates for disturbances in the entire range of variation of the object's parameter values and has successful approximating and extrapolating properties. But it should be noted that despite the achieved success, the capabilities of the neural network system are limited in production terms due to the complexity of its setup and training.

Based on this, it is proposed to consider a promising task of scientific research - the analysis of the effectiveness of the joint (parallel) work of the NNC and the fuzzy regulator.

Let us carry out a comparative analysis of the operation of parallel neural systems with a PID - controller and a fuzzy controller under the influence of deep parametric disturbance along the reference channel (Fig. 4).

Note that the values of the parameters of the transfer function of the object along the task channel significantly go beyond the range of the NNC training sample. This experiment should be considered a test for the stability of the ACS of the microclimate system at a minimum peak load (emergency operation).

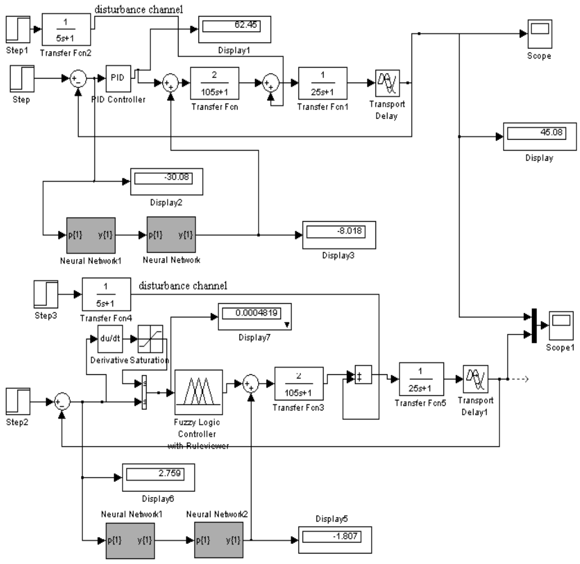


Figure 4: Computer models of NNC on PID and fuzzy controllers for obtaining comparative characteristics for different operating modes

Two parallel operating regulators (NR and NMC) successfully compensate for changes in the parameters in the range of uncertain a priori information. Analysis of transient processes (Fig. 5) demonstrates the advantage of an NNC working in tandem with a fuzzy regulator, which can be explained, by the presence of universes of membership functions of a nonlinear type, as well as by a knowledge base of 49 production rules that describe in detail the deviations of error and its derivative on a significant range.

The aperiodic process of NNC with fuzzy controller has a regulation time: $T_r = 100$ sec, the neurosystem with PID is also stable, but has a significant first deviation $\Delta Y_1 = 2$, and $T_r = 250$ sec. Experiments have shown the success of the proposed self-organizing management system.

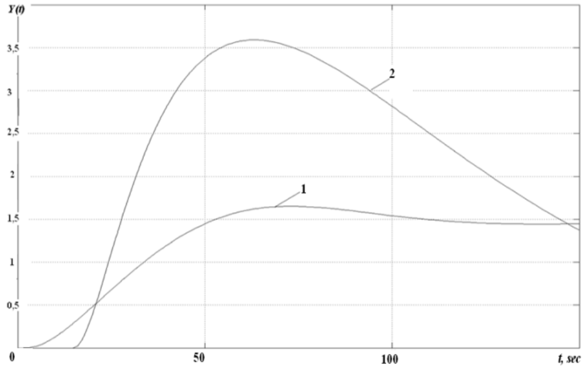


Figure 5: Transient processes along the reference channel: 1- NNC process with PID and fuzzy controller, 2 - NNC and PID process

The proposed control system can be improved by developing and using new algorithms, as well as expanding the functionality of the system (regulation of air exchange, humidity, chemical and biological composition of air, etc.).

7. Conclusions

Experiments using MatLab with varying values of the object transfer function parameters have demonstrated that the traditional adaptive approach is ineffective in a number of cases (ineffective ACS) in contrast to the algorithm of HN, the transient processes of which proved to be optimal. From the obtained results it can be concluded that the proposed intelligent adaptive ACS of temperature has the following advantages over traditional methods of adaptation according to CFC, which are currently used in ACS of HPS:

1) fast process of finding optimal regulator settings for cascade ACS with the ability of their approximation and extrapolation as well as under the action of uncertain disturbances;

2) smaller first deviation and shorter transient process of regulation;

3) the possibility of the ACS optimal operation under all the modes of the operation;

4) the possibility to use different ACS with PI and PID regulators during the adaptation processes in thermal power engineering.

Analysis of intelligent control methods allows us to conclude that the use of neuro-fuzzy logic is promising to improve the quality of control of microclimate systems. Such intelligent systems have the properties of self-learning and the ability to change depending on the circumstances. The scientific task of searching for new, more effective methods of adaptive regulation using the experience of an expert from the ACS remains open. The use of the apparatus of neural networks, in which conclusions are made on the basis of the theory of fuzzy sets, makes it possible to successfully identify an object - a microclimate system that functions in dynamic modes, especially characteristic of mobile objects and abrupt changes in climatic conditions.

8. References

[1] Widrow, B., Lehr, M. A.: 30 years of adaptive neural networks: perceptron, madaline, andbackpropagation. Proceedings of the IEEE. 78(9):1415–1442. (1990) doi:10.1109/5.58323

[2] Mohiuddin, M. S. (2014). Performance Comparison of Conventional Controller with Fuzzy Logic Controller using Chopper Circuit and Fuzzy Tuned PID Controller. Indonesian Journal of Electrical Engineering and Informatics (IJEEI), 2(4). doi:10.11591/ijeei.v2i4.120.

[3] Cheng, H. X., Zhang, D. S., & Cheng, L. (2013). Comparative Study on Fuzzy PID Controller and Conventional PID Controller. Applied Mechanics and Materials, 328, 112–116. doi:10.4028/www.scientific.net/amm.328.112

[4] Srinivas, P. (2012). Comparative Analysis of Conventional Pid Controller and Fuzzy Controller with various Defuzzification Methods in a Three Tank Level Control System. International Journal of Information Technology, Control and Automation, 2(4), 75–86. doi:10.5121/ijitca.2012.2407.

[5] Liu, X., Zhang, H., Niu, Y., Zeng, D., Liu, J., Kong, X., & Lee, K. Y. (2020). Modeling of an ultra-supercritical boiler-turbine system with stacked denoising auto-encoder and long short-term memory network. Information Sciences, 525, 134–152. doi:10.1016/j.ins.2020.03.019.

[6] Mikhailenko, V. S., Shcherbinin V. A., Leshchenko, V. V. Kharchenko, R. Yu., N. V. Lozhechnikova (2020) Modeling the process of hazardous emissions formation in the exhaust gases of ship's steam boilers Informatics and Mathematical Methods in Simulation Vol. 10 (2020), No. 3-4, pp. 154-166. doi 10.15276/imms.v10.no 3-4.154.

[7] Mikhaylenko, V. S., Kharchenko, R. Y., & Shcherbinin, V. A. (2020). Analysis of the Predicting Neural Network Person Recognition System by Picture Image. Automatic Control and Computer Sciences, 54(3),249–258. doi:10.3103/s0146411620030037.

[8] Mikhailenko, V. S., & Kharchenko, R. Y. (2014). Analysis of traditional and neuro-fuzzy adaptive system of controlling the primary steam temperature in the direct flow

steam generators in thermal power stations. Automatic Control and Computer Sciences, 48(6),334–344. doi:10.3103/s0146411614060066.

[9] Kharchenko, R.Yu., & Kochetkov, O. V. (2020). Efficiency increasing of the ergatic control systems on ships. Trends in the development of modern scientific thought : The X th International scientific and practical conference, Vancouver, 23-26 November 2020 : proceedings. – Canada: 2020. - P. 682-685. DOI - 10.46299/ISG.2020.II.X

[10] Mikhailov, S. A. & Kharchenko, R. Y. (2020). Intelligent climate management in vessels (2020). Ships' electrical engineering, electronics and automation : Materials of the 10th international scientific and practical conference Odessa, 24.11.2020 – 25.11.2020 : proceedings. – Ukraine, NU «OMA»: 2020. – P. 217 – 219. DOI:10.31653/2706-7874. SEEEA-2020.11.1-245

[11] Kharchenko, R.Yu. (2020). Modern possibilities of intelligent climate control for enclosed spaces (2020). About the problems of science and practice, tasks and ways to solve them : The VI th International scientific and practical conference, Milan, 26-30 October 2020: Library of Congress Cataloging-in-Publication Data. – Italy, 2020. - P. 487-490. DOI - 10.46299/ISG.2020.II.VI

[12] Mikhailov, S. A., & Kharchenko, R. Y., (2019). Improving the efficiency of functioning the ship ergatic microclimate control systems of ship spaces speed. Automation of Ship Technical Facilities, 25(1), 63–71. doi:10.31653/1819-3293-2019-1-25-63-71.

[13] Mikhailov, S. A., & Kharchenko R. Yu. (2020). Intelligent climate management in vessels. Materials of the 10th international scientific and practical conference «Ships' electrical engineering, electronics and automation» 24.11.2020 – 25.11.2020, Odessa, Ukraine NU «OMA», p. 217 – 219. DOI:10.31653/2706-7874.SEEEA-2020.11.1-245

[14] Mikhailov, S. A., & Kharchenko, R. Y. (2019). Improving the efficiency of functioning the ship ergatic microclimate control systems of ship spaces speed. Automation of Ship Technical Facilities, 25(1), 63–71. doi:10.31653/1819-3293-2019-1-25-63-71.

Implementation of Recursive Deep Learning Algorithms for Natural Language Processing

Yana Daniiel [1], Kostiantyn Onyshchenko [2]

[1] *Kharkiv National University of Radio Electronics, Ave. 14, Kharkiv, 61166, Ukraine*
[2] *Kharkiv National University of Radio Electronics, Ave. 14, Kharkiv, 61166, Ukraine*

Abstract
The new model family introduced in this work is summarized under the term Recursive Deep Learning. The models in this family are variations and extensions of unsupervised and supervised recursive neural networks (RNNs) which generalize deep and feature learning ideas to hierarchical structures. The RNN models of this work obtain state of the art performance on paraphrase detection, sentiment analysis, relation classification, parsing, image-sentence mapping, and knowledge base completion, among other tasks.

Keywords
Natural language processing, recursive deep learning, recurrent neural network, text data, learning models, human language, dependency trees

1. Introduction

To date, information has become the unavoidable link among humankind. The amount of unstructured text data produced daily keeps growing, and humanity has faced a new challenge on how to intelligently process it for further usage. Development of learning models is intended to ensure automatic human language representation and extraction of different types of knowledge from it.

Natural language processing (NLP) is widely used for extracting information from different sources, grammatical analysis, or sentiment analysis of unstructured data [13]. The existing solutions, bases on different machine learning models, have a number of shortcomings. The goal in this work is to analyze these solutions and develop general and scalable algorithms that can solve such tasks.

2. Problem statement

The first shortcoming of some existing algorithms is Simplifying Language Assumption. The common first step of text classification requires the data to be formatted beforehand to be compatible with the algorithm. The word order and grammatical structure of sentences is ignored. The text is represented as unordered word list. This creates the comprehensive difficulties in understanding whether the overall meaning of a sentence is positive or negative. The other simplification example is linkage of the word to the previous word or a fixed group of words next to this word to define the context and speech tag.

The other shortcoming is Feature Representation. Modern learning systems have different features integrated which enhance text processing but slow down the runtime of the final algorithm. For instance, usage of part-of-speech tags, parsing tree features or the relationship of words in a large taxonomy (WordNet) and application of special labels for each named entity [1].

This work is dedicated to developing new deep models that mitigate the shortcomings illustrated above. The idea of the new model family is summarized under the term of Recursive Deep Learning. The unsupervised and supervised recursive neural networks will parse natural language and define the grammatical structure of a sentence. On this basis, the networks will align

EMAIL: yana.daniiel@nure.ua (A. 1);
kostiantyn.onyshchenko@nure.ua (A. 2)
ORCID: 0000-0002-3895-0744 (A. 1);
0000-0002-7746-4570 (A. 2)

their neural architecture. The recursion will apply the same neural network to every node of the grammatical structure. The presence of grammatic structure instead of breaking the sentence into 'bag of words' will allow to properly capture the semantic context [3].

Recursive application of the same neural network will allow to investigate how words compose the meaning in complex sentences. They will address the fundamental issue of learning feature vector representations for variable sized inputs without ignoring structure or word order [2]. Discovering this structure will help us to characterize the units of inputs and how they compose to form a meaningful whole. This is a prerequisite for a complete and plausible model of language.

3. Tree Structured Propositions

3.1. Dependency Tree RNNs - For Sentence-Image Mapping

This section is dedicated to introduction of a Dependency Tree RNNs model. The model learns to map sentences and images into a common embedding space. In this space, the model can retrieve one from the other. The model for mapping sentences into multimodal space is based on RNNs.

The dependency trees are the basis of DT-RNN model to learn vector representations for respective sentences [4]. The outputs of convolutional neural networks applied to images are mapped into the same space.

The multi-modal representations, illustrated on Figure 1 are the result of comparison of sentences and images.

This provides the ability to query images with a sentence and give sentence descriptions to images.

Compositional vector representations inside dependency trees is computed by the model. The compositional vectors computed by DT-RNN capture the meaning of sentences better comparing to the models based on constituency trees (CT-RNNs) or Recurrent Neural Networks, which are focused more on sentence's action and its agents. The meaning of sentence captured by RT-RNN is defined in terms of similarity to a "visual representation" of the textual description. The vectors induced by DT-RNN are more robust to changes in the syntactic structure [5].

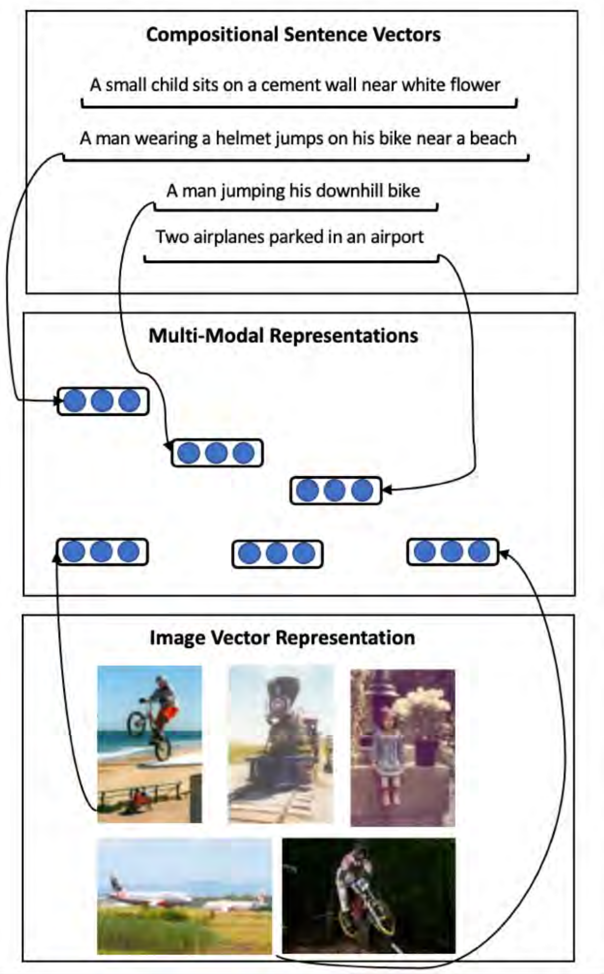


Figure 1: DT-RNN Multi-Modal Representations formation

3.2. Dependency-Tree Recursive Neural Networks

The representations for longer phrases from single word vectors can be built from simply linearly average the word vectors. This method is that it can provide reasonable performance in some tasks [7]. The disadvantage of this method is that the words are given equal weight and therefore the method cannot distinguish important differences in visual descriptions.

The RNN models combine word vectors for longer phrases [7]. The way of combination is not a simple averaging. The vectors are combined with an RNN in binary constituency trees. These trees have potentially several hidden layers [8].

One more issue is that induced vector representations capture a lot of syntactic structure of the sentence. The task of mapping images with sentence description requires the higher level of invariance to syntactic differences.

The model described below is focused on recognizing actions and agents.

3.2.1. DT-RNN Inputs: Word Vectors and Dependency Trees

The DT-RNN model can compute a vector representation for an ordered list of *m* words, represented by phrase or sentence. The single words are mapped to a vector space. The sentence is parsed from the vector space [9].

Each word is mapped to a *d*-dimensional vector. The word vectors are initialized with the unsupervised model. Both local and global contexts are used by the model to learn single word vector representation.

The *50*-dimensional vector is used in all experiments. The word embedding matrix X is used by finding the column index i of each word: $[w] = i$ and retrieving the corresponding column xw from X. Henceforth, an input sentence s is represented as an ordered list of (*word,vector*) pairs:

$$s = ((w_1, x_{w_1}), \ldots, (w_m, x_{w_m})) \quad (1)$$

3.2.2. Comparison to RNN Models

There are several important differences of DT-RNN from the other RNN models [10]. The RNN models are based on constituency trees (CT-RNNs) and use the standard composition function to compute a hidden parent vector h from exactly two child vectors $(c1,c2)$ in a binary tree:

$$h = f(W\,[c_1\ c_2]) \quad (2)$$

This can be rewritten to show the similarity to the DT-RNN as:

$$h = f(W_{l1}c_1 + W_{r1}c_2) \quad (2)$$

However, there are several important differences.

Firstly, in previously mentioned RNN models, the parent vectors had the same dimensionality. This allowed them to be recursively compatible. This also allowed them to be used as an input to the next composition.

The new model can map single words into a hidden space. The parent nodes can be composed from the hidden vectors. The new model provides higher capacity representation, which is effective for nodes with many children.

Secondly, the DT-RNN is applicable for *n*-ary nodes in the tree. This enhancement can be applied for constituency tree CT-RNNs [12].

Thirdly, previously mentioned models apply larger weight in the final sentence representation to the words merged last in a tree. This issue is related to computing parent nodes in constituency trees. On the contrary, the DT-RNN captures the most important effect by pushing the central content words to be merged last so its final sentence representation is more robust to less important modifiers [11].

Fourthly, DT-RNN allows some untying of weights. This depends on either how far away a constituent is from the current word or what its semantic relationship is.

4. Experimental results

A dataset of 1000 images, each with 5 sentences is used for the experiment. The Figure 2 illustrates the examples from the dataset of image-sentence mapping.

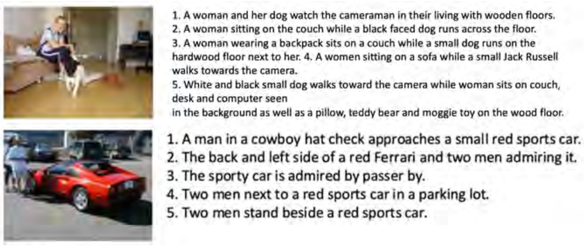


Figure 2: Examples from the dataset of images and their sentence descriptions

Three different experiments are conducted on DT-RNN model.

The first experiment is the analysis of quality of sentence vectors that capture similarity in visual meaning.

The second experiment is the analysis of Image Search with Query Sentences. Each model is queried with a sentence. The experiment is conducted to find an image showing the visual meaning of the sentence.

The third experiment relates to Describing Images by Finding Suitable Sentences. The reverse search is performed – the model is queried with an image. The experiment is conducted to find the closest textual description in the embedding space.

4.1. Similarity of Sentences Describing the Same Image

The experiment is conducted on 500 sentences from the test set. The sentences are mapped into the multi-modal space. Each sentence has its nearest neighbor found based on inner products. The sentences are sorted. The rank or position of the nearest sentence that describes the same image is recorded.

If all the images were unique and the visual descriptions close-paraphrases and consistent, I would expect a very low rank. However, usually a handful of images are quite similar and sentence descriptions can vary greatly in detail and specificity for the same image.

Table 1 shows comparison of methods for sentence similarity judgments. Lower numbers are better since they indicate that sentences describing the same image rank more highly (are closer). The ranks are out of the 500 sentences in the test set.

Table 1
Sentences Similarity for Image

Model	Mean Rank
Random	101.1
BoW	11.8
CT-RNN	15.8
Recurrent NN	18.5
kCCA	10.7
DT-RNN	11.1
SDT-RNN	**10.5**

Table 2 describes comparison of methods for image search with query sentences. Shown is the average rank of the single correct image that is being described.

Table 2
Image Search

Model	Mean Rank
Random	52.1
BoW	14.6
CT-RNN	16.1
Recurrent NN	19.2
kCCA	15.9
DT-RNN	13.6
SDT-RNN	**12.5**

Table 3 shows average rank of a correct sentence description for a query image.

Table 3
Describing Images

Model	Mean Rank
Random	92.1
BoW	21.1
CT-RNN	23.9
Recurrent NN	27.1
kCCA	18.0
DT-RNN	19.2
SDT-RNN	**16.9**

The Table 1 shows the results. Averaging the high-quality word vectors already captures a lot of similarity. The chain structure of a standard recurrent neural net performs worst since its representation is dominated by the last words in the sequence which may not be as important as earlier words.

4.2. Image Search with Query Sentences

The experiment allows to measure the quality of how the model finds images that display the visual meaning of a sentence. A query sentence is mapped into the vector space. The images are found in this space using simple inner products. The new DT-RNN has a higher performance comparing to the other models, see Table 1-3.

Figure 3 describes images and their sentence descriptions assigned by the DT-RNN. Green sentences are correct, red ones are close to the image vector but incorrect.

Figure 3: Images and their sentence descriptions assigned by the DT-RNN

5. Conclusions

In this work, a new family of recursive deep learning algorithms for natural language processing was introduced. The proposed recursive deep learning family allows to solve several natural language processing tasks, including the word and sentence-level predictions of continuous and discrete nature.

The variants of tree structures were explored in this work. The dependency trees were given to the RNN model as an input. The focus of the RNN models was solely put on the semantic content of a sentence and the prediction task. The dependency trees were explored as the underlying structure. The main action of a sentence was brought into focus for the final representation generated by DT-RNN. This has been particularly effective for grounding semantics by mapping sentences into a joint sentence-image vector space.

RNN models' potential, illustrated on dependency trees, is applicable for many downstream semantic tasks and can be computed with high efficiency rate. The DT-RNNs are not restricted to only binary trees and can process multiple children. The RNNs are also promising for sentence processing and 3d object classification.

One of the improvements to be implemented is the development of a general model of natural language processing. The model should jointly capture continuous, fuzzy, logical, and discrete nature of language. The captured data should be connected to the set of facts. Capturing the fuzzy nature of language allows to obtain an understanding of vague sentiment descriptions. This is also applicable for semantic parsing to retrieve information from knowledge bases.

The further area of improvements includes the provision of inter-sentence information flow for discourse structure understanding. For instance, solving concrete tasks as coherence and anaphora resolution. One more issue for deep models is first order logical reasoning. This may be applicable for retrieving the relevant information from datasets using natural language questions.

6. References

[1] Y. Bengio. Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2009.

[2] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. Object recognition as machine translation. ECCV, 2002

[3] D. Erhan, A. Courville, Y. Bengio, and P. Vincent. Why does unsupervised pre-training help deep learning? JMLR, 11, 2010

[4] K. Erk and S. Pad ́o. A structured vector space model for word meaning in context. EMNLP, 2008.

[5] P. Blunsom. K.M. Hermann. The role of syntax in vector space models of compositional semantics. ACL, 2013.

[6] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A.Y. Ng. Multimodal deep learning. ICML, 2011.

[7] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A.Y. Ng. Multimodal deep learning. ICML, 2011.

[8] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. Manning, A. Ng, and C. Potts. Recursive deep models for semantic compositionality over a sentiment treebank. EMNLP, 2013.

[9] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. IEEE, 86(11):2278–2324, 1998.

[10] R. Socher, C. D. Manning, and A. Y. Ng. Learning continuous phrase representations and syntactic parsing with recursive neural networks. In NIPS-2010 Deep Learning and Unsupervised Feature Learning Workshop.

[11] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A.Y. Ng. Multimodal deep learning. ICML, 2011.

[12] G. E. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Van-houcke, P. Nguyen, T. N. Sainath, and B. Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process. Mag., 29(6), 2012, p.82–97.

[13] I. Afanasieva, N. Golian, O. Hnatenko, Y. Daniiel, K. Onyshchenko. Data exchange model in the internet of things concept. Telecommunications and Radio Engineering 78 (10), 2019.

Subsystem for diagnostics of the reliability of pipeline systems and its probabilistic model

Irina Gavrilenko [1], Anatoliy Litvinov [2]

[1,2] O.M. Beketov National University of Urban Economy in Kharkiv, 17, Marshal Bazhanov Street, Kharkiv, 61002, Ukraine

Abstract

The issues of functional reliability assessment within the framework of the pipeline system reliability diagnostics subsystem are considered. The basis for the development of the functional reliability diagnostics subsystem is the breakdown of the pipeline system into emergency repair zones, as well as the replacement of the structure with the macro structure of emergency repair zones, which completely inherits the functional reliability of the system. On the basis of emergency repair zones probabilistic models for calculating the functional reliability of pipeline systems have been developed, which are used in the algorithm for calculating the reliability of the entire system. The reliability of pipeline systems at the operational stage is supported by timely prevention and elimination of failures. A probabilistic model has been developed that makes it possible to assess the characteristics of the process of prevention and elimination of failures, to select the optimal parameters. Corresponding numerical calculations have been carried out.

Keywords

Criterion, flow, model, pipeline, probability, reliability, service, system

1. Introduction

The reliability of pipeline systems can be assessed by various indicators, but among them two main can be distinguished two main ones that affect the functioning of the pipeline system and which we call functional reliability [1]:

– maintainability, which determines the property of the system to continuously transport the target product to consumers during repair and maintenance work;

– the probability of uninterrupted supply of the target product to a particular consumer or group of consumers over a certain period of time, which determines the objective possibility of the system to meet its purpose [2]. Increasing the functional reliability of pipeline systems is one of the most important tasks facing operators and in solving which all subjects of the transportation system are interested.

The reliability of pipeline systems is set at the design stage and is maintained at the stage of construction and commissioning. But over time, the reliability of individual elements of structures decreases, which leads to the need for operational diagnostics of the reliability of pipeline systems in order to make appropriate management decisions and measures to maintain reliability at a given level.

This can be implemented by developing a pipeline systems reliability diagnostic subsystem. It is part of the automated control system of the pipeline systems. In this subsystem, data on the state of individual elements of the pipeline systems is collected and functional reliability is calculated, using the appropriate methods and models.

Analysis of recent investigations has revealed that existing methods of calculating the reliability of pipeline networks cover the calculation of indicators that characterize their technical condition [2, 3, 4]. The study [5] is focused on the accuracy of hydraulic calculation for delivery and distribution of the target product. The issues of functional reliability of pipeline networks are either not considered at all, or concern only its assessment [6]. Therefore, new methods and models are needed for assessing the functional reliability of pipeline systems and taking them into account when designing an automated process control system.

EMAIL: i.gavrilenko@ukr.net (A. 1); litan6996@gmail.com (A. 2)
ORCID: 0000-0001-5296-4344 (A. 1); 0000-0001-7063-7814 (A. 2)

2. Formulation of the problem

Pipeline systems are durable, repairable systems. Their reliability is laid down at the design stage and is maintained through the implementation of preventive measures and thorough repairs, as well as repair and restoration work in case of various failures that occur at random times. Over time, the reliability of individual structural elements decreases, which leads to the need for prompt diagnostics of pipeline system reliability in order to make appropriate management decisions and measures to maintain reliability at a given level. This can be realized only by developing a diagnostic system, collects data on the state of individual pipeline systems elements and calculates functional reliability using appropriate methods and models, the development of which is the goal of this work.

3. Main part

In the general case, the automated system of reliability diagnostics (ASDR) should be a subsystem of the automated process control system of the whole pipeline system [7].

ASDR is designed to provide a single information space for timely and coordinated support for management decisions in the workplaces of technicians and repair personnel of structural units of the pipeline system.

Since the automated process control system of pipeline system is generally a multi-level system, the functional structure of ASDR is considered as a system to support management decisions at each level of management.

It includes the implementation of a sequence of the following steps: data entry; data analysis; data processing: development and selection of optimal plans; performance control.

Factory catalogs and technical passports of equipment, executive documentation, operating experience of the existing pipeline, retrospective statistical material, information coming from control sensors installed at critical points of the pipeline network are the actual basis for the development of the reliability diagnostics subsystem of pipeline system. The main goal of creating of the reliability diagnostic subsystem of pipeline system is :

– planning and control of execution of current, medium and major repairs;

– monitoring and forecasting of the technical condition of the pipeline and its individual elements;

– monitoring of pipeline resource and reservoir parks.

The basis for the development of the functional reliability diagnostics subsystem is the breakdown of the pipeline system into emergency repair zones, as well as the replacement of the structure with the macro structure of emergency repair zones (ERZ), which completely inherits the functional reliability of the system.

The basis for the development of the functional reliability diagnostics subsystem is the division of the pipeline system into emergency repair zones and the replacement of the structure with the macrostructure of emergency repair zones (ERZ), which completely inherits the functional reliability of the system. Such a division of the pipeline system is gradually reflected in the analytical method for calculating the functional reliability of the system.

According to it, seven consecutive stages must be carried out to obtain a mathematical model of the functional reliability of the pipeline network [8].

1. Formation of a mathematical model of a pipeline transport network with a complex topological structure in the form of a weighted graph.

2. Dividing the initial weighted graph of the pipeline network into subgraphs, each of which corresponds to one ERZ.

3. Calculation of technical reliability of ERZ as an independent macroelement in the operation of pipeline system.

4. Converting the initial weighted graph of a large-dimension network to a weighted macrograph ERZ of small-dimension.

5. Construction of a simplified ERZ macrograph in relation to a specific consumer of the pipeline network.

6. Construction of a calculation model of the functional reliability of the pipeline network for a specific consumer.

7. Formation of a mathematical model of functional reliability of the network relative to a particular consumer using the classical methods of the theory of reliability of technical systems and direct calculation of functional reliability.

A mathematical model is formed for each consumer O_i of the pipeline system, where i=1,2,…,K and is the total number of consumers. If several consumers receive the target product from only one zone, then the corresponding mathematical models of functional reliability coincide.

The following initial data are used to form a mathematical model of functional reliability in relation to the consumer O_i:

– calculation model of functional reliability in relation to the consumer O_i;

– weight function p at the vertices of the ERZ graph, which determines the technical reliability of each ERZ network;

– weight function pa on the edges of the graph ERZ, which determines the technical reliability of the shut-off valves of the network.

If the calculated model of functional reliability for an arbitrary consumer O_i consists only of series-connected and parallel-connected elements (no bridge connections), then the process of forming a mathematical model corresponds to the algorithm shown in the Figure 1.

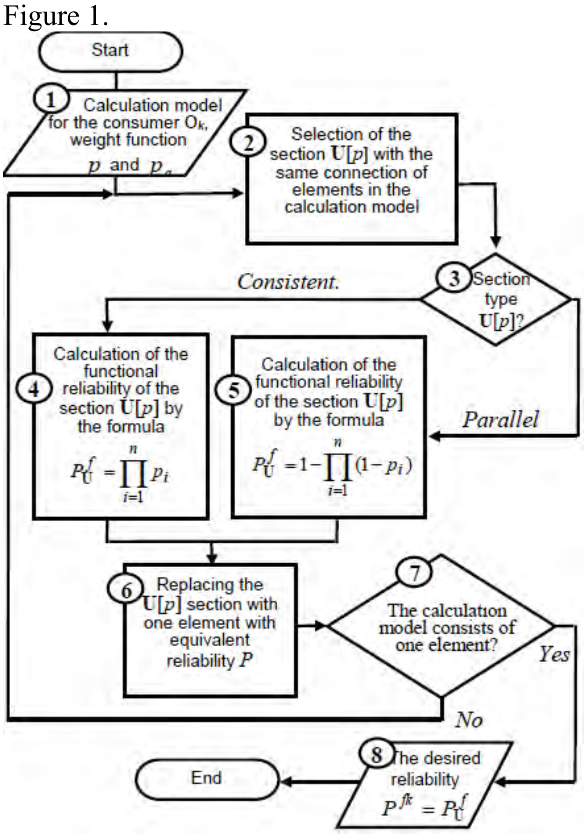


Figure 1: Scheme of the algorithm for building a mathematical model of functional reliability of the network in relation to a particular consumer

As follows from the algorithm, the formation of a mathematical model of the functional reliability of the network in relation to the consumer O_i is a cyclical process of replacement in the calculation model of sections U[p] with the same connection of elements by one element with equivalent reliability.

Equivalent reliability is calculated by the formula $P_U^f = \prod_{i=1}^{f} p_i$ (in the case of series connection of elements) or by the formula $P_U^f = 1 - \prod_{i=1}^{f}(1 - p_i)$ (in the case of parallel connection of elements). Here f is the number of elements in the same type of fragment; p_i-the probability of trouble-free operation of the element of the pipeline network corresponding to the i-th element of the fragment. The value p_i is selected according to the weight functions P and P_a of the ERZ graph.

The cyclic replacement process continues until the computational model consists of only one element. The calculated formula P_U^n for the reliability of this element will be the desired mathematical model of the functional reliability of the network in relation to the k-th consumer of the system:

$$P_i^f = P_U^f, i = 1,2,\dots,K \qquad (1)$$

The value P_i^n indicates the probability that the pipeline network provides the target product from the source to the k-th consumer of the system for a certain period of time (usually one year).

Under this example, the initial data are as follows:

– calculation models of functional reliability, which are determined in the previous stages;

– weight function p on the vertices of the ERZ graph;

– weight function p_a on the edges of the ERZ graph.

The work of the algorithm for constructing mathematical models in the example gives the following results (generated mathematical models):

– relating the first consumer O_1:

$$P_1^f = p_{a1}p_{a2} \times \left[1 - \begin{pmatrix} 1 - p_1 p_{a3} p_3 p_{a5} p_5 p_{a7} \times \\ p_{a9} p_8 p_{a16} p_{a11} p_{10} p_{a19} \times \\ p_{a13} p_9 p_{a18} p_{a12} p_7 p_{a15} \end{pmatrix} \times \right. \qquad (2)$$
$$\left. \times \left(1 - p_2 p_{a4} p_4 p_{a8} p_{a7} p_6 p_{a14}\right)\right] p_{a10}$$

– relating the second consumer O_2:

$$P_2^f = p_{a1}p_{a2}\left[1 - \left(1 - p_1 p_{a3} p_3 p_{a5} p_{a6}\right) \times \right. \qquad (3)$$
$$\left. \times \left(1 - p_2 p_{a4} p_4 p_{a8} p_{a7}\right)\right]$$

– relating the third consumer O_3:

$$P_3^f = p_{a1}p_{a2}\left[1 - \left(1 - p_1 p_{a3} p_3 p_{a5} \cdots p_{a9} p_{a8} p_{a16}\right) \times \right. ; \qquad (4)$$
$$\left. \times \left(1 - p_2 p_{a4} p_4 p_{a7} p_{a8} p_6 p_{a14} p_{a10} p_7 p_{a15} p_{a12} p_9\right)\right]$$

– relating the fourth consumer O_4:

$$P_4^f = p_{a1}p_{a2}\left[1-\left(1-p_1p_{a3}p_3p_{a6}p_{a5}p_5p_{a17}\right)\times\right.$$
$$\times\left(1-p_2p_{a4}p_4p_{a7}p_{a8}p_6p_{a14}p_{a10}p_7p_{a15}p_{a12}\times,\right. \quad (5)$$
$$\left.\left.\times p_9p_{a18}p_{a13}p_{10}p_{a19}p_{a11}p_8p_{a16}\right)\right]p_{a9}$$

where $p_{a1}, p_{a2}, \dots, p_{a19}$ – the value of the weight function p_a; $p_1, p_2, \dots, p_{10}$ – the value of the weight function p.

Calculations of functional reliability by mathematical models (2) – (5) in the example give the following values: $p_1^f = 0{,}9351$; $p_2^f = 0{,}9756$; $p_3^f = 0{,}9232$; $p_4^f = 0{,}9316$.

For the successful operation of the pipeline system reliability diagnostics subsystem, operational monitoring of the pipeline system condition, preventive and repair work, timely detection of critical areas and emergency areas are required. These works are carried out by emergency teams. In this regard, the task of calculating their optimal number arises. Since applications for repair work come in random moments of time, it is necessary to involve mathematical methods of queuing theory to perform the task of calculating the optimal number of repair crews.

From the database "Prehistory" it is possible to estimate the intensity of receipt of requests for call to repair crews λ and the average time of elimination of another accident τ. Let the number of repair crews n. Then the process of eliminating emergency situations on the pipeline system can be described by a multi-line queuing system. The total flow of service requests for repair requests at the entrance of a specialized organization is formed as a superposition of many service flows from emergency areas.

Therefore, according to the queuing theory it is Poisson (simpler) [9], in which the probability that during the τ time will arrive exactly k queries is described by Poisson's formula

$$P_k(\tau) = \frac{(\lambda\tau)^k}{k!}e^{-\lambda t}$$

and the distribution of time between the arrival of neighboring applications is exponential with the distribution function $A(t) = 1 - \exp(-\lambda t)$. The distribution of the time of liquidation of pipeline accidents can also be chosen exponential with the distribution function $B(t) = 1 - \exp(-\mu t)$, where $\mu = 1/\tau$ is the intensity of service requests. With the exponential distribution law, the whole system will rely on the most difficult mode of operation .

That is, we obtained the so-called Markov multilinear queuing system, the study of which was conducted in [10]. If we denote the workload of one team by $\rho = \lambda/\mu$, then its main characteristics of the developed model are as follows:

- the probability that all crews are free:

$$p_0 = \left[\sum_{i=0}^{n-1}\frac{\rho^i}{i!} + \frac{\rho^n}{(n-1)!(n-\rho)}\right]^{-1}, \frac{\rho}{n} < 1 \quad (6)$$

- the probability that all crews are busy and there are no queues for service:

$$p_n = \frac{\rho^n}{n!}\cdot\left[\sum_{i=0}^{n-1}\frac{\rho^i}{i!} + \frac{\rho^n}{(n-1)!(n-\rho)}\right]^{-1}, \frac{\rho}{n} < 1 \quad (7)$$

The average waiting time the start of service:

$$T_q = \frac{p_n}{\mu n(1-\frac{\rho}{n})}, \frac{\rho}{n} < 1 \quad (8)$$

Accordingly, the load factor of teams to service requests (as a percentage) can be represented as:

$$k_z = \left(1 - \frac{1}{n}\sum_{i=1}^{n-1}\frac{n-i}{i!}\rho^i \cdot p_0\right)\cdot 100\% \quad (9)$$

Having the basic indicators of system, it is possible to carry out tasks concerning optimum service of requests for a call of repair crews.

From the standpoint of a specialized organization that serves pipeline system, you need to minimize costs, while adhering to certain restrictions, namely: the deviation of the actual time of repair from the normative should not exceed the value , and all requests to eliminate accidents must be fulfilled. This means that the organization has a certain margin of capacity to service emergency calls.

Minimize costs within the developed model can be solely due to the number of teams n_{min}, aiming at their maximum loading. So, we got the mathematical programming problem, which is: to find n_{min}, which maximizes the function k_z when the constraints $T_q < T_{kr}$ and $\frac{\lambda}{(n\mu)} < 1$ are met.

The problem can be formulated differently: to find n_{min} which minimizes function T_q at performance of restrictions $k_z > k_{min}$ and $\frac{\lambda}{(n\mu)} < 1$ where k_{min} is the minimum level of loading of crews on service of accidents on pipeline system. Both problems can be solved by numerical optimization methods.

Given the low accuracy of the approximation of distribution laws, the tasks can also be solved graphically. Figure 2 shows graphs of the delay time of the beginning of repair (in hours) depending on the number of repair crews and the load per crew, with a normal repair time of 10 hours.

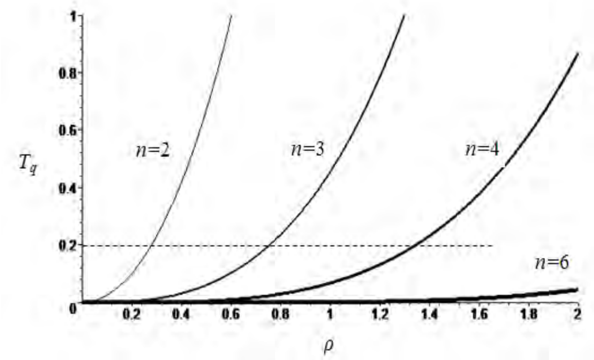


Figure 2. Graphs of delay time of the beginning of repair

4. Conclusions

The obtained mathematical models can be used not only to calculate the specific value of the indicator of functional reliability, but also to analyze them in order to further improve the design of the network. Only the general appearance of the model can lead to conclusions about the feasibility of new changes in the structure of the network. Thus, all models (2) – (5) show a weak point in the design of the network associated with the latches a_1 and a_2. If the failure of any section does not lead to the cessation of transportation of the target product to at least one consumer, the failure of the latch or leads to a general cessation of transportation of the target product in the network. According to model (2), if the latch a_{10} fails, the transportation of the target product to the consumer C_1 (and only to him) becomes impossible. Similarly, models (4) and (5) indicate that failure of the latch a_{13} or a_9, leads to the cessation of transportation of the target product accordingly to the consumer O_1 or O_4.

Therefore, the analysis of only the type of models (2) – (5) leads to the conclusion that it is necessary to preserve the latches a_1 and a_2, as well as the desired reservation of latches a_9, a_{10}, and a_{13}. In the latter case, to correct the situation, it is enough to attach to each latch in series another, the same.

The pipeline system system reliability diagnostics system and the ERZ method in particular bring the greatest benefit to the design of new pipeline system, which allows for a comparative analysis of alternative structures of the pipeline network by the criterion of functional reliability.

The use of queuing models allows you to minimize the operating costs of specialized organizations for the maintenance of pipeline systems while adhering to the specified constraints such as minimum time of delay of repair.

5. References

[1] M. I. Samoilenko, I. O. Gavrilenko, Functional Reliability of Pipeline Transportation Systems, Kharkiv, KHNUM named after O. M. Beketov, 184 p., 2009.

[2] Petrosov V.A. Management of regional water supply systems, Kharkov: Osnova, -- 320 p., 1999.

[3] E. M. Galperin, V. A. Zayko, P. A. Gorshkalev, Reliability Standards of Complex Engineering Systems. Conference Series: Materials Science and Engineering, 262 (1) : 012093, 2017.

[4] V. V. Rybalko, Mathematical models of reliability control of energy objects : St.-Petersburg, 151 p., 2010.

[5] N. Y. Yakhyaev, A. V. Korablin, Reliability Theory and Diagnostics basics (Moscow: Center Academy), 256 p., 2009.

[6] S. G. Konesev, R. T. Hazieva, Methods for assessing reliability of complex components and systems. Cur. Prob. of Sc.& Ed. Elec. Scien.Jour. 1., 2015.

[7] E. A. Bogdanov, Fundamentals of Oil and Gas Equipment Technical Diagnostics. Moscow, Higher School, 279 p., 2006.

[8] N. I. Samoylenko, I. A.Gavrilenko, T. S. Senchuk, “Creation of mathematical models for ordering graph edges of the pipeline distribution network”, Eastern-European Journal of Enterprise Technologies, Vol. 3, No. 4 (75), 2015, pp. 21–25. doi: 10.15587/1729-4061.2015.42811

[9] Sanjay K. Bose. An Introduction to Queueing Systems. – New York : Kluwer Academic/Plenum Publishers, 2002.

[10] J. F. Shortle, J. M. Thompson, D. Gross, C. M. Harris. Fundamentals of queueing theory. – Hoboken, USA: John Wiley and Sons, Inc, 548 p., 2018.

Intelligent measurement and control systems

Lala Bakirova [1]., Aladdin Bayramov [2]

[1] *Azerbaijan State Oil and Industry University, Azadlig avenue 20, Baku, AZ-1010, Azerbaijan*
[2] *«Azerbaijan Railways» Dilara Alieva 230, Baku, AZ-1010 Azerbaijan, CJSC, Department of İnfrastructure, Head of signaling and communication service, Azerbaijan*

Abstract
As the modern level of equipment and technology development enables to develop intelligent systems based on intelligent sensors and devices, a structural, functional and physical model of intelligent measurement and control system has been developed to safely determine the location and condition of the research object, taking into account various conditions and factors. Along with the measurement and processing of information, a software model of an intelligent system has been developed that allows the storage of information in a special purpose database and the management of the facility on the basis of the right decisions according to the state of the object, its interaction with other facilities.

Keywords
Intelligent device, database, control system, intelligent system, diagnostics, monitoring

1. Introduction

The use of information transmission and receiving modules in accordance with the parameters of the facility can be considered as a standard structural solution of remote measurement and control systems. Only the methods and means of obtaining information differ from various structural solutions. During the monitoring of the facility state, it is required to use manned and unmanned tools that allow to conduct remote research in order to collect information in a database, observe, analyze, evaluate and predict the effectiveness of the control. Moving facilities of the same or different status are used as an autonomous system that receives information from static and dynamic facilities and knows the location of itself and another facility in relation to itself. Thus, by performing the exchange of information, facilities can have control functions, while ensuring the security of themselves and other moving facilities. The information obtained in this case is stored in memory in the onboard system of the carrier and transmitted to the central control system at certain time intervals, at the request of the ground center. It is also considered to interfere in the processes taking place at the facility from the central remote control.

From this point of view, the diagnosis of problems based on the measurement of characteristic parameters for the monitoring of the facility itself and its interaction raises the development of a system of measurement-control and control from a wide range of functional capabilities for each facility.

A system of adaptive selection of interest based on the perception of confidential interactions between facilities using Gestalt psychology was applied. Due to selecting and sharing the appropriate interests, the observer can adapt to the changes around the robots and realize the situation around the working robot. Experience has shown that the proposed system allows to reduce the number of errors and operating time, as well as increase the probability of successful operation. However, the lack of use of the vision system and the lack of accurate information about the changing parameters manifests itself as a shortcoming of the system. In order to improve the remote control process, observer robots should not only provide visual

EMAIL: lala_bekirova@mail.ru (A. 1); bayramov-aladdin@mail.ru (A. 2);
ORCID: 0000-0003-0584-7916 (A. 1); 0000-0003-3714-0210 (A. 2)

support at appropriate angles, but also move smoothly at speeds in response to changes in the environment. Thus, speed control and adaptation to territorial changes are of great importance in terms of placement in real conditions [1].

Today, various solutions have been created in the field of Car-to-Car and Internet of Things (loT) forming the basis of wireless technologies. The paper provides a sample of the application of the Internet of Things on "Development of smart living quarters based on information measurement control systems." Existing pilot systems of thermal control and distributed data and control measurement systems are developed through dial-up lines, wireless network of GSM networks, global Internet channels and solve special problems related to situation monitoring and remote control of heating equipment points. Systen users will receive information from a single center, via mobile phone, television, voice control, as well as remotely through the site system. The system combined existing solutions and a set of features required from the Internet of Things, as well as new exciting technical solutions. The project "Smart system of living quarters" is a set of software and a number of modules required for the implementation of this technical device. The intelligent system of living quarters can be implemented both in the maximum configuration and in the form of separate modules. The system can be controled and monitored from any device and from anywhere with Internet access [2].

The presented system, in addition to detecting the facility at a certain distance, has the ability to evaluate its movement and functionality, compare the actual situation with previous data in the database, and, if necessary, perform self-control and control by making corrections. The results obtained and the functional capabilities of the system are intended to be applied to prevent accidents in terms of safety of people and facilities. The functionality of the system can be expanded on the basis of various additional technical capabilities to address higher security issues. In accordance with future requirements, the possibility of reworking hardware and technology, software is intended. All this can allow to eliminate the existing limitations, such as inconsistencies in the technical parameters of existing sensors, short measurement distances, environmental impact, based on the use of high-tech intelligent sensors and advanced software [3].

It was shown that it can be used to control a mobile pocket robot through simple Android and Arduino devices and by adding an Android application. There are other projects to control a mobile robot using an Android device, but the main purpose of this project is to show that open source platforms will be combined to create a platform for studying and teaching the program. The robot simulator obtains live video from an external camera for navigation and object detection, and the robot successfully finds the target and is able to visually determine the situation. The Bluetooth technology used for control makes it impossible to control after a certain distance due to the short distance, and no signal can be exchanged with the moving facility [4].

Some autonomous intelligent robots were started in various difficult situations. Remote-controlled robots are mainly used when control accuracy and precision are required. In 2015 DARPA robotics competition, most of the robots were semi-automatic. Autonomous operation has already been carried out in position control and route planning; however, general human control remains irreplaceable [6].

Various researchers worked on the increase in working efficiency by providing the necessary data to the operator to reduce operating time and errors. Errors are usually caused by distance uncertainty or blind spots (blockages caused by the location of the sensor or surrounding objects). For this reason, the remote-controlled robot is aimed to being provided additional data to to prevent blind spots and increase measurement accuracy [7].

Control systems are becoming more complex and confused due to the ability to distribute data processing across various devices exchanging information through industrial network systems. Therefore, at present a method is required to be determined for the design of intelligent distributed automated control systems. The proposed method requires a functional design model being research input to select the hardware architecture and evaluate the early reliability of the design. One of the most important pieces of equipment is industrial network systems, which are the nervous system of modern automation systems. Such network-related constraints as temporal compatibility and traffic characteristics are then analyzed and after that failure modes and causes and effects are identified for each level of industrial network systems. A number of prevention tools are introduced to prevent or minimize the harmful effects of failure modes, impact analysis and occurrence of failure modes. This data is beneficial in the design of automation systems and can be used in design methodologies based on functional division [8]. However, it should be noted that while the facility is moving the movement safety does not depend only on its technical characteristics. In this case, the state of the traffic line, internal and external factors, the level of regulation of traffic nodes, the state of other facilities in motion and likely to encounter, information processing, analysis and exchange for accurate evaluation of the state of the facility and other interacting facilities emerge as important issues.

2. Problem statement and the purpose of the work

Such issues as control and remote control of various parameters between static stationary facilities or continuous information exchange between sequentially moving facilities with the same status, minimizing the interval, increasing line capacity and security remain relevant today and make new ways of solution important to be found. It is necessary to ensure the exchange of information and security between these facilities, as well as between other moving facilities intersecting with them at the same level. Taking into account such main issues as development of two-way remote measurement and control system, adaptive iteration methods, error correction, partial and in some cases independent control function, self-education, linearization, a system enabling to protect and control the facilities is intended to be developed by determining the condition of the facility. Self-diagnostics of the intelligent system facilitates the rapid elimination of malfunctions by determining the location, cause of malfunctions, errors in the program, automatic self-checking and self-calibration of the measuring channel and determining the probability of failure of elements.

3. Basic information

The proposed system differs from the existing system with infrared and bluetooth information exchange in that the information exchange channel operates in a radio frequency band with a specific frequency and the control distance can be significantly increased. Selecting the specified diopos in real time increases the efficiency of the remote control. If moving objects with the same status are moving on the same line, if there is a problem with one of the objects, the objects send information to each other and to the center. The results of the experiment show the effectiveness of the proposed system in terms of execution time and number of errors.

In industry, there is a wide demand for the condition of the facility and the management of technical facilities, depending on the requirements for tracking vehicles, making decisions based on the results obtained by observing the condition of objects, measuring, controlling the parameters and comparing the given parameters. The condition of the object is constantly monitored by temperature, humidity, carbon monoxide, pressure, etc. can be controlled according to the indicators. In addition to the condition of the facility, it is possible to monitor its location and location using satellite communications using GSM, GPRS modules and sensors installed in local radio towers (active and passive). Different types of sensors can be used to measure these parameters.

The greater amount of data, the greater the accuracy of the information. Accuracy of information is the quality of information. In Figure 1, the temperature sensor T01 in the object is marked with the sensor P indicating the location of the object N01 for humidity (in addition, transmitter, receiver, GPS, GSM modules are supposed to be connected) and the nominal limit parameters T01n, N01n, P to the VBPM database, respectively is recorded in the software, and if the received signals go beyond the specified measurement parameters, the object is immediately alarmed by a visible signal indication and an audible signal, and at the same time it is possible to send information to other objects via wireless communication.

During the measurement process, a sensor or multi-variable sensor is connected to the input of the MK-microcontroller via the M-multiplexer. The T-timer determines the frequency of inspections. To collect all the data, an external storage device ESD is added to the system, and any information can be read here. The same system can be applied to objects with the same status, where the process from one point to another can be interrupted by an interrupt signal. Objects are constantly interacting with each other and can exchange information on the specified parameters. The advantage of the system is that it can operate as an autonomous system and has the ability to make decisions on issues such as speed reduction when the approach distance is short on the basis of written software. Here, the system has features such as an intelligent system - traffic safety, self-diagnostics, its inclusion in the system in the event of a fault on the trajectory, notification of the location of the fault to the local control point and other future object, the ability to make autonomous decisions. happens.IMCS

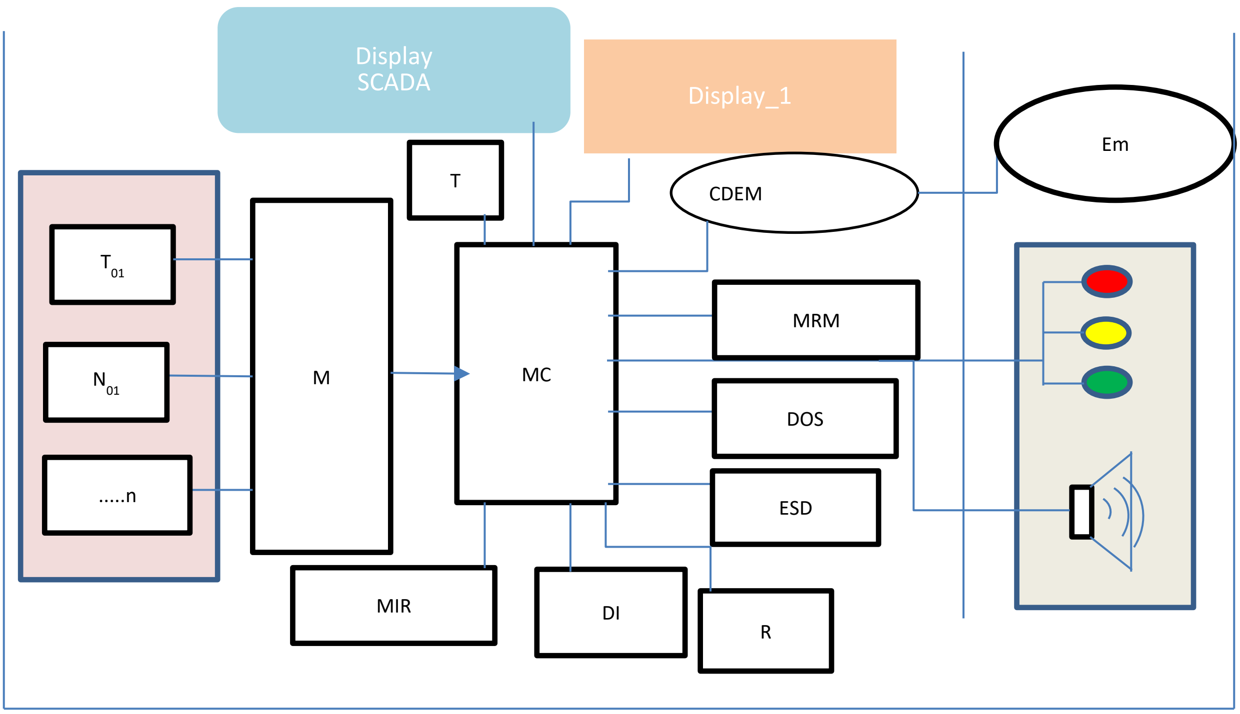


Figure 1: Structural model of the proposed IMCS module

At the same time, the objects can transmit and receive information about each other's condition remotely with the help of MRM - measuring receiver (antenna, module). The system in the object can send commands to the IMI - the control device of the executive mechanism, as well as to the environment in which the system is located, and thus to the IMIC in another object, thus remotely interfering with processes in another object. With the help of direct intervention (Dİ) in the remote control system, it is possible to control the operation of the valve, magnetic actuator, brakes, engine, relay, fire extinguishing system, respirator. Objects moving on the line with the same status can receive information from passive sensors, static objects installed on the edge of the line, get information about the coordinates of other objects and provide information about their location relative to them. The resulting intelligent system performs the following functions:

execution of measurement algorithm and obtaining measurement results, information processing, calibration of the measurement channel, linearization of the conversion characteristics of the measurement channels, scaling, calculation of measurement errors and presentation with measurement results, measurement channel correction, measurement statistical processing of results, etc.

In the block diagram reflecting the functional connections of the proposed intelligent measurement and control system (Figure 2), the Central Remote Measurement Control System - CRMCS, IMCS 1, 2 ..... n intelligent measurement and control modules placed in dynamic objects

(DO), Static measuring point 1, 2 ... .n –SMP is shown.

The structural model of the proposed IMCS module can be configured and used as both a receiver and a transmitter.

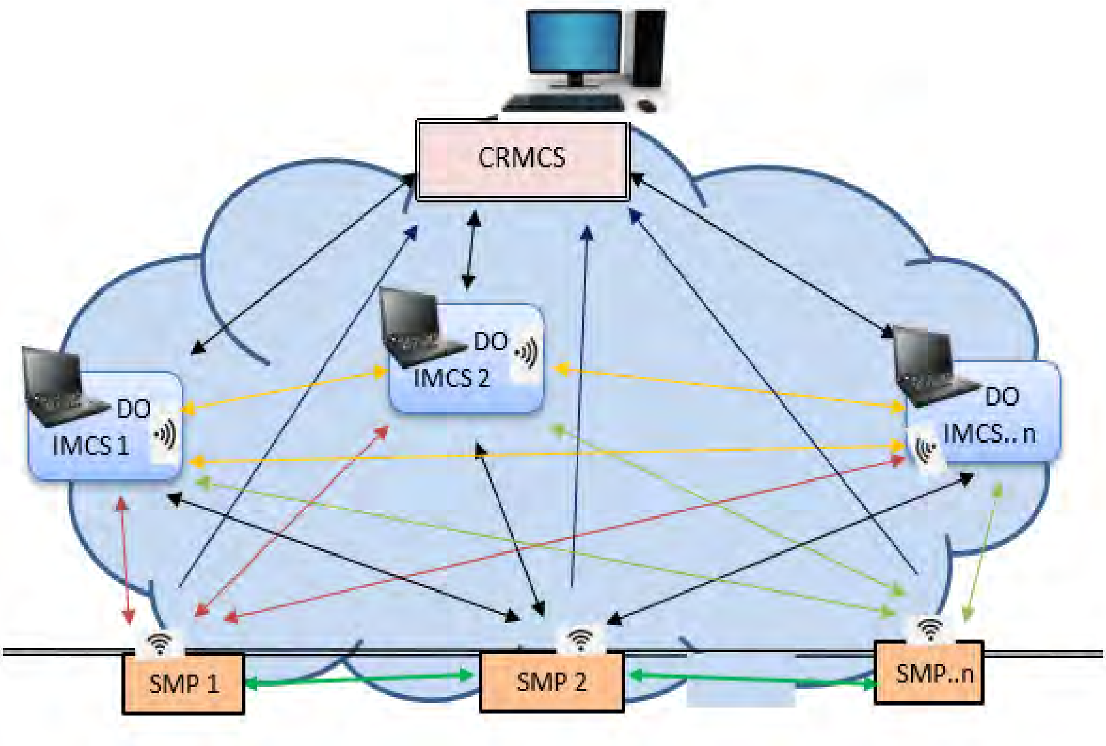


Figure 2: Block diagram showing the functional relationships of the proposed intelligent measurement and control system

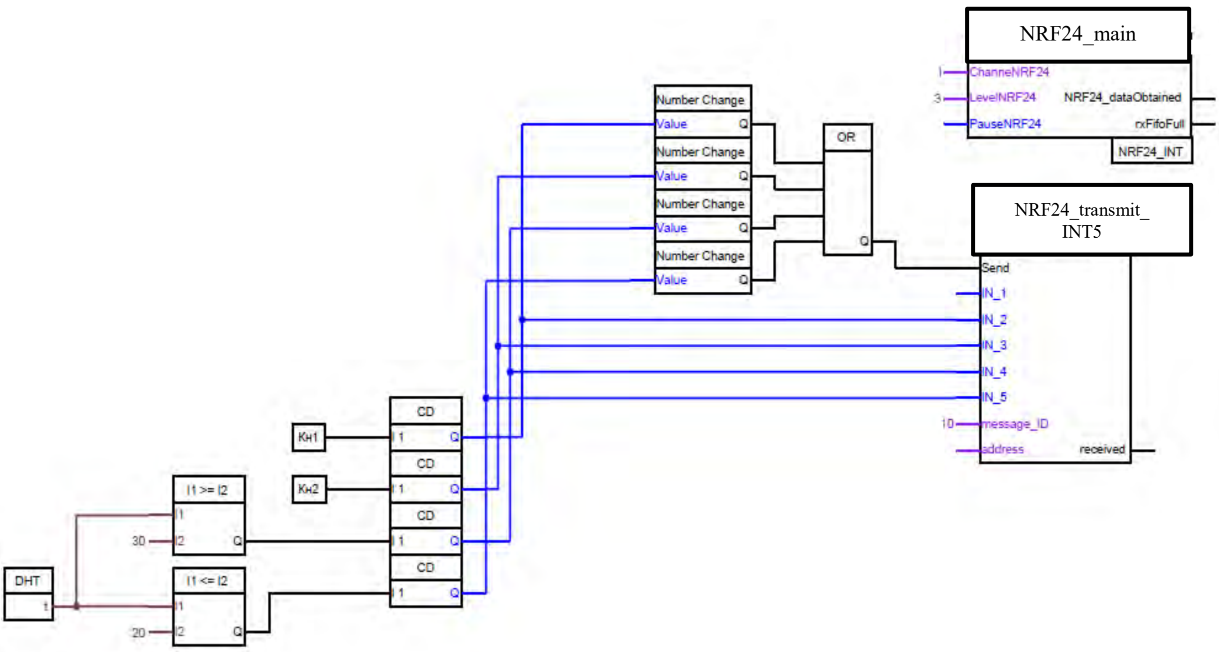


Figure 3: Functional block diagram (FBD) of the transmission module

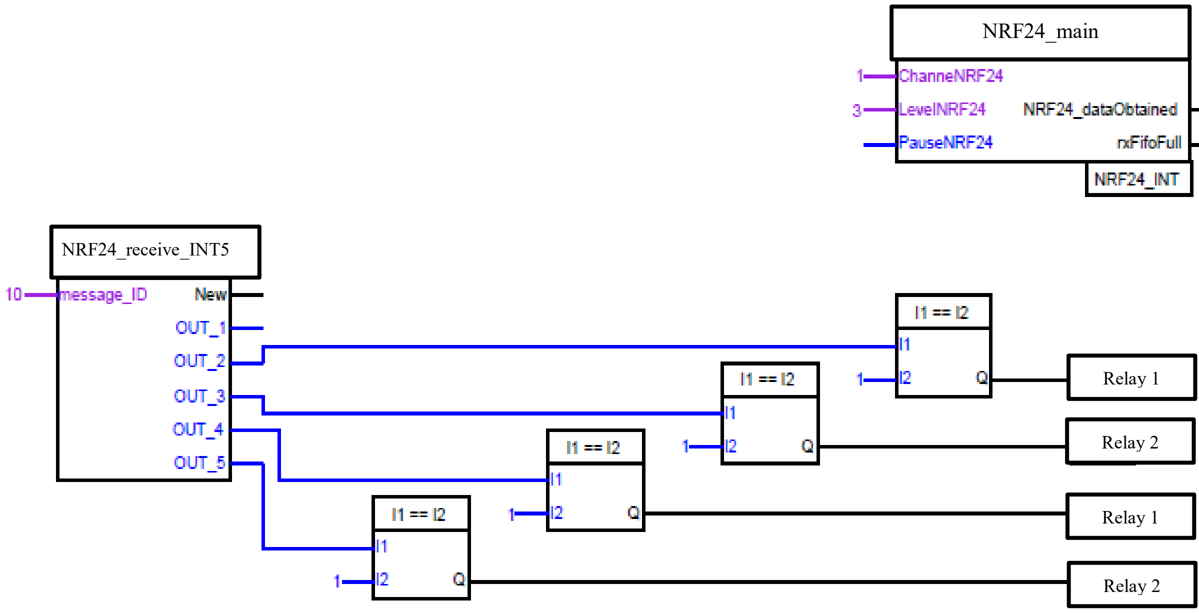


Figure 4: Functional block diagram (FBD) of the receiver module

In the functional block diagram (FBD) of the transmission module (Fig. 3), a DHT-11 sensor is connected as a transmitter to measure temperature and humidity. The sensor settings dialog shows the digital input to which the sensor is connected and the type of sensor. The comparison block compares the analog signal supplied to inputs I1 and I2 according to the selected condition, and if this condition is met, output Q yields a logical 1 (one) unit. The types of conditions are as follows:

1. Equal ("I1" = "I2");
2. I1 is greater than I2 ("I1"> "I2");
3. I1 is smaller than I2 ("I1" <"I2");
4. I1 greater than I2 or equal to I2 ("I1"> = "I2");
5. I1 can be less than I2 or equal to I2 ("I1" <= "I2").

In the considered sample, if I1> = I2, i.e. if the temperature is equal to or greater than 30^0 C, appropriate signal is sent to the receiver, and relay 1 of execution mechanism turns on a red light to indicate that the temperature exceeds the set norm. If I1 <= I2, it means that the temperature is equal to or less than 20^0C and the execution mechanism on the receiver side turns on relay 2 green light. Kn1 and Kn2 buttons are used to test and verify the operation of the system. It should be noted that the selected temperature can be changed according to the nature of any facility and the above-mentioned conditions can be selected. Input number - CD block is used to identify an input with a high level of logic. The output of the block will have a value corresponding to the input number. If there is no high level at any input of the block, the output will be 0. If there is a high level signal at several inputs, the block output will have a value corresponding to the highest output with the highest number. At the program stage, the number of inputs can be configured in the block editor.

The Number change detection block is used to detect the changes in the number supplied to the “EN” input. When the value at output “Q” changes, a pulse is generated over the length of a program cycle. For a block, it is possible to determine a dead zone separately for each direction of the figure change. These parameters are set in the block editor. There are no parameters in logic OR block. Block 1 is used to send information from nRF24L01 + to another. Several such blocks can be used in a project, but in order to work properly, message_ID must be unique throughout the network. Inputs: Send - (Send information) sends information at the top of the communication channel. IN_1-IN_5- sent numbers. message_ID - (0-255) - data packet identifier has to correspond to the message_ID of NRF24_Receiving_INT5 block on the receiving side for it to work properly and be unique throughout the network. Address - the address where the data set will be copied. In the “Device Address” parameter of block NRF24_Main (NRF24_Home), it is necessary to enter the number according to the rules. Outputs: received - confirmation of receiving the data packet by the addressee. A single impulse. If there are more than one device on the network with the same address, the signal is invalid.

The signal only says that the receiving nRF24L01 + module has received this data packet, but does not indicate that the controller has read and operated it. If 1 appears on the receiver side at the rxFifoFull output of the NRF24_Main block, or is always there, there is a high probability of data loss.

The receiver module (Figure 4) only works with the NRF24_Main and NRF24_INT blocks. Block 1 is used to receive int from one nRF24L01 + to another. Several such blocks can be used in a project, but in order to work properly, message_ID must be unique throughout the network. Inputs: message_ID - (0-255) - data packet identifier has to correspond to the message_ID of the NRF24_Receiving_string block has to correspond to the message_ID and has to be unique throughout the network. Outputs: numbers with data OUT_1 - OUT_5.

New - Updated message on OUT_1 - OUT_5 outputs. A single impulse. The comparison block operates under the conditions corresponding to the block on the transmitter. The controller's output block is used to control the physical output of the controller. There are two types of output: digital and analog (transverse pulse modulation). The receiver is programmed within the temperature parameters of the transmitter, and if the data obtained from the transmitter is above the specified temperature, a command is given to the relay with the execution mechanism and the red light is activated. If the temperature is below the set limit, the green light is switched on. In order to check the transmission and reception of signals in the system, relay 1 and relay 2 having execution mechanism on the receiver side, by pressing the buttons Kn1 and Kn2 on the transmitter side. By means of block NRF24_Main (NRF24_Home) on both sides, the frequency of the channel, the distance, its power depending on the distance are configured for data exchange. In addition, it is possible to solve any problem by attaching additional modules to the system. The physical model of the built-in system is shown in Figure 5.

Figure 5: Physical model of the system based on two modules

4. Conclusion

One of the important conditions of the control is taking into account the condition and movement status of other moving (dynamic) and stationary (static) facilities moving along a defined trajectory or freely, but in this case the facility shares these directions of movement along with its internal and technical condition. Obtaining physical parameters based on intelligent sensors and devices reflecting the internal condition and technical characteristics of a moving facility is not enough for intelligent control. Appropriate sensors and devices should be applied, as well as, the on-board systems of other facilities sharing the movement trajectory should regularly receive information from these sensors and devices to take into account the conditions and factors affecting the movement of the facility to increase traffic safety and reduce risks, and the on-board system of the moving facility. In order to make the right decisions, systems that can interact in terms of trajectory and territory should be able to exchange information about each other's internal situations and critical situations. In this case, it is possible to set up an on-board measurement and control system which allows to classify information about the internal condition, technical parameters of each facility, the state of the areas in which it is moving and along with other external factors, facility and facilities it may interact with according to various situational risk criteria and safe control with the right decisions on the base of the obtained results.

On the base of comparative analysis, free decision making, self-diagnostics and on-site correction of errors, exchange of important information with the on-board system of another object, hardware and software allowing autonomous decision-making is a characteristic feature of the created intelligent control system, taking into account not only the traffic safety of the controlled facility, but also traffic safety of the facility it is interacting with by adding other functions to these functions, the system can be used not only for traffic safety, but also for environmental protection, emergency information transmission and decision-making, especially for the safety of human life.

5. Reference

[1] Visual support system for remote control by adaptive ROI selection of monitoring robot. Shota Samejima, Khusniddin Fozilov and Kosuke Sekiyama. *ROBOMECH Journal* **volume 5**, Article number: 6 (2018) Published: 15 March 2018.

[2] Open source hardware and software platform for robotics and artificial intelligence applications Liang S Ng, K O Tan, Clement T H Lai, S K Ng, Mohammed A H Ali, Musa Mailah, Wan Azhar Yussof, Zamzuri Hamedon, Zulkifli Yussof. IOP Conf. Series: Materials Science and Engineering 114 (2016) 012142 doi:10.1088/1757-899X/114/1/012142.

[3] Object detection using ultrasonic sensor. Arun Francis G, Arulselvan M, Elangkumaran P, Keerthivarman S, Vijaya Kumar J. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6S, April 2019.

[4] Piasco N, Marzat J, Sanfourche M (2016) Collaborative localization and formation flying using distributed stereo-vision. In: 2016 IEEE international conference on robotics and automation (ICRA), Stockholm, Sweden, pp 1202–1207.

[5] Detection of Lock on Radar System Based on Ultrasonic US 100 Sensor And Arduino Uno R3 With Image Processing GUI. F Baskoro, B R Reynaldo. IOP Conf. Series: Materials Science and Engineering 336 (2018) 012016 doi:10.1088/1757-899X/336/1/012016

[6] Lim J, Lee I, Shim I, Jung H, Joe HM, Bae H, Sim O, Oh J, Jung T, Shin S et al (2017) Robot system of DRC-HUBO+ and control strategy of team KAIST in DARPA robotics challenge finals. J Field Robot 34(4):802–829

[7] Barnes DP, Counsell MS (1999) Haptic communication for remote mobile manipulator robot operations. In: American Nuclear Society, proc. 8th topical meeting on robotics & remote systems. Citeseer

[8] Design of intelligent distributed control systems: A dependability point of view Laurent Cauffriez1*, Joseph Ciccotelli2, Blaise Conrard3, Mireille Bayart4 & the members of the working-group CIAME5. L. Cauffriez, J. Ciccotelli, B. Conrard, M. Bayart & al. / Journal of Reliability Engineering and System Safety, pp. 19-32, Vol. 84/1, April 2004, Elsevier, ISSN 0951-8320

Module of Current Control and Forecasting of IT Project Management Tasks

Yulia Zhmaeva [1], Serhii Udovenko [2], Larysa Chala [3] and Vadim Shergin [4]

[1, 3, 4] *Kharkiv National University of Radio Electronics 1, Nauky ave. 14, Kharkiv, 61166, Ukraine*
[2] *Simon Kuznets Kharkiv National University of Economics 2, Nauky ave. 9-A, Kharkiv, 61166, Ukraine*

Abstract

An approach to the current control of the tasks to be implemented in the management of IT projects is proposed. This approach involves the implementation of a software module that analyzes the relationship between total and work performed during the iteration (sprint), based on the use of Agile methodology. This module allows you to perform the current correction of the chart by expanding the input parameters to increase the accuracy of the prediction of the characteristics defined as the weighted average of the results of previous sprints.

Keywords

IT project, Agile, task diagram, software module

1. Introduction

Flexible methodologies for creating software are focused on the use of dynamic formation of requirements and ensuring their implementation as a result of permanent interaction of working groups consisting of specialists of different profiles. Such methodologies as extreme programming, DSDM (Dynamic Systems Development Method), Scrum, FDD (Feature driven development) have become widespread. These methods are usually based on minimizing risks by reducing the development process to a series of short iterations (sprints). Each iteration should contain the tasks necessary for the issuance of a mini increase in functionality: planning, analysis of requirements for quality, as well as and requirements for duration of design. At the end of each iteration, the design team reevaluates the development priorities. The task diagram is the most common tool for applying the Agile methodology, which reflects the dynamics of design and is used to monitor progress on any iteration of the sprint [1].

Modern IT project management systems (including Jira and RTC) support the automatic generation of development progress charts. Jira is a commercial bug tracking system designed to interact with users. The main element of accounting in this system is the task (ticket), which contains the project name, topic, type, priority, components and content. The set of tasks forms a database for constructing a diagram. The Rational Team Concept (RTC) system provides a schedule of development progress and manages the software development lifecycle, providing contextual teamwork for distributed teams in real time. Coherent processes based on the use of these systems help to improve the quality of software developed, but not all the factors influencing the schedule are taken into consideration. In addition, the diagram shows the overall speed of the team, but often teams consist of specialists of different profiles (analysts, developers, managers, testers etc.). In this case, it is unclear in which part of the process there are difficulties. Thereby the chart needs to be improved by expanding the input parameters to increase the accuracy of the results of monitoring its implementation.

The purpose of this work is creating and program implementation of the module for monitoring execution of IT project tasks, which analyzes the relationship between planned and performed work during the iteration (sprint), based on the use of Agile methodology [2, 3].

When developing the module, both standard indicators characterizing human resources (working hours, days, iteration duration, overtime

EMAIL: yuliia.zhmaieva@nure.ua (A. 1); udovenkosg@gmail.com (A. 2); larysa.chala@nure.ua (A. 3); vadim.shergin@nure.ua (A. 4)
ORCID: 0000-0002-7396-9280 (A. 1); 0000-0001-5945-8647 (A. 2); 0000-0002-9890-4790 (A. 3); 0000-0002-4388-8180 (A. 4)

hours, degree of participation of a team member in the project, maximum team speed) and additional indicators to increase the clarity of performances results (focus factor, the amount of unplanned workload, the amount of work remaining, the difference between the estimates of task performance) should be taken into account.

2. Task performance chart creating

The task performance chart should contain a set of parameters that determine the trend for plotting. This chart consists of a few characteristics used in Agile methodologies. The reference trend presented in Fig. 1, considers the speed of the team for each calendar week and distributes the landmark accordingly [4]. The reference trend curve reflects the amount of work that the team must perform at each checkpoint in order to complete all tasks on time.

Figure 1: An example of a landmark chart

In the first and last calendar weeks of the sprint, the expected amount of work performed is lower than in the second, as the sprint begins and ends in the middle of the week. Indicators of the reference point are estimated in man-hours, i.e. in units of working time, which corresponds to the hour of work of one person [5]. This allows us to estimate the work of planning more accurately, comparing the number of employees and deadlines. The reference point is determined by the formula:

$$E = \{E_{start}, E_{week1}, \dots, E_{weekN}\}, \qquad (1)$$

where E_{start} – estimated scope of work at the beginning of the iteration; $E_{week1}, \dots, E_{weekN}$ – scopes of work at control points.

The components in (1) are defined as follows:

$$E_{start} = S + AVG\Delta_{start,} \qquad (2)$$

where S – scope of work during the iteration; $AVG\Delta_{start}$ – the scope of unplanned work at the beginning of the iteration;

$$E_{weekN} = \frac{E_{weekN-1} - E_{start} \times C_{weekN}}{C_{sprint}}, \qquad (3)$$

where C_{weekN} – team speed during the week N; C_{sprint} – team speed during the sprint.

Each task scheduled for execution for a certain period (sprint) is evaluated in ideal hours. This parameter allows us to consider all processes needed to calculate the resource use of the project at the stage of sprint planning. Time of the tasks summarizes, forming a Team Capacity metric (capacity, scope of work):

$$C_{period} = 8 \times WD * EN - V - SN + O, \qquad (4)$$

where WD – number of working days; EN – number of project participants; V – number of vacations days; SN – time spent on sick leave; O – overtime hours.

Thus, when calculating the scope of work, many indicators are considered [6]. Thereby, the team is given more accurate indicator of timestamps.

In addition, it is necessary to use the trend scope that reflects the actual performance of tasks.

Each task indicates how much time has already been spent on its implementation (spent) and how much is left to spend (remaining). The sum of all remaining indicators is determined by the formula:

$$S = \sum RV, \qquad (5)$$

where RV – number of man-hours.

In the classical form, such a model is not indicative enough. If the team is guided only by the current state of the team-board, the team's performance will not be fully reflected, because of the team-board that contains the task for the current sprint does not reflect a number of unscheduled activities to be performed (e.g. debugging from expert comments on code quality control and concomitant regression testing) [7]. The executant also may deviate from the assessments determined previously. Thus, the work scope by the end of the sprint is usually greater than initially.

If the team does not consider the forecast of the scope of unplanned tasks presented in Fig. 2, it may lead to risks of loss of product quality.

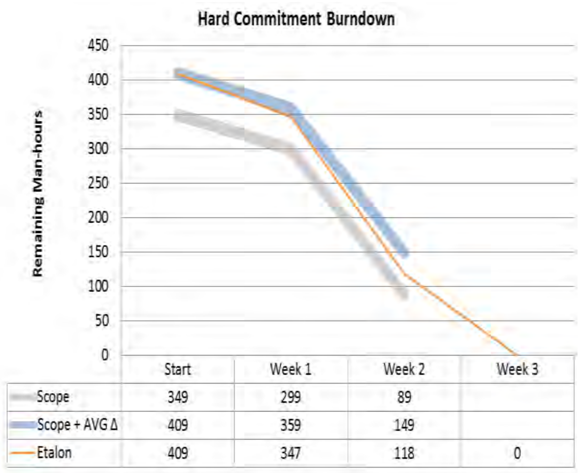


Figure 2: The scope of unplanned work (normalized)

The indicator of the normalized scope of the remaining work, which includes a forecast for unplanned tasks and exceeding the estimates, is determined as follows:

$$S + AVG\Delta = S * UTC * EDC, \qquad (6)$$

where UTC – number of unplanned tasks; EDC – measure of excess the estimates.

The forecast of unplanned tasks and excess of estimations is carried out for calculation of the most exact expected result of sprint. Modern error tracking systems do not consider these indicators, which can significantly affect the work of the team. If the forecast for the unplanned scope of work is not taken into account, it may contribute to the risks that will entail: reduced quality of the product being developed; excessive increase in the assessment of the required resources, which in turn will affect the financial component of the project; non-compliance with Agile methodology.

The projected performance characteristics of the IT project are defined as a weighted average of the results of previous sprints. The weight of each sprint is determined by the consolidated assessment and adjusted by a set of adaptive rules. This indicator includes the weight of all unplanned tasks. The forecast of unplanned scope of tasks is determined by the formula:

$$US_{forecast} = \frac{S \times (US_{AVG} - US_{current})}{100\%}, \qquad (7)$$

where US_{AVG} – weighted running average of the sprints results; $US_{current}$ – scope of the unplanned works during the current sprint.

Scope of the unplanned works is determined as:

$$US_i = \frac{\sum RV_i}{S + AVG\Delta_i \times 100\%}, \qquad (8)$$

where $\sum RV_i$ – number of unplanned tasks in current sprint.

The weighted running average of the sprints results can be calculated as:

$$US_{AVG} = \frac{US_i \times USW_i}{\sum USW_i}, \qquad (9)$$

where USW_i – coefficient of unplanned tasks.

The second indicator that significantly affects the forecast is the difference between the initial assessment of the task and the actual time spent.

The sum of these indicators reflects the excess of the initial total estimate. The procedure for adjusting the impact factor is determined for a specific sprint individually by the following criteria: determination of the most anomalous exceedances of the assessment; identification of the reasons for the excess; definition of preventive steps.

The forecast of exceeding the estimates is determined by the formula:

$$ED_{forecast} = \frac{(S + US_{forecast}) \times ED_{AVG}}{100\%}, \qquad (10)$$

where ED_{AVG} – the weighted running average of the previous sprints.

The error in the estimate in the current sprint is determined by the formula:

$$ED_i = \left(\frac{\sum SV + RV - OV}{S_i}\right) \times 100\%, \qquad (11)$$

where SV – spent time; RV – remaining time; OV – scheduled time.

The deviation from the estimate in the current sprint is determined by the formula:

$$ED_{AVG} = \frac{ED_i \times EDW_i}{\sum EDW_i}, \qquad (12)$$

where EDW_i – the coefficient of excess of estimates.

Based on the comparison between the reference and actual trends, the main indicators of team effectiveness are calculated [8]. The first indicator shows the difference between the planned and executed scope of work, i.e. how far the team lags the schedule or ahead of it (considering the forecast for deviations).

This parameter (in man-hours) is determined by the following sum:

$$LAG = S + AVG\Delta - E. \quad (13)$$

If the team performs the task uniformly, following the plan, the deviation from the schedule will be zero. If the lag parameter (13) is negative, it signals that the team is ahead of schedule. If it is positive, it may indicate that the team does not have time to complete the work on time. This situation can be explained by the following reasons: increasing the complexity of one or more user stories; reduction of resources; errors in the initial assessment at the sprint planning stage; finding a large number of defects after the testing phase; increasing the amount of related functionality that was involved at the time of development; lack of experience and knowledge in the team to solve problems.

This indicator is necessary for the team to eliminate risks at the stage of its formulation. The causes of risks can be identified using indicators such as focus factor and the number of man-hours in the current sprint. However, these metrics allow the team to monitor the execution of tasks in the sprint, but do not reflect the cost-effectiveness of the team's time resources.

It is advisable to use the so-called focus factor – an indicator that shows how much time the team spends on the direct implementation of planned tasks. This factor can be noised by various noises (e.g. the inability or unwillingness of the team to record the time spent at each stage of development).

The focus factor is calculated by the formula:

$$FF = \frac{WD}{C} \times 100\%, \quad (14)$$

where WD – the number of man-hours that were spent; C – team speed.

Usually the lower allowable value of the command focus factor in the Agile methodology is taken to be 75%. This means that the team must spend more than 75 percent of the time on the team board to successfully implement the IT project (Fig. 3).

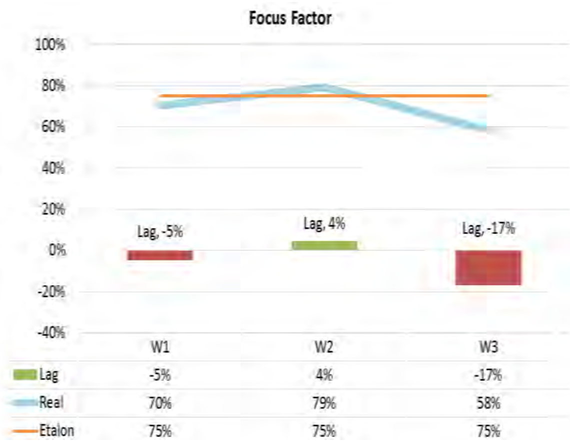


Figure 3: Example of graphical visualization of command focus factor

To identify the cause of the low command focus factor most accurately, this factor should be considered for developers, testers, and other team members separately. For example, the overall focus-factor may be less than 75% due to insufficient efficiency of business analysts. In this case, the manager will pinpoint the problem that is causing the lag and eliminate it. An additional indicator that reflects the effectiveness of the team is the speed of execution of user stories. This parameter should be considered considering the maximum, reference and actual team speed. The maximum speed that the team can count on corresponds to the full involvement of third-party developers, testers and other project participants. The reference benchmark considers only current resources (excluding underemployment or the absence of some team members). The actual speed is one in the current sprint, which includes sick leave, vacations, overtime, etc.

3. Module implementation

The process of "Sprint Management" is the most important in the module. The brief description of this process is given in table 1.

In addition, the module provides the following options: registration and authorization, adding new data (users, task cards, teams, sprints); data deletion; task board management. Use cases are designed to adjust the functional requirements for the management of IT project development processes. This allows contractors, as well as the customer and potential users to assess in real time the prospects of the project.

Table 1
Characteristics of the business process "Sprint Management"

Business process	Sprint management
The main participants	Manager, performers
Incoming event	Project order
Resulting event	Actual allocation of resources in the project
Source documents	Adjusted task diagram
Management	Specification of requirements for the implementation of the IT project
Business process client	IT project customer

A description of the use-case of the "Creating sprints" is given in Table 2.

Table 2
Use-case "Creating sprints"

Characteristic	Value
Context of use	Creating a new sprint every week
Precondition	The manager has administrator rights
Trigger	The manager wants to create a new or parallel sprint
The script	Log in to the system with administrator rights; go to the "Iterations" tab; click "New"; enter the necessary data; click "Create"

If the use case is successful, a new sprint will be created.

The description of the "Sprint Management" use-case is shown in Table 3.

The description of the Task Board Management use case is shown in Table 4.

Table 3
Use-case "Sprint Management"

Characteristic	Value
Context of use	Control sprints, add new commands or tasks, end sprints
Precondition	The manager has administrator rights
Trigger	According to the results of the analysis of indicators, it is necessary to change, complete or delete the sprint
The script	Select one of the created iterations (sprints) and click "Edit" / "Destroy"; if "Edit" is selected make changes and click "Save"

Table 4
Use-case «Task Board Management»

Characteristic	Value
Context of use	Changing the position of the task cards on the board (transfer to the categories "Must be completed", "In progress", "Completed")
Precondition	The user must be registered as a manager in the system
Trigger	The status of the task has changed, it must be assigned to the appropriate category
The script	Log in to the system; go to the "Iterations" tab; choose the right sprint; if necessary, add a card (click "Add card"); For changing the position of the card, select it and drag to the desired position

If the use case is successful, the task will be moved to the board, and if the task is assigned to the "completed" state, the diagram will show the real time of the task, as well as the difference between the declared and real time.

The storyboard of the options for using the module (graphical representation of the location of its functions on the module site) is presented in Figures 4-10.

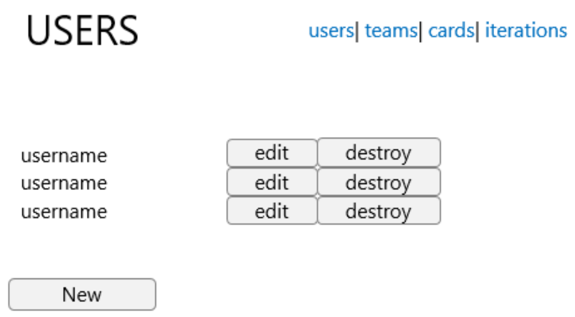


Figure 4: Home page and user management page

Figure 5: Team management page

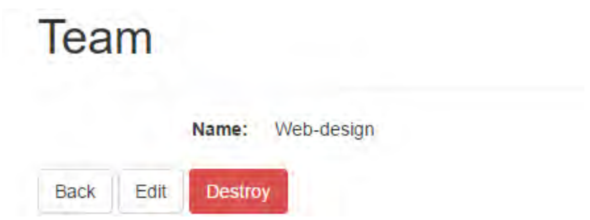


Figure 6: Team information

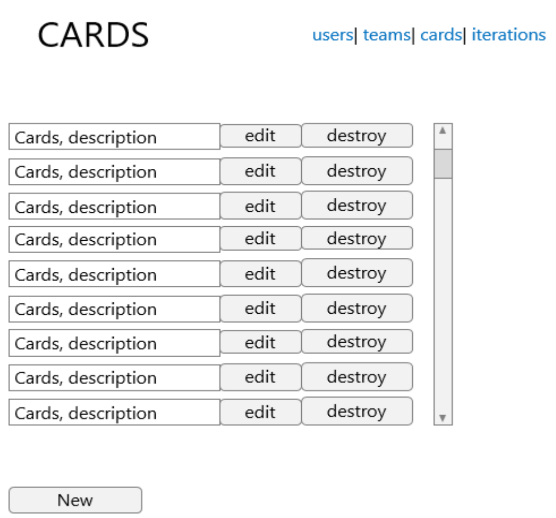


Figure 7: Task card management page

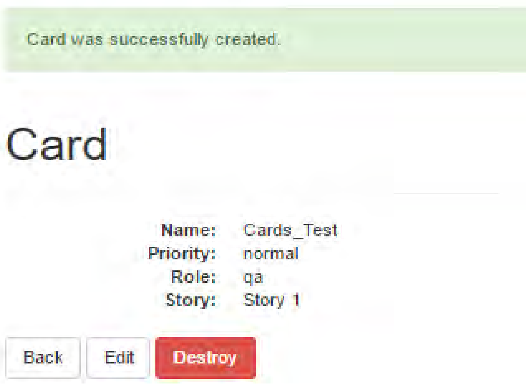


Figure 8: Card information (opens after creation)

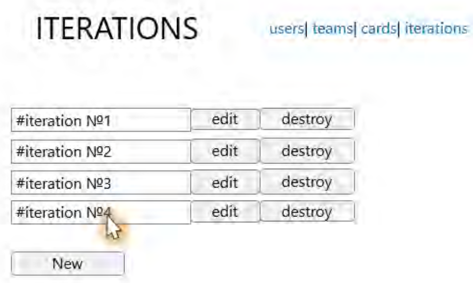


Figure 9: Sprint management page

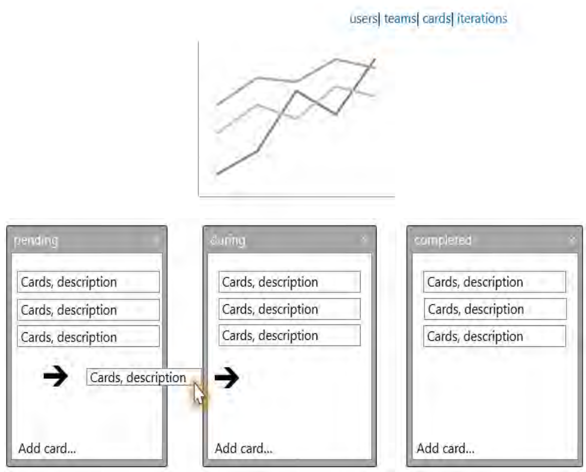


Figure 10: A page with a task board and a schedule of time spent

During the testing process, the functions of placing cards on the board were checked, as well as adding new cards and changing their status. To add a new card to the sprint, click "Add card" and specify the required fields to fill, after which the card will be added to the board. To change the status of the task, simply hold the cursor on the desired card and drag the card to the column with the current status.

The module is hosted and deployed on free Heroku hosting. To successfully deploy the application, the user needs Internet access, a web browser, and a host address.

The hardware and operating system (MUC OS / Windows 7/8/10 / Linux OS), as well as the Internet browser (Google Chrome / Opera / Mozilla / MS edge / Internet explorer) do not affect the quality of the module can be selected by personal user preferences.

When working with the module, the user is given the opportunity to use different forms to fill in data: the number of team members, role and employment, number of working days, dates of checkpoints and more. The "Graphics" page displays an adjusted task performance diagram, in the construction of which the attributes and parameters described above are involved (Fig. 11).

Figure 11: Example of an adjusted task performance diagram

In Fig. 11 the ordinate axis shows the number of man-hours, the abscissa axis - the control points that were determined when creating the sprint according to the working weeks (the first control point is the first day of the sprint, the fourth – is the last one). The user is given the opportunity to adjust the control points to monitor performance at any stage of the sprint.

The task chart consists of four graphs. The bottom graph shows the actual execution of tasks. This figure corresponds to the speed with which the team performs project tasks according to the board. The reference benchmark is presented on the average graph, which shows the scope of work the team must complete at each checkpoint to have time to complete all tasks on time. The reference trend takes into account the speed of the team for each calendar week and distributes the benchmark accordingly. The upper schedule corresponds to the normalized scope of work remaining (it includes a forecast for unplanned tasks and exceeding estimates).

The Lag indicator, represented by rectangles, makes it possible to determine how far the team is lagging behind or ahead of the schedule, taking into account the forecast for deviations.

The user can to save the schedule in different formats: PNG, JPEG, PDF, SVG. Thus, the document can be formed.

4. Conclusions

Modern IT project management systems should provide a forward-looking way to coordinate the stages of software development and information technology. In particular, the promising methodology is dynamic formation of the task diagram using the Agile.

The task diagram helps the teams of performers to self-organize during the sprint. If this chart shows that the development process is behind schedule, team members will act appropriately to prevent risks. For example, a team can reallocate tasks to identify bottlenecks by expanding resources to speed up the process.

The proposed module allows to improve the chart. Predicted characteristics are defined as weighted average results of previous sprints. The weight of a specific iteration is determined by a consolidated score, which includes the weight of all unscheduled tasks.

The hardware and operating system of the module, as well as the Internet browser can be chosen according to personal preferences of users, because they do not have a significant impact on the quality of the module.

A promising direction for future development of this approach is the expansion of its functionality based on the use of templates MVC and Spring Framework, which provides solutions to many problems for creating IT projects using Java.

5. References

[1] M. Cohn, Succeeding with Agile: Software Development, Washington, Pearson Education, 2009, 114 p.

[2] R. Martin, Agile Software Development, Principles, Patterns, and Practices, New York, Pearson Education, 2002, 458 p.

[3] K. S. Rubin, Essential Scrum: A Practical Guide to the Most Popular Agile Process, New York, Addison-Wesley Professional, 2012, 281 p.

[4] A. Maurya, Running Lean: Iterate from Plan A to a Plan That Works, Washington, O'Reilly Media, 2012. – 129 p.

[5] E. Ries, The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses, Washington, Crown Business, 2011, 90 p.

[6] A. Croll, Lean Analytics: Use Data to Build a Better Startup Faster, Washington, O'Reilly Media, 2013, 370 p.

[7] J. Rasmusson, Agile Samurai: How Agile Masters Deliver Great Software, Washington, Pragmatic Bookshelf, 2010, 237 p.

[8] Yu.V. Zhmayeva, S.G. Udovenko, L.E. Chalaya, Adaptivnoye prognozirovaniye optimal'nogo kolichestva resursov IT proyekta po metodologii Agile, Avtomatizirovannyye sistemy upravleniya i pribory avtomatiki, (2015), 173, 4 - 13.

THE SYSTEM CONCEPT OF CLIMATE MANAGEMENT

Nataliya Fil [1] , Leonid Nefedov[2]

[1,2] *Kharkiv National Automobile and Highway University, Yaroslava Mudrogo St 25, Kharkiv, Ukraine, 61000*

Abstract

The typical climate management system and climate monitoring and assessment systems were analyzed. A conceptual model of climate management has been developed in the form of a "black box", which reflects the interaction of the initial information, informative, resource, and methodological support for obtaining the output data.

Keywords

climate, risk, climate management system, conceptual model, monitoring, information management.

1. Introduction

In recent years, the entire world community and Governments around the world have begun to reflect on the climate change impact on various human activities, national infrastructure, and productive activities.

The climate system is a complex system whose behavior is determined by the interaction between the atmosphere, the ocean, the cryosphere, the biosphere, and the active terrestrial layer [1].

Changes in average climatic conditions cause variations in the frequency, intensity, spatial coverage, duration, and timing of extreme weather and climate events. All of these changes will affect climate in the future. Various extreme events (hurricanes, storm surges, floods, droughts, and heatwaves), as well as changes in the regimes of certain climatic systems, such as monsoons, on a limited spatial and temporal scale, may have a stronger effect on the performance of information technical systems of different countries. One of the most pronounced trends is the increase in the frequency and intensity of heavy rains. The year 2021 is a further confirmation of this. In mid-July, Western European countries received record rainfall, with rains in Germany flooding the Rhine highlands, destroying thousands of German homes and killing at least 179 people. Then the floods came to the southern lands of Germany, Switzerland, and Austria, reached the border with the Czech Republic. Iran and Kazakhstan are suffering from the drought. In the central province of Henan (China), floods caused by the typhoon killed 69 people, rains flooded the subway and eroded the highways, and hundreds of thousands of people were affected by evacuation. Experts argue that the summer of 2021 is the point at which humankind, after years of ignoring global warming, has faced the consequences of climate change [2]

According to the United Nation (UN) estimates, there are a number of sectoral risks [2].

One of the main causes of the warming observed is considered to be the increase in atmospheric concentrations of greenhouse gases, which absorb the thermal radiation reflected by the Earth surface and thereby increase the accumulation of heat in the Earth's system.

The increasing concentration of greenhouse gases significantly raises the risks of change associated with freshwater resources. Climate change is an uneven process, with warming occurring more rapidly in circumpolar regions than at the equator.

Renewable sources of surface water and groundwater resources will decline in the 21st century. These trends are likely to accelerate in the future: e. g., forecasts of precipitation in the Eastern Mediterranean during the decade 2020–2029 suggest that it could decline by up to 25 % compared to 1990–1999. Snow cover in the Northern Hemisphere also tends to decrease. In some regions, such as the Alps and Scandinavia, snow depths are steadily decreasing in the lowlands, but increasing in the mountains, while in other mountain areas (such as the Carpathians, the Pyrenees, and the Caucasus), no pattern can be observed [3].

As a consequence of global climate change, it is natural disasters that will become the major component of the threat list in the near future. The UN estimates that most of Europe will soon experience a significant increase in the frequency of flooding, from once every 100 years to once every 5–15 years [4].

Shortly, accidents and natural disasters will increasingly threaten the world economy, population, and balanced development.

Many Governments have developed plans for the safe and sustainable development of their countries. But emergencies related to climate change increase the risks of human casualties, especially for underdeveloped countries.

In 2021, countries began to recover their economies from the COVID-19 crash.

The UN Secretary-General has proposed a number of climate-related activities [5].

First of all, it is the transition to a "green economy", i.e., the transition to a less carbon-intensive economy [5].

Ukraine is part of the global system to combat climate change. The Law of Ukraine «On Principles of Monitoring, Reporting and Verification of Greenhouse Gas Emissions» came into force in Ukraine on 1 January, 2021 [6]. The adoption of that instrument was an important step in the fight against climate change, based on the experience of the European Union.

Ukraine adopts the best European practices based on the European Union legal framework. As early as the beginning of the 2000s, a market for greenhouse gas emission allowance trading (EU ETS) was created under the EU directive [6]. This is one of the most effective market mechanisms for reducing harmful emissions into the atmosphere. Thus, in 2019, the European Union has reduced their volume by 3.8% [6].

Reducing climate risk within the framework of purely national interests cannot be fully achieved; it is a global challenge.

2. Literature review

There is much research on climate change issue. Thus, it has become known that the average annual temperature has increased by 0.3–0.6ºC over the last hundred years, and in Europe by about 1.2ºC [5].

Work [6] notes that there is a rapid trend towards significant climate change in Ukraine today. The temperature change resulted in a discrepancy in the existing regionalization of Ukraine.

Climate change, increasing variability of weather conditions, and extreme weather events on the territory of Ukraine negatively affect the health, working capacity, comfort, and life expectancy of people [7].

A possible recurrence scenario of comfortable and uncomfortable weather situations for cold and warm periods in the year in the Kyiv region by the middle of XXI in conditions of further climate changes is proposed. Bio-meteorological data are effective indicators of the thermal structure of the environment surrounding humans, especially under climate change conditions. Their assessments are necessary for application in various areas.

Work [8] indicates that generalization and evaluation of data on climate change consequences, available in the scientific literature for natural and socio-economic systems, on human health is a relevant scientific challenge. Specific methodological developments are needed to address this problem correctly.

The problem of climate change requires the modernization of legislation on environmental protection and energy conservation. Work [9] reviews best practices in modernizing the European Union legislation and policies on climate protection and energy conservation.

3. Aim of the work

The present work aims to develop a conceptual model for a climate management system that will allow eliciting the interaction of the initial information, informative, resource, and methodological support for obtaining the necessary initial information.

4. Development of a model for climate management

As abovementioned, the reduction of climate risk within the framework of purely national interests is not possible to a full extent and is a global challenge. Accordingly, the issue has been given increased attention in various international events, such as symposiums, summits, conferences, etc.

Each of the countries participating in such events is committed to making its full-valued contribution to climate risk reduction, taking into account its national interests.

The development of a modern concept for climate risk reduction is based on the idea of preventing and redressing damage to the environment, health, and property of citizens.

Climate risk reduction takes place at different levels (Fig. 1). The global level of climate management involves predicting and monitoring the processes and conditions of the planetary ecosystem and its component subsystems.

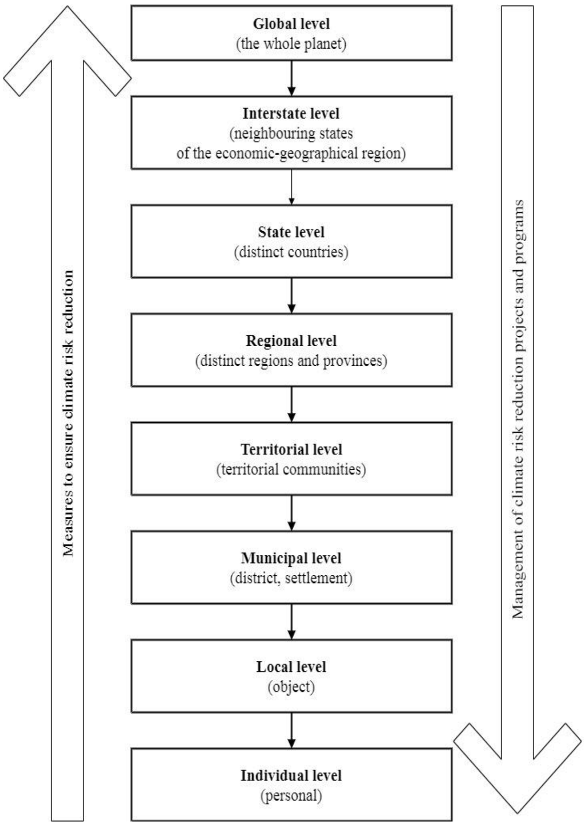


Figure 1: Climate management levels

It is above all the monitoring and management of the main factors characterizing the climate crisis. Global climate management is the prerogative of international organizations: the United Nations, UNESCO, UNEP, and others.

Key steps in addressing global climate change and reducing climate risk are the UN Periodic Conference on Environment and Development. Management methods include the adoption of international acts on environmental protection, the implementation of interstate climate programs, and the creation of intergovernmental forces to eliminate environmental disasters □10□.

Thus, at the global level, climate management has addressed a series of international climate challenges, including the introduction of international Red Books on Biodiversity, the adoption of a declaration to ban the production of ozone-depleting substances, and the adoption of the Kyoto Protocol.

The interstate, regional, and territorial levels of climate management are oriented towards solving the problems of ensuring the climate risk reduction of large geographical or economic zones.

This level of climate management includes the following:

- □ greening the economy;
- □ development and introduction of new environmentally friendly technologies;
- □ maintaining an economic growth rate that does not impede the restoration of environmental quality and contributes to environmental management [11].

At the municipal level, climate management is carried out by local governments. They develop and implement appropriate spatial plans, detect facilities that pose a particular risk to climate change, and regularly inform citizens about the climatic situation.

Ministries and agencies are responsible for developing projects and implementing state climate management programs; maintaining state records and monitoring the activities of facilities that pose a special threat to climate change; identification of factors, objects, and zones of increased climatic risk; carrying out operational measures in case of threat of occurrence or natural dangerous situations.

Climate management of industrial production at the local level is based on the assessment of climate hazards of individual industrial facilities. An industrial facility is defined as a separate industrial site, industrial establishment, or a group of industrial units, which may be considered as the sole source of technological impact.

Comprehensive and effective monitoring should be an integral part of the climate management mechanism. At the same time, monitoring should cover all parts of the system – from economic entities to government agencies.

The main principles of climate management are:

- □ development of climate policy taking into account the peculiarities of existing technology;
- □ making decisions aimed at reducing climate risks;
- □ organization of control over all stages of the technological process (from a continuous determination of air and water quality to a selective biological analysis of the most sensitive species of fauna or local population)

and environmental monitoring in the facility area.

In this regard, priority tasks for climate management can be identified as:

☐ conservation of raw materials and energy;

☐ minimization of waste and environmental pollution;

☐ assessing the degree of climate risk and costs associated with the proposed activity for further estimation of possible financial investments;

☐ informing the public and the population about the nature of organization production activity and environmental nature state in the area where production is located.

In order to implement these principles, it is necessary to apply adequate methods to achieve high results in economic practice and significantly reduce anthropogenic climatic impacts on the environment.

Technology of strategic climate management is graphically presented in Fig. 2.

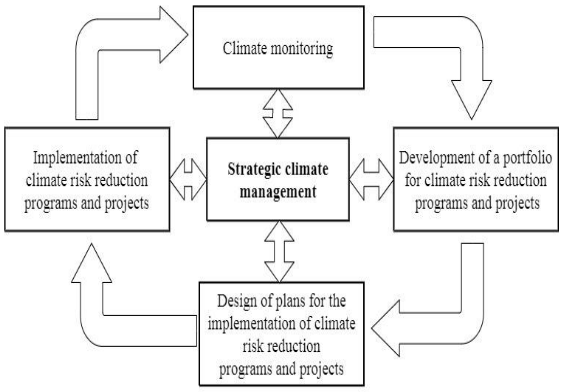


Figure 2: Technology of strategic climate management

One of the important elements of climate management is climate monitoring.

Conceptually, a climate management system (CMS) can be conceived as a "black box" (Fig. 3), which is characterized by inputs (initial information) and outputs (CMS output), informative, resource (bottom arrows), and methodological support (arrows on top).

The structure of the task of synthesizing a climate monitoring system consists of the following main sub-tasks: determination of climate control points; selection of climate monitoring indicators at control points; choice of measurement tools; determination of location of climate data collection devices; determination of a control point list for each of the climate data collection devices; choice of climate data collection devices in each location.

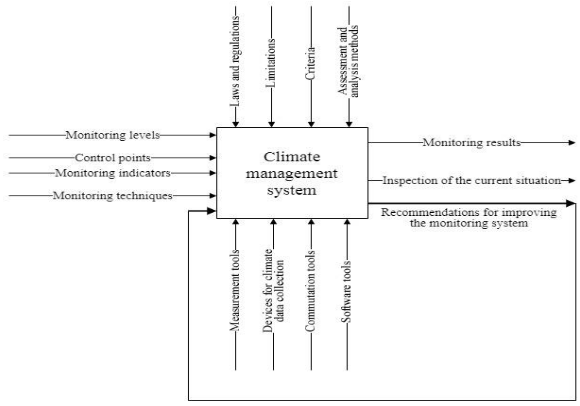


Figure 3: Conceptual model of climate management system in the form of a "black box"

The tasks of CMS synthesis include the organization and planning of CMS.

Let's assume that the locations of installation of automated workstations dispatcher and server are already known.

The tasks of structural, topological, and parametric synthesis are as follows. The problem of structure choice for a climate management monitoring system can be represented in such a way: considering that sets of typical climate data collection devices, as well as their functions and characteristics, are specified, it should be determined the number of levels of the monitoring system, set of elements at each level, and their interrelationships.

The task of selecting a topology consists in determining the location and connection between the system elements. The problem of parametric synthesis is to choose the functional characteristics of the elements, subsystems and connections. The problem is solved for specific structural, topological and technological characteristics of the system. In addition, the procedure for synthesizing a climate management system as a mandatory step includes the analysis of the monitoring results and inspection of the current situation. The climate management system should provide information support for solving the problems of synthesizing the climate monitoring system.

The climate management system should use computational intelligence technologies, mathematical models of observed natural phenomena and be based on formal structural descriptions of multifaceted tasks of climate management. Methods of system analysis and general system theory, statistical analysis, discrete optimization with boolean variables, neural networks, fuzzy set theory, and cluster analysis methods should be used in solving the assigned tasks.

5. Conclusions

The paper analyzes a climate management system. Based on the analysis, the conceptual model of the climate management system is proposed, which allows determining the input and output information, the sequence of tasks for the synthesis of the climate management system.

The conceptual model of climate management is presented in the form of a "black box" reflecting the interaction of initial information, informative, resources, and methodological support to obtain the necessary output information. The created conceptual model of the climate management system allows developing a multilevel hierarchical climate management system from a single system position. Further research will focus on building a climate management system in the form of intellectual information technology. The use of a climate management system will make it possible to take into account the specificities of the problem area.

6. References

[1] Climate Risk Assessment and Management/ 2021. URL: https://www.adaptationcommunity.net/climate-risk-assessment-management.

[2] N. Kondratyev Flooded Western Europe, Russia and China (July 28, 2021)/ Electronic periodical "Novaya Gazeta". Moscow, Russia. URL: https://novayagazeta.ru/articles/2021/07/27/zatopilo-zapadnuiu-evropu-rossiiu-i-kitai.

[3] A.O. Kokorin Climate Change: A Review of the State of Scientific Knowledge on Anthropogenic Climate Change. Moscow: World Wide Fund for Nature, 2005. P. 20. URL: https://www.elibrary.ru/download/elibrary_26449350_29291665.pdf.

[4] The Role of Hydrometeorological Services in Disaster Risk Management. Proceedings from the joint workshop coorganized by: the World Bank, the United Nations International Strategy for Disaster Reduction, and the World Meteorological Organization. Washington, D.C. (March 12, 2012). URL: https://www.unisdr.org/files/27645_webrestheroleofhydromet.pdf.

[5] Goal 13: Take urgent action to combat climate change and its impacts. URL: https://www.un.org/sustainabledevelopment/climate-change.

[6] Monitoring of greenhouse gas emissions begins in Ukraine. (January 04, 2021). URL: https://www.ukrinform.ru/rubric-economy/3165873-v-ukraine-nacinaut-monitoring-po-vybrosam-parnikovyh-gazov.html.

[7] S. G.Boychenko, O. G Zabarna. Estimation of comfort of weather conditions and trends of their changes for the Kyiv region in the conditions of climate change. Geofizicheskiy Zhurnal 41(6) (2019) 128–143. https://doi.org/10.24028/gzh.0203-3100.v41i6.2019.190071.

[8] G.E. Insarov, C.L.Méndez, S.M.Semenov, P.Z. Yanda Concept of risk and visualization of its changes in the reports of the intergovernmental panel on climate change. Fundamental and Applied Climatology 2 (2020) 100-119. doi:10.21513/2410-8758-2020-2-06-34.

[9] Kremer, L. Policy and law of climate protection in the European Union / L. Kremer // Modernization of the legislation of the European Union on climate protection and energy conservation: Sat. scientific. tr. / RAS. IHP. Sector of Environmental and Legal Research, INION. Center for social. scientific-inform. issled. Department of Jurisprudence; Resp. ed. Dubovik O. L., Alferova E. V. (2014) 10 – 42.

[10] Assessing the adaptive capacity of the Finnish environment and society under a changing climate. URL: https://www.syke.fi/en-US/Research__Development/Research_and_development_ projects/Projects/Assessing_the_adaptive_capacity_of_the_Finnish_environment_ and_society_under_a_changing_climate_FINADAPT.

[11] Demyanenko A. F. Development of the project of environmental management systems in hierarchical organizational structures. Science and technology of transport: scientific, technical and production journal (2003) 74-79.

Analysis of inconsistencies in publications of functional requirements in the form of mathematical expressions

M.V. Yevlanov [1], I.Yu. Panforova [2]

[1,2] *Kharkiv National University of Radio Electronics, Nauki ave. 14, Kharkiv, 61166, Ukraine*

Abstract

Functional requirements descriptions are the first descriptions of information systems that are formed during the life cycle of these systems. Therefore, all subsequent delays in the creation of an information system are in one way or another associated with errors made during the publication of functional requirements for the system. One of these delays is the correction of the information system, caused by the need to eliminate the detected duplication of functional requirements. Therefore, there is a constant need to eliminate duplication of individual functions of the information system even at the stage of formation and analysis of functional requirements for this system. The task is complicated by the fact that in a number of cases scenarios for fulfilling functional requirements are mathematical models or expressions. However, the existing functional requirements analysis methods are not intended for the analysis of such publications.

In the course of the study, it was proposed to use an improved method for synthesizing descriptions of architecture options for the information system being created to search for duplicated functional requirements published in the form of mathematical expressions. This method allows to compare knowledge-based models of publications of functional requirements for an information system. To test the efficiency of the method, it was proposed to use the functional requirements for the task "Formation and maintenance of an individual plan for a scientific and pedagogical employee of the department." These requirements contain a number of scenarios for calculating the desired parameters of an individual plan.

The studies carried out have shown that the usual publications of functional requirements in the form of mathematical expressions do not allow identifying situations of duplication of such requirements. The reason for this in the study is proposed to consider excessive detailing of mathematical expressions, which are publications of functional requirements. Recommendations have been developed and experimentally tested to eliminate redundant detailing of functional requirements publications.

Keywords

Functional requirement, requirements analysis, inconsistencies of requirements.

1. Introduction

According to modern concepts of the life cycle of systems [1], many requirements for the system are chronologically the first and basic representation of this system. Therefore, it is the requirements for the system that should contain all the information necessary for the creation and operation of the system throughout the entire life cycle. The experience of IT project management shows: elimination of mistakes made during the formation of requirements requires the highest costs of all works of a similar purpose [2]. This is especially pronounced during the implementation of IT projects for the creation or modification of information systems (IS).

To identify errors in system requirements, it is proposed to evaluate requirements publications against a variety of requirements characteristics. Thus, [1] states that a “good” requirement must be necessary, freely realizable, consistent, meaningful, complete, reflect the specifics, feasible, traceable, verifiable, acceptable and have boundaries. This approach allows to assess the degree of quality of requirements for the system as the degree of approximation of the characteristics of each specific requirement to the optimal or acceptable values.

However, in practice, it is difficult to apply these characteristics to assess the quality of IP requirements. The main reason for this difficulty should be considered the lack of formal descriptions of these characteristics. It is this

EMAIL: maksym.ievlanov@nure.ua (A.1); iryna.panforova@nure.ua (A.2);
ORCID: 0000-0002-6703-5166 (A.1); 0000-0001-7032-9109 (A.2)

drawback that makes it difficult to use standard analysis methods to identify the values of the characteristics of the IS requirements. Such an analysis becomes especially important in the course of determining the value of such a characteristic of IS requirements as consistency. This characteristic is the main one for making a decision on the inclusion of a functional requirement in the set of formulated functional requirements that must be implemented as part of IS creation or modification.

Another problem is the use of this requirement to assess the consistency of a functional requirement of publication. In most cases, such publications are natural language texts, visual models, or software prototypes. However, there is a large class of tasks, scenarios for fulfilling functional requirements for which are represented in the form of mathematical expressions. The analysis of such publications of functional requirements was practically not considered in detail. Therefore, research in the field of analysis of publications of IS requirements in the form of mathematical expressions for consistency should be recognized as relevant.

2. Modern research in the field of analysis of system requirements

Currently, the main methods for analyzing functional requirements are the methods developed by the beginning of the 2000s and described in [3]. However, these methods do not make it possible to eliminate errors made in the course of the formation of IS requirements. The main reason for these errors in [4] is the gap between the perception of the generated requirement by the customer and the analyst. Attempts to bridge this gap through the use of modern Agile methods did not give the desired effect [5].

The need to quantify the consistency of functional requirements for IS necessitates research into formal descriptions of system requirements. At the same time, for a formal description of requirements, a variety of devices are proposed – from the theory of probability [6] to the theory of categories [7]. However, most researchers prefer to use knowledge-based models to formally describe requirements. Visual models [8, 9] are often used as publications of requirements for the formation of such models. At the same time, instead of developing fundamentally new methods for analyzing requirements, the focus is on improving existing methods [10]. Therefore, this work aims to study the possibility of using existing methods for analyzing the functional requirements for IS for consistency. At the same time, the publications of these requirements are mathematical expressions that determine the scenario for fulfilling the published functional requirement for IS.

3. Materials and methods of research

As a method for studying functional requirements, an improved method for analyzing individual frames of knowledge-oriented descriptions of functional requirements for IS for consistency is used. This method was proposed by the authors in [11] and improved in [12]. This method is based on the use of a knowledge-based formal description of functional requirements, which has the form

$$K_i^f = \{D_{fr}^i, D_{if}^i, D_{fr\_rel}^i\}, \qquad (1)$$

where K_i^f – designation of knowledge-oriented description of the i-th formulated functional requirement, $K_i^f \in K^f$; K^f – a set of functional requirements; D_{fr}^i – a set of frames describing the i-th functional requirement, $\langle d_n^{im}, \langle d_{el\_fr}^{imn}, d_{el\_fr\_t}^{imn} \rangle \rangle \in D_{fr}^i$, m=1,2,3,…,u, n=1,2,3,…,x; u – the number of frames in the description of the i-th requirement; x – the number of elements in the description of the m-th frame of the i-th requirement; d_n^{im} – description of the name of the m-th frame of the i-th requirement; $d_{el\_fr}^{imn}$ – description of the name of the n-th element of the m-th frame of the i-th requirement; $d_{el\_fr\_t}^{imn}$ – description of the type of the n-th element of the m-th frame of the i-th requirement; D_{if}^i – a set of interfaces describing the i-th functional requirement, $\langle d_g^{im}, \langle d_{el\_if}^{imn}, d_{el\_if\_t}^{imn} \rangle \rangle \in D_{if}^i$, m=1,2,3,…,v, n=1,2,3,…,y; v – the number of interfaces in the description of the i-th requirement; y – the number of elements in the description of the m-th interface of the i-th requirement; d_g^{im} – description of the name of the m-th interface of the i-th requirement; $d_{el\_if}^{imn}$ – description of the name of the n-th element of the m-th interface of the i-th requirement; $d_{el\_if\_t}^{imn}$ – description of the type of the n-th element of the m-th interface of the i-th requirement; $D^{i}_{fr\_rel}$ – a set of links describing the i-th functional requirement, $\langle d^{im}_{fr\_rel}, \langle d^{imn}_{el\_fr\_rel}, d^{imn}_{el\_fr\_rel\_t}\rangle\rangle \in D^{i}_{fr}$, m=1,2,3,…,w, n=1,2,3,…,z; w – the number of links in the description of the i-th requirement; z – the number of elements in the description of the m-th connection of the i-th requirement; $d^{im}_{fr\_rel}$ – description of the name of the m-th connection between interfaces and / or frames of the i-th requirement; $d^{imn}_{el\_fr\_rel}$ – description of the name of the n-th element of the m-th link of the i-th requirement; $d^{imn}_{el\_fr\_rel\_t}$ – description of the type of the n-th element of the m-th link of the i-th requirement.

The improved method consists of the following stages and steps [A12].

Stage 1. Choose a frame $fr^{ia} \in K^{f}_{i}, K^{f}_{i} \in K^{f}$, which was not considered earlier.

Stage 2. Choose $fr^{jb} \in K^{f}_{j}, K^{f}_{j} \in K^{f}$, which was not considered earlier.

Stage 3. Assess the degree of mismatch between the frames fr^{ia} and fr^{jb} by performing the following steps.

Step 3.1. Check the condition $(d^{ia}_{n} \subseteq d^{jb}_{n}) \vee (d^{ia}_{n} \supseteq d^{jb}_{n})$. If the condition is not met, go to Stage 4.

Step 3.2. Calculate the cardinalities of the sets of elements of the frames fr^{ia} and fr^{jb}.

Step 3.3. Select the minimum value from the results of Step 3.2 and assign it to the variable $\min(fr^{ia}, fr^{jb})$.

Step 3.4. Form an array of test results res, the number of elements in which is equal to $\left\lceil \min(fr^{ia}, fr^{jb})/2 \right\rceil + 1$.

Step 3.5. Determine the value of the variable $k = 0$.

Step 3.6. Check inequality

$$\left|\langle d^{ian}_{el\_fr}, d^{ian}_{el\_fr\_t}\rangle \cap \langle d^{jbn}_{el\_fr}, d^{jbn}_{el\_fr\_t}\rangle\right| \leq k. \quad (2)$$

If inequality (2) is satisfied, determine the value of the element $res(k) = 1$. Otherwise, define the value of the element $res(k) = 0$.

Step 3.7. Increase the value of the variable k by one. If $k \leq \left\lceil \min(fr^{ia}, fr^{jb})/2 \right\rceil$, then go to Step 3.6. Otherwise, complete Stage 3.

Stage 4. Exclude the frame fr^{jb} from further consideration. If all frames $fr^{jb} \in K^{f}_{j}$ are not considered, then go to Stage 2.

Stage 5. Exclude the representation K^{f}_{j} from further consideration. If all representations of the set $\left\{K^{f}_{j}\right\}$ are not considered, then choose the previously not considered representation $K^{f}_{j} \in \left\{K^{f}_{j}\right\}$ and go to Stage 2.

Stage 6. Exclude the frame fr^{ia} from further consideration. If all frames $fr^{ia} \in K^{f}_{i}$ are not considered, then go to Stage 1.

Stage 7. For detected cases of mismatch of frames fr^{ia} and fr^{jb} publish the values of res arrays and complete the application of the method.

The expression $\left\lceil \min(fr^{ia}, fr^{jb})/2 \right\rceil$ means rounding the number $\min(fr^{ia}, fr^{jb})/2$ up to the nearest integer (ceiling number $\min(fr^{ia}, fr^{jb})/2$ by Kenneth Anderson).

The application of the improved method will allow for each case of complete or partial coincidence of the names of the frames fr^{ia} and fr^{jb} to obtain tables with estimates of the degree of inconsistency of the descriptions of the elements of these frames. In this case, the number of zero elements of the res array will be the initial data for the subsequent quantitative assessment of the degree of inconsistency of the analyzed frames.

To describe the functional requirements in the form of a mathematical model, the model of the problem "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" was used. This model was considered in [12] and has the following form

$$\alpha_e C_e = \sum_{a=1}^{k}\sum_{b=1}^{l}\sum_{c=1}^{m} t_{abc} + \sum_{d=1}^{p_1} t_d + \sum_{f=1}^{p_2} t_f + \sum_{g=1}^{r} \frac{a_g t_g}{q_g} + \sum_{i=1}^{s} t_i, \quad (3)$$

where α_e – is part of teacher's rate; C_e – is quantity of hours allocated for one teacher's rate for the planned academic year; t_{abc} – is quantity of hours of academic work that is planned to groups of students and to academic discipline; t_d – is quantity of hours of methodological work that is planned to successful completion of the academic work; t_f – is quantity of hours of methodological work that is planned to improve quality of academic work; t_g – is quantity of hours of types scientific work that is planned for the academic year; a_g – is planned number of results of the type of scientific work; q_g – is planned number of co-authors for the result of the type of scientific work; t_i – is quantity of hours of organizational and educational work that is planned for the academic year.

Equation (2) must be met subject to the conditions [12]:

$$\begin{cases} 0 < \alpha_e \leq 1; \\ t_{abc}, t_d, t_f, t_g, t_i, a_g > 0; \\ q_g \geq 1; \\ t_{abc}, t_d, t_f, t_g, t_i, a_g, q_g \in Z. \end{cases}, \quad (4)$$

The result of the analysis of the functions of the task "Formation and maintenance of an individual plan of the scientific and pedagogical employee of the department" in [12] are the following precedents:

- Forming document section "position and stake";
- Quantification of hours of education work planned for the academic year;
- Forming document section “education work”;
- Quantification of hours of methodological work that is planned to successful completion of the academic work;
- Quantification of hours of methodological work that is planned to improve quality of academic work;
- Forming document section “educate-methodical work”;
- Quantification of hours of scientific work planned for the academic year;
- Forming document section “scientific work”;
- Quantification of hours of organizational and educational work planned for the academic year;
- Forming document section “organizational and educational work”;
- Forming document section “final distribution of time by type of work for the academic year”.

Some of these precedents describe the functions that form individual sections of the document "Individual plan of the scientific and pedagogical staff member of the department." However, another part of these use cases describes the functions for calculating the number of hours for various types of employee work. It is difficult to estimate the degree of duplication for such functions in the usual way, since their descriptions are elements of the model (3). Therefore, recommendations should be developed on the use of an improved method for analyzing individual frames of knowledge-based descriptions of functional requirements for IS to identify conflicting functional requirements, which are presented in the form of mathematical expressions.

4. Statement of the main research results

Approbation of the above improved method for analyzing individual frames of knowledge-based descriptions of functional requirements for IS was considered in [12]. The initial data in this approbation were frames of individual parts of the document "Individual plan of the scientific and pedagogical employee of the department", which described the input and output flows of individual functions of the corresponding task. The analysis of computational functions in [12] was not carried out.

Therefore, in this work, it is possible to first highlight the precedents that describe the computational functions of the task "Formation and maintenance of an individual plan of the scientific and pedagogical staff member of the department". These use cases include:

1. Quantification of hours of education work planned for the academic year;
2. Quantification of hours of methodological work that is planned to successful completion of the academic work;
3. Quantification of hours of methodological work that is planned to improve quality of academic work;
4. Quantification of hours of scientific work planned for the academic year;
5. Quantification of hours of organizational and educational work planned for the academic year.

The publications of functional requirements for IS elements designed to implement the listed precedents are given in Table 1.

When forming frames that form a knowledge-oriented description (1) based on such publications, it is proposed to proceed from the following recommendations:

- Each mathematical expression describing a specific use case must be described by one and only one frame;
- As the name of this frame, the name of the use case is used, processed by Porter's stemmer with the subsequent removal of stop words;

Table 1
Functional requirements publications

Requirement designation	Requirement publication
f_{IS_1}	$\sum_{a=1}^{k}\sum_{b=1}^{l}\sum_{c=1}^{m} t_{abc}$
f_{IS_2}	$\sum_{d=1}^{p_1} t_d$
f_{IS_3}	$\sum_{f=1}^{p_2} t_f$
f_{IS_4}	$\sum_{g=1}^{r} \frac{a_g t_g}{q_g}$
f_{IS_5}	$\sum_{i=1}^{s} t_i$

- The frame slots are mathematical symbols for the elements of the publication, the values of which can change during calculations.

The mathematical expressions themselves define the scenario of the use cases and represent the description of the attached procedures associated with the frame as a whole [14]. Their analysis requires additional research and is not considered in this work.

In accordance with the proposed recommendations, the initial data for the method will be five knowledge-oriented descriptions of functional requirements, each consisting of one frame. The list of descriptions and names of frames are given in Table 2.

Table 2
Frame names for knowledge-based functional requirements descriptions

Description designation	Frame name designation	Frame name value
$K_1^{f_{IS}}$	d_n^{11}	Quantif hour educ work plan academ year
$K_2^{f_{IS}}$	d_n^{21}	Quantif hour methodolog work plan success complet academ work
$K_3^{f_{IS}}$	d_n^{31}	Quantif hour methodolog work plan improv qualiti academ work
$K_4^{f_{IS}}$	d_n^{41}	Quantif hour scientif work plan academ year
$K_5^{f_{IS}}$	d_n^{51}	Quantif hour organiz educ work plan academ year

Description of frame elements and their types is given in Table 3.

Table 3
Functional requirements publication frame element descriptions

Frame element designation	Frame element value	Frame element type notation	Frame element type value
$d_{el\_fr}^{111}$	a	$d_{el\_fr\_t}^{111}$	Numerical
$d_{el\_fr}^{112}$	k	$d_{el\_fr\_t}^{112}$	Numerical
$d_{el\_fr}^{113}$	b	$d_{el\_fr\_t}^{113}$	Numerical
$d_{el\_fr}^{114}$	l	$d_{el\_fr\_t}^{114}$	Numerical
$d_{el\_fr}^{115}$	c	$d_{el\_fr\_t}^{115}$	Numerical
$d_{el\_fr}^{116}$	m	$d_{el\_fr\_t}^{116}$	Numerical
$d_{el\_fr}^{117}$	t_{abc}	$d_{el\_fr\_t}^{117}$	Numerical
$d_{el\_fr}^{211}$	d	$d_{el\_fr\_t}^{211}$	Numerical
$d_{el\_fr}^{212}$	p_1	$d_{el\_fr\_t}^{212}$	Numerical
$d_{el\_fr}^{213}$	t_d	$d_{el\_fr\_t}^{213}$	Numerical
$d_{el\_fr}^{311}$	f	$d_{el\_fr\_t}^{311}$	Numerical
$d_{el\_fr}^{312}$	p_2	$d_{el\_fr\_t}^{312}$	Numerical
$d_{el\_fr}^{313}$	t_f	$d_{el\_fr\_t}^{313}$	Numerical
$d_{el\_fr}^{411}$	g	$d_{el\_fr\_t}^{411}$	Numerical
$d_{el\_fr}^{412}$	r	$d_{el\_fr\_t}^{412}$	Numerical
$d_{el\_fr}^{413}$	a_g	$d_{el\_fr\_t}^{413}$	Numerical
$d_{el\_fr}^{414}$	t_g	$d_{el\_fr\_t}^{414}$	Numerical
$d_{el\_fr}^{415}$	q_g	$d_{el\_fr\_t}^{415}$	Numerical
$d_{el\_fr}^{311}$	i	$d_{el\_fr\_t}^{311}$	Numerical
$d_{el\_fr}^{312}$	s	$d_{el\_fr\_t}^{312}$	Numerical
$d_{el\_fr}^{313}$	t_i	$d_{el\_fr\_t}^{313}$	Numerical

In the course of applying the improved method described above, it is possible to see that the condition in Step 3.1 will be satisfied only for frames fr^{11} and fr^{51} (the name of the frame d_n^{11} is a subset of the name of the frame d_n^{51}). Therefore, further steps of Stage 3 of the improved method should be carried out only to determine the degree of mismatch between the elements of the frames fr^{11} and fr^{51}.

As a result of Step 3.2 and Step 3.2 of Step 3 of the refined method, $\min(fr^{ia}, fr^{jb}) = 3$ was determined.

The results of the implementation of Steps 3.4, 3.5 and 3.6 of the improved method are shown in Table 4.

Table 4
Values of array elements of check results for frames $\boldsymbol{fr^{11}}$ and $\boldsymbol{fr^{51}}$

Array element designation	Array element value
$res(0)$	0
$res(1)$	1
$res(2)$	1
$res(3)$	1

Table 4 allows to draw the following conclusion: despite the strong coincidence of the names of the frames fr^{11} and fr^{51}, the descriptions of their elements and types do not correspond to each other.

This discrepancy is explained by the ambiguity of the term "educational work". This ambiguity leads to the use of homonyms to designate educational work and educational work. With the elimination of homonymous descriptions, the revealed discrepancy between the first and fifth of the precedents listed above will also be eliminated.

Thus, the possibility of detecting inconsistent knowledge-oriented descriptions of functional requirements on the basis of their publication in the form of mathematical expressions is confirmed.

5. Conclusions

The study of the possibility of finding a discrepancy between knowledge-oriented descriptions of functional requirements, published in the form of mathematical expressions, has been carried out. Recommendations for the formation of frame descriptions of publications of requirements in the form of mathematical expressions are proposed. The results obtained allow to assert the possibility of using the improved method for the analysis of individual frames of knowledge-oriented descriptions of functional requirements for IS, which are published in the form of mathematical expressions. However, this conclusion needs additional verification on various types of requirements and mathematical expressions.

6. References

[1] GOST R 57193–2016. System and program engineering. System life cycle processes (ISO/IEC/IEEE 15288:2015, NEQ). (2016). Introduced 31–10–2016. M.: Standartinform, 98.

[2] Project Management Body of Knowledge: fifth ed. (2013). Newton Square: Project Management Institute, Inc., 586.

[3] M. Maguire, N. Bevan (2002). User requirements analysis. A review of supporting methods. Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, 25-30 August 2002, 133-148. doi: 10.1007/978-0-387-35610-5_9

[4] Stowell F., Cooray S. (2017). The appreciative system, learning, and its impact on information systems design. Communications of the Association for Information Systems, 40, 93-119.

[5] H.F. Martins, A.C. de Oliveira, E.D. Canedo etc. (2019). Design thinking: Challenges for software requirements elicitation. Information, 10, 12, 371. doi: https://doi.org/10.3390/info10120371

[6] Yu Y.-J., Liu C. (2017). Little Model in Big Data: An Algebraic Approach to Analysing Abstract Software Behaviours. Ruan Jian Xue Bao/Journal of Software, 28 (6), 1488-1497. doi: 10.13328/j.cnki.jos.005229

[7] Asteasuain F., Braberman V. (2017). Declaratively building behavior by means of scenario clauses. Requirements Engineering, 22, 2, 239-274. doi: 10.1007/s00766-015-0242-2

[8] Kaiya, K. Adachi, Y. Chubachi (2019) Requirements Exploration by Comparing and Combining Models of Different Information Systems. In: M. Virvou, F. Kumeno, K. Oikonomou (eds) Knowledge-Based Software Engineering: 2018. JCKBSE 2018. Smart Innovation, Systems and Technologies, vol 108. Pp. 64-74. doi: 10.1007/978-3-319-97679-2_7.

[9] S.-K. Kim, T. Myers, M.-F. Wendland, P.A. Lindsay (2012) Execution of natural language requirements using state machines synthesised from Behavior Trees. Journal of Systems and Software. 85, 11, 2652-2664. doi: 10.1016/j.jss.2012.06.013.

[10] Ambreen T., N. Ikram, M. Usman, M. Niazi (2018). Empirical research in requirements engineering: trends and opportunities.

Requirements engineering, 23, 1, 63-95. doi: 10.1007/s00766-016-0258-2
[11] Ievlanov, M., Vasiltcova N., Panforova I. (2018).Development of methods for the analysis of functional requirements to an information system for consistency and illogicality. Eastern-European Journal of Enterprise Technologies, 1/2 (91), 4-11. doi: 10.15587/1729-4061.2108.121849
[12] Vasiltcova, N., Panforova, I., & Neumyvakina, O. (2020). Improving a method to analyze the requirements for an information system for consistency. Eastern-European Journal of Enterprise Technologies, 3/2 (105), 17–27. https://doi.org/10.15587/1729-4061.2020.205518
[13] Vasyltsova N., Panforova I., Kuzma Ye. Formation of Function Use Cases Based on Its Mathematical Model // Fourth International Scientific and Technical Conference “COMPUTER AND INFORMATION SYSTEMS AND TECHNOLOGIES”. Kharkiv: DISA PLUS LLC, 2020: 59-60.
[14] Levykin V.M., Kernosov M.A. (2008). Research and development of a frame model of the document structure. New technologies, 1 (19). 149-154.

Generalized model for managing the operation of a web-based information system

Levykin V.M. [1], Petrichenko O.V. [2] and Neumyvakina O.Ye. [3]

[1,2,3] *Kharkiv National University of Radio Electronics, Nauki ave. 14, Kharkiv, 61166, Ukraine*

Abstract

The modern methodology for managing the operation of web-based information systems has a number of features that determine its contradictions and shortcomings. These shortcomings determine the necessity and urgency of conducting a study to improve the existing methodology for managing the operation of Web-based ISs and create, as a result of this improvement, models for managing the operation of Web-based ISs. The aim of this study is to develop a generalized model for managing the operation of a web-based information system.

The main method for conducting this study is to recommend a process for determining the needs and requirements of a stakeholder. The choice of this method is due to the existing representation of the concepts, models and methods of managing the operation of the system as requirements for the service management system. In the course of the study, recommendations were identified for improving the existing methodology for managing the operation of web-based IS. On the basis of the recommendations highlighted, it is proposed to expand the set of basic terms of the subject area of the methodology, and the semantics of each new basic term was determined. The use of an improved set of basic terms made it possible to develop the basic category-theoretical descriptions of the operated IS and elements of the management processes for the operation of the information system. The resulting category-theoretical descriptions are used in the development of a generalized category-theoretical model for managing the operation of a web-based information system. This generalized model reflects the relationship of the formulated category-theoretical descriptions at each stage of the system operation management and during the transitions between these stages. The use of these models makes it possible to subsequently determine the main data structures of information and software support of modern information technologies for automated management of the operation of web-based information systems.

Keywords

Information system, operation management, IT-service, theory of categories.

1. Introduction

The modern methodology for managing the operation of web-based information systems (IS), on which theoretical research and applied developments are based on automating the solution of management problems, has the following features [1]:

• Operated IS should be considered as a set of separate IT-services, implemented in the form of IT-services (in this case, any IT-service can be implemented as one or many interconnected IT-services);

• Managing influences on multiple IT services should be viewed as management decisions related to continuous improvement of IT service capabilities and any need to change both the service management system and individual services (services);

• Use of the “efficiency” criterion as the main quantitative criterion for service management, which is defined as the degree of implementation of planned activities and achievement of planned results in service management;

• The methodological framework for service management should be considered as a set of requirements put forward for a service management system and its individual processes and based on a set of basic terms.

This set of basic terms of the subject area, on which the requirements for the service

EMAIL: viktor.levykin@nure.ua (A.1); oleksandr.petrychenko@nure.ua (A.2), olga.neumyvakina@nure.ua (A.3)
ORCID: 0000-0002-7929-515X (A.1); 0000-0002-1319-5041 (A.2); 0000-0001-6936-6543 (A.3)

management system are based, are divided in [1] into two such subsets:

1. Terms specific to control systems as a whole (21 terms);

2. Terms specific to service management (29 terms).

The considered features of the modern methodology for managing the operation of web-based IS determine the main contradiction of this methodology. The essence of this contradiction is as follows. The existing methodology uses one presentation level to describe the exploited IS. However, for the Provider and the Consumer of IT services of this IS, the description of the operated IS presupposes the simultaneous existence of several levels of presentation - in particular, the business level, the level of IT services, the level of IT services and the level of IT infrastructure [2, 3].

Another drawback of the modern methodology for managing the operation of web-based ISs is a large number of indicators characterizing the operated IS and the executable operation processes. Each IT service and each Web-based IS IT service is defined by separate sets of these indicators, which leads to a rapid increase in the number of indicators by which the effectiveness of the operation of the Web-based IS as a whole should be determined.

Therefore, studies on improving the existing methodology for managing the operation of web-based ISs and creating, as a result of this improvement, models for managing the operation of web-based ISs are relevant in theoretical and applied aspects.

2. Methods of research

In accordance with the provisions of the methodology for managing the operation of web-based ISs, the development of a generalized model for managing the operation of such a system is proposed to be considered as the development of a new requirement for the IT service management system of the operated system. The main difference of this requirement from others will be that this requirement will be formulated not in the form of text or a visual model, but in the form of a formal description of the model being developed, which is based on many basic terms of the methodology.

It is currently recommended to formulate the requirements for the system in accordance with the description of the process for determining the needs and requirements of the interested party. Moreover, in our case, of all the activities and tasks of this process, it is necessary to perform only the activities and tasks shown in Table 1 [4].

Table 1
Activities and objectives of the process for determining the needs and requirements of the stakeholder

Process activities	Tasks to be solved during the execution of the process activity a
Identifying stakeholder needs	Determination of the context of use within the concept of functioning (operation) and the initial concepts of life cycle
	Determination and justification of stakeholder needs
Development of operational concepts (operational concepts) and other life cycle concepts	Determination of a representative set of scenarios to identify all the required capabilities consistent with the expected operational and other lifecycle concepts

The choice of these activities and tasks is due to the following considerations:

• The activity “Preparing to identify the needs and requirements of interested parties” is not performed in this case, since the generalized model for managing the operation of a web-based IS should be the same for all interested parties - participants in the creation of a service management system;

• The task “Prioritization of the selected needs from top to bottom” is not fulfilled, since it is possible to tell about the formation of only one need;

• The task "Define the interaction between users and the system" is not met, since the expected result of the study is a generalized model for managing the operation of a web-based IS, and not its presentation by individual stakeholders;

• The activities "Converting needs into stakeholder requirements", "Analyzing the needs and requirements of stakeholders" and "Managing the Determination of Needs and

Requirements of Stakeholders" in this study are not performed due to the need for additional research on the constraints and critical characteristics of the quality of the generated requirement due to the results development of a generalized mathematical model.

In the general case, the implementation of those indicated in Table 1 activities and tasks of the process of determining the needs and requirements of the interested party should ensure the formation of such results [4]:

• The context of the use of opportunities and concepts in the life cycle stages, including operational concepts;

• The identified needs of the stakeholders.

At the same time, it is indicated in [4] that the formulations of descriptions of these needs can be performed in a variety of ways and, as a rule, are very far from formal templates for representing the requirements of stakeholders.

3. Results of the development of a generalized management model for a web-based IS operation

The developed model is based on the results of improving the methodology for managing the operation of web-based IS. These results in [3] are presented in the form of the following initial assumptions and recommendations:

• The exploited IS should be considered as a set of interrelated representations at various levels (business level, level of IT services, level of IT services, level of IT infrastructure);

• The main conventional "unit" for describing data processing actions in the operated IS at any level of presentation is a transaction;

• It is possible to reduce the dimension of the problem of managing the operation of the IS, provided that the individual properties of the IS are identified, which characterize the manifestations of the individual qualities of this system;

• Each individual property of the IS is determined by a multitude of requirements (functional and non-functional) and a multitude of requests to change these requirements and individual elements of the system.

3.1. Determination of the context of use of the model within the improved methodology

Based on the above initial assumptions and recommendations, in [3] it is proposed to supplement the existing basic terms with the following:

• The term "Presentation Layer", which defines the presentation levels of the exploited IS;

• The term "Transaction", which defines the features of a unified description of processes in the operated IS;

• The term "Indicator", which defines the features of a unified formal description of any indicators that characterize the operating IS or the processes of operating this system;

• The term "Library of Indicators", which defines the peculiarities of grouping individual indicators into a single library;

• The term "Property", which defines the peculiarities of grouping individual indicators into a formal description of a specific property of a web-based IS at various levels of the system presentation.

Based on the results of the expansion of the set of basic terms, it became possible to determine the basic descriptions of the IS and the processes of managing its operation, which are subject to formalization. These descriptions include [3]:

1. Description of the exploited IS at various levels of presentation;
2. Description of the states of the operated IS;
3. Description of indicators characterizing the operated IS and its configuration elements;
4. Description of the efficiency of the IS operation by its individual properties and in general.

The description of the exploited IS at various levels of presentation is based on the basic terms "Information System", "Provider of IS", "Customer of IS", "Requirement of IS", "RFC IS", "Configuration of IS", "Presentation Layer", "Configuration Item" and "Transaction".

The description of the states of the operated IS is based on the basic terms "Configuration of IS", "Presentation Layer", "Configuration Item", "Transaction", "State of IS", "Planned state" and "Current state".

The description of the indicators characterizing the exploited IS and its configuration elements are based on the basic terms "State of IS", "Indicator" and "Library of Indicators".

Description of the efficiency of the IS operation by its individual properties and in general is based on the basic terms "State of IS", "Planned state", "Current state", "Indicator", "Presentation Layer", "Property", "Efficiency" and "RFC IS ".

3.2. Determination and formalization of highlighted descriptions as stakeholder needs

The description of the operated IS at various levels of presentation is proposed to be presented in the form of the L_IS, category, which has the following form

$$L_{IS} = \begin{bmatrix} Ob_{IS}, Ob_{Pr}, Ob_{Cust}, Ob_{rec}, Ob_{RFC}, \\ Ob_{Conf}, Ob_{P\_L}, Ob_{CI}, Ob_{Tr}, \\ H_{Ob_{IS}}^{Ob_{Pr}}, H_{Ob_{Cust}}^{Ob_{IS}}, H_{Ob_{rec}}^{Ob_{IS}}, H_{Ob_{rec}}^{Ob_{Pr}}, \\ H_{Ob_{IS}}^{Ob_{Cust}}, H_{Ob_{Conf}}^{Ob_{rec}}, H_{Ob_{Conf}}^{Ob_{P\_L}}, \\ H_{Ob_{Conf}}^{Ob_{CI}}, H_{Ob_{Conf}}^{Ob_{Tr}}, H_{Ob_{Tr}}^{Ob_{CI}}, H_{Ob_{RFC}}^{Ob_{Pr}}, \\ H_{Ob_{RFC}}^{Ob_{Cust}}, H_{1_{Ob}} \end{bmatrix}, \quad (1)$$

where Ob_{IS} – a subset of L_{IS} category objects describing the basic term "Information System"; Ob_{Pr} – a subset of L_{IS} category objects describing the basic term "Provider of IS"; Ob_{Cust} – a subset of L_{IS} category objects describing the basic term «Customer of IS»; Ob_{rec} – a subset of L_{IS} category objects describing the basic term «Requirement of IS»; Ob_{RFC} – a subset of L_{IS} category objects describing the basic term «RFC IS»; Ob_{Conf} – a subset of L_{IS} category objects describing the basic term «Configuration of IS»; Ob_{P_L} – a subset of L_{IS} category objects describing the basic term «Presentation Layer»; Ob_{CI} – a subset of L_{IS} category objects describing the basic term «Configuration Item»; Ob_{Tr} – a subset of L_{IS} category objects describing the basic term «Transaction»; $H_{Ob_{IS}}^{Ob_{Pr}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Pr} and Ob_{IS}; $H_{Ob_{Cust}}^{Ob_{IS}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{IS} and Ob_{Cust}; $H_{Ob_{rec}}^{Ob_{IS}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{IS} and Ob_{rec}; $H_{Ob_{rec}}^{Ob_{Pr}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Pr} and Ob_{rec}; $H_{Ob_{IS}}^{Ob_{Cust}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Cust} and Ob_{IS}; $H_{Ob_{IS}}^{Ob_{Cust}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Cust} and Ob_{IS}; $H_{Ob_{Conf}}^{Ob_{rec}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{rec} and Ob_{Conf}; $H_{Ob_{Conf}}^{Ob_{P\_L}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{P_L} and Ob_{Conf}; $H_{Ob_{Conf}}^{Ob_{CI}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{CI} and Ob_{Conf}; $H_{Ob_{Conf}}^{Ob_{Tr}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Tr} and Ob_{Conf}; $H_{Ob_{Tr}}^{Ob_{CI}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{CI} and Ob_{Tr}; $H_{Ob_{RFC}}^{Ob_{Pr}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Pr} and Ob_{RFC}; $H_{Ob_{RFC}}^{Ob_{Cust}}$ – a subset of L_{IS} category morphisms defined between the subsets Ob_{Cust} and Ob_{RFC}; $H_{1_{Ob}}$ – a subset of unit morphisms that are defined on the corresponding subsets of category objects.

The description of the states of the operated IS is proposed to be presented in the form of the category L_{St}, which has the following form

$$L_{St} = \begin{bmatrix} Ob_{Conf}, Ob_{P\_L}, Ob_{CI}, Ob_{Tr}, Ob_{St}, \\ Ob_{Pl\_St}, Ob_{C\_St}, H_{Ob_{Conf}}^{Ob_{P\_L}}, H_{Ob_{Conf}}^{Ob_{CI}}, \\ H_{Ob_{Conf}}^{Ob_{Tr}}, H_{Ob_{Tr}}^{Ob_{CI}}, H_{Ob_{St}}^{Ob_{Conf}}, H_{Ob_{Pl\_St}}^{Ob_{St}}, \\ H_{Ob_{C\_St}}^{Ob_{St}}, H_{1_{Ob}} \end{bmatrix}, \quad (2)$$

where Ob_{St} – a subset of L_{St} category objects describing the basic term «State of IS»; Ob_{Pl_St} – a subset of L_{St} category objects describing the basic term «Planned State»; Ob_{C_St} – a subset of L_{St} category objects describing the basic term «Current State»; $H_{Ob_{St}}^{Ob_{Conf}}$ – a subset of L_{St} category morphisms defined between the subsets Ob_{Conf} and Ob_{St}; $H_{Ob_{Pl\_St}}^{Ob_{St}}$ – a subset of L_{St} category morphisms defined between the subsets Ob_{St} and Ob_{Pl_St}; $H_{Ob_{C\_St}}^{Ob_{St}}$ – a subset of L_{St} category morphisms defined between the subsets Ob_{St} and Ob_{C_St}.

The description of the indicators characterizing the operated IS and its configuration elements is proposed to be presented in the form of the L_{St} category, which has the following form

$$L_{Ind} = \begin{bmatrix} Ob_{St}, Ob_{Ind}, Ob_{Lib}, H_{Ob_{St}}^{Ob_{Ind}}, \\ H_{Ob_{Lib}}^{Ob_{Ind}}, H_{Ob_{C\_St}}^{Ob_{St}}, H_{1_{Ob}} \end{bmatrix}, \quad (3)$$

where Ob_{Ind} – a subset of L_{Ind} category objects describing the basic term «Indicator»; Ob_{Lib} – a subset of L_{Ind} category objects describing the basic term «Library of Indicators»; $H_{Ob_{St}}^{Ob_{Ind}}$ – a subset of L_{Ind} category morphisms defined between the subsets Ob_{Ind} and Ob_{St}; $H_{Ob_{Lib}}^{Ob_{Ind}}$ – a subset of L_{Ind} category morphisms defined between the subsets Ob_{Ind} and Ob_{Lib}.

The description of the efficiency of the IS operation by its individual properties and in general is proposed to be presented in the form of the category L_{Eff}, which has the following form

$$L_{Eff} = \begin{bmatrix} Ob_{St}, Ob_{Pl_{St}}, Ob_{C_{St}}, Ob_{Ind}, Ob_{P\_L}, \\ Ob_{Prop}, Ob_{eff}, Ob_{RFC}, H_{Ob_{Pl\_St}}^{Ob_{St}}, \\ H_{Ob_{C\_St}}^{Ob_{St}}, H_{Ob_{St}}^{Ob_{Ind}}, H_{Ob_{Prop}}^{Ob_{St}}, \\ H_{Ob_{Prop}}^{Ob_{eff}}, H_{Ob_{Pl\_St}}^{Ob_{eff}}, H_{Ob_{C\_St}}^{Ob_{eff}}, \\ H_{Ob_{RFC}}^{Ob_{St}}, H_{1_{Ob}} \end{bmatrix}, \quad (4)$$

where Ob_{Prop} – a subset of L_{Eff} category objects describing the basic term «Property»; Ob_{eff} – a subset of L_{Eff} category objects describing the basic term «Efficiency»; $H_{Ob_{Prop}}^{Ob_{St}}$ – a subset of L_{Eff} category morphisms defined between the subsets Ob_{St} and Ob_{Prop}; $H_{Ob_{Prop}}^{Ob_{eff}}$ – a subset of L_{Eff} category morphisms defined between the subsets Ob_{eff} and Ob_{Prop}; $H_{Ob_{Pl\_St}}^{Ob_{eff}}$ – a subset of L_{Eff} category morphisms defined between the subsets Ob_{eff} and Ob_{Pl_St}; $H_{Ob_{C\_St}}^{Ob_{eff}}$ – a subset of L_{Eff} category morphisms defined between the subsets Ob_{eff} and Ob_{C_St}; $H_{Ob_{RFC}}^{Ob_{St}}$ – a subset of L_{Eff} category morphisms defined between the subsets Ob_{St} and Ob_{RFC}.

3.3. Development of a generalized model as a scenario for fulfilling a needs

Management of the operation of the web-based IS in accordance with the provisions outlined in [1] is based on the "Plan - Do - Check - Act" cycle. In this loop:

1. At the “Plan” stage, plans for the operation of the IS are developed;
2. At the “Do” stage, the planned works on the operation of the IS are carried out;
3. At the “Check” stage, work is carried out to check the results of the implementation of plans for the operation of the IS;
4. At the “Act” stage, work is carried out to correct the plans for the operation of the IS.

Then a generalized model for managing the operation of a web-based IS, taking into account the above-mentioned features of the "Plan - Do - Check - Act" cycle, can be represented as a supercategory, which has the following form [3]:

$$M_O = \begin{bmatrix} L_P, L_D, L_{Ch}, L_A, F_{L_D}^{L_P}, F_{L_{Ch}}^{L_D}, F_{L_D}^{L_{Ch}}, \\ F_{L_A}^{L_{Ch}}, F_{L_P}^{L_A} \end{bmatrix}, \quad (5)$$

where L_P – supercategory that formally describes the "Plan" stage; L_D – supercategory, formally describing the "Do" stage; L_{Ch} – supercategory, formally describing the "Check" stage; L_A – supercategory, formally describing the "Act" stage; $F_{L_D}^{L_P}$ – одноместный ковариантный функтор, который устанавливает связь между надкатегориями L_P и L_D; $F_{L_{Ch}}^{L_D}$ – one-place covariant functor that establishes a relationship between the supercategories L_D and L_{Ch}; $F_{L_D}^{L_{Ch}}$ – one-place covariant functor that establishes a relationship between the supercategories L_{Ch} and L_D; $F_{L_A}^{L_{Ch}}$ – one-place covariant functor that establishes a relationship between the supercategories L_{Ch} and L_A; $F_{L_P}^{L_A}$ – one-place covariant functor that establishes a relationship between the supercategories L_A and L_P.

Functors $F_{L_D}^{L_P}$, $F_{L_{Ch}}^{L_D}$, $F_{L_A}^{L_{Ch}}$ and $F_{L_P}^{L_A}$ link the supercategories L_P, L_D, L_{Ch} and L_A according to control cycle of the operated IS. The functor $F_{L_D}^{L_{Ch}}$ establishes a connection between the supercategories L_{Ch} and L_D to describe situations when the IS operation is carried out in accordance with the developed plans and there is no need to correct these plans.

To detail the formal description of the supercategories L_P, L_D, L_{Ch} and L_A pet’s divide the previously introduced category of description of the states of the operated IS L_{St} nto two subcategories: the subcategory of the description of the planned state of the operated IS L_{St}^P and the subcategory of the description of the current state of the operated IS L_{St}^C. Subcategory L_{St}^P will be of the form

$$L_{St}^P = \begin{bmatrix} Ob_{Conf}, Ob_{P\_L}, Ob_{CI}, Ob_{Tr}, Ob_{St}, \\ Ob_{Pl\_St}, H_{Ob_{Conf}}^{Ob_{P\_L}}, H_{Ob_{Conf}}^{Ob_{CI}}, \\ H_{Ob_{Conf}}^{Ob_{Tr}}, H_{Ob_{Tr}}^{Ob_{CI}}, H_{Ob_{St}}^{Ob_{Conf}}, H_{Ob_{Pl\_St}}^{Ob_{St}}, \\ H_{1_{Ob}} \end{bmatrix}, \quad (6)$$

and the subcategory L_{St}^C will have the form

$$L_{St}^{P}=\begin{bmatrix} Ob_{Conf}, Ob_{P\_L}, Ob_{CI}, Ob_{Tr}, Ob_{St}, \\ Ob_{C\_St}, H_{Ob_{Conf}}^{Ob_{P\_L}}, H_{Ob_{Conf}}^{Ob_{CI}}, \\ H_{Ob_{Conf}}^{Ob_{Tr}}, H_{Ob_{Tr}}^{Ob_{CI}}, H_{Ob_{St}}^{Ob_{Conf}}, H_{Ob_{C\_St}}^{Ob_{St}}, \\ H_{1_{Ob}} \end{bmatrix}. \quad (7)$$

Then, taking into account the previously developed categorical descriptions of the exploited IS, the supercategory L_P can be described as follows [3]:

$$L_P=\left[L_{IS}, L_{St}^{P}, L_{Ind}, F_{L_{St}^{P}}^{L_{IS}}, F_{L_{Ind}}^{L_{St}^{P}}\right], \quad (8)$$

where $F_{L_{St}^{P}}^{L_{IS}}$ – a one-place covariant functor that establishes a connection between the supercategories L_{IS} and L_{St}^{P}; $F_{L_{Ind}}^{L_{St}^{P}}$ – a one-place covariant functor that establishes a connection between the supercategories L_{St}^{P} and L_{Ind}.

The supercategory L_D can be described as follows [3]:

$$L_D=\left[L_{IS}, L_{St}^{C}, L_{Ind}, F_{L_{St}^{C}}^{L_{IS}}, F_{L_{Ind}}^{L_{St}^{C}}\right], \quad (9)$$

where $F_{L_{St}^{C}}^{L_{IS}}$ – a one-place covariant functor that establishes a connection between the supercategories L_{IS} and L_{St}^{C}; $F_{L_{Ind}}^{L_{St}^{C}}$ – a one-place covariant functor that establishes a connection between the supercategories L_{St}^{C} and L_{Ind}.

The supercategory L_{Ch} can be described as follows [3]:

$$L_{Ch}=\begin{bmatrix} L_{IS}, L_{St}^{P}, L_{St}^{C}, L_{Ind}, F_{L_{St}^{P}}^{L_{IS}}, F_{L_{St}^{C}}^{L_{IS}}, \\ F_{L_{Ind}}^{L_{St}^{P}}, F_{L_{Ind}}^{L_{St}^{C}} \end{bmatrix}. \quad (10)$$

The supercategory L_A can be described as follows [3]:

$$L_A=\left[L_{IS}, L_{Eff}, F_{L_{Eff}}^{L_{IS}}\right]. \quad (11)$$

where $F_{L_{Eff}}^{L_{IS}}$ – a one-place covariant functor that establishes a connection between the supercategories L_{IS} та L_{Eff}.

4. Conclusions

A generalized category-theoretical model for managing the operation of a web-based IS has been developed as a scenario for fulfilling the need for a management system for IT services of an operated IS (5). This model establishes formal descriptions of a web-based IS at individual stages of the management cycle and the relationship between these descriptions. Also, category-theoretical models of formal descriptions of web-based IS at each stage of the management cycle were developed. These models make it possible to formally represent the interaction of the selected descriptions of the IS and the processes for managing its operation, which are to be formalized within each specific stage of management, as well as the interaction of a set of these descriptions during the transition from one stage of the management cycle to another. The use of these models makes it possible to determine the basic data structures of information and software support of modern information technologies for automated management of the operation of web-based ISs.

The results obtained determine the prospects for further research as the development of models and methods that allow detailed descriptions of individual elements of models (5) and (8) - (11). Such models and methods will determine the features of solving specific problems of managing the IS operation at different stages of the management cycle.

5. Acknowledgements

The authors thank Maksym Evlanov, Doctor of Technical Sciences, Professor of the Department of Information Control Systems of the Kharkiv National University of Radio Electronics for active assistance in carrying out this research.

6. References

[1] ISO/IEC 20000-1. Information technology – Service management – Part 1: Service management system requirements. Third edition 2018-09. (2018). Introduced 2019-01-01. Geneva: ISO Copyright Office, 96 p.

[2] V. Levykin, M. Yevlanov, M. Kernosov, Pattern planning of requirements to the informative systems: design and application, The «Kompaniya «Smit LTD», Kharkov, Ukraine, 2014.

[3] Petrichenko O.V. Improving enterprise IT-service management methodology, Management Information System and Devices 177 (2021) 4–12.

[4] GOST R 57193–2016. System and program engineering. System life cycle processes (ISO/IEC/IEEE 15288:2015, NEQ). (2016). Introduced 31–10–2016. M.: Standartinform, 98.

Comparison of Distributed Transactions Execution in Different Types of DBMS

Mazurova Oksana [1], Naboka Artem [1], Shirokopetleva Mariya [1] and Sydorenko Ganna [2]

[1] Kharkiv National University of Radio Electronics, Nayka av., 14, Kharkiv, 61166, Ukraine
[2] National Technical University "Kharkiv Polytechnic Institute", str. Kyrpychova 2, Kharkiv, 61002, Ukraine

Abstract
The work is devoted to the study of methods for implementing distributed transactions on the ACID principle using replication technology. The paper highlights some of the most commonly used methods for replicated distributed databases. Experimental research of methods and their comparison was carried out on time of execution of various transactions, and also on resources necessary for performance of these queries. The results of the experiments are given in absolute values, as well as the functions of change of these metrics. Based on the obtained results, applied recommendations for the use of a particular method are formed.

Keywords
DBMS, distributed database, transaction, ACID, two-phase commit, SQL, NoSQL, NewSQL, database loading, efficiency metric, performance time, MongoDB, MS SQL Server, VoltDB
.

1. Introduction

Databases have become an integral part of modern applications, and in some is a central element. Along with the increasing popularity of databases, increased volume and most of the data should be stored. That is why there is a problem of scaling large databases to ensure adequate response speed of the application.

There are two main methods of scaling databases, horizontal and vertical [1]. Horizontal scaling involves partitioning the entire database into multiple integrated network of physical servers that store a portion of the data used in the majority of cases.

Together with horizontal scaling implementation there is a problem in distributed databases transactions that meet the ACID properties [2]. This task is not trivial, because the nodes of the database cluster are usually located on different physical servers, which are connected using the TCP/IP protocol, which in itself has no tools for implementing transactionality

The reason for creating a distributed database in most cases is to ensure high availability of application data, which means the need to quickly receive and process data. In turn, the most popular solutions that implement ACID transactions, at one time were the classic relational database such as SQL Server, Oracle, MySQL, Postgres. But as the load on the database increases and the stored data increases, the classic relational databases scale vertically and do not have built-in mechanisms for horizontal scaling to shards and replicas. So mostly classical relational DBMS can not provide high availability to large data while the large number of users.

On the other hand, with the need to deploy distributed databases, NoSQL databases have become popular [3]. NoSQL databases support transactionality only at the level of one record (therefore there is no possibility to carry out transactions over the data located on various nodes) [4]. That is why NoSQL databases are used in search and analytics systems [5], but cannot be used in applications where transactionality is critical.

There are several approaches to distributed transactions. Most of them are not focused on the transactionality of the database itself, and the transactionality of distributed software applications of the system, for example, for microservice architecture uses the pattern Saga

EMAIL: oksana.mazurova@nure.ua (A. 1); artem.naboka@nure.ua (A. 2); marija.shirokopetleva@nure.ua (A. 3); Ganna.Sydorenko@khpi.edu.ua (A. 4)
ORCID: 0000-0003-3715-3476 (A. 1); 0000-0003-2178-7984 (A. 2); 0000-0002-7472-6045 (A. 3); 0000-0002-0761-2793 (A. 4)

[6], which allows distributed transactions, which can be implemented by two methods: choreography and orchestration.

All of the above solutions are quite effective in their field, but they do not directly relate to transactionality at the database level. Other approaches, such as Local Transaction, can be used directly to ensure transactionality at the database level. Its essence is the parallel execution of classical transactions on each of the cluster nodes.

Another approach is the Flexible Transaction approach, which is essentially very similar to the Saga microservice template, in which classic transactions are opened (but not approved) on each node and queries are executed [7]. Only after the requests have been successfully executed on all nodes, transactions begin to be approved one by one on each node. Obviously, this approach provides more efficient distributed transactions than Local Transaction, but it still keeps transactions open for a long time, which will reduce efficiency with high concurrency.

2. Research task

The aim of the study is to increase the efficiency of ACID transaction implementation for replicated databases by comparing the methods of their implementation based on the use of relational, NoSQL and NewSQL databases, as well as to develop appropriate recommendations for their use. The following global challenges need to be addressed during the study:

- to analyze and select relevant methods for implementing distributed transactions that could be used in modern applications;
- develop a plan for experimental research of selected methods (database scheme, transactions, constraints, etc.);
- to conduct an experimental study of the effectiveness of selected methods taking into account the following parameters: the size of disk space for database storage; execution time of queries with a variable number of users, replicas and data; RAM required by the database server; CPU time spent by the database server;

based on the results of the experiment to develop recommendations for the use of transaction support methods.

It is necessary to determine the dependence of the execution time of requests to read data and distributed transactions, as well as the number of resources consumed for the selected DBMS by the following parameters, for which we introduce the notation:

- V - the amount of data in the database;
- C - the number of simultaneous connections;
- RP - the number of replica node

Resources consumed are many different types of resources, such as storage space, RAM, CPU time:

$$R = R_{HDD} \cup R_{RAM} \cup R_{CPU} \qquad (1)$$

This means that the following dependencies must be defined:

$$\begin{aligned} R_{HDD} &= f(V, C, RP) + \varepsilon \\ R_{HDD} &= f(V, C, RP) + \varepsilon \\ R_{HDD} &= f(V, C, RP) + \varepsilon \end{aligned} \qquad (2)$$

where ε - deviation in the measurement.

3. Analysis of methods of implementing of distributed transactions in different types of DBMS

When the first horizontally scalable databases appeared, there were no ready-made solutions for conducting distributed transactions directly from the DBMS box. So the first proposed method of implementing distributed transactions was the implementation of transactions at the application level, which at that time was based on the Two-phase-commit template [8]. The basis of this pattern is a layer between the application and the TransactionManager database, which implements a distributed transaction in two steps: checking whether the transaction can be committed for all nodes in the cluster, and then in case of a positive result directly isolated transaction commit on all nodes.

The advantage of this method is the independence of the database type and the ability to fully customize the transaction system for the needs of the application.

NoSQL databases were not originally designed to perform complex transactions, so they could not be fully used in applications where transactionality is critical (money transfers, reservations, etc.). Understanding this problem, NoSQL database developers are trying to add ACID transaction support for their products.

Logically, it is necessary to choose the most popular use NoSQL databases, which include the following: MongoDB; Cassandra; HBase; Redis;

Couchbase. To select a NoSQL database used the following criteria:

- popularity of the database (to determine the popularity of the resource db-engines with available estimates);
- CAP theorem principles (the number of principles supported by the DBMS);
- the presence of a built-in transactional system;
- type of license;
- bandwidth - the maximum number of read and write operations per second (for a cluster of two nodes).

To solve this problem will be choosing NoSQL database using linear additive reduction of weights, which is widely used for solving problems in different areas of choice [9], according to the formula:

$$Z^* = max \sum_{j=1}^{n} \alpha_j \, \beta_j a_{ij} \qquad (3)$$

where α_j – normalizing factors, β_j – weights.

Normalize all the values of the criteria for which we reduce them from different scales to one in the range [0; 1] given min and max:

$$f = \frac{f_{measure} - f_{min}}{f_{max} - f_{min}} \qquad (4)$$

Next, using the proportional method to determine the weights. As already mentioned, it is very important that the database is used by as many people as possible, and since we are primarily researching transactionality, it is also very important to maximize the functionality of the built-in transactionality. That is why we believe that the popularity of the database and built-in transactionality are 4 times more important than the license and bandwidth, and 2 times more important than following the CAP principles (because they are still quite important in terms of database architecture). The values of the weights and the results of the additive convolution are given in table.1.

Table 1

NoSQL database with criteria values in numeric format Popularity CAP Principles Built - in Transactionality License Bandwidth

	Popularity	CAP Principles	Built - in Transactionality	License	Bandwidth	Z^*
Weight.	0.33	0.17	0.33	0.08	0.08	
MongoDB	457.73	4	10	2	13,462.51	0.83
Cassandra	118.84	4	1	3	31,144.24	0.34
HBase	46.92	3	1	3	23,373.93	0.11
Redis	153.63	3	5	3	74,239.05	0.4
Couchbase	31.82	4	10	3	26,140.82	0.6

Relatively recently, a new type of NewSQL database was developed, the main purpose of which was to preserve the relational data model, but at the same time provide flexible scaling inherent in NoSQL database [10]. Therefore, when studying the implementation of distributed transactions, it is necessary to consider this type of database. You must also select a specific NewSQL database for the experiment. Although the choice of NewSQL is very narrow, the two most popular databases of this type are Clustrix and VoltDB. These databases generally have the same functionality, especially in terms of transactions and scaling, but differ only in internal implementation. VoltDB will be used for the experiment only because of its greater popularity and detailed documentation, as well as integration with a large number of programming languages.

Therefore, it is proposed to choose the following methods for experimental studies:

- study of the implementation of distributed transactions using the template Two-Phase Commit on the example of MS SQL Server;
- exploration and use of built-in capabilities of distributed transactionality in NoSQL databases on the example of MongoDB;
- research and use of NewSQL databases on the example of VoltDB.

4. Experiment planning

The planning of the experiment began with the selection of the subject area to demonstrate the applied nature of the experiment. It was decided that such a market segment as e-commerce was chosen as the subject area. This subject area is mostly associated with heavy workload and big data [11]. Also, most systems that sell goods or services online require transactional database, because they usually provide the functionality of creating orders and paying for them, which in themselves must be performed in a single transaction.

For the experiment, you need to create a database cluster with a variable number of replicas, which must be in a consistent state after the transaction.

These measurements should be made with different amounts of data and parallel users who will access the database at the same time. For greater accuracy, each request or transaction must be measured several times.

4.1 Constraints of the experiment

The first of these constraints is the use of Read Committed isolation Constraints by default for each approach. For MongoDB, this level of isolation is achieved by setting the *read concern* (with a snapshot value to read only matched data) and *write concern* (a value equal to the number of nodes in the cluster) [12].

Another constraint is the type of distributed database. Replication was chosen for the study, because it is on such a distributed basis that it is more appropriate to check the atomicity and consistency of transactions, because these transactions must be atomically and consistently recorded at all nodes of the cluster.

Measurements will be made for a cluster of virtual database server machines. Each server will run in the Azure cloud on a B2s virtual machine with 2 CPUs, 4 GB of RAM and 8 GB of hard disk space. Windows Server 2016 will be used as the operating system.

The final limitation is the amount of data, the number of simultaneous connections and the number of replicas for which measurements will be made. It is obvious that all these quantities can be infinite, and therefore for experiment it is necessary to allocate certain discrete quantities on which measurements will be carried out. It was decided to allocate certain modes that will represent specific discrete values of the above values.

For each load factor (number of data connections or nodes) will be conducted two types of measurements:

- only the values of this factor change, and the values of other factors remain static. This approach is needed to determine the dependence of execution time and resource consumption on a particular factor
- all factors change simultaneously. This approach is needed to determine the total impact of all factors. The values of all factors will increase by one level at a time.

For a more detailed experimental study, it was decided to use different volumes of databases and modes of interaction with them. Accordingly modes were identified from the point of view of different data indicators [13]:

- the number of entities in the database: Basic (122 064); Basic+ (568 064); Basic++ (1 266 064); Medium (2 420 064); Intensive (8 050 064)
- number of simultaneous connections: Basic (10); Basic+ (30); Basic++ (50); Medium (100); Intensive (300)
- number of nodes: Basic (2); Basic+ (3); Medium (4); Intensive (6).

4.2 Domain and business process modeling

As a result of the analysis of domain, its ER-diagram was developed (Fig. 1) and database structures for the relational data model, as well as for the BSON format used in MongoDB, which are given in [13].

E-commerce systems offer different functionality to their users. The main queries for measurement are transaction queries, but also for a more comprehensive analysis of the effectiveness of a database, it is also appropriate to measure read operations that do not require transactions. Consider the basic queries that conduct measurements.

The first such request was decided to read the order data by its ID. Next, it would be appropriate to implement a more complex read query with a complex predicate and grouping. In terms of application, such a request could be to obtain the three most popular categories of goods in the last month. The next queries will be adding an order with a variable number of order items. The number of order items to insert is logical to make equivalent to the number of items for this mode. Query to create an order is one of the most common in e-commerce, because this is the operation performed by customers of the system.

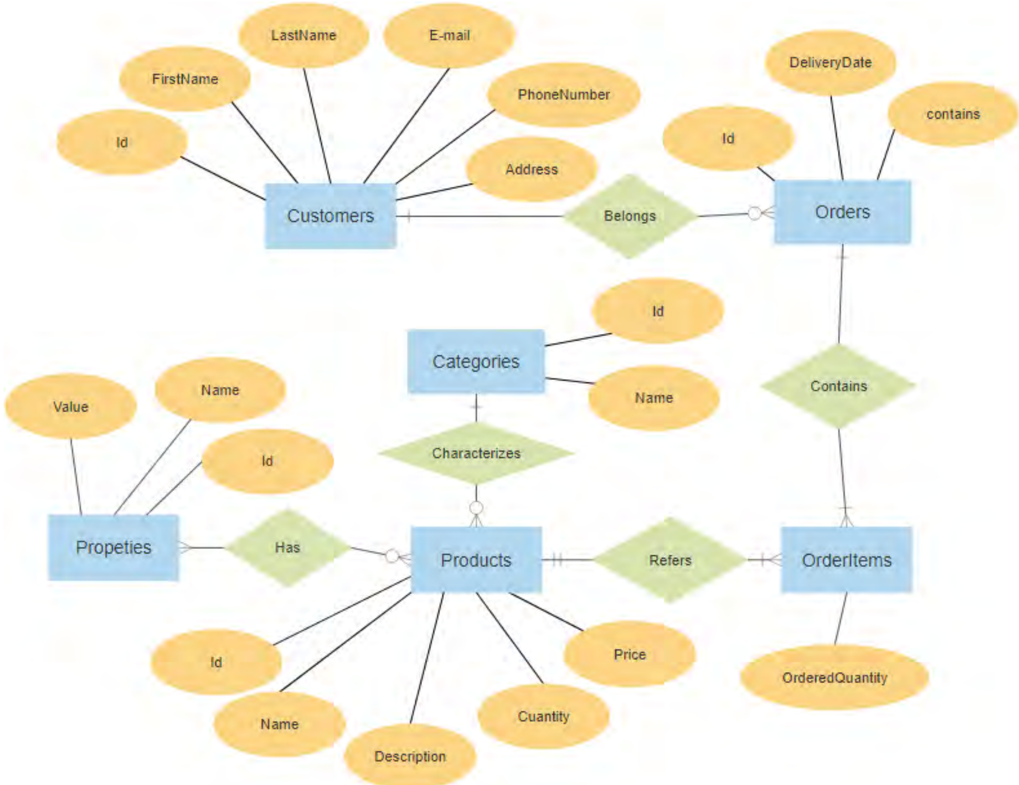


Figure 1: ER-diagram of the e-commerce subject area

The next query was decided to implement the operation of updating the goods, namely to reduce the price of those goods by 10%, the number of which in the warehouse is less than 10 units. This request makes sense, because very often the same e-shops, as usual, reduce the price of the finished product in order to sell the entire batch as soon as possible..

These queries are written in T-SQL languages and BSON format queries that will be used for measurements.

The results for MongoDB NoSQL database and VoltDB NewSQL database described in [13]. DBMS SQL Server is selected to implement the Two-Phase Commit.

5. Results of the experiment

After obtaining the measurement results, we present the main most interesting performance trends based on experiments to support ACID transactions for NoSQL DBMS MongoDB, NewSQL DBMS VoltDB and application-level transactions for SQL Server and C #.

First of all, a comparison of resource consumption of each of the DBMS is shown. The following is a comparative histogram that shows how many megabytes of memory the database uses for each mode for each load mode. In this case, the same number of entities is stored in each mode for all three DBMSs (Fig.2).

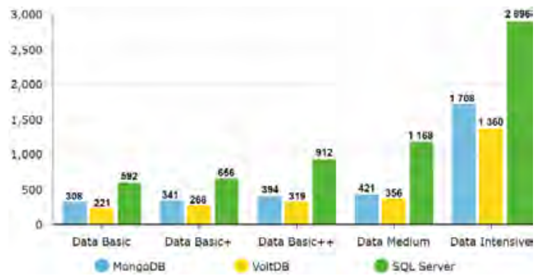
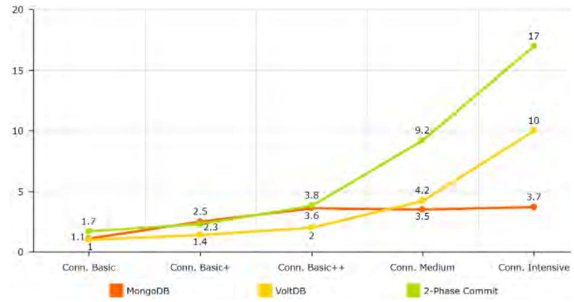


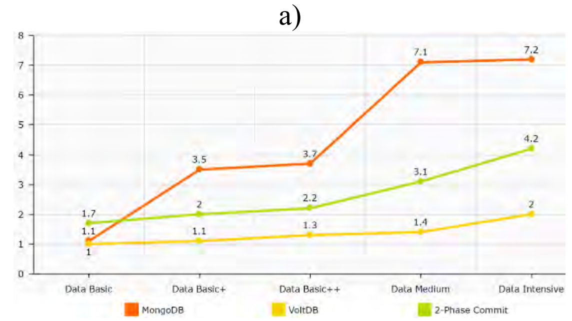
Figure 2: Comparative histogram of database size

Then give comparing the number of CPU time consumed by different methods when performing queries to different load modes. Note that CPU time consumption for all methods depends on the amount of data in the database. Figure 3a shows a graph comparing the growth of CPU time consumption with a gradual increase in the amount of data stored in the database.

The results of measurements with a variable number of simultaneous connections to the database are shown in Fig.3b. In this case, on the contrary, MongoDB shows a very efficient use of CPU time, while for SQL methods, the consumption of this resource increases quite rapidly with increasing connections.



a)



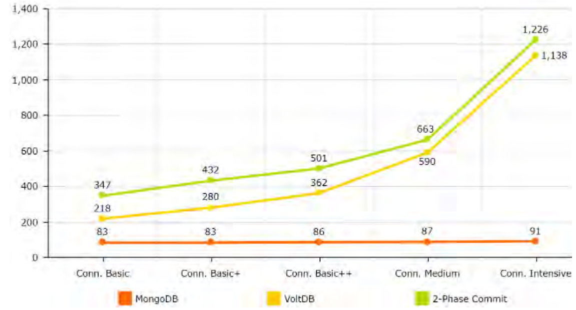
b)

Figure 3: Comparative graph of CPU time consumed when changing: a) the amount of data; b) the number of concurrent connections

A comparison of RAM usage is shown in Figure 4a. From the above data, we can see that MongoDB consumes several times less RAM than VoltDB and Two-Phase Commit. For example, when changing the amount of data, MongoDB consumes from 91 MB of RAM for the lowest level and up to 99 MB for the highest level, and when changing the number of connections from 83 MB to 97 MB. Even in the presence of a statistical error shows that growth, if there is, it is very slow. At the same time, VoltDB consumes when changing data from 218 MB to 560 MB, and when changing connections from 218 MB to 1138 MB. And the most needed implementation through Two-Phase and SQL Server, which consumes from 347 MB to 581 MB when changing data and from 347 MB to 1226 MB when changing connections number.

Changing the number of concurrent connections remain the same trends as in the previous case. MongoDB consumes up to 100 MB of RAM at all load levels, which is several orders of magnitude lower than other methods. VoltDB and Two-Phase Commit have very similar values and grow over time at about the same rate, but all in absolute numbers, VoltDB uses RAM more efficiently. The corresponding results are shown in Figure 4b.

a)



b)

Figure 4: Comparative charts of RAM usage by changing: a) the amount of data; b) the number of concurrent connections

Similar results are obtained by changing all load factors at the same time, but in this case the memory consumption for MongoDB still increases significantly, however, still much less than for the other two DBMS. Consumption and growth trends for VoltDB and Two-Phase Commit remain the same, but the latter consumes more RAM in absolute numbers. The corresponding results are shown in the histogram in Figure 5.

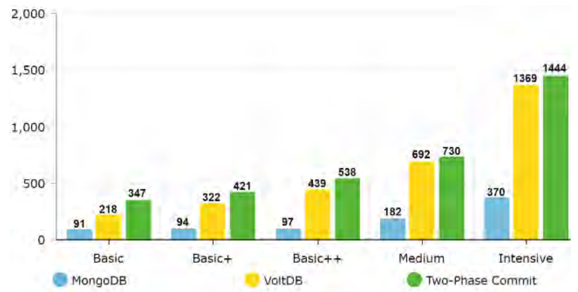
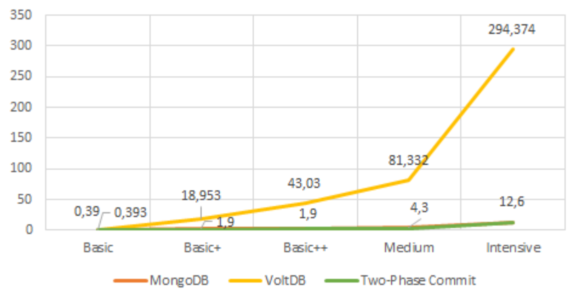


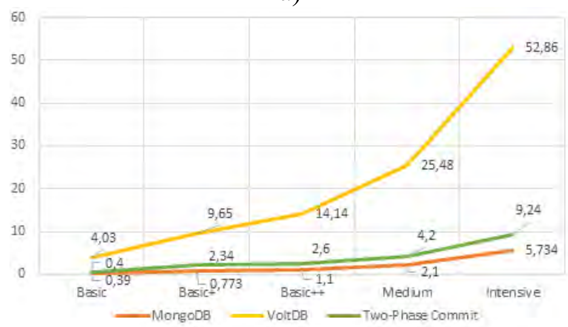
Figure 5: Comparative histogram of RAM consumption by changing all load factors

It should also be noted that all 3 methods are very weakly dependent on the number of replicas in the cluster in terms of RAM consumption.

Let's analyze the results for data read requests. Note that for all three DBMSs, the query execution time does not depend on the amount of data in the database, because the index read operation has a logarithmic complexity, and depends on the number of replicas in the cluster, because you only need to connect to one database.Consider the read request for obtaining the three most popular categories of goods, which includes operations of grouping, sorting and limiting data. When you run this query, all methods have a characteristic increase in runtime from the amount of data and the number of concurrent connections. On the other hand, the execution time for all methods depends on the number of simultaneous connections, as shown in Figure 6a. Slightly different situation with the execution time of this request when changing the factor of the number of simultaneous connections. As in the previous case, VoltDB executes the request at all levels more slowly than the other two DBMS. But on the other hand, MongoDB executes the request faster than the Two-Phase Commit, and this difference becomes more noticeable with each subsequent level. This once again demonstrates the effectiveness of MongoDB at high loads from simultaneous user connections. The results of measurements are shown in Figure 6b.



a)

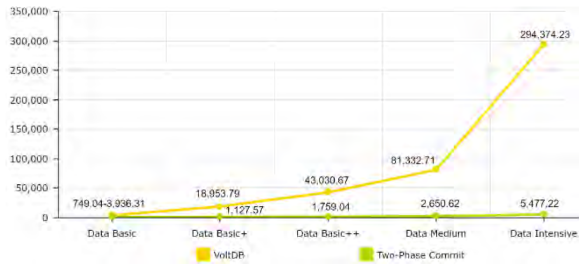


b)

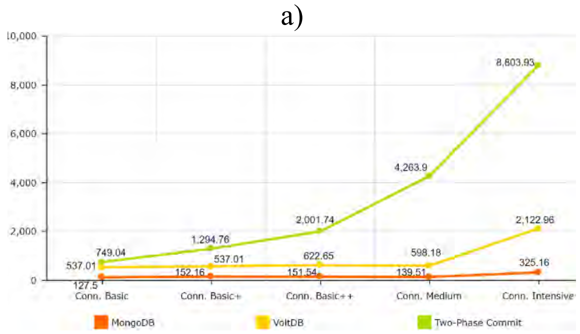
Figure 6: Comparative charts of execution time of the query for obtaining categories of goods when changing a) the amount of data; b) number of connections

In the simultaneous change of all factors VoltDB also remains the least efficient DBMS for this complex read query, while the Two-Phase Commit method via SQL Server gives better results at the first 4 levels, but at peak load MongoDB is significantly inferior (113.86ms vs. 80.73ms for SQL Server).

We present the result of the analysis for the transaction of insertion of the order with its elements. The experiment showed that MongoDB executes this transaction several times faster than VoltDB and Two-Phase Commit when changing any factor. It was also found that the insertion time for MongoDB does not depend on the amount of data in the database and varies from 86.76 ms to 133.09 ms at different levels. At the same time, the VoltDB and Two-Phase Commit methods depend on the amount of data in the database at runtime, but in absolute terms and growth rates, VoltDB loses because the queries for this method take much longer (Fig. 7a). And at change of quantity of simultaneous connections execution time for all DBMS steadily increases. The longest transaction is performed by Two-Phase Commit, followed by VoltDB, and the fastest is performed by MongoDB (Fig. 7b).



a)



b)

Figure 7: Comparative charts of the execution time of the transaction on the insert when changing a) the amount of data; b) number of connections

It should be noted that with the simultaneous change of all load factors, the absolute values are slightly higher, but the general trend remains the same as in the case of a change in the number of connections.

Only MongoDB and Two-Phase depend on the number of replicas factor. VoltDB does not depend on this factor and executes the transaction when this factor changes from 771.9 to 847.57ms.

Data update operations were also investigated. To complete this transaction, you will need to scan the entire table or collection of products. With a variable amount of data in the database, the time for all methods increases. At the same time, this request is performed most quickly by VoltDB, and Two-Phase Commit works faster than MongoDB only at the initial load levels, but has a much faster growth of this time. the situation when the number of simultaneous connections to the database will gradually increase is considered. The results are shown in Fig. 8.

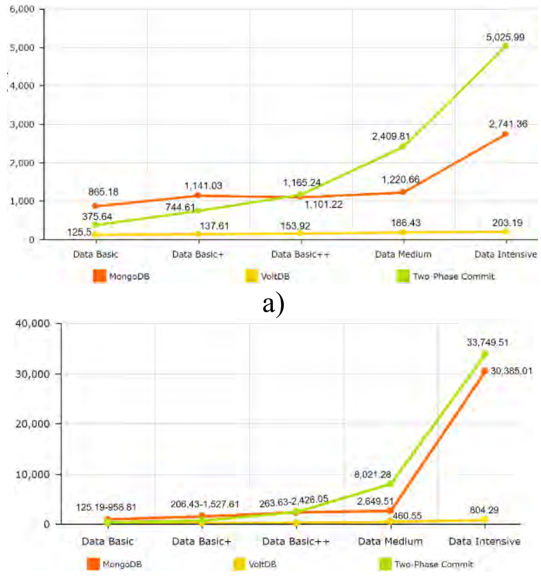


Figure 8: Comparative graph of the execution time of the transaction for the update when changing a) the amount of data; b) number of connections

Although the values have increased, but the trends have remained: the query is executed the fastest by VoltDB, then at the initial levels the Two-Phase Commit works faster, but has a faster growth of time.

It was found that VoltDB is almost independent of the number of replicas in the database cluster when changing the number of replicas. At the same time, the use of the other two databases requires significantly more time, but when you change the number of replicas, it executes the SQL Server query faster, and has a time growth rate less than MongoDB.

6. Recommendations

After a comparative analysis of all methods of implementing distributed transactions and identifying the weaknesses of each of them, it is possible to develop practical recommendations for the use of a method in a particular situation [13]. Recommendations are given in table.2.

Two-Phase Commit pattern was one of the first methods of implementing distributed ACID transactions, and therefore has many drawbacks. And given the performance of query and transaction time and resource consumption, we can conclude that when developing an application of any type it is best not to use this method of implementing distributed ACID transactions, because this implementation will be quite expensive and will not give the expected performance. Two-Phase Commit is implemented using SQL Server database and C # programming language.

Table 2
Recommendations for the use of the studied methods

Methods Conclusions	Two-Phase Commit	MongoDB	VoltDB
regarding the use of resources	Requires much more resources of all types: storage space on storage media, amount of CPU time and RAM	Consumes less RAM than other methods	Consumes less storage space than other methods and requires slightly less CPU time than MongoDB; the load consumed several times more RAM in any mode
regarding transaction processing	Showed the best performance only on read requests only at a certain type of load, but	Performs a simple read requests by index and complex queries is much faster;	Performs orders of magnitude faster transactions in which update operations prevail over other operations;

	even then the gain in time compared to the nearest competitor was minimal	very effectively handles a large number of concurrent connections for most data change operations	the factor of the number of replica nodes in the cluster does not affect the speed of most requests
regarding the scope and modes of application	Can be used in domains where there is an urgent need to configure the algorithm for conducting and recording transactions in a distributed DB, but implementation in the project will require funds for the rental of powerful virtual machines to host DBMS servers and applications	More efficient when the same virtual machine must host one or more applications that require a lot of RAM to run continuously; other repositories that require a lot of RAM must be on the same virtual machine.	It is recommended if the database should store tens of gigabytes of data, which are not very often or intensively processed, and the speed of this processing isn’t a priority; effective if one machine stores data from other DB or files, or one or more applications are deployed on the same virtual machine with a database server that requires constant CPU and multi-threaded computing.t
regarding examples of application systems	Systems in which you need to implement an exceptional level of transaction isolation that is not supported by a particular database, or when you need to record transactions on nodes in a specific order (for example, depending on the geolocation of the server)	Web servers that cache the client's response to expedite the response; centralized data logging systems; control systems for various environmental metrics via the IoT, ticket booking; banking applications with the function of money transfer; applications for building graphical objects, graphs and charts based on numerical data; e-commerce systems and online competition systems	Storing backup copies of databases or outdated data; for solving classification problems in computer vision systems; antivirus systems or search bots that are constantly active; applications for co-editing documents; applications for financial exchanges; systems displaying various data and metrics in real time, from which obsolete data should be removed as quickly as possible to avoid taking up disk space.

In a situation where it is not known in advance for the projected application which type of load will prevail, it will be more appropriate to use MongoDB, because when all load factors change for the vast majority of requests and transactions, this method executes them faster than any other, considered.

7. Conclusions

The article examines the methods of implementing distributed ACID transactions based on the use of relational DBMS MS SQL Server, NoSQL DBMS MongoDB and NewSQL DBMS VoltDB in terms of performance.

Selected DBMS are relevant tools for the implementation of distributed transactions and represent popular areas in the field of databases, that are widely used in modern applications.

For the experiments ER-model for the popular field of e-commerce was developed and appropriate database structures for the relational and BSON models, as well as transactions to these databases, were designed.

Nonlinear dependences of growth of the consumed resources on simultaneous change of all factors at performance of all types of transactions for all DBMS are established. As a result of the study, a linear dependence of the increase in the execution time of transactions for data update when changing the number of replicas for Mondodb and Voltdb was established/

Experimental studies of the effectiveness of selected methods were carried out with increasing load, allowing to compare not only the absolute values of metrics, but their trends. During the experiments used metrics on query execution time and resource required to fulfill these requests. The analysis of these metrics have been identified comparative advantages and disadvantages of each approach. The analysis of these metrics have been identified comparative advantages and disadvantages of each approach.

Thus, based on the results of comparing different aspects of performance of methods of implementation of distributed ACID transactions, practical recommendations for their use for a specific type of application were developed. These recommendations can be used in the design and development of of more efficient application software.

8. References

[1] Tamer Özsu M. Principles of Distributed Database Systems. Springer International Publishing, 2020. 674 p.

[2] Blokdyk G. ACID Transactions Second Edition. 5STARCooks, 2018. 282 p.

[3] Kuzochkina A., Shirokopetleva M., Dudar Z., Analyzing and Comparison of NoSQL DBMS. Proceedings of the International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 2018, pp. 560–564. doi: 10.1109/INFOCOMMST.2018.86 32133.

[4] Sadalage P., Fowler M. NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, 1st Edition. Addison-Wesley Professional, 2012. 192 p.

[5] Moniruzzaman A B M, Hossain S. A. NoSQL Database: New Era of Databases for Big data Analytics – Classification, Characteristics and Comparison. Int J Database Theor Appl (2013). 6.

[6] Saga distributed transactions. Documentation. URL: https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/saga/saga

[7] Bhiri, Sami & Gaaloul, Khaled & Perrin, Olivier & Godart, Claude. Overview of Transactional Patterns : Combining Workflow Flexibility and Transactional Reliability for Composite Web Services (2005). 440-445. doi:10.1007/11538394_37.

[8] Teresa K. Abuya. Handling distributed transaction failures: A practical approach. – LAP LAMBERT Academic Publishing, 2016.

[9] Mazurova O., Samantsov O., Topchii O., Shirokopetleva M., A Study of Optimization Models for Creation of Artificial Intelligence for the Computer Game in the Tower Defense Genre, Proceedings of 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 2020, pp. 491-496, doi: 10.1109/PICST51311.2020.9468057.

[10] Astrova I., Koschel A., Wellermann N., Klostermeyer P. Performance Benchmarking of NewSQL Databases with Yahoo Cloud Serving Benchmark. Proceedings of the Future Technologies Conference (FTC) 2020, Volume 2. FTC 2020. Advances in Intelligent Systems and Computing, vol 1289. Springer, Cham. doi: 10.1007/978-3-030-63089-8_17

[11] Leshchynskyi V., Leshchynska I. (2020). Principles of explanation in e-commerce system based on sales dynamics. (2020). URL: http://csitic.nure.ua/article/view/201452. doi: 10.30837/IVcsitic2020201452.

[12] MongoDB Causal Consistency and Read and Write Concerns. Documentation. URL: https://docs.mongodb.com/manual/core/causal-consistency-read-write-concerns/

[13] Mazurova, O., Naboka, A. and Shirokopetleva, M. Research of ACID transaction implementation methods for distributed databases using replication technology. Innovative Technologies and Scientific Solutions for Industries. 2 (16) (Jun. 2021), 19-31. doi: 10.30837/ITSSI.2021.16.019.

Information Technology for Environmental Monitoring Systems in Conditions of Complex Borders of Territories

Kyrychenko Iryna [1], Tereshchenko Glib [2]

[1, 2] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract
As the problem of environmental protection has become international, one of the main tasks of regional governments is to increase the level of environmental and man-made safety in the region, which means, above all, the need to prevent emergencies and reduce their material and social consequences. Therefore, the creation of effective environmental monitoring systems should provide the solution of two problems – the creation of effective systems of information and intellectual support for decision-making and modeling of the environmental situation based on monitoring data.

Keywords
Approximation, Association, GIS System, Environmental Monitoring, Pollution Locus, Interpolation Function, R-Functions

1. Introduction

Nowadays, the rapid progress of science and technology, on the one hand, has allowed to meet all the needs of human society, but on the other hand, it has worsened the conditions of its existence. The solution of environmental assessment tasks is necessary to develop new models and methods focused on solving the problems of computational geometry [1-3] and mathematical physics [4], where there are problems of modeling turbulent movement in the atmosphere [5] and the ground layer [6].

The information and analytical system (IAS) for risk management of emergency situations (ES) should be an organizationally ordered set of software and hardware, information resources and technologies that enables the integration of data of knowledge about regions of Ukraine for their effective use in solving research and administrative tasks associated with analysis, modeling, forecasting and risk management of natural and man-made disasters. IAS manages emergency risks, integrates in its composition technologies for monitoring the natural, man-made, and social environment in the interests of emergency prevention and response; transmission, collection and storage of information on the sources of emergencies and life support facilities, analysis, modeling and forecasting the consequences of emergencies.

Information technology should be the basis for analytical support for decision-making in the Situational Territorial Center for Monitoring and Managing Emergencies. The functional basis of mais to be a set of its information processes, each of which solves a certain task of collecting, analyzing, integrating, disseminating, displaying spatial-oriented data, as well as integrating data and knowledge to develop management decisions.

The main technologies of this type include geoinformation systems (GIS), connected with Internet data collection systems from stationary and mobile sources of information (satellites, aviation data collection systems, meteorological stations, etc.), which contain electronic maps of the area, demographic, meteorological and other data, the use of which is necessary to obtain adequate solutions [7-11].

2. Use of geoinformation systems for data collection and forecasting of ecological situation

When examining the process of environmental pollution by emissions of industrial waste, it is necessary to assess the impact of harmful

EMAIL: iryna.kyrychenko@nure.ua (A. 1);
hlib.tereshchenko@nure.ua (A. 2)
ORCID: 0000-0002-7686-6439 (A. 1);
0000-0001-8731-2135 (A. 2)

pollution on the biological environment. Therefore, an important task at this time is the forecast of changes in environmental systems under the influence of natural and anthropogenic factors [12].

The development of methods for preliminary calculation of areas of possible placement of industrial enterprises, residential areas and recreation areas in compliance with sanitary pollution standards is relevant.

Mathematical models and algorithmic support of modern emergency decision-making system should include:

- system of deciding on the fact before an emergency or accident, based on comparison in real time of ecosystem information and reference data;
- system of accumulation of knowledge about accidents and before emergencies, which is necessary for the development of the expert system;
- emergency or accident elimination scenario development system.

Such a decision support system can be based on knowledge bases, which is explained by the need to overcome the barrier of human perception. The Knowledge Base will offer the dispatcher various ready-made decision-making scenarios. Also, the decision support system can be based on special rules, frames or logical rules. By integrating operating systems — both digital and conventional — can achieve significant success, especially in planning and at the stages of monitoring and eliminating the consequences of the ES. Some of the advantages are that [8]:

- the use of GIS modeling reduces field research work;
- saves time and cost of project discussion;
- accelerates the receipt of income due to more prompt elimination of the consequences of the ES;
- modern, easily accessible satellite data is used to solving emerging problems, especially for large areas;
- spatial databases on the environment and socio-economic aspects are created, which can be easily updated and changed, as well as shared with various services of the Ministry of Emergency Situations;
- the time to create programs that include spatial data is reduced compared to using more traditional methods;
- the digital spatial data used in the dialog mode makes it very easy to select a track to the place of the ES and solving other issues;
- when solving emergencies, the cost of indemnification is reduced;
- property databases are created to help in solving future prospects of use.

Since in GIS graphic objects are associated with tabular data, both graphic and semantics need to be processed simultaneously and at the same time efficiently processed, which complicates the gigantic amount of information. It becomes clear that the creation of a good GIS - the process is quite complex, requiring greater effort and the use of innovative solutions.

3. Mathematical Tools

The presence of a geoinformation system and a distributed database of the territorial information and analytical system is a necessary, but insufficient condition for the creation of a regional system and, moreover, aninterregional environmental monitoring system.

For a comprehensive solution to the problems of large-scale system management, it is advisable to use an intelligent integrated environment that combines means of intelligent data analysis [13] and geoinformation system.

Intelligent Data Analysis (IDA) is a new direction in the field of information systems. Its purpose is to identify hidden patterns and rules in super-large arrays of heterogeneous data and the possibility of their application to solve decision support tasks [14].

This process consists of three main stages:

- research (detection of patterns, previously unknown relationships);
- use of detected patterns for model construction;
- analysis of exceptions to identify and explain deviations in the found patterns.

The findings turn into information that is characterized as knowledge that can continue to be applied to new data. Finding new knowledge by means of IDA is a direction that uses the methods of artificial intelligence, mathematics and statistics.

When modeling the ecological situation in regions with a complex border shape and the most significant sources of pollution, we will use normal equations of drawings, R – equilibrium and the Inter-location formula of Lagrange. The form of the boundary of the region in which we will conduct research on the ecological situation, we approximinate with the help of a certain number of segments, selected experimentally from the conditions of the performance of the task and the result obtained, and present a closed broken one.

The normal equation of a closed polygonal with vertices in points $M_1(x_1,\ y_1)$, $M_2(x_2,\ y_2)$, $\ldots$, $M_n(x_n,\ y_n)$, we will write as follows:

$$f(x,\ y) \equiv \underset{i\,=\,1}{\overset{n\,-\,1}{\sim}} \varphi(x,\ y,\ x_i,\ y_i,\ x_{i+1},\ y_{i+1}) \sim$$
$$\sim \varphi(x,\ y,\ x_n,\ y_n, x_1,\ y_1). \quad (1)$$

Due to the associativity of the equivalence operation, we can include in the general construction the function of the region boundary normal functions of the boundary areas in any sequence. At the same time, the effect of symmetry of the region's picture will be achieved if the shape and boundary conditions have the same type of symmetry.

With the help of this approach, the equation of the boundary of Ukraine is constructed $\omega_{Ukraine} = 0$. In this case, 310 straight segments were used [5].

The functional reflection of the boundary of Ukraine by a single analytical expression allows us to further include its (or equations of boundary areas) in the solution of the algorithm.

When constructing the pollution function of the region, it is advisable to use the Inter-location formula:

$$u(x,\ y) = u_0 + u_1 = \frac{\overset{n}{\underset{i=1}{\sim}} \omega_i}{\overset{n}{\underset{i=1}{\sim}} \frac{\omega_i}{\phi_i}} + \left(\underset{i\,=\,1}{\overset{n}{\sim}} \omega_i \right) \Phi, \quad (2)$$

where $\omega_i(x,\ y) = 0$ - equation of the i-th source with a known pollution level.

Φ – undefined component included in the residual member. It can be appproximed by some polynomy with undefined coefficients that are at a minimum of functionality (if it is known) on many functions that adopt the specified values of elements on the locuses, which are the areas of sources of pollution.

In this way, it has been proven that to build a mathematical model for the spread of contaminants in a region with a complex shape of the boundary, you can use the equations of R-Functions theory to build normal locus equations. In this case, it is possible to use the interlocation formula for constructing the pollution function considering the available experimental data, as well as expanding the problem to be solved on the basis of introducing new data into it, for example, considering the presence of wind, the nature of attenuation, the function of the pollution level, etc.

The tasks of localization of a point, or the task of checking the belonging of a point of a certain area, arise, for example, when it is necessary to check that some settlement falls (or does not fall) into the zone of influence of the factors of defeat when this zone is specified by a certain area.

4. Intelligent data analysis in environmental monitoring tasks

The development of the software is designed for interactive work in real time and should ensure the display of data, calculations, and display of results with an accuracy of up to the minutes of the arc. All basic data and processing results, to facilitate their perception and improve the quality of analysis, are reflected in graphical and textual form. The software package should be open to expansion on the part of users.

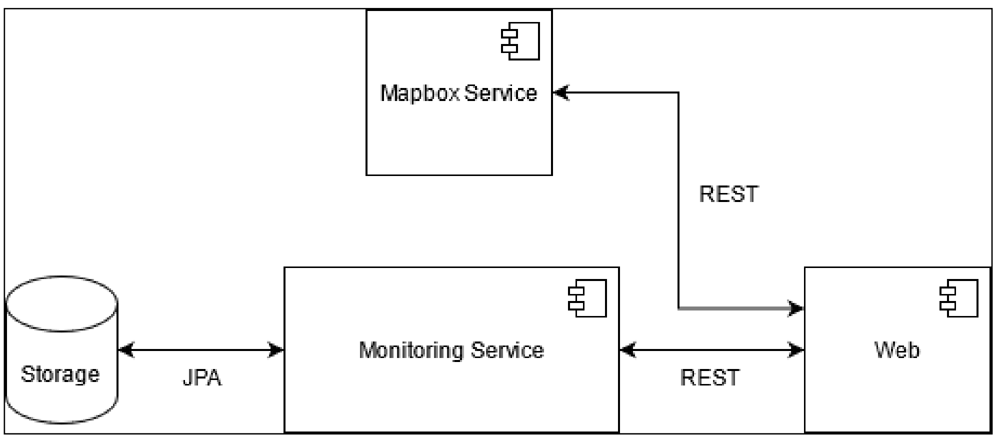


Figure 1: Architecture of development of the solution

The main modules of the designed software system are as follows:

- geographical map of the territory in question – allows you to get acquainted with the peculiarities of the terrain and the location of settlements.
- database about settlements and others andobjects – allows you to display, drive and edit information about settlements and other objects.
- database about the centers of influence – allows you to display, drive and edit information about the centers of influence.
- construction of areas of influence cells – allows to build and display: the area of influence of the cell on settlements and other objects (additionally, a list of settlements under the influence of the cell is displayed); convex shell and level lines (an approximating polygon for arbitrary set of impact cells, displaying them on the map along with polygon level lines, specifying the boundary of the image area and prohibiting the zone).

Construction of regions with the display of impact cells is the main mode of work with the software complex. The main form is designed to be built and displayed on the map: areas of influence on settlements and other objects (additionally, a list of affected settlements is displayed); a convex shell of influence cells (an appeximating polygon for an arbitrary set of influence cells).

You can perform the following actions: display, enter and edit information about settlements and parameters of impact cells; map settlements and impact cells; build an area of influence and an approximation polygon for arbitrary set of impact cells, display them on the map along with polygonal level lines, set the boundary of the image area and prohibit the zone to display (if necessary for drawing).

The software complex is designed for interactive work in real time and provides data display, calculations and display of results with accuracy up to the minutes of the arc.

All basic data and processing results, to facilitate their perception and improve the quality of analysis, are displayed in graphical and text form.

The main modules of the designed software system are as follows:

• Service processor - provides an API for working with data storage and calculation logic (back-end);

• UI web application (front-end);

• service provider of online maps for websites and applications.

The construction of areas of influence cells allows you to build and display: the area of influence of the campfire on settlements and other objects (additionally, a list of settlements under the influence of the campfire is displayed); Convex shell and level lines (an approximating polygon for an arbitrary set of impact foci, displaying them on the map along with polygon level lines, specifying the boundary of the image area, and prohibiting zone display).

The user is provided with a set of service capabilities for working with record tables.

5. Conclusions

To solve the actual problem of modeling the ecological situation in Ukraine, an analysis of the problem was carried out, on the basis of which the main tasks of the study were formulated and solved. It has been shown that in order to solve the problems of modeling the environmental situation, it is necessary to create a GIS, the use of which, together with a distributed database of the territorial information and analytical system of environmental monitoring, ensures an increase in the efficiency of data collection and forecasting of the environmental situation, as well as a reflection of the results of the solution directly on the terrain map. On this basis, an experimental version of GIS and its corresponding database is proposed, which are used to provide initial data when modeling the situation and displaying the results of the solution on the map of Ukraine.

To assess the environmental situation anywhere in Ukraine, modeling methods based on the interpolation of environmental monitoring and forecasting of the intensity distribution of individual emissions based on the cone of the spread of pollutant, representing the solution of the diffusion equation, are proposed.

The proposed methods allow predicting levels of environmental pollution based on two-dimensional interpolation methods. Depending on the task, they provide modeling of the ecological situation locally and globally (for the whole territory of Ukraine) — due to the possibility of linear interpolation by the values of pollution levels in certain places, as well as polynomial Lagrange interpolation, by building a regular grid of pollution levels on this primary data, and building level lines for the function of pollution intensity. In this case, the user can set the interpolation accuracy, according to the density of the sensors, based on the choice of order that interpolates the polynomy in the range from linear interpolation to polyonomy.

The proposed modeling algorithms are implemented in the form of a software package that provides real-time modeling of environmental pollution levels.

6. References

[1] I. Shubin, Development of Conjunctive Decomposition Tools, in: Systems Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems, COLINS 2021, Volume. I, yyy. 890-900.

[2] V. V. Akimenko, D. V. Kayani, A. B. Symonenko, Means of computer graphics in the task of analysis and control of the ecological state of the atmosphere of the region, Software products and systems 3 (1995) 10-12.

[3] Advanced remote sensing: extraction and application of ground information, 2018, p. 248. URL: http://edutechsoft.blogspot.com/2018/09/advanced-remote-sensing-terrestrial.html.

[4] F. Preparatu, M. Sheymos, Obchyslyuval′na heometriya, Myr, Moskva, 1989.

[5] R. A. Uvarov, Modeling of the ecological situation taking into account turbulent motion in the atmosphere, Radioelectronics and Informatics 3 (2001) 129 – 134.

[6] R. A. Uvarov, Mathematical modeling of ecological situation in Ukraine, Radioelectronics and Informatics 4 (2000) 125 – 129.

[7] S. Solonska, V. Zhyrnov, Adaptive semantic analysis of radar data using fuzzy transform (Book Chapter). Springer, Lecture Notes on Data Engineering and Communications Technologies. 48 (2020) 157–179.

[8] S.Solonskaya, V. Zhirnov, Intelligent analysis of radar data based on fuzzy transforms // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika) 77 (15) (2018) 1321–1329.

[9] N. Sharonova, I. Kyrychenko, G. Tereshchenko. Application of Big Data Methods in E-Learning Systems, in: CEUR Workshop Proceedings 2870 of the 5rd International Conference on Computational Linguistics and Intelligent Systems, COLINS-2021, Volume I, pp. 1302-1311.

[10] I. Shubin, V. Skovorodnikova, A. Kozyriev, Mining Methods for Adaptation Metrics in E-Learning, in: Proceedings of the 3rd International Conference Computational Linguistics and Intelligent Systems CoLInS 2019, NTU “KhPI”, Kharkiv, (2019) 288–300.

[11] T. Imielinski, W. Jr. Lipski, The Relational Model of Data and Cylindric Algebras, J. of Comp. and Syst. Sci. 28 (1984) 80–102.

[12] M. Makaruk, A. Nazarov, I. Shubin, N. Shanidze, Knowledge Representation Method for Object Recognition in Nonlinear Radar, in: Proceedings of the 5th International Conference on The Computational Linguistics and Intelligent Systems, COLINS 2021, V. I, pp. 948-958.

[13] G. Proniuk, N. Geseleva, I. Kyrychenko, G. Tereshchenko, Spatial Interpretation of the Notion of Relation and Its Application in the System of Artificial Intelligence, in: Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Systems, COLINS-2019, Kharkiv, Ukraine, pp. 266-276.

[14] Ye. G. Klimenko, Programmno-algoritmicheskiye sredstva intellektual'nogo analiza dannykh, Radioelektronika i informatika 3 (2001) 64-67.

Advantages and Disadvantages of Tools for Parallelism Implementation in The Ruby Programming Language

Dudar Zoia [1], Kyrychenko Iryna [2], Titova Margarita [3] and Kobziev Volodymyr [4]

[1, 2, 3, 4] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract

The paper considers the methods of parallelization in the Ruby programming language, namely parallelization in the MRI Ruby environment and parallelization in the JRuby environment, analyzes the execution time of parallel programs in different ways, reveals the difference between sequential and parallel program execution, compares with parallelization in C ++ programming language. The peculiarities of using Global Interpreter Lock and its connection with the results of a multithreaded program are considered. This paper presents the results of programs running on Windows 10 (x64) and Intel Core i5-8265U (1.60GHz). Multiple threads within a single process have significantly lower overhead than the corresponding number of processes because they share address space and memory. Global Interpreter Lock is a mechanism used in computer language interpreters to synchronize thread execution so that only one thread can run at a time. An interpreter that uses GIL always allows only one thread and one thread to run at a time, even if it runs on a multi-core processor. Ruby-MRI and CPython are the two most common examples of popular GIL translators. However, if you have the ability to use a version other than CRuby, you can use an alternative Ruby implementation, such as JRuby or Rubinius, because they don't have GIL and they support true parallel streaming Ruby. Improved multi-threaded performance can lead someone to believe that we can keep adding more threads - mostly indefinitely - to keep making our code faster and faster. It would be nice if it were true, but the reality is that streams are not free, so sooner or later you will run out of resources.

Keywords

Parallelization, Sequential Execution, Multithreaded Programs, Parallel Execution

1. Introduction

Recently, the problem of using parallel approach has been facing developers, because if you want to write code for anything but the smallest and weakest systems, you will write parallel code. However, the importance of using parallel programming in the Ruby language has not been sufficiently appreciated [3-5], since many developers bypass this issue due to the presence of GIL. Therefore, consideration of this problem is relevant and will help in the further development of complex problems.

In recent years, worldwide attention to increasing the use of high-performance computing. First, the number of users of personal miniclusters is growing, and secondly, there is a concentration of powerful computing resources in the centers of collective use and the development of remote access infrastructure using telecommunications. Finally, everyone is now aware of the fact that a further increase in computer performance only by improving the performance of electronics elements has reached the limit determined by physical laws. Further increase in productivity is possible only by parallelizing information processing processes.

Nowadays, computer technology is actively developing and many new tasks that require complex calculations are emerging, such as neural network training, computer modeling and cryptography tasks, telescope data processing, object recognition tasks, and more. Of course, in order to keep up with this development, it is necessary to change the methods of solving and calculating these problems. Reducing the time to

EMAIL: zoia.dudar@nure.ua (A. 1); iryna.kyrychenko@nure.ua (A. 2); marharyta.titova@nure.ua (A. 3); volodymyr.kobziev@nure.ua (A. 4)
ORCID: 0000-0001-5728-9253 (A. 1); 0000-0002-7686-6439 (A 2); 0000-0001-5951-3036 (A. 3); (0000-0002-8303-1595 (A. 4)

perform a large amount of work by breaking it down into separate jobs that can be performed independently and simultaneously is used in many industries. In the field of computer technology, this is achieved by increasing the number of simultaneously running processors. However, it is known that the transfer of a conventional program to a multiprocessor computer system may not give the expected performance gain. Moreover, as a result of such transfer the program can work more slowly.

Experience has shown that it is extremely difficult to write an effective parallel program or to adapt an existing sequential program for parallel calculations. This is due to the fact that the transition to the use of multiprocessor systems is characterized by a fundamentally new content. The purpose of this paper is to find out whether the Ruby programming language is suitable for creating parallel programs, whether it is comparable to other languages (in particular, the high-level C++ language), whether there is a difference between sequential and parallel program execution and how significant it is, what tools promote parallelism in Ruby and what are their strengths and weaknesses.

2. Literature overview

Concurrency and parallelism are not the main strengths of Ruby, but even in this department the language does offer tools that are probably good enough to deal with most use cases. Many Ruby developers ignore threads, although this is a very useful tool.

Ivan Shikhalev in his article “Ruby and multitasking” gives an overview of threads and processes in the default implementation of the language. The author highlights the following advantages and disadvantages: “When using streams, we remain within the same process, which determines both the pros and cons: on the one hand, the possibility of using shared variables, on the other, mutual dependence. An additional disadvantage of the ruby implementation is global. a blocker due to which only one thread is actually executing at any given time.” [1]. If we consider the article by Gernot Gradwohl “Scalable Ruby — Concurrency and Parallelism Explained”, then the purpose of his work is to show how to make use of concurrency in Ruby within one process, and to show the difference between parallelism and concurrency. The author explains concurrency with the example of Node and compares it to parallelism in the Ruby programming language. Disadvantages compared to Node concurrency in his opinion: “Every thread maps to one OS thread. Therefore, it is not possible to create hundreds, or even thousands, of threads. The developer has to do more ceremony. In other words, he has to do a lot of the stuff more explicitly.” [2]. Both authors say that there is no parallelism in Ruby since there is a Global Interpreter Lock (if we are talking about the default implementation of the Ruby language - MRI). But the VM has special optimizations for some operations, including I/O. Where a thread is waiting for I/O and then Ruby passes control to the next thread. In addition, threads can invoke non-Ruby code, such as native C extensions, and they can run in parallel. Therefore, in this article I want to show that parallelism in Ruby is available and not inferior to other programming languages, such as, for example, C++.

3. The relevance of research

The demand for computers operating at ever higher speeds is constantly growing. Pharmacists are constantly developing new drugs and medicaments using computers. Astronomers are trying to reproduce the history of the universe, from the big bang to the present day. Aircraft designers could get more accurate results without building huge wind tunnels, but simply designing their designs on a computer. Despite the increasing power of computers, their capabilities will never be enough to solve many non-trivial tasks (industrial, scientific, technical).

Despite the increasing growth in the speed of processors, the problem of information transfer speed is still encountered. This is due to the fact that the speed of propagation of electromagnetic waves in copper wires and light in optical fiber cables still remains equal to 20 cm/ns. In addition, the faster the processor works, the more it heats up, so the task of protecting it from overheating arises.

Computer designers are working to improve the performance of their machines. One of the ways to make processors speed up is to increase their frequency, but this method has technological limitations. Therefore, to improve performance at a given processor clock speed, most developers use parallelism, which consists in the simultaneous execution of two or more operations.

Consider two main forms of parallelism: instruction-level parallelism and process-level parallelism. In the first case, several processes are simultaneously working on one task, while in the second case, parallelism is implemented by running many commands every second. Both approaches have their own advantages.

Concurrency can be used at different levels. At its lowest level, it can be used in a processor through pipelining and a superscalar architecture with multiple function blocks.

At the next level, it is possible to integrate external CPU cards with improved computing capabilities into the system. As a rule, plug-in processors implement special functions: processing of network packets, cryptography, processing of multimedia data. With the help of these functions, it is possible to increase the performance of specialized applications by 5-10 times.

Recently, entire organizations have been integrating over the Internet, resulting in loosely coupled distributed computing grids, or grids.

If processors or processing elements that are far apart and exchange small amounts of data with high latency, then they are called loosely coupled. If processors or processing elements are nearby and exchange large amounts of data with low latency, then they are called tightly coupled.

The orientation of the Von Neumann architecture is aimed at sequential execution of program commands. At present, the requirements for productive computing technology are increasing more and more, and, as a result, the limitation of this approach becomes obvious. The future lies in the transition to parallel computing, both within the framework of one VM, and through the creation of multiprocessor systems and networks that will unite both separate processors and separate computers. The term "computing system" is more appropriate instead of the term "computing machine". The difference between a computing system and a computing machine lies in the means that implement parallel processing by building parallel branches in computations. The use of parallel computing systems is an urgent direction in the development of computers. This circumstance is caused not only by the physical limitation of the maximum possible speed of operation of ordinary sequential computers [5], but also by the existence of computational problems, for the solution of which the capabilities of the existing computer technology are always insufficient. As a result, we can conclude that knowledge of modern trends in the development of computers and hardware to achieve parallelism, the ability to develop models, methods and programs for parallel solution of data processing problems should be attributed to the important qualification characteristics of a modern programmer. In this article, we will cover the basic aspects of parallel computing.

Parallel computing can be understood as problem solving processes in which several computational operations can be performed simultaneously at the same time. This definition is quite general, but nevertheless, some conclusions can be drawn. First of all, we can conclude that parallel computations are not reduced to multiprocessor computing systems, in addition, the operations performed in the definition should be directed to a common task. What caused this interest in parallel computing?

Firstly, this is a constant outstripping of computing needs, the speed of existing computers. The so-called “big challenge” problem is known in the literature, including climate modeling, genetic engineering, drug development, and the like. All of these listed problems are computationally time consuming. According to experts, solving these problems requires computers with a performance of about 1 teraflop per second.

Also, the development of parallel computing systems is facilitated by the theoretical limited growth of the speed of sequential computations and a sharp decrease in the cost of multiprocessor computing systems, as a result of which many computers have become multi-core. What tasks should a parallel programming language solve, and what tasks are solved when realizing parallel computations?

The potential of a programming language should provide an opportunity to express the principle of a parallel program implementation scheme, to ensure adequate data structuring, the formation of a program structure that meets the given principles of distribution of procedural fragments for distributed execution.

The implementation of the tasks of distributing command / procedural program fragments among the execution devices of the computing facility, as well as the tasks of distributing data and integrating the entire event series ensure the execution of computations. The solution to these questions directly depends on what the ideological basis of data processing is. In general, this is a two-pronged dialectical problem: the structuring of data in a certain sense predetermines the procedural basis and at the same time the procedural principle sets the direction for an adequate structuring of data.

The most important component for almost all parallel architectures is the ability to migrate data and program fragments. Data migration can be present explicitly or implicitly and is extremely important primarily for architectures with explicitly or implicitly shared memory. At the same time, the structural representation of the algorithm and the corresponding program largely determines the possibility of distributing the program among the executing devices. The requirement to be able to migrate program fragments leads to the idea of representing a program and its fragments as a data structure. The ability to represent programs as data structures provides another important property - the ability to access program components as data, and, therefore, the ability to change programs during computation, including self-change.

Imperative parallel programming (launching parallel processes on demand by a program instruction) also requires special approaches to program structure issues. This is primarily due to the choice of parallelization method. The question is extremely subjective and depends on many factors. This considers, first of all, the scheme and the speed of parallelization on a given computing device. Fine-grained parallelization (the level of parallelization is not higher than instructions / commands) requires more frequent switching on of the parallelization unit than with coarse-grained parallelization (level of procedures and higher). Today, in most cases, parallelization at the module level is used, or in-command parallelization - hardware decomposition of one instruction into subcommands and their distribution among specialized subprocessors (vector-pipeline computers such as Cray and others, or hardware division of machine instructions into specialized processors (DSP processors, pipelined processors) However, fine-grained parallelization is inherent in systems with automatic parallelization at the level of high-level compilers or at the level of assembly languages or autocode. Automatic distribution has two stylistics of the solution: static, based on the compilation of the source code, and dynamic, using the means of identifying potentially parallel executable fragments of the executable code and distributing these fragments among the processing elements. Dynamic parallelization makes it possible to identify parallel executable components more efficiently, but it is extremely time-consuming in terms of execution time: as a rule, continuous monitoring of linear (or linearizable) program sections is carried out to identify possible simultaneously executable program instructions.

But parallel computing is not universal, and not all tasks can be accelerated with it. For example, the simple task of summing numbers is the same for both sequential and parallel computations. On the other hand, for optimization tasks, parallel computations provide huge performance gains. Therefore, the use of parallel computing systems is relevant today. What knowledge should a specialist have to cope with the tasks of parallel programming? The architecture of parallel computing systems, computation models and methods of complexity analysis, parallel computation methods, parallel programming (languages, development environments, libraries). A deep knowledge of parallel programming also provides many advantages when applying for a job.

4. Experiments

To create parallel programs, you must consider the architecture of modern processors, programming languages, optimization methods, environments and libraries for parallel programs. Therefore, in this scientific work we will consider the environment of the programming language Ruby, consider the methods of parallelization in it and analyze the implementation of parallel programs in comparison with the programming language C++.

Experiment 1. Consider the following problem: Given a number. You must reduce this number by one with each iteration until it reaches 0.

Example 1.1. Implementation of sequential execution in the C++ programming language.

```
void count (int n) {
    while (n > 0) {
      n -= 1;
   }
   }
void single () {
   Timer timer ("single thread time");
   count (200000000);
}
```

Example 1.2. Implementation of sequential execution in the Ruby programming language for MRI Ruby and JRuby [5].

```
def count(n)
  while n > 0 do
```

```
    n -= 1
  end
end
```

We measure time [4]:

```
puts Benchmark.measure {
  count (200000000)
}
```

Example 1.3. Implementation of parallel execution in the C ++ programming language using threads.

```
void multi () {
    Timer timer2("multi thread time");
    std::thread first(count, 100000000);
    std::thread second(count, 100000000);
    first.join();
    second.join();
}.
```

Figure 1 shows the measurement results for C++ executed in Visual Studio.

```
single thread time: 0.321730 seconds
multi thread time: 0.179569 seconds
```

Figure 1: Measurement results for C++

Example 1.4. Implementation of parallel execution in the programming language Ruby using threads for MRI Ruby and JRuby [5].

```
Puts Benchmark.measure {
  t1 = Thread.new {count (100000000)} # n/2
  t2 = Thread.new {count (100000000)} # n/2
  t1.join
  t2.join
}
```

Figure 2 shows the measurement results for JRuby executed in JRuby environment.

```
true
:count
 17.471000   0.000000  17.471000 ( 17.463250)
nil
 10.552000   0.000000  10.552000 ( 10.543745)
nil
```

Figure 2: Measurement results for JRuby

Figure 3 shows the measurement results for MRI Ruby executed in VS Code.

```
2.579000   0.000000   2.579000 (   2.570048)
2.546000   0.000000   2.546000 (   2.541701)
```

Figure 3: Measurement results for MRI Ruby

Now we will make a pivot table where we will present all the measurement results.

Measurement results – Table 1 shows the calculation method and its execution time in seconds.

Table 1
Measurement results

Sequential C++	Parallel C++	Sequential MRI Ruby	Parallel MRI Ruby	Sequential JRuby	Parallel JRuby
0.3217	0.1796	2.570	2.542	17.463	10.544

As we can see from the table, in C++ and JRuby the time has decreased by about half.

Why did multiple threads reduce execution time by almost 41% in JRuby, but had no effect in MRI Ruby? The reason is that Ruby MRI uses GIL (Global Interpreter Lock) to ensure that only one thread is running at a time. This prevents it from using both available cores in our system. Therefore, both threads are likely to run one after the other, or the interpreter may be switching the context of the threads. This is simultaneity. In the case of JRuby, there is no GIL, so the interpreter uses both cores to run threads simultaneously and thus reduces execution time. This is parallelism.

If we haven't achieved any multithreading advantage in MRI Ruby, what's the point of using threads at all?

The answer lies in the nature of the task. Note that the counting method written above is not inherently blocking, which means that it does not wait for any system resources, I/O, and so on. Let's make a simple change and create a blocking version [5].

```
def count(n)
   while n > 0 do
    sleep(0.5)
    n -= 1
  end
end
```

Let's take n = 20

The output for MRI Ruby:

Single-thread: 10.122942.

Multi-thread: 5.080528.

As you can see, now the time has also halved, but for large numbers it will take too long.

Let’s make another experiment 2. Consider the following problem: You have an array. You need to sum the first halve, the second halve and sum the results of both halves.

Example 2.1. Implementation of sequential execution in the C++ programming language and using threads.

```
template <typename It> void sum_partial(It a, It b, int32_t& sum) { sum = std::accumulate(a, b, 0); }
template <typename Iter>
void test(Iter begin, Iter end, const size_t num_of_threads_orig) {
    const size_t num_of_threads = num_of_threads_orig + 1;
    const size_t length = std::distance(begin, end);
    const size_t size = length / num_of_threads;

std::vector<std::thread> thr(num_of_threads - 1);
    std::vector<int32_t> res(num_of_threads);
    size_t begin_ = 0;
    for (size_t i = 0; i < num_of_threads - 1; ++i) {
        size_t end_ = begin_ + size;
        thr[i] = std::thread(sum_partial<Iter>, begin + begin_, begin + end_,
            std::ref(res[i]));
        begin_ = end_;
    }
    sum_partial(begin + begin_, begin + length, res[num_of_threads - 1]);

   const double start = omp_get_wtime();
    for (size_t i = 0; i < num_of_threads - 1; ++i)
        thr[i].join();

  const double end_time = omp_get_wtime();
    std::cout << "Time: " << std::fixed << end_time - start << "\n";
}
int main() {
    std::vector<int32_t> vec{ 1, 2,  3,  4,  5,  6,  7,  8, 9, 10, 11, 12, 13, 14, 15, 16 };
    std::cout << "Single "; test(vec.begin(), vec.end(), 1);
    std::cout << "Multi ";  test(vec.begin(), vec.end(), 2);
    return 0;
}
```

Figure 4 shows the measurement results for C++ executed in Visual Studio.

```
Single Time: 0.009292
Multi Time: 0.001827
```

Figure 4: Measurement results for C++

Example 2.2. Implementation of sequential execution in the Ruby programming language for MRI Ruby and JRuby.

```
def sumArr(n)
  sum = 0
  for i in n do
    sleep(0.5)
    sum += i
  end
 return sum
end
array = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]
half = array.length/2
first = array.first
last = array.last
puts Benchmark.measure {
  sumArr(array)
}
```

Example 2.3. Implementation of parallel execution in the programming language Ruby using threads for MRI Ruby (with blocking version) and JRuby.

```
puts Benchmark.measure {
  t1 = Thread.new { sumArr(first..half)  }# n/2
  t2 = Thread.new { sumArr(half..last)  }# n/2
  t1.join
  t2.join
}
```

Figure 5 shows the measurement results for JRuby executed in JRuby environment.

```
 0.002000   0.000000   0.002000 (  0.001457)
ı1
 0.001000   0.000000   0.001000 (  0.000070)
```

Figure 5: Measurement results for JRuby

Figure 6 shows the measurement results for MRI Ruby executed in VS Code.

```
0.000000   0.000000   0.000000 (   8.139256)
0.000000   0.000000   0.000000 (   4.565987)
```

Figure 6: Measurement results for MRI Ruby

Measurement results – Table 2 shows the calculation method and its execution time in seconds.

Table 2
Measurement results

Sequential C++	Parallel C++	Sequential MRI Ruby	Parallel MRI Ruby	Sequential JRuby	Parallel JRuby
0.009292	0.001827	8.1393	4.5660	0.001457	0.000070

If you look at the resulting numbers you can see that the results are no worse than in the first experiment, which once again confirms the fact that parallelism in the Ruby language exists, is possible, and is effective.

5. Conclusions

The hardware capabilities of computing devices are reaching their limit, so the relevance of methods and tools for optimizing software to solve modern problems is only growing. All programming languages have their advantages and disadvantages. The ease of use and semantics of the Ruby language [15] makes it a great entry into the world of software engineering. Some examples of Ruby projects that you already know that use streams for performance reasons are the Puma and Bundler web servers. You may know multiprocessor processing from Chrome - each tab exists for security reasons in a separate process.

In a Ruby environment, creating new child processes can increase productivity, but it also entails certain limitations. First of all, new processes impose additional responsibilities on the developer. Additional caution is required to perform them. Creating a multiprocessor program is much more complicated than creating a multithreaded program. It makes sense when the number of new processes is not too large, their execution takes a long time (creating a process is a bit expensive - especially in MS Windows), we have a multi-core processor, we do not want to exchange data between processes (or if we know how safe share them) and when we don't care about returning data from the process (which is a bit problematic). In general, each process must be independent, and the parent process must be the controller of these processes. Everyone should take some time to consider the question - do my project really need several processes? Multiprocessor programs can cause many more problems and are more difficult to implement. Make sure you understand what you are doing and why you are doing it. It is also good to know a little about the operating system - how will new processes be planned? Why are they designed this way? But if you want to try, you should always check whether the pros and cons of multiprocessor processing meet the business and technological requirements. Thread.new seems safer and has fewer potential problems. Be careful when using processes, and keep in mind that MRI Ruby uses GIL, so if you want to use streams, you should consider using JRuby or Rubinius.

Thus, high parallelism is not only achievable in Ruby, but also simpler than you might think. One viable approach is to simply branch out the running process to multiply its processing power. Another trick is to take advantage of multithreading. Although threads are lighter than processes, requiring less overhead, you can still run out of resources if you run too many threads at once. At some point, you may need to use a thread pool. Fortunately, many of the intricacies of multithreading are alleviated by using any of the many available Ruby Gem, such as Celluloid and its Actor model. The decision as to which one to use depends on the nature of your program, your work environment, and your requirements.

6. References

[1] Ivan Shikhalev. Ruby and multitasking. URL: https://shikhalev.org/2015/04/ruby-multitasking.html (date of access: 03.09.2021)

[2] Gernot Gradwohl. Scalable Ruby - Concurrency and Parallelism Explained. URL: https://betterprogramming.pub/scalable-ruby-concurrency-and-parallelism-explained-68b09a7aeb53 (date of access: 03.09.2021)

[3] Jesse Storimer. Does the GIL Make Your Ruby Code Thread-Safe? URL: http://www.rubyinside.com/does-the-gil-make-your-ruby-code-thread-safe-6051.html (date of access: 16.05.2021)

[4] James Britt, Neurogami. Benchmark. URL: https://ruby-doc.org/stdlib-2.5.3/libdoc/benchmark/rdoc/Benchmark.html (date of access: 10.05.2021)

[5] Neda Tavakoli, Dong Dai, Yong Chen, Client-side straggler-aware I/O scheduler for object-based parallel file systems, Parallel Computing, Volume 82, 2019, Pages 3-18, ISSN 0167-8191, https://doi.org/10.1016/j.parco.2018.07.001.
[6] Tan, G., Gao, G.R. Guest Editorial: Special issue on Network and Parallel Computing for Emerging Architectures and Applications. Int J Parallel Prog 49, 625–627 (2021). https://doi.org/10.1007/s10766-021-00720-3
[7] Awad A. Younis, Rajshekhar Sunderraman, Mike Metzler, Anu G. Bourgeois, Developing parallel programming and soft skills: A project-based learning approach, Journal of Parallel and Distributed Computing, Volume 158, 2021, Pages 151-163, ISSN 0743-7315, https://doi.org/10.1016/j.jpdc.2021.07.015.
[8] Guo, J., Yi, Q. & Psarris, K. Enhancing the Effectiveness of Inlining in Automatic Parallelization. Int J Parallel Prog (2021). https://doi.org/10.1007/s10766-021-00722-1
[9] Mehmet Deveci, Christian Trott, Sivasankaran Rajamanickam, Multithreaded sparse matrix-matrix multiplication for many-core and GPU architectures, Parallel Computing, Volume 78, 2018, Pages 33-46, ISSN 0167-8191, https://doi.org/10.1016/j.parco.2018.06.009.
[10] Danelutto, M., Mencagli, G., Torquati, M. et al. Algorithmic Skeletons and Parallel Design Patterns in Mainstream Parallel Programming. Int J Parallel Prog 49, 177–198 (2021). https://doi.org/10.1007/s10766-020-00684-w
[11] Anatoliy Doroshenko, Eugene Tulika & Olena Yatsenko (2021) Enhancing parallelism of distributed algorithms with the actor model and a smart data movement technique, International Journal of Parallel, Emergent and Distributed Systems, DOI: 10.1080/17445760.2021.1971665
[12] Inna V. Stetsenko, Alexander A. Pavlov & Oleksandra Dyfuchyna (2021) Parallel algorithm development and testing using Petri-object simulation, International Journal of Parallel, Emergent and Distributed Systems, DOI: 10.1080/17445760.2021.1955113
[13] Ruo Ando, Youki Kadobayashi & Hiroki Takakura (2021) Choice of parallelism: multi-GPU driven pipeline for huge academic backbone network, International Journal of Parallel, Emergent and Distributed Systems, DOI: 10.1080/17445760.2021.1941009
[14] Abhishek Sarkar. Ruby Concurrency and Parallelism - With a Practical Example. URL: http://abhisheksarka.github.io/ruby-concurrency-and-parallelism-with-a-practical-example/ (date of access: 03.05.2021)
[15] Editorial Team. PHP vs Ruby vs Python: The Three Programming Languages in a Nutshell. URL: https://1stwebdesigner.com/php-vs-ruby-vs-python/ (date of access: 17.05.2021)

Using the Characteristic of Sample Data Distribution Law in the Machine Learning Tasks

Lyudmyla Kirichenko [1], Oksana Pichugina [2] and Volodymyr Kobziev [1]

1 Kharkiv National University of Radio Electronics, 14 Nauki ave., 61166, Kharkiv, Ukraine
[2] National Aerospace University "Kharkiv Aviation Institute", 17 Chkalova Street, 61070 Kharkiv, Ukraine

Abstract
The paper introduces a new characteristic of a random variable with a bounded support titled a distribution indicator, defined as the ratio of the squared support length to the random variable variance. It is shown that this characteristic of the whole family of random variables of the same distribution law type. Computer simulation of random time series signals and evaluation of the distribution indicator had demonstrated the possibility of detecting the distribution law of the series observations by analyzing this characteristic only.

Keywords
Sample random variable, probability distribution, distribution law indicator, time series clustering

1. Introduction

In the development of intelligent measuring, control and diagnostic instruments, the tasks of automatic recognition of objects, represented by a time series, arise. In the case when the initial data are random by nature, the recognition may be reduced to just in identifying the probability distribution type for these data. Respectively, the knowledge of the distribution allows choosing optimal algorithms for modeling or information processing [1]-[3], in particular for infocommunication technologies of information transfer [4]-[6].

The quantitative characteristic of the distribution law acquires special significance in tasks related to the time series classification or clustering by machine learning [7]-[8]. In this case, a quantitative assessment can be used as a feature that is an input for the classifier or a metric in the case of clustering [9]-[12].

So, in this paper, we pose a problem of experimental detecting of the probabilistic distribution law of sample data, including the time series, based on the calculation of a quantitative value.

2. Related Work

The classical method for identifying the distribution law of a random variable (RV) is the construction of a histogram and it analysis [13], [14]. When using this method, human participation is required to determine the distribution law. Another method is the analysis of such numerical characteristics as the mathematical expectation, variance, mode and median by evaluating the corresponding sample statistics, such as mean, sample variance, etc. This stage is followed by validating statistical hypothesis on a particular distribution law with a specific expectation and a standard deviation. Respectively, detecting a type of distribution such as normal, uniform, Poisson etc. might require significant effort since it is closely related with rather precise estimates of the distribution parameters. These methods are described in detail in [1], [15].

Most of the considered tests for determining the distribution law of the observed data have been implemented in modern software tools [16]-[18]. We introduce a new numerical characteristic of a continuous RV with a bounded support, i.e.

EMAIL: lyudmyla.kirichenko@nure.ua (A 1);
oksanapichugina1@gmail.com (A. 2);
volodymyr.kobziev@nure.ua (A. 3)
ORCID: 0000-0002-2780-7993 (A.1); 0000-0002-7099-8967 (A.2); 0000-0002-8303-1595 (A. 3)

defined on a bounded interval $[a, b]$ titled a distribution indicator.

It will be shown that this characteristic depends only on the RV distribution law, it can be considered as an identifier of the type of distribution. Also, it can be qualitatively assessed by a sample from the distribution by a sample identifier.

3. Indicator of Distribution Law

3.1. Distribution type indicator

By the type of distribution we will mean a family of RVs ξ with variance σ^2 and mathematical expectation m obtained from a RV ξ_0 with zero mathematical expectation and unit variance by linear transformation

$\xi = k\xi_0 + b$, where $k \in R^1_{>0}$, $m \in R^1$.

If ξ_0 has a probability density function (PDF) $p_{\xi_0}(x)$ and is defined on $[a_{\xi_0}, b_{\xi_0}]$, then the PDF of the transformed RV ξ determined on the interval $[a,b]=[a_\xi, b_\xi]$ will be equal to

$$p_\xi(x) = \frac{1}{\sigma} p_{\xi_0}\left(\frac{x-m}{\sigma}\right).$$

It is easy to see that, in order the conditions

$$M[\xi]=m, Var[\xi]=\sigma^2$$

hold given that $\xi = k\xi_0 + b$,

$$M[\xi_0]=0, Var[\xi_0]=1,$$

the coefficients k,b should be taken as follows: $k=\sigma$, $b=m$. Thus

$$\xi = \sigma\xi_0 + k.$$

Definition. The distribution type indicator $Id[\xi]$ is calculated by the formula

$$Id[\xi] = \frac{(b-a)^2}{Var[\xi]},$$

where $[a,b]$ is the RV support, and $Var[\xi]$ is its variance.

The following statement can be proved.

Proposition. The indicator $Id[\xi]$ is constant for all RVs of the same type.

Proof. Let us use the properties of the variance of the RV:

a) if the random variables X,Y are independent, then
$$Var[X+Y]=Var[X]+Var[Y];$$
b) if k is a constant, then
$$Var[kX]=k^2Var[X];$$
c) the variance of the constant random variable C is equal to zero –
$$Var[C]=0.$$

By assumption $M[\xi_0]=0, Var[\xi_0]=1$ and $\xi_0 \in [a_{\xi_0}, b_{\xi_0}]$. From that,

$$[a,b]=[\sigma a_{\xi_0}+m, \sigma b_{\xi_0}+m],$$

wherefrom $b-a=\sigma(b_{\xi_0}-a_{\xi_0})$. Hence it follows that $Id[\xi_0]=(b_{\xi_0}-a_{\xi_0})^2$. On the other side,

$$Id[\xi]=Id[\sigma\xi_0+\mathrm{m}]=\frac{(b-a)^2}{Var[\xi]}=$$

$$=\frac{(\sigma(b_{\xi_0}-a_{\xi_0}))^2}{\sigma^2}=(b_{\xi_0}-a_{\xi_0})^2=Id[\xi_0].$$

Thus, the proposition is true.

Since ξ was chosen in an arbitrary way in the family of RVs of the same type, which follows from the linear dependence of ξ on ξ_0, this statement is true for an arbitrary random one of the same type.

Remark. Theoretically, it is possible to expand the family of RVs with a certain probability type indicator $Id[\xi_0]$. For example, with the same support and having mirror-symmetric PDF will have the same distribution type indicator. In this case, in order to distinguish the RV from the mirror one, it is necessary to analyze other characteristics, such as mode, coefficient of asymmetry and kurtosis. However, in practice, such situations are quite few.

Table 1 shows the values of the distribution type indicator $Id[\xi]$ for three common probability distributions [19] with support $[a,b]$ and the PDF $p_\xi(x)$:

$p(x)=\dfrac{1}{b-a}$ - Uniform;

$p(x)=\dfrac{1}{\pi\sqrt{(x-a)(b-x)}}$ - Arcsine;

$p(x)=\dfrac{2}{b-a}-\dfrac{2}{(b-a)^2}|a+b-2x|$ - Symmetric triangular.

Table 1
$Id[\xi]$ **values** for **the probability distributions**

Distribution type	$Var[\xi]$	$Id[\xi]$
Uniform	$\frac{(b-a)^2}{12}$	12
Arcsine	$\frac{(b-a)^2}{8}$	8
Symmetric triangular	$\frac{(b-a)^2}{24}$	24

3.2. Practical application and evaluation of the distribution type indicator

In practical applications, the distribution type indicator can be used to identify the type (law) of RV distribution by values $X^0 = \{x_i\}_{i=\overline{1,n}}$ of a time series of length n [20].

As an estimate of the indicator $Id[\xi]$, we propose to use a numerical characteristic (further referred to as the sample distribution type indicator of the RV ξ)

$$\hat{Id}_n = \frac{(X_{\max} - X_{\min})^2}{S^2},,$$

where $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of the series X^0, $(X_{\max} - X_{\min})$ is the range of the sample X^0, S^2 is an unbiased sample variance. The latter means that S^2 is an unbiased and consistent estimate of $Var[\xi]$, i.e.

$$\lim_{n\to\infty} S^2 = Var[\xi], M[S^2] = Var[\xi].$$

It follows that the estimate $\hat{Id}_n$ is also consistent:

$$\lim_{n\to\infty} \hat{Id}_n = \frac{\lim_{n\to\infty}(X_{\max} - X_{\min})^2}{\lim_{n\to\infty} S^2} = \frac{(b-a)^2}{Var[\xi]} = Id[\xi].$$

However, it can be biased and shifted to the left:

$$M(\hat{Id}_n) = \frac{M(X_{\max} - X_{\min})^2}{M(S^2)} = \frac{M(X_{\max} - X_{\min})^2}{Var[\xi]} \le$$

$$\le \frac{(b-a)^2}{D[\xi]} = Id[\xi],$$

whence it can be seen that it is unbiased estimate if and only if the range of the sample coincides with the range of the random variable $X_{\max} - X_{\min} = b - a$, otherwise it will be biased to the left.

For an experimental evaluating the distribution type indicator $Id[\xi]$, a numerical simulation of values of the RV $\xi \in [a,b]$, where $[a,b]$ is bounded was done for the probability distributions listed in Table 1.

The RV values were presented as a time series of length n. Figure 1 shows plots of dependence of $\hat{Id}_n$ on n. The figure shows the convergence of the estimate $\hat{Id}_n$ of the distribution type indicator to the corresponding theoretical value $Id[\xi]$.

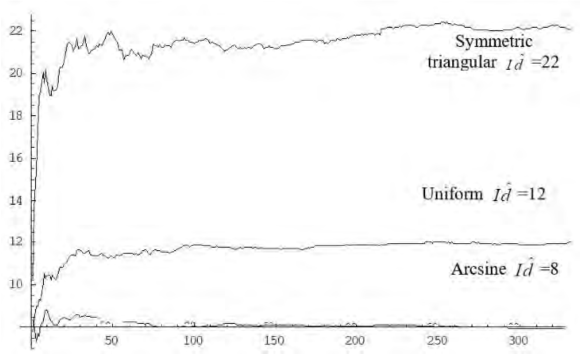


Figure 1: Dependence $\hat{Id}$ on n

The consistency of the estimate $\hat{Id}$ implies the mean square convergence of the series $\hat{Id}_i, i = 1,2,...$ to $Id[\xi]$. The average value of n for which the condition is satisfied

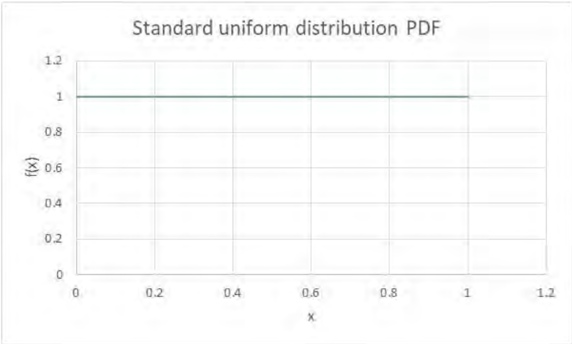
$$\frac{1}{n-n_1}\sum_{i=n_1}^{n}\left(Id[\xi] - \hat{Id}_i\right)^2 < \varepsilon$$

for an arbitrary $\varepsilon > 0$ on the interval $[n_1, n]$, $n_1 < n$ characterizes the rate of this convergence. The values n_1 calculated for the different values of ε are given in Table 2.

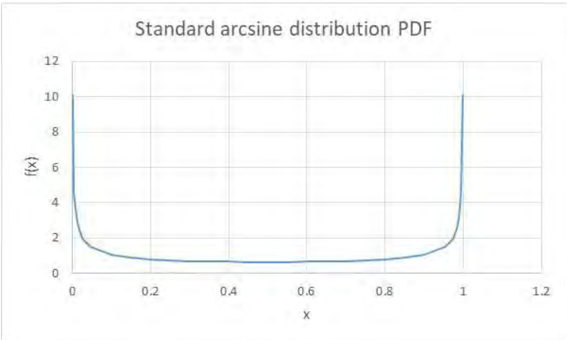
Table 2
Convergence of the estimates

Probability distribution	ε =0,3	ε =0,5	ε =1
uniform	53	48	30
arcsine	19	16	12
Symmetric triangular	351	315	204

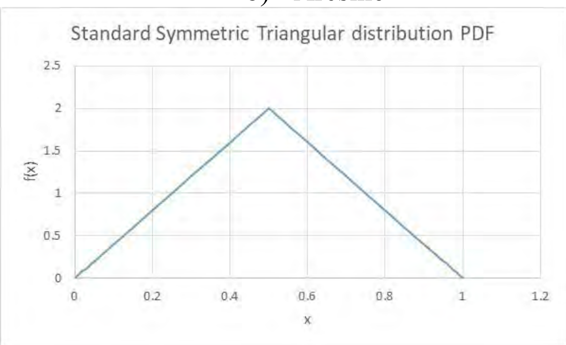
The convergence rate is closely related to the values of the PDF in the vicinity of the boundaries of the RV support $[a,b]$.



a) Uniform



b) Arcsine



c) Symmetric triangular

Figure 2: PDF of the standard distributions with support $[0,1]$

Figure 2 shows that the probability of falling out of numbers near the ends of the interval $[0,1]$ for the standard arcsine distribution is greater than that of the standard uniform distribution, so the range $\left(X_{\max} - X_{\min}\right)$ approaches its true value faster as n increases. For the symmetric triangular distribution, the probability of falling close to the bounds of the support $[0,1]$ is relatively small compared to the one of falling near the mean value 0.5, so the convergence to is slow. The values given in Table 2 can be considered as the size of the time series required for identifying the probability distribution type with the accuracy ε .

Also in the work, a numerical experiment was carried out on the clustering of model data. Three samples of independent random variables, which were discussed above, were generated. Clustering was performed using the k-means method; sample lengths were different: 20, 30, and 3 values. Figure 3 shows the results of clustering.

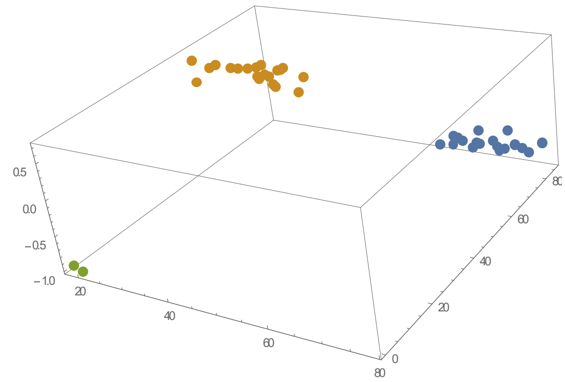


Figure 3: Results of clustering various independent random variables

4. Conclusions

The paper proposes and theoretically substantiates a method for identifying the probability distribution type of a continuous RD with a bounded support. The distribution type indicator $Id[\xi]$ introduced for this purpose provides a numerical characteristic of the distribution. Numerical studies have shown that the proposed method makes it possible to identify the distribution type of a RV using a statistical characteristic $\hat{Id}_n$ playing a role of an estimate of $Id[\xi]$ value and called the sample distribution indicator is the latter is evaluated on samples of a relatively small size. The contributions can be used for the development of measuring instruments, control, diagnostics and many tasks of object recognition, where speed and quality of the recognition is important.

5. References

[1] S. Brandt, Data Analysis: Statistical and Computational Methods for Scientists and Engineers. Cham Switzerland ; New York: Springer, 2014

[2] O. Pichugina, New Approaches to Modelling Covering Problems in Monitoring Optimization, in: 2019 IEEE

International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PICST), 2019, pp. 300–304. doi:10.1109/PICST47496.2019.9061386
[3] A. Srivastava and E. P. Klassen, Functional and Shape Data Analysis. New York: Springer-Verlag, 2016. doi: 10.1007/978-1-4939-4020-2
[4] L. Kirichenko and T. Radivilova, "Analyzes of the distributed system load with multifractal input data flows," 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), Lviv, 2017, pp. 260-264, doi: 10.1109/CADSM.2017.7916130.
[5] I. Ivanisenko, L. Kirichenko and T. Radivilova, "Investigation of multifractal properties of additive data stream," 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2016, pp. 305-308, doi: 10.1109/DSMP.2016.7583564.
[6] O. Pichugina, N. Muravyova, Data Batch Processing: Modelling and Applications, in: 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S T), 2020, pp. 765–770. doi: 10.1109/PICST51311.2020.9467928.
[7] Esling, P., Agon, C.: Time series data mining. ACM Computing Surveys 46(1) (2012).
[8] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., Muller, P. A.: Deep learning for time series classification: a review. Data Mining and Knowledge Discovery 33(4), 917-963 (2019).
[9] Ben D.: Feature-based time-series analysis, https://arxiv.org/abs/1709.08055
[10] Kirichenko, L., Radivilova, T., Bulakh, V. (2018) Machine Learning in Classification Time Series with Fractal Properties. Data 4(1):5. doi: 10.3390/data4010005.
[11] Trovero M. A., Leonard M. J. (2018) Time Series Feature Extraction. Paper SAS2020-2018. SAS Institute Inc 1-18. URL: https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2020-2018.pdf
[12] Kirichenko, L., Radivilova, T., Tkachenko, A. (2019) Comparative Analysis of Noisy Time Series Clusterin. CEUR Workshop Proceedings 2362:184-196. http://ceur-ws.org/Vol-2362/paper17.pdf
[13] Greenwood, P.E.; Nikulin, M.S. (1996). *A guide to chi-squared testing*. New York: Wiley.
[14] Montgomery, Douglas C.; Runger, George C. (2003). Applied Statistics and Probability for Engineers . John Wiley & Sons, Inc. p. 104.
[15] G. Cowan, Statistical Data Analysis, 1st edition. Oxford : New York: Clarendon Press, 1998.
[16] Deisenroth, Marc Peter; Faisal, A. Aldo; Ong, Cheng Soon (2020). Mathematics for Machine Learning. Cambridge University Press. p. 181.
[17] P. S. P. Cowpertwait and A. V. Metcalfe, Introductory time series with R. New York : Springer-Verlag, 2009.
[18] C. M. Mastrangelo, J. R. Simpson, and D. C. Montgomery, ‘Time Series Analysis’, in Encyclopedia of Operations Research and Management Science, S. I. Gass and M. C. Fu, Eds. Springer US, 2013, pp. 1546–1552. doi: 10.1007/978-1-4419-1153-7_1045
[19] S. Kotz and J. R. V. Dorp, Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. Singapore ; Hackensack, NJ: World Scientific Publishing Company, 2004.
[20] Time Series Analysis: With Applications in R, 2nd edition. New York: Springer, 2010.

Methods of mathematical apparatus developing based on logical models in information systems

Ihor Shubin[1], Stanislav Snisar[2] and Victoria Skovorodnikova[3]

[1] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*
[2] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*
[3] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract

The aim of this research is to expand the mathematical apparatus used for formalization in multi-purpose automated control systems and intelligent information recognition systems. Attention is paid to the advantages and disadvantages of different models of knowledge representation. The aim of the research is also to use algebraic methods in automated systems.

Keywords

Mathematical Support, Knowledge Representation, Logical Model, Production System, Frame System, Semantic Grid, Predicate Algebra, Quantifier Equation, Logic Algebraization

EMAIL: igor.shubin@nure.ua (A. 1); sn.stanislav4@gmail.com (A. 2); viktoriia.skovorodnikova@nure.ua (A. 3)
ORCID: 0000-0002-1073-023X (A. 1); 0000-0003-0260-3027 (A. 2); 0000-0003-0436-4197 (A. 3)

1. Introduction

Currently available intelligent information systems are able to perform functions that were previously considered exclusively the prerogative of man: to prove mathematical theorems, translate texts from one language to another and perform many other functions. However, in the long run the ideal computer must surpass the ability of man to think logically, to analyze obtained information, to solve the most complex problems, to interact with the environment. Relational and logical ways of presenting knowledge play an important role in the development of mathematical support of information systems [1-2]. One of the effective universal mathematical ways to describe information is the algebra of predicates and predicate operations. In the language of this algebra it is easy and convenient to describe various formalized information, to form queries in databases and to model human activity [3].

In various computerized industries there is a need to process information displayed in natural language. In automated control systems (ACS), which include man as his organic link, the main form of information transmission are documents that contain a significant amount of textual information. Computer modeling of word processing will automate many types of human intellectual activity, expand its capabilities. The basis of ACS are automated information systems (AIS), the purpose of which is to automate the processes of accumulation, retrieval and generalization of information. The effectiveness of AIS is recognized by their ability to process non-formalized or weakly formalized information.

2. Analysis of formal methods of intelligent systems modeling

The form of information presentation in the computer has a great influence on the speed and quality of information processing by intelligent systems [4]. In general, different information systems use different ways of presenting knowledge depending on specific applications. [5]. A characteristic feature of modern intelligent systems is that the data processing process is based on the necessary knowledge of the problem area, pre-entered into the system's knowledge base, while previously created systems worked with data that were simply processed by various programs.

Knowledge is a complexly organized types of data that differ from traditional presentations of data by four main features:

- knowledge contains in the record not only the information part, but also the descriptive one - it records all the data about the informational unit that may be needed in the user's work with the system;
- knowledge in the knowledge base creates complex, usually hierarchical structures, which is achieved by introducing different relationships on the information units entered in the knowledge base;
- informational units that reflect knowledge can be arranged in more complex units and decomposed into simpler ones;
- as part of the informational units that characterize some knowledge, can be attached or built-in procedures, which allows you to activate these procedures as a result of the appearance in the knowledge base of different information units or links between them.

The peculiarity of knowledge representation systems is that they model human activities, often carried out in an atypical variety. Thus, an important step in the development of intelligent systems is to create an optimal model of knowledge about the subject area of application systems. Obviously, the choice of a particular type of knowledge representation depends on the areas of formalization [6].

However, among the different ways of presenting knowledge, generated by the specifics of knowledge structures diversity, we can distinguish the logical model, frame and production systems, semantic grids. Each method of presentation has its advantages and disadvantages, and is associated with a certain structure and areas of application of knowledge.

The advantages of logical models are the singularity of the theoretical substantiation and the possibility of implementing a system of formally accurate definitions and results. For this reason, intelligent systems that use a logical model of knowledge have become quite widespread.

The disadvantages of using the logical construction of systems include unconstructiveness and semantic limitations. At the same time, human logic is often not limited to the usual formalism of logical languages and is an intelligent model with a indistinct structure.

However, the limits of formal logic is expanding, as evidenced by the appearance of modal, multivalued and probabilistic logics. This allows to expand the application of logic in informational systems.

In production models, knowledge is represented by a set of "if-then" rules. Such systems have a direct or reverse result. The advantages of production systems are:

- simplicity of creating and understanding individual rules;
- simplicity of replenishment and modification;
- simplicity of realization of logical result mechanism.
- The disadvantages of such systems include:
- ambiguity of mutual relations between rules;
- the complexity of assessing the holistic reflection of knowledge;
- difference from the human structure of knowledge;
- lack of flexibility in the logical result.

With the growth of knowledge, it is reasonable to use different methods of pre-grouping and structuring knowledge. The production system can be used with a frame model, which gives good results.

The model of knowledge representation, based on frame theory, is a synthesized single model of human memory and consciousness. A key concept in this theory is the concept of a frame as a data structure for representing an object. Frames consist of slots that store information about the frame, and are linked into a single frame system. Knowledge representation languages based on a frame model are quite effective in situations where it is necessary to apply different result rules depending on the situation.

The relationship between different concepts and entities of specific objects can be described using a grid consisting of nodes expressing some concepts and arcs describing the interdependence between objects. This kind of formalism of knowledge representation is called the semantic grid. Methods of realization of a specific system of concepts and relations between them depend on the forms of representation of knowledge in systems.

For example, a semantic grid can be broken down into a set of conceptual graphs, which in turn are logical formulas. This allows you to make more clear the various relationships that arise from the description of objects by logical formulas. Thus, semantic grids can be found as some graphical version of the difference of predicates.

The peculiarity of the semantic grid is in the integrity of the system made on its basis, which does not allow to separate the knowledge base and the mechanism of results in such a system. On the basis of the semantic grid, specific application systems were created, such as the SIR natural language, the SCHOLAR logic learning system and the GASNET diagnostic system.

The aim of the research is algorithmic and software implementation of the method of solving quantifier linear equations on the basis of algebra of linear predicate operations, formal apparatus of linear-logical operators and methods of solving logical equations.

The main objectives of the research:

- analysis of formal ways of intelligent systems;
- analysis of methods of logic algebraization;
- study of the apparatus of linear predicate operations algebra;
- study of the linear-logical operators theory;
- further software implementation of quantifier linear equations solving method.

The developed algorithms and software can be used in the development of linguistic software for automated information systems, in information retrieval systems, in solving problems of logical inference in databases and expert systems, as well as in solving problems of recognition and classification of objects, including in intelligent radar systems for detecting inconspicuous objects. The method of solving logical equations, implemented in the software system, will greatly reduce the amount of necessary calculations [7].

3. Application of algebraic methods in automated systems

Numerous practical additions to modern abstract algebra in databases and intelligent systems have recently led to increased interest in the possibility of information algebraic description. At the same time, practice suggests unexpected new structures that enrich algebra

[8]. Based on the application of algebraic methods in programming theory, various translators from high-grade languages and various algorithmic algebras have been developed.

Automation of software systems development and computer design is an important and actual problem of computer technology, which requires the development of utilitarian theory of algorithms. One of the main tasks of this theory is the problem of optimal translator from one language to another.

There are two algorithmic languages and some algorithm implemented in one of them. It is necessary to find the optimal implementation of this algorithm in another language according to the given criteria. When performing applied tasks, as a rule, the first language is a language of a high degree, focused on a certain range of tasks, and the second is the internal language of the machine.

Thus, it is necessary to translate from the programming language to the machine language while optimizing the source program. The solving process of such problem is divided into a set of intermediate stages, each of which is a partial algorithm optimization and translation into an intermediate language corresponding to this stage.

Let S = { f } be a set of languages and $P = \left\{ \overset{m}{f', f'', f} \right.$ be a set of translators, each of which is a program in f and translates programs from the input language f ' to the source language f ". The translator can be considered as a unary operation, with a domain representing a language f ', and the range of values of f ". Let there be translators

$$\left\{ \overset{m^1}{f', f'', f_1} \right. \text{ and } \left\{ \overset{m^2}{f_1', f_2'', f} \right. \tag{1}$$

We can calculate the recoding operation as follows

$$\left[\overset{m^1}{f', f'', f_1} \right] = \overset{m^1}{f', f'', f_2} \tag{2}$$

On the translators $\overset{m}{f_1, f_2, f_1}$ and $\overset{m}{f_2, f_3, f}$ the composition operation is performed:

$$\overset{m}{f_1, f_2, f_1} * \overset{m}{f_2, f_3, f} = \overset{m}{f_1, f_3, f} \tag{3}$$

With the help of these operations, you can formalize the processes often used in programming.

To perform equivalent transformations of algorithms, it is necessary to build an algebra of algorithms that would allow to perform transformations using clear algebraic language.

By database we mean an information system that stores, processes information and is able to respond to requests. Moreover, it should be possible to obtain not only information directly stored in the database, but also derived information obtained on the basis of basis information. The problem of obtaining derived information is directly related to the result problem in intelligent systems. Database queries can be written with the formulas of some logical language, for example, using the language of statements differences or predicates differences of different order, and the expressive possibilities of these differences are different.

There are all sorts of differences for classical and non-classical logics. Boolean algebras, for example, correspond to the classical statements difference, and special Gating algebras correspond to the intuitionistic statements difference.

4. Solving the problem

After processing this data by a software system, during which the consistency of logical equations system is determined and the missing features are disabled, the output is the whole array of logical equations system solutions.

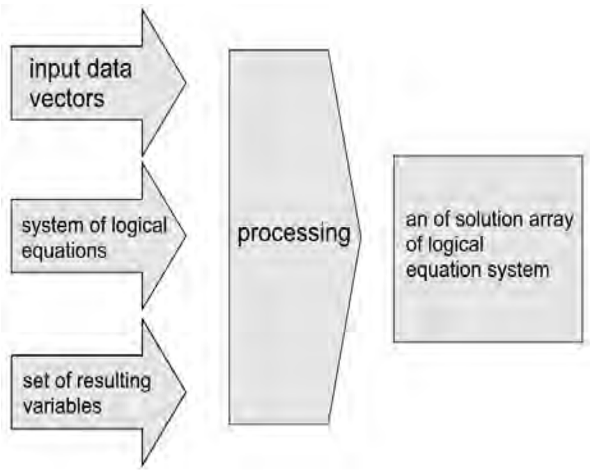


Figure 1: Object model of the developed system.

The example given illustrates the possibility of using the linear-logical operators theory and the method of solving a quantifier predicate equation for processing and storing information in databases.

Assume that the database contains information about four factories that produce parts for machines. Let factory z_1 produce parts d_1 and d_2, factory z_2 produce parts d_2 and d_3, factory z_3 produce parts d_1 and d_4, factory z_4 produce parts d_3 and d_4. The existing "factory-part" relationship is easily described by the following binary predicate:

$$P_1(z,d) = \begin{cases} 1, \text{ if the factory } z \text{ produces detail } d, \\ 0, \text{ otherwise} \end{cases} \quad (4)$$

where $z \in \{z_1, \ldots z_4,\}$ and $d \in \{d_1, \ldots d_4,\}$. Thus, information about factories can be stored in the form of a formula record of the predicate P_1 *(z, d)*. Next, let's get information about which factories produce part d_1. The corresponding predicate disclosing this requirement is written in the second form:

$$P_2(d) = \begin{cases} 1, \text{ if } \mathrm{d} = \mathrm{d}_1, \\ 0, \text{ otherwise} \end{cases} \quad (5)$$

As a result, the predicate $P_3(z)$, corresponding to the required information, is denoted by a quantifier equation of the form:

$$\exists d \; P_1(z, d) \wedge P_2(d) = P_3(z) \quad (6)$$

and specifies the following relationship:

$$P_3(z,d) = \begin{cases} 1, \text{ if the factory } z \text{ produces detail } d_1, \\ 0, \text{ otherwise} \end{cases} \quad (7)$$

The solution of this quantifier predicate equation is obtained from the solution of the corresponding operator equation A*X=Y in the linear-logical space $E_{\wedge}^{n}$. The matrix of operator A looks like this:

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix} \quad (8)$$

and the vector X=(1 0 0 0). As a result of the action of the operator A on the vector X we get the vector Y is equal to (1 0 1 0) and means that the part d_1 is made by factories z_1 and z_3. Thus, the searching operation for information of interest in the database is replaced by the operation of multiplication operator. Now let's consider which factories produce parts d_1 or d_3. Using the additive property of the linear logical operator A, we have

$$A*X_1 \vee A*X_3 = A(X_1 \vee X_3) = A*X_4 \quad (9)$$

The vectors X_1 and X_3 are created by predicates that formulate the details d_1 and d_3, respectively. Thus, the answers to more complex queries in the database also come from solving the operator equation.

Using the algorithm for solving the operator equation, described in the previous section, you can find the specified factories parts that they produce. For example, let's consider what parts the z_2 factory makes. Therefore, the logical vector Y=(0 1 0 0). As a result of solving the operator equation of the form

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix} * \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{vmatrix} \quad (10)$$

with respect to X we obtain the vector (0 1 0 0) and (0 0 1 0). Therefore, the factory z_2 produces parts d_2 and d_3.

5. Conclusions

The paper analyzes the formal ways of intelligent systems developing: ways of presenting knowledge depending on the specific areas of systems application; formal languages that allow you to represent knowledge in computer memory. The application of algebraic methods in automated systems is considered.

The problem of obtaining derived information is directly related to the problem of result in intelligent systems, and in the case of applying an algebraic approach to the description of derived information, there is a certain algebraic system - query algebra, in terms of which derived information is written through the base.

Practical applications of modern abstract algebra in databases and intelligent systems are found. Based on the application of algebraic methods in theoretical programming, various translators from high-grade languages and various algorithmic algebras have been developed.

The prospects of using the considered method of solving logical equations in information systems, in particular, in databases are shown. It provides the ability to obtain not only information directly stored in the database, but also derived information obtained on the basis of basis information

6. References

[1] Y.A. Shreyder, A.A. Sharow, “Sistemy i modeli [Systems and models],” Moscow: Radio and communication, 1982, 152 p. (in Russian)

[2] B.I. Plotkin, “Universalnaya algebra, algebraicheskaya logika i bazy dannykh [Universal algebra, algebraic logic and databases],” Moscow: Mir, 1990, 322 p. (in Russian)

[3] V.B. Klimushev, “Modelyuvannya korelyatsiynoho alhorytmu rozpiznavannya obraziv na bazi formalnoyi hramatyky [Modeling of correlation algorithm of image recognition on the basis of formal grammar],” Automated control systems and automation devices, no. 115, pp. 54-74. (in Ukrainian)

[4] M.F. Bondarenko, I.O.Leshchynska, N.P.Kruglikova, N.E.Rusakova and Yu.P.Shabanov-Kushnarenko, "About relational networks," Bionics of Intelligense Sci. Mag, 2010

[5] A.S. Kutsenko, E.E. Ignatchenko, "A Semantic-Statistical Approach to Identifying Key Phrases in the Problem of the Thematic Classification of Tests," Visitor of the National Technical University "KhPI", Kharkiv: NTU "KhPI", no. 6, 2002, pp. 149-154.

[6] I. Kyrychenko, G.Tereshchenko, G.Proniuk, N.Geseleva, Spatial Interpretation of the Notion of Relation and Its Application in the System of Artificial Intelligence, in: Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Systems, COLINS-2019, Kharkiv, Ukraine, pp. 266-276.

[7] N. Sharonova, I. Kyrychenko, G. Tereshchenko, Application of Big Data Methods in E-Learning Systems, in: CEUR Workshop Proceedings 2870 of the 5rd International Conference on Computational Linguistics and Intelligent Systems, COLINS-2021, Volume I, pp. 1302-1311.

[8] M.F. Bondarenko, Z.V. Dudar, N.T. Protsay, V.V. Cherkashyn, V.A. Chykyna, Y.P. Shabanov-Kushnarenko, “Alhebra predykativ i predykatnykh operatsiy [Algebra of predicates and predicate operations],” Radio electronics and informatics, no. 1, 2004, pp. 5-22. (in Ukrainian).

A KNOWLEDGE-BASED APPROACH TO ENGINE CONTROL UNIT PARAMETERS CLASSIFICATION

Svitlana Kuznichenko[1], Tetiana Tereshchenko [2], Iryna Buchynska [3], Olena Makovska [4], Daria Bulysheva [5], Viacheslav Sedov [6]

[1, 2, 3] *Odessa State Environmental University, 15 Lvivska Str, Odesa, 65016, Ukraine*
[4] *Kyiv National University of Technologies and Design, Nemyovycha-Danchenka str. 2, Kyiv, 01011, Ukraine*
[5, 6] *Odesa State Agrarian University, Panteleimonivska str.13, Odesa, 65012, Ukraine*

Abstract

The papers considers the issues of improving the efficiency of the classification process of cards of electronic control units of a car engine. A knowledge-based classification algorithm has been developed that uses characterizers to calculate a given characteristic of the input data matrix. The algorithm is implemented in the C # programming language using the .NET Framework. The results of testing on the family of blocks of Mercedes-Benz Bosch EDC16C31/EDC16CP31 cars showed that in the case of using a large set of educational images for training then the system determines the class of the input card with high accuracy. The tests carried out determined the optimal time and number of images for training the system.

Keywords

Classification with learning, pattern recognition, knowledge base, software system, electronic engine control unit.

1. Introduction

Rising living standards in developed countries have led to a sharp increase in the number of road transport. Government authorities and public organizations are constantly working to change environmental standards, motivate and force the motor-vehicle industry to develop and implement technologies for neutralization of exhaust gases [1]. Under these conditions, manufacturing plants are forced to produce a large number of software options that are standardized for different operating conditions of road transport. At the same time, it is difficult enough to take into account all possible conditions because the number of options that can be set in the program settings when releasing the car is limited. This has led the third-party companies to develop ways to change the calibration parameters of injection systems to fine-tune an engine, to remove the limitations imposed by the manufacturing plant for the engine running characteristics. Many companies around the world are developing software and tools to adjust the software of electronic vehicle control units.

There are several software products that are designed to edit the calibration tables in the electronic vehicle control units. There are two of them, which have a built-in system for calibration and data recognition: WinOLS and Swiftec [2], [3]. The analysis of existing software has shown that systems of this type are not adaptive and capable of learning. Classification is based on well-defined algorithms or dictionaries created by developers. We can identify the following disadvantages of modern card recognition systems represented in the market:

- the lack of self-learning, adaptability;
- the lack of ability to create dictionaries independently;
- the low speed of work;
- a small number of calibration table classes provided.

The analysis of publications in this field of research has shown that the most common classification system in such systems is artificial

EMAIL: skuznichenko@gmail.com (A. 1); tereshchenko.odessa@gmail.com (A. 2); buchiskayira@gmail.com(A. 3); olenam2011@ukr.net (A. 4); bu.dasha.bu@gmail.com (A. 5); vjacheslavsedov@gmail.com (A. 6)
ORCID: 0000-0001-7982-1298 (A. 1); 0000-0001-7691-6996 (A. 2); 0000-0002-0393-2781 (A. 3); 0000-0001-6199-4377 (A. 4); ORCID: 0000-0003-2907-9413 (A. 5); 0000-0001-7974-3760 (A. 6);

neural networks (ANNs) [4-7]. The use of neural networks implies that the prepared data is submitted to the input. The processing of this data requires complex calculations, which significantly reduces the speed of classification. At the same time, ANNs are difficult to learn; the learning outcome depends on the number and quality of images in the learning sample; it is difficult to solve classification problems; they are able to give images of one class for images of another class, as well as quickly lose experience with conflicting input data. In the presented study, the use of ANNs as the main recognition mechanism requires the creation of its network topology with additional transformations of input data, as well as the development of its own network learning mechanism. This requires a lot of labour, and the result does not meet the requirements of extensibility and self-learning of the system in the course of operation. In addition to ANNs the image recognition processes use the spectral signal decomposition and subsequent analysis of the signal spectra in order to obtain data on the signal characteristics. There are several standard mathematical methods for decomposing a signal into spectra, such as the Fourier transform and wavelet transform [8-10]. The analysis of the test results has shown that the system based on such methods requires complex transformations of the results of spectral decomposition. The use of spectral decomposition to solve the classification problem is possible in case of determining some characteristics of input data and using them as data for classification. The use of only spectral decomposition in the problem considered requires complex calculations and has an unpredictable result.

The purpose of the work is to develop a system for classifying the parameters of the electronic control unit of a car engine, which provides high speed processing of input data, has the ability to learn and serialize learning data in separate dictionaries of the knowledge database.

2. The main research material

As a result of reading the program, a binary file - a dump, which contains the area of calibrations, the data from which must be recognized, is taken from the memory of the engine control unit. In the fig. 1, each value in the dump is represented as points on a two-dimensional field, where in the horizontal line the addresses in the memory are set, and in the vertical line - the values.

At the same time, calibration cards are visible. But even an experienced specialist is not always easy to determine exactly where each of the cards begins, what size it has, and what this a characteristic is. The classification system is needed to determine all these indicators.

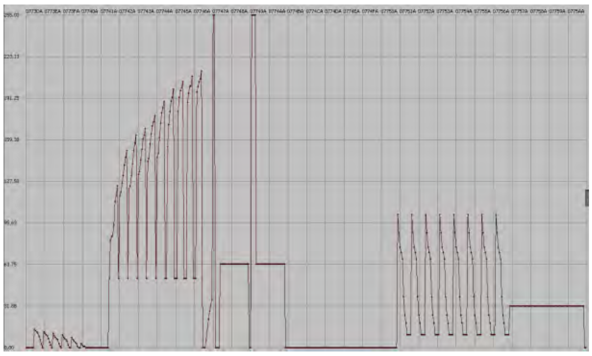


Figure 1: Graphical representation of the dump

Before classifying cards, it is necessary to provide the generated matrix of values to the input of the system. To do this, we need to split the dump into tables. This problem is quite difficult to solve, except for a few cases where the information about the structure of the calibration area is directly in this dump. For example, in case of the Delphi systems, the dump contains the addresses of all cards with sizes in one table - the problem is to find such a table automatically and get data from it. The Bosch electronic control unit is also easy to split into cards. After the first stage of dump processing is completed, we obtain a set of ready-to-classify cards, which are a table of values. The input data for the system are matrices of different dimensions M.

2.1. The statement of the classification problem

The analysis of the properties of the input matrices M requires a preliminary step to “recognize” and provide the essential properties in the form of a vector for each class. Let’s consider the statement of the classification problem. Classification is the problem of assigning the output matrix m* to some class of properties. In this case, the possible classes C are given, a set of other matrices M is available, as well as the essential features F that characterize these matrices: F = Ø or F ≠ Ø are known (or not known). The general statement of the problem of classification of properties is defined in the following form:

- a set of matrices M is given. A set of matrices can be represented by a subset of similar matrices, which are called classes C;
- information about classes C, information about the output matrix m*, the belonging of which to a certain class is unknown, exists or is extracted;
- it is necessary to construct an algorithm ψ from the description of the matrix m* to determine the belonging of m* to some class c*.

A formal statement of the classification problem with unknown properties F, which must be pre-extracted: M ≠ Ø, C ≠ Ø, M={mi}, i = 1, 2,…, Mk; m*∉ M, C={cj}, j = 1, 2, …, Ck; are given; the ratio R ⊂ M × C and some similarity criterion dist are known. We need to construct an algorithm ψ: m* → c*, c*∈ C, such that

$$
\begin{gathered}
c^* = \psi(m^*, M, R1, C, F, dist),\\
\psi = \langle \psi1, \psi2 \rangle, \psi1: M \to F, \psi1: m^* \to f^*,\\
\text{де } F = \{f_n\}, n = 1,2,\dots,N_k;\ f^* \in F;\\
\psi2: f^* \to c^*, c^* \in C, R1 \subset F \times C.
\end{gathered}
\qquad (1)
$$

For example, the algorithm ψ can be represented as follows:

$$
\begin{aligned}
&c^* = c_z, z = \arg\min(g_i);\\
&g_i = dist(f^*, f_i),\\
&f_i = \psi1(m_i),\\
&f^* = \psi1(m^*), i = 1,2,\dots,M_k.
\end{aligned}
$$

2.2. Classification algorithm

The classification unit consists of a set of simple characterizers. The characterizer will be a separate element that determines a certain property of the input matrix by its own algorithm ψ1 and compares its value with the reference data (obtained earlier in the learning process) by the algorithm ψ2.

The classification algorithm is given in the Fig. 2. At classification the test block receives a set of reference features for each of the characterizers corresponding to a certain class. If at least one of the tests gives a negative answer, there is a transition to the next class for testing. If the compliance check does not give a positive result for the whole set of classes, then the class is considered indeterminate.

The algorithm represented in the fig. 2 contains a double cycle; at the upper level there is an iterative passage through the known classes that have been taught. At this level, the algorithm works with the card dictionary; only the list of classes is important to it. If we implement the algorithm in this form, the number of characterizers will be:

$$N_\psi = n \cdot m, \qquad (2)$$

where n – the number of characteristics; m – the number of known classes.

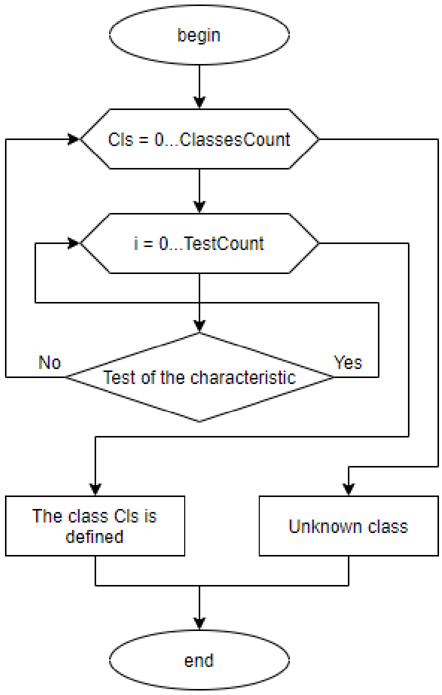


Figure 2: Flowchart of the classification algorithm

With this approach, the system implementation will require a large amount of memory, even if the characterizers are quite simple and contain little learning data.

To reduce the amount of memory, as well as to speed up the algorithm, it is possible to create only one set of characterizers, because it is basically the same for all classes of cards. At the stage of checking a particular class, each characterizer receives learning data for this class, after which there is a check. This method is quite easy to implement, based on the fact that the data is stored all together in one dictionary, and significantly improves the system. The second possible optimization is based on the verification algorithm. After all, the answer for each class is:

$$R = \wedge_{i=0}^{N} C_i(M), \qquad (3)$$

that is, the total result R for this class of cards is the conjunction of the results of each of the characterizers Ci(M). If at least one of them gives a negative answer, we can go immediately to the next class of cards (fig. 2). One of the possibilities of optimization is the priority of the characterizers themselves. In this case, there are two options:

- as a result of careful testing to determine the most appropriate sequence of characterizers;
- to implement a priority of characterizers for each class, i.e. the sequence will be different for each class.

The first option was used to optimize the testing process. However, we can assume that as the number of classes increases, the need to implement the second option will increase. In order to predict the transition to the second option, an additional property - priority was assigned for each characterizer.

2.3. Types of characterizers

The characterizer is an independent unit that calculates the characteristics of the card, for which it is responsible using its own algorithm ψi. In general, the characterizer is a software module that:

- calculates the characteristics of the card defined by it;
- provides an algorithm for storing taught data;
- is able to download stored taught data;
- answers the question of whether the card fits the taught data set in advance.

In addition, each characterizer has a priority - a natural number assigned to it from the outside when initializing the test block.

A feature for each characterizer is a set of real numbers that are calculated by a given algorithm ψi inside this characterizer. Thus, each characterizer has its own format for storing reference features. In the general case, the feature can be represented by the vector P = {p1, p2, …, pn}. The sets of reference features that are taught for each class are a dictionary. The storing of real numbers is quite simple to implement, so the issue of serialization of learning data can be considered fundamentally resolved.

The characterizer is determined by the following set of properties:

$$\Psi = \left\langle P, W, f_{\psi}(m^*), f_t(m^*) \right\rangle, \quad (4)$$

where P – the format of the feature; W – the algorithm for learning and storing learning data; fψ(m*) – the function for calculating features under the input matrix m*; ft(m*) – the function for testing the matrix m* based on taught data.

Here is a brief description of some types of characterizers developed for the proposed system.

The size characterizer determines the size of the input table and checks the size validity in this class. The feature format: P = {K, L}, where K is the number of lines, L is the number of columns. The learning algorithm is based on determining the smallest and largest of the sizes. So for a set of matrices M = {mi| i = 1,2,…,n} the feature vectors are calculated as follows:

$$\begin{aligned} P_{\min} &= \arg\min_i (size(m_i)); \\ P_{\max} &= \arg\max_i (size(m_i)); \end{aligned} \quad (5)$$

The testing algorithm checks the size availability in the list and determines whether the size is within the valid value range: (size(m*) ≥ Pmin) AND (size(m*) ≤ Pmax).

The value range characterizer determines the minimum and maximum valid values in the calibration table. The feature format: P = {min, max}. The learning is based on finding the minimum Pmin and maximum Pmax values from all matrices of the learning sample. The test checks the condition of the found minimum and maximum value in the range stored at the learning stage: (min(m*) ≥ Pmin) AND (max(m*) ≤ Pmax).

The finite difference characterizer [11] of the selected fragment. At initialization the function for selecting elements from a matrix defining the selected fragment D(m) = {x1, x2,…,xn} is set. The final differences of the selected fragment Δxi = xi+1 – xi are calculated and compared with the taught ones, based on the allowable residual vector.

2.4. Algorithm for determining the degree of similarity of sequences

An algorithm based on the calculation of approximation coefficients was used to determine the degree of similarity of feature vectors. The input data is defined as a set of sequences: Xi = {x0, x1, x2,…, xn} and is given to the average scale according to the following scaling algorithm. The midrange of values for all tested sequences is determined:

$$\begin{aligned} P_{\min} &= \frac{1}{n}\sum_{i=0}^{n} \min(X_i); \\ P_{\max} &= \frac{1}{n}\sum_{i=0}^{n} \max(X_i). \end{aligned} \quad (6)$$

The coefficient of proportionality is calculated:

$$k = \frac{P^i_{\max} - P^i_{\min}}{P_{\max} - P_{\min}}. \qquad (7)$$

The value of each element of sequence is calculated:

$$x'_{ij} = k(x_{ij} - P^i_{\min}) + P_{\min}, \qquad (8)$$

where Pimin – the minimum value of the ith sequence, Pmin – the minimum value of the midrange.

For scalable sequences, an approximation [12] of each sequence is performed on the basis of the polynomial: y = ax2 + bx + c. MNC was used for approximation in the system. The coefficients a, b, c are the basis for the study of similarity. In this case, the feature format: P = {a, b, c} or ΔP = {Δa, Δb, Δc}, where P is an average value of the vector of approximating coefficients, and ΔP is the maximum deviation of a, b, c from the average (vector P). For ease of presentation, the vectors can be combined into one 6-timer vector containing 3 coefficients and 3 of their maximum deviations.

The use of approximation to study the similarity has shown a good result, given the fairly high speed of its work.

For some characterizers to work, we need to store learning data more compactly. To do this, when learning in a sample with different sizes of features, key sizes are selected, and others are adjusted to such key sizes. To adjust the sequence to a key (or nodal) size, it is necessary not just to interpolate, but to scale, taking into account that the nodal points do not match. For this purpose the algorithm of interpolation with floating nodal points is offered. The fig. 3 shows an example of interpolation with floating nodal points, where blue points are nodes of the output function to be adjusted to the function shown by the red points.

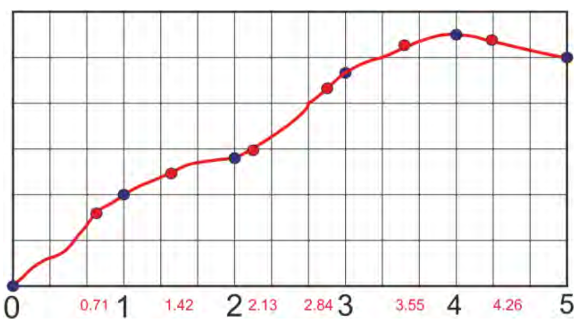


Figure 3: Interpolation with floating nodal points

Only the abscissas of the first and last nodes coincide. Therefore, for this kind of interpolation it is necessary to know only the new values of the abscissa. The algorithm consists of the following stages:

1. The scale, which is equal to the ratio of the number of intervals of the output function to the number of intervals of the objective function k = (n-1)/(m-1), where n is the number of points in the output function, m is the required number of points, is calculated.
2. The value of the abscissa is determined based on the scale by multiplying the node number by the found scale Xi=i×k.
3. For each new value of the abscissa, the value of the function is calculated by interpolation between adjacent points of the output function. Any of the existing interpolation algorithms can be used for this (in this case the linear interpolation algorithm was used).

3. The results of the study

The system contains two components: a test block, which contains all necessary characterizers, and a dictionary manager, which manages the test block. The system is built from components on the principle of single responsibility, when each component is responsible for only one action. In this case, the task of the test block is to answer whether the image corresponds to the current class being tested, based on the taught reference features provided to the test block. The role of another component is to manage the dictionary of the taught data and perform the classification task.

The software classification system was developed using the .NET Framework platform. The programming language is C#. To test the system availability, a console application was developed, as well as a small sampling of cards - 3-5 for each class was peformed, and only 3 different classes were selected. The cards were sorted by folders, in one folder - a separate class of cards; the size of cards was stated in the file name. First, a test block is created, in which the following characterizers are marked:

- size analyzer;
- spectrum analyzer;
- analyzer of finite differences of the first line;
- analyzer of finite differences of the last line;
- sequence similarity analyzer.

Then, for each class, a matrix is loaded into the array and added to the dictionary, having previously received the taught data. After initializing the data, one random card with the provided sample is loaded and the ability of the system to determine its class is checked. To debug the system, the values of the finite differences and the residual vector are displayed

The results of the preliminary testing are given in the table 1.

Table 1
The speed of learning in the preliminary testing

1 sample	1000 samples	1000000 samples
3 ms	28 ms	49 ms
4 ms	32 ms	55 ms

The analysis of the results of the preliminary testing has shown that the system is built satisfactorily. The next stage is to build the system of testing and to analyse the system efficiency. For this task, it was decided to be limited to modular testing. The Bosch EDC16C31/EDC16CP31 unit family for Mercedes-Benz cars were selected for testing. At the learning stage, a learning sample was created, consisting of more than 100 different cards, which were saved from 6 different calibration files of the same family of engine control units. This learning sample presents 10 main classes of cards, i.e. about 10 learning standards for each class. After that, several cards were selected for each class to test the accuracy. The table 2 shows the test results.

Table 2
The results of testing the accuracy of the system recognition

Card class	Number of standards	Number of test	Timems	Test result
Advance	19	5	55	100%
Boost Limiter via PAtm.	8	2	14	50%
Boost Limiter via Temp.	6	1	8	100%
Boost Pressure	16	6	82	100%
Driver's Wish	10	4	228	60%
Geometry Control	14	5	293	60%
Rail Pressure	20	5	310	100%
Smoke limiter via Boost	8	3	195	60%
Smoke Limiter	12	2	133	100%
Total	118	33	1318	74%

.

4. Conclusions

According to the results of testing the software system for classifying the parameters of the electronic engine control unit, the following can be noted: in cases where the number of learning images is large (12 or more), the result meets the requirements, for a small number of images - the classes begin to intersect.
On average, 5 tests take 150 ms. The tests were performed in the debug mode, so the speed indicators are overstated. When we run the test program from the EXE file in the Release mode, the time indicators will be lower.

The accuracy of classification after debugging the system was 75 - 82%.

The vector of further studies can be aimed at solving the problem of increasing the speed and accuracy of classification by implementing the priorities of characterizers, as well as the introduction of additional features and types of characterizers to eliminate intersections of

classes in case of classification by a small number of learning images

5. References

[1] European Parliament. (2009). Directive 2009/33/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of clean and energy-efficient road transport vehicles.

[2] Swiftec - Forget The Limits. Vcpowerteam.com. (2020). Retrieved 11 November 2020, from https://www.vcpowerteam.com/products/swiftec/.

[3] Software WinOLS. EVC electronic. (2020). Retrieved 11 November 2020, from https://www.evc.de/en/product/ols/software/

[4] de Nola F., Giardiello G., Gimelli A., Molteni A., Muccillo M., Picariello R. Volumetric efficiency estimation based on neural networks to reduce the experimental effort in engine base calibration. Fuel, 2019. Vol. 244. P. 31-39.

[5] Makarov A. (2011). Algorithms for control and diagnostics of aircraft GTE control systems based on neural network models and fuzzy logic (candidate of technical sciences). Ufa State Aviation Technical University.

[6] Jian-Da W., Chiu-Hong L. An expert system for fault diagnosis in internal combustion engines using wavelet packet transform and neural network. Expert Systems with Applications, 2009. Vol. 36. P. 4278-4286.

[7] Lujan J. M., Climent H., García-Cuevas L. M., Moratal A. Volumetric efficiency modelling of internal combustion engines based on a novel adaptive learning algorithm of artificial neural networks. Applied Thermal Engineering, 2017. Vol. 123. P. 625-634.

[8] Shatnawi Y.; Al-khassaweneh M. Fault Diagnosis in Internal Combustion Engines Using Extension Neural Network. IEEE Transactions on Industrial Electronics, 2013. Vol. 61. P. 1434 - 1443.

[9] Jian-Da W., Peng-Hsin Ch., Yo-Wei Ch., Yao-jung Sh. An expert system for fault diagnosis in internal combustion engines using probability neural network. Expert Systems with Applications, 2008. Vol. 34. P. 2704-2713.

[10] Jian-Da W., Chen J.-Ch. Continuous wavelet transform technique for fault signal diagnosis of internal combustion engines. NDT & E International, 2006. Vol. 39. P. 304-311.

[11] Soloveichik, Yu.G., Royak, M.E., & Persova, M.G. (2007). Finite element method for scalar and vector problems. NSTU.

[12] Laurent, P.J. (1975). Approximation and Optimization. Peace.

Intelligent Model of Business Ecosystem with Competitive Behavior

Halyna Padalko[1], Volodymyr Kobziev[2], Iryna Yakovleva[3] and OlgaYarovaya[1]

[1] *National Aerospace University "Kharkiv Aviation Institute", 17 Chkalova st., 61070, Kharkiv, Ukraine*
[2] *Kharkiv National University of Radio Electronics, 14 Nauki ave., 61166, Kharkiv, Ukraine*
[3] *National University of Urban Economy in Kharkiv, 17 Marshala Bazhanova st., 61002, Kharkiv, Ukraine*

Abstract

Mathematical and simulation modeling plays an important role in economics. Not only for forecasting, planning and solving other applied problems, but also for the study of socio-economic processes. The aim of the article was to develop a model of society, divided into three types of players, which reflect competitive behavior in a closed system. Hawks show aggressive strategy, doves passive and law-abiding demonstrative. This set of players expands on the classic "hawk-dove" problem in game theory. This simplest model of game theory describes the competitive relationship in a particular animal population and the development of an evolutionarily stable strategy. Using the MatLAB Simulink environment, strategies for behavior in a competitive environment have been developed. The system has been tested for equilibrium. The rules for winning players belonging to different types have been formulated.

Keywords

Competitive system, game theory, sustainability, economical modelling, business system.

1. Introduction

The coronavirus pandemic has had a huge impact not only on human life. The world economy has suffered from it [1]. World governments are looking for a way out of a difficult economic situation. According to forecasts, GDP in most countries is expected to decline by 4-7% in 2022 [2]. After the announcement of a global pandemic, the approach of the global economic crisis became very tangible. The risks associated with the spread of the coronavirus have negatively affected the value of shares and assets of many companies [3].

These problems include:

- Decreased production activity.
- Decrease in forecast indicators of the economy.
- Shrinking the service sector.
- Stock market problems.
- Decline in oil prices.
- Decrease in world tourism.

Also, the COVID-19 pandemic and the restrictions caused by it shortened the period of digital transformation of the economy, forcing companies to rapidly rebuild their business processes [4]. The business transformation due to the pandemic will lead to a radical change in the approach both to working with human capital and to a change in the required skills of employees. Already today, focal digitalization is accelerating en masse. Businesses need people who can embed digital technologies in companies, work digitally with each other, with customers and suppliers, and add value [5]. Over the years preceding the pandemic, we have seen a widespread adoption of digital approaches in various fields of activity, such as education [6-7], business [8-9], security [10-12], insurance [13-14], project management [15-16], urban economy [17-18], medical diagnostics [19-20], public health [21-22], etc.

Analyzing current economic data related to the coronavirus epidemic is a major challenge. Investors, analysts and traders should accumulate positive and negative changes in the industry and

EMAIL: galinapadalko95@gmail.com (A. 1), volodymyr.kobziev@nure.ua (A. 2) yakovlevairine@gmail.com (A. 3), helga.yarovaya@gmail.com (A. 4)
ORCID: 0000-0001-6014-1065 (A. 1);
0000-0002-8303-1595 (A. 2); 0000-0001-5204-9779 (A. 3);
0000-0003-4969-1677 (A. 4)

individual companies, check the reliability of information, group it, analyze it. This tool is the modeling of economic systems [23].

Mathematics makes it possible to uniformly describe a wide range of facts and observations, carry out their detailed quantitative analysis, help predict how the object under study will behave in different conditions, that is, justify the development forecast, and justifying predictions are already the pride of any science. Mathematization of any field of activity is carried out through the development and penetration of three elements of applied mathematics – mathematical models, computational algorithms and computers [24]. The mathematical model is based on a certain simplification of the original, it is not a copy, but always an approximate reflection of it [25]. By replacing a real object with a corresponding model, it becomes possible to formulate the problem of studying the original as a mathematical one and solve it with a universal mathematical apparatus that does not depend on the specific area of the subject being studied [26].

The use of mathematical methods in economics refers to the period of the birth of economic sciences; stages of development, ups and periods of decay can also be distinguished. But the situation changed dramatically with the advent of computers – economic and mathematical research received an additional impetus [27]. First of all, let us note that computers themselves are developing rapidly, drawing ever new branches into this process and, undoubtedly, acting as a catalyst for scientific and technological progress. The process of mathematization of the science of economics demanded the training of highly qualified specialists capable of realizing those enormous opportunities, the truly enormous potential that the study of economics gives by the methods of mathematics and cybernetics in a single symbiosis [28]. Thus, the use of the mathematical apparatus and simulation modeling in the study of economic processes makes it possible to identify new patterns for effective decision-making [29].

2. Problem statement

The rules for the competitive behavior of players in the economic system are defined. It is necessary:

- to formulate the rules for winning the players of the original model;
- to develop a software model of competitive behavior in a business system;
- to determine the equilibrium points and system stability characteristics.
- to determine the equilibrium points and system stability characteristics.

The aim of the paper is to develop model of competitive behavior in the business system.

Object of research is business processes in closed competitive business system.

Subject of research is social and economic simulation approach.

3. Methods and models

Let's assume that society is divided into three groups according to their behavior in business environment. The strategies are defined according to expanded classical game theory model "Hawk-Dove" [30] and consist of three types: "Hawk", "Dove", "Law-abiding".

The Hawk-Dove model is as follows [31]. When two animals fight for prey, as a rule, both act aggressively and try to injure the enemy. When the fight is about to begin, two options are possible: to retreat, losing prey, but keeping life (this is how pigeons do), or to fight until victory and, possibly, lose life (this is how hawks act). Suppose a small group of hawks appears in a pigeon community. Initially, the number of hawks will grow, since their strategy is more profitable (whenever a dove and a hawk meet, the hawk will be the winner). As the number of hawks grows, the number of skirmishes between them will increase, which means that the damage they will inflict on each other will also increase. Over time, a certain balance will be formed between hawks and pigeons. This is exactly what happens in the real world.

In the proposed model, the classic types of players are supplemented with one more – "Law Abiding". So we have three possible behaviors: demonstration, fighting, run, and each individual act on one of the strategies in table 1.

An individual applying a strategy $i = \overline{1,3}$ to an opponent applying a strategy $j = \overline{1,3}$ gets a win a_{ij}. It is believed that the gain a_{ij} affects the behavior of the economic system. Suppose that only pure strategies are applied, i.e. each individual always belongs to the same type and applies the same strategy, and that the offspring inherits the strategy of the parent.

Number i	Strategy	Initial tactic	Tactics
1	Hawk	Fight	Fight
2	Dove	Demonstration	Run
3	Law-abiding	Fight	Fight

Table 1. Tactics of players for given strategies.

Let x_i is part of the players of the economic system, applying the strategy i. Then

$$\sum_{i=1}^{3} x_i = 1 \qquad (1)$$

where $x_i \geq 0$.

Winning for players using strategy i against all others is

$$\sum_{j=1}^{3} a_{ij} x_i = (Ax)_i \qquad (2)$$

where A is the payment matrix. At the intersection of the two strategies, you can see the winnings that the players will receive. The player's win is

$$\sum_{i=1}^{3} x_i (Ax)_i = x^T Ax \qquad (3)$$

Therefore, the profit from applying the strategy i is equal to

$$(Ax)_i - x^T Ax \qquad (4)$$

The player's reproduction rate for a group applying strategy i is considered to be proportional to the advantage of this strategy, which gives

$$\dot{x}_i = x_i((Ax)_i - x^T Ax) \qquad (5)$$

Equation (5) makes sense only for those points of the space R^3 that satisfy condition (1). That is, for the area of possible strategies, we can get a payment matrix by setting "points" for the result of each collision.

For example:

victory = 6,
defeat = 0,
injury = -10,
loss of time = -1.

Those specific values that are given here are irrelevant, their sign and the order of absolute values are important.

If a "hawk" meets with a "dove" or "law-abiding", then it wins, so that $a_{12} = a_{13} = 6$. If there are two "hawks", they fight until one of them gets injured.

Both "hawks' win with equal probability, and the gain is equal to $a_{11} = 0.5 * (6-10) = - 2$.

If the "dove" meets the "hawk" or "law-abiding", then it loses, therefore $a_{21} = a_{23} = 0$, but two "doves" continue their demonstrations to each other until one of them surrenders, so $a_{22} = 0.5 * (6 + 0) -1 = 2$.

Finally, "law-abiding" lose to "hawks" ($a_{31} = 0$), win against "doves" ($a_{32} = 6$) and have a 50% chance of winning from their own kind ($a_{33} = 0.5 * (6 + 0) = 3$). In this way

$$A = \begin{bmatrix} -2 & 6 & 6 \\ 0 & 2 & 0 \\ 0 & 6 & 3 \end{bmatrix} \qquad (6)$$

It is also useful to note that the advantage of the strategy does not change if any column of the matrix A is added with a constant value. Using such a transformation, the matrix A can be simplified by making its diagonal elements equal to zero. The dynamic equations (5) will not change. Therefore, we can assume

$$A = \begin{bmatrix} 0 & 4 & 3 \\ 2 & 0 & -3 \\ 2 & 4 & 0 \end{bmatrix} \qquad (7)$$

4. Results

The developed model is implemented in the MatLAB Simulink environment. The software implementation of the competitive behavior model in the business system in the MatLAB Simulink environment is shown in Figures 1.

The software implementation of this model has four subsystems. $x^T \cdot A \cdot x$ (Fig. 2), $(Ax)_1$ (Fig. 3), $(Ax)_2$ (Fig. 4), $(Ax)_3$ (Fig. 4).

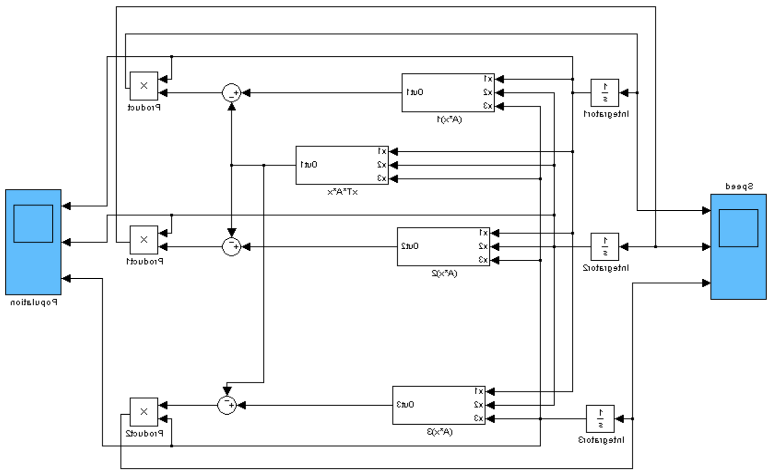


Figure 1: Model of competitive behavior.

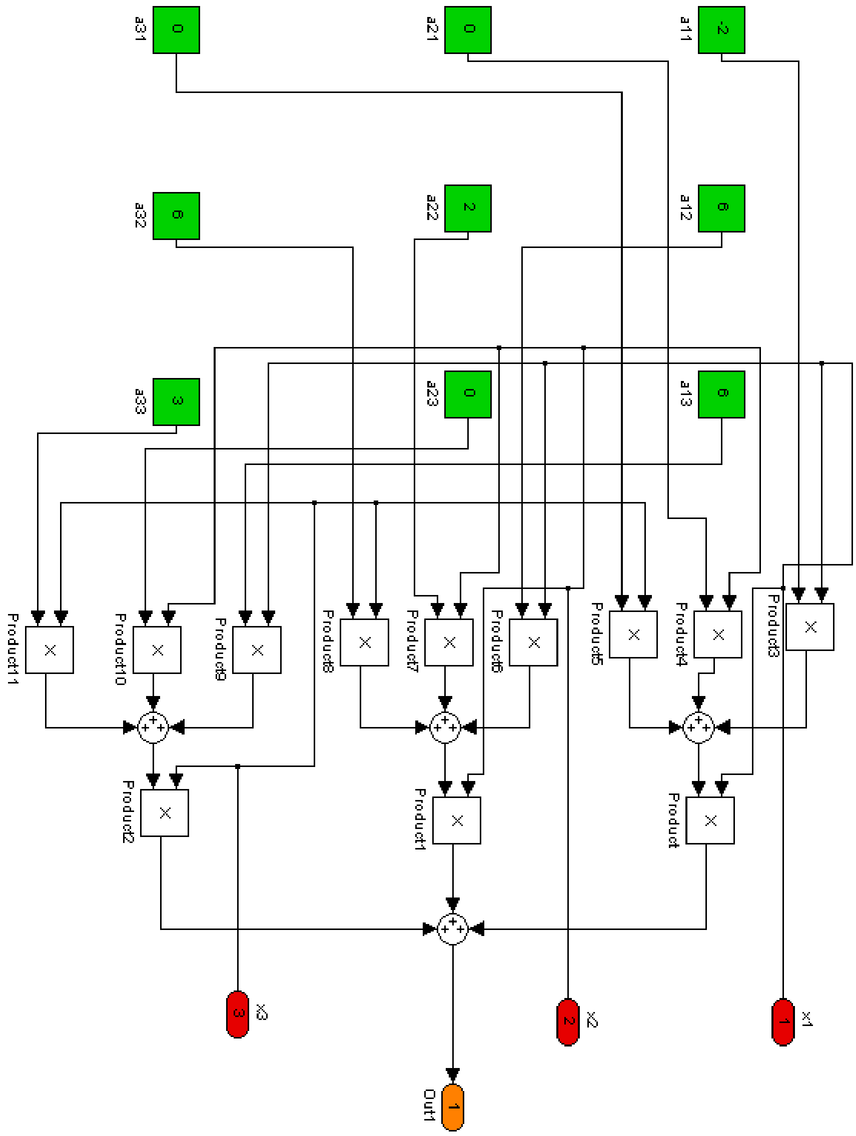


Figure 2: Subsystem x^TAx.

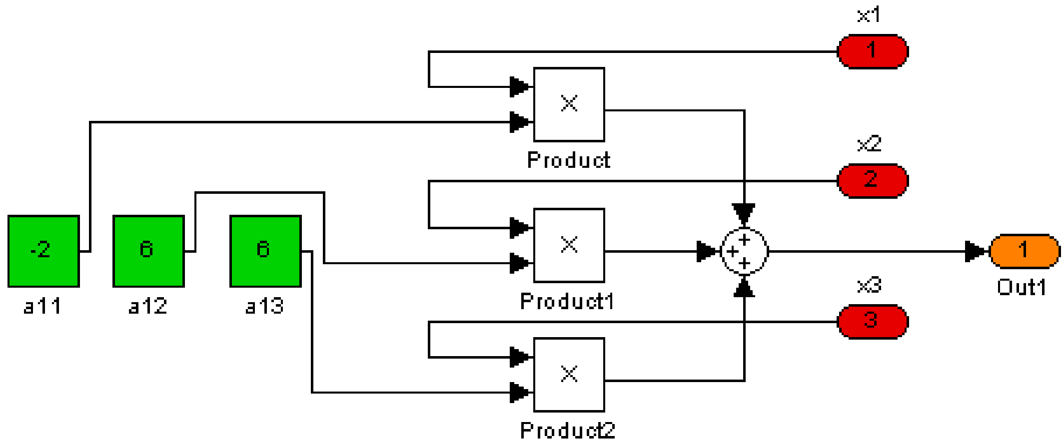


Figure 3: Subsystem Ax_1.

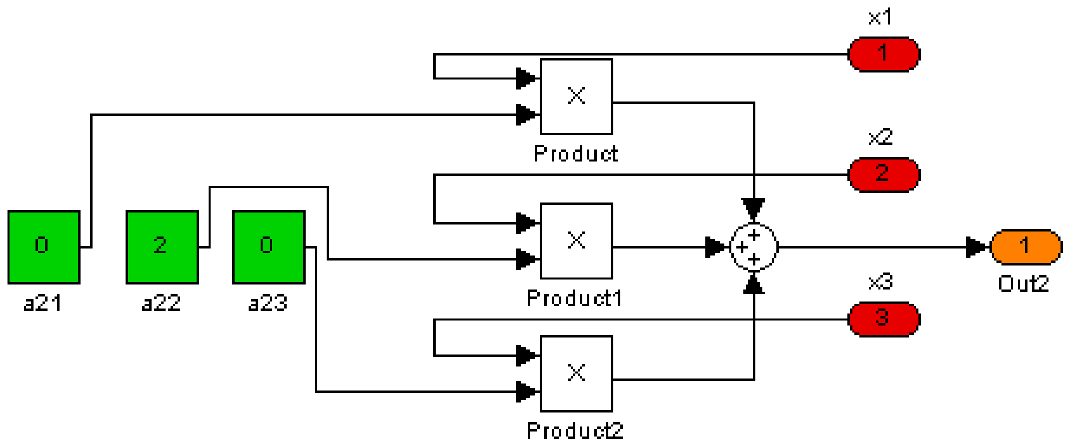


Figure 4: Subsystem Ax_2.

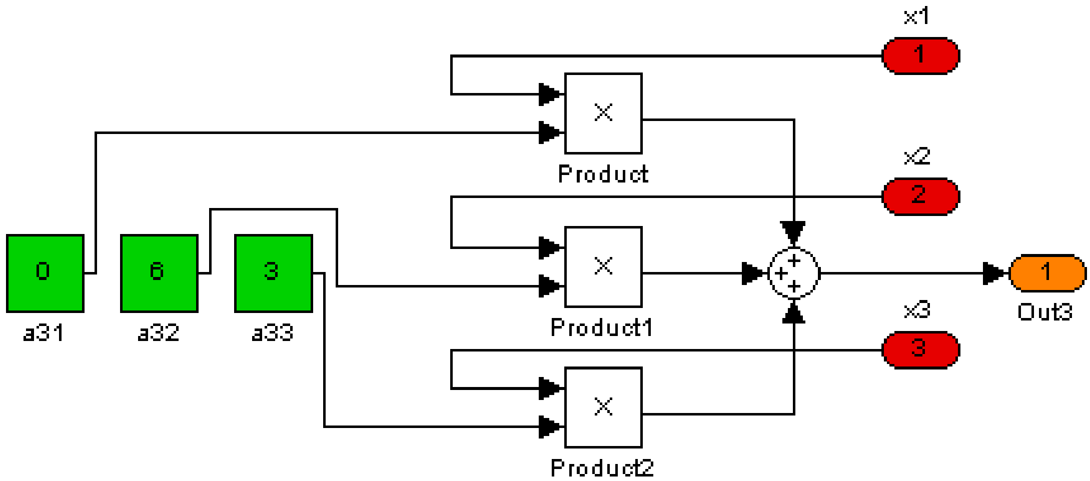


Figure 5: Subsystem Ax_3.

Results of simulation has shown the following results. Figure 6 represents dynamics of the population with each strategy.

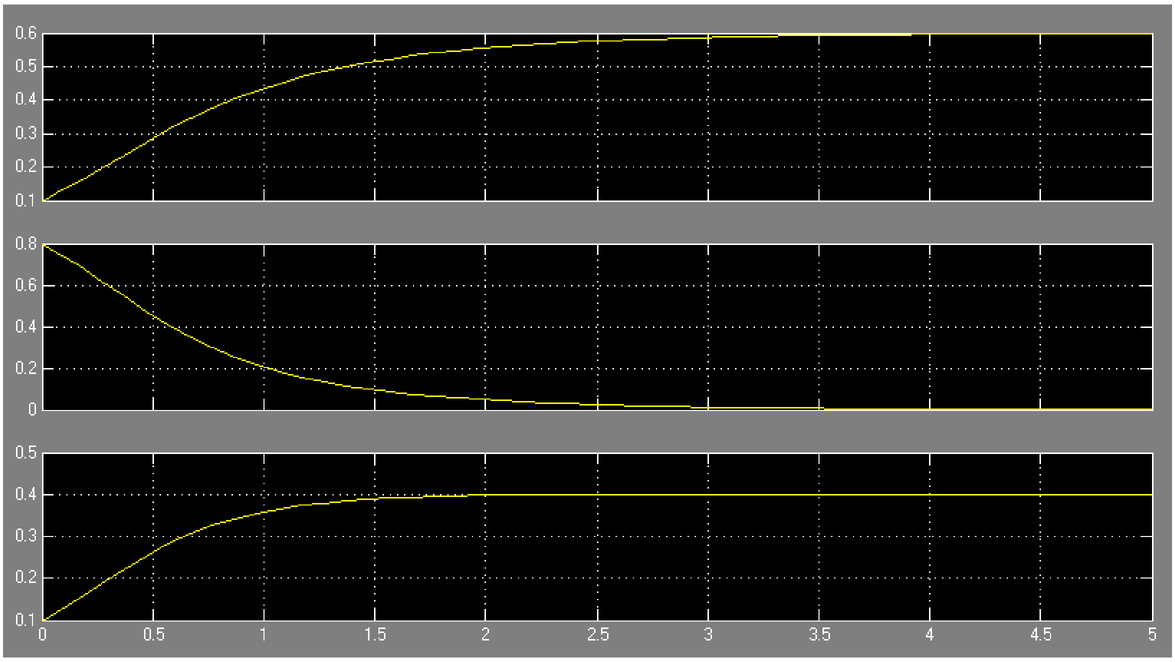


Figure 6: Dynamics if the population development with each strategy (“Hawk”, “Dove”, “Law-abiding”).

Figure 7 represents the reproduction rate of individuals with each strategy.

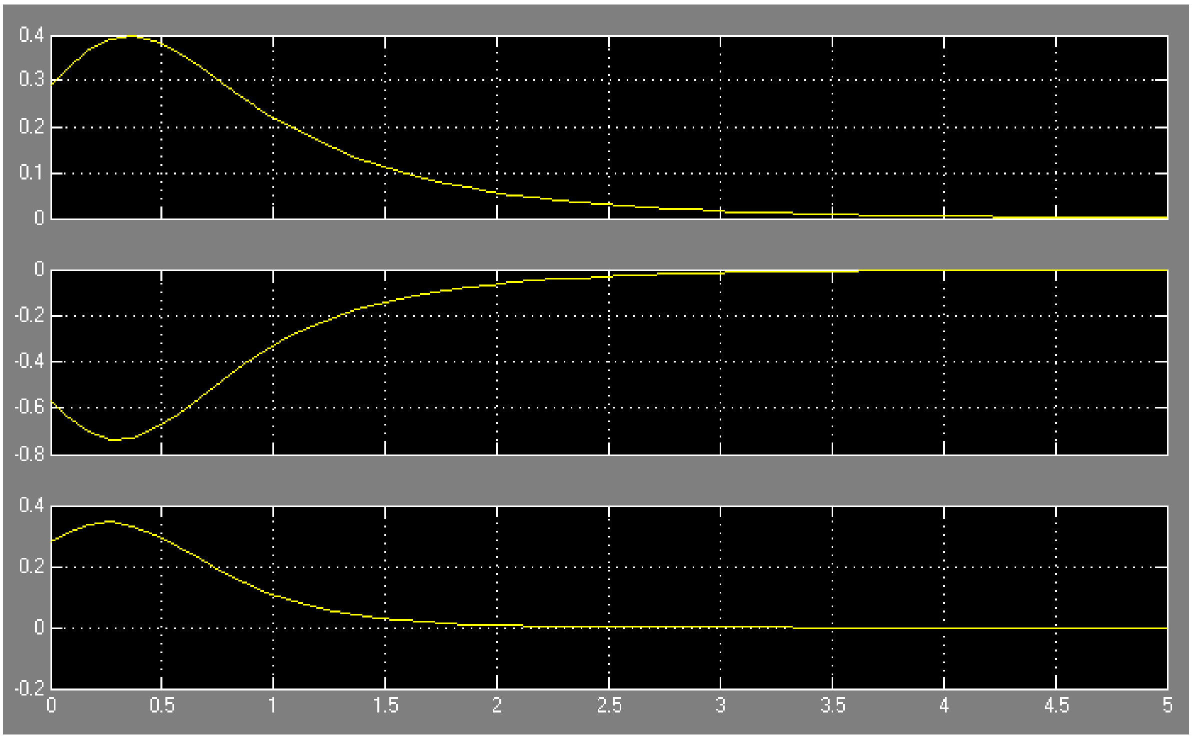


Figure 7: Reproduction rate for each strategy (“Hawk”, “Dove”, “Law-abiding”).

5. Conclusions

Under given research model of competitive behavior in business system has been developed for three-strategy competitive game.

In the described model, it is obvious that neither of the two strategies of the classic game "Hawk-Dove" is satisfactory: hawks always prevail over pigeons, but lose in fights with each other, and pigeons do not get injured in fights with each other, but always yield to hawksю The independent party "Law-Abiding" helps to reduce the number of fights between hawks and at the same time prohibits them from taking advantage of the behavior of pigeons, reducing the number of aggressive encounters.

A promising continuation of research is the study of the system for stability.

References

1. D. Chumachenko, et. al.: On-Line Data Processing, Simulation and Forecasting of the Coronavirus Disease (COVID-19) Propagation in Ukraine Based on Machine Learning Approach, Communications in Computer and Information Science 1158 (2020) 372-382. doi: 10.1007/978-3-030-61656-4_25
2. D. Cutler D.: How Will COVID-19 Affect the Health Care Economy? Journal of American Medical Association. 323 (22) (2020) 2237-2238. doi: 10.1001/jama.2020.7308.
3. Y. Xiao, G. Chen, X. Nie, Z. Bai, X. Zhu: Intelligent Prediction of Anti-risk Repayment Ability of Small and Medium-sized Enterprise Commercial Loan Data Based on Ant Colony Algorithm. 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) (2021) 1653-1656. doi: 10.1109/ICOEI51242.2021.9453015.
4. P. Carrijo, B. Alturas, I. Pedrosa: Analysis of Digital Transformation Maturity Models. 2021 16th Iberian Conference on Information Systems and Technologies (CISTI). (2021) 1-6. doi: 10.23919/CISTI52073.2021.9476644.
5. F. Murillo-Cabezas, E. Vigil-Martín, N. Raimondi, J. Pérez-Fernández: Pandemia de Covid-19 y transformación digital en Cuidados Intensivos [Covid-19 pandemic and digital transformation in critical care units]. Medicina Intensivna (English Edition) 44 (7) (2020) 457-458. doi: 10.1016/j.medin.2020.04.004.
6. D. Chumachenko, et. al.: Intelligent expert system of knowledge examination of medical staff regarding infections associated with the provision of medical care. CEUR Workshop Proceedings. 2386 (2019) 321-330.
7. N. Iivari, S. Sharma, L. Ventä-Olkkonen: Digital transformation of everyday life – How COVID-19 pandemic transformed the basic education of the young generation and why information management research should care? International Journal of Information Management. 55 (2020) 102183. doi: 10.1016/j.ijinfomgt.2020.102183.
8. P. Piletskiy P., et. al.: Development and Analysis of Intelligent Recommendation System Using Machine Learning Approach. Advances in Intelligent Systems and Computing 1113 (2020) 186-197. doi: 10.1007/978-3-030-37618-5_17
9. B. Trenerry, et. al.: Preparing Workplaces for Digital Transformation: An Integrative Review and Framework of Multi-Level Factors. Frontiers in Psychology. 12 (2021) 620766. doi: 10.3389/fpsyg.2021.620766.
10. D. Chumachenko, S. Yakovlev: On Intelligent Agent-Based Simulation of Network Worms Propagation. 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM). (2019) 11-14, doi: 10.1109/CADSM.2019.8779342.
11. S. Kernebeck, et. al.: Impact of mobile health and medical applications on clinical practice in gastroenterology. World Journal of Gastroenterology. 26 (29) (2020) 4182-4197. doi: 10.3748/wjg.v26.i29.4182.
12. D. Chumachenko, K. Chumachenko, S. Yakovlev: Intelligent Simulation of Network Worm Propagation using the Code Red as an Example. Telecommunications and Radio Engineering 78 (5) (2019) 443-464. doi: 10.1615/TelecomRadEng.v78.i5.60
13. K. Bazilevych, et. al.: Stochastic modelling of cash flow for personal insurance fund using the cloud data storage. International Journal of Computing 17 (3) (2018) 153-162.
14. N. Dotsenko, et. al.: Modeling of the processes of stakeholder involvement in command management in a multi-project environment. 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings. 1 (2018) 29-32.

doi: 10.1109/STC-CSIT.2018.8526613
15. V.P. Mashtalir, S.V. Yakovlev: Group structures on quotient sets in classification problems. Cybernetics and Systems Analysis. 50 (4) (2014) 507-518. doi: 10.1007/s10559-014-9639-z
16. N. Dotsenko, et. al.: Project-oriented management of adaptive teams' formation resources in multi-project environment. CEUR Workshop Proceedings 2353 (2019) 911-923.
17. N. Davidich, et. al.: Advanced Traveller Information Systems to Optimizing Freight Driver Route Selection. 2020 13th International Conference on Developments in eSystems Engineering (DeSE) (2020) 111-115.doi: 10.1109/DeSE51703.2020.9450763.
18. N. Davidich, et. al.: Monitoring of urban freight flows distribution considering the human factor. Sustainable Cities and Society. 75 (2021) 103168. doi: 10.1016/j.scs.2021.103168
19. M. Mazorchuck, et. al.: Web-application development for tasks of prediction in medical domain. 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings. 1 (2018) 5-8. doi: 10.1109/STC-CSIT.2018.8526684
20. D. Chumachenko, T. Chumachenko: Intelligent Agent-Based Simulation of HIV Epidemic Process, Advances in Intelligent Systems and Computing. 1020 (2020) 175-188. doi: 10.1007/978-3-030-26474-1_13
21. Y. Polyvianna, et. al.: Computer aided system of time series analysis methods for forecasting the epidemics outbreaks. 2019 15th International Conference on the Experience of Designing and Application of CAD Systems, CADSM 2019 - Proceedings (2019) 8779344. doi: 10.1109/CADSM.2019.8779344
22. D. Chumachenko, et. al.: On agent-based approach to influenza and acute respiratory virus infection simulation, 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018 – Proceedings. (2018) 192-195. doi: 10.1109/TCSET.2018.8336184
23. S. Yakovlev, et. al.: The concept of developing a decision support system for the epidemic morbidity control, CEUR Workshop Proceedings 2753 (2020) 265–274.
24. S. Jabeen, T. S. Malik, A. Abdullah, M. H. Hasan: Lockdown Strategy to Control Covid-19 Pandemic Using Automata. 2021 International Conference on Computer & Information Sciences (ICCOINS) (2021) 384-389, doi: 10.1109/ICCOINS49721.2021.9497149.
25. V.P. Mashtalir, V.V. Shlyakhov, S.V. Yakovlev: Group structures on quotient sets in classification problems. Cybernetics and Systems Analysis 50 (4) (2014) 507-518. doi: 10.1007/s10559-014-9639-z
26. C. Pu, Z. Quan, D. Daud: Research on Enterprise Economic Management Mode under Market Economy. 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME) (2020) 243-246. doi: 10.1109/ICEMME51517.2020.00054.
27. Shih-Yang Chiao, C. S. Xydeas: Modeling the behaviors of players in competitive environments. IEEE/WIC International Conference on Intelligent Agent Technology, 2003. IAT 2003. (2003) 566-569. doi: 10.1109/IAT.2003.1241146.
28. D. Chumachenko, O. Sokolov, S. Yakovlev: Fuzzy recurrent mappings in multiagent simulation of population dynamics systems. International Journal of Computing 19 (2) (2020) 290-
29. S.N. Gerasin, V.V. Shlyakhov, S.V. Yakovlev: Set coverings and tolerance relations. Cybernetics and Systems Analysis 44 (3) (2008) 333-340. doi: 10.1007/s10559-008-9007-y
30. D. Chumachenko, I. Meniailov, K. Bazilevych, T. Chumachenko: On Intelligent Decision Making in Multiagent Systems in Conditions of Uncertainty. 2019 11th International Scientific and Practical Conference on Electronics and Information Technologies, ELIT 2019 – Proceedings. (2019) 150-153. doi: 10.1109/ELIT.2019.8892307
31. T. Sakiyama: Temporary Suspension brings about the Cooperative Evolution in the Hawk-Dove Game: Temporary Suspension in the Hawk-Dove Game. 2019 IEEE Symposium Series on Computational Intelligence (SSCI) (2019) 3221-3224, doi: 10.1109/SSCI44817.2019.9002950.

Automated System Development for the Printed Circuit Boards Optical Inspection Using Machine Learning Methods

Igor Nevliudov[1], Iryna Botsman[1], Olena Chala[1] and Kirill Khrustalev[1]

[1] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract
The problem of printed circuit boards (PCB) quality optical inspection at their production stage is considered. The automated method of PCB optical inspection on the machine learning methods basis is proposed. The necessary neural network parameters to develop an automated PCB inspection method are calculated. The main capabilities of the created artificial neural network for identifying PCB under test defects are analyzed. The results of the conducted neural network testing that confirm its operability and possibility of use for PCB inspection at the stage of production are presented. The software program was developed that is used for transformations over images, such as converting an image to a grayscale color space and image binarization, which speeds up the neural network by reducing the size of the input matrix to a binary value per pixel of the image. The accuracy of finding each of the PCB defects types was also investigated.

Keywords
Automated system, optical inspection, printed circuit boards, machine learning, intelligent manufacturing, neural networks.

1. Introduction

Today in the instrument-making industry one of the most responsible stages for production of printed circuit boards (PCBs) as components of electronic devices for a wide applications range is quality control [1]. At the same time, the task of developing and implementing of new high-precision and universal inspection methods is relevant to ensure a high percentage of accepted products, taking into account the constant complexity of the modern PCBs design, including the conductive topologies density increasing and their weight reducing [2].

Optical PCBs inspection allows to quickly identify errors that occurred during their manufacture, for example:

An example of bulleted list is as following.

- mechanical damages to the board, e.g. cracks in the dielectric base, splits, distortions, etc., which can be caused by defects in the base material and occur in case of some technological operations modes violation during the PCB structure manufacturing process;
- deviation of tolerances during the PCB tracing formation;
- occurrence of current-carrying elements short circuits or breaks;
- manufacture defects of mounting holes in the PCB, for example, in case of non-compliance with the relevant drilling modes or filling the hole with solder or solder mask, which will prevent further work with this product;
- violation of the contact pads geometry, etc.

During the technological process of the finished printed module assembly and installation and after its completion the quality control of installation of electronic components (EC) is performed, in particular, the accuracy inspection of positioning the EC packages pins on contact pads or into holes, the quality of solder or welded joints, etc [3-6].

EMAIL: igor.nevliudov@nure.ua (A. 1); irina.botsman@nure.ua (A. 2); olena.chala@nure.ua (A. 3), kirill.khrustalev@nure.ua (A. 4)
ORCID: 0000-0002-9837-2309 (A. 1); 0000-0003-1110-9602 (A. 2); 0000-0003-2454-3774 (A. 3), 0000-0002-0687-5153 (A. 4)

2. Research Objective

One of the possible ways to improve automated methods of optical PCBs inspection is the use of artificial neural networks to classify objects in the images of the products under test [7-9].

Therefore, the aim of the work is to develop a new automated method of printed circuit boards optical inspection based on the creation and training of a neural network.

The research object is the process of printed circuit boards inspection, the subject of the study is PCB with a printed conductor pattern. For research performing the methods of theoretical researches, comparative analysis and methods of machine learning are used.

To recognize the objects on PCB images it was decided to use convolutional neural networks, the architecture of which is based on the knowledge that the input information is an image. It significantly reduces the required number of calculations.

In particular, the YoloV4 network, which is based on the DarkNet architecture, was chosen as the basis for the necessary neural network building. The use of this architecture allows neural networks to work with high accuracy at high bit rates or even when used with video.

3. PCB Images Processing Steps

The search for defects on the PCB requires its image to be large, usually these dimensions are close or larger than a square with a side of 4096 pixels with three bits of color, and an example is shown in Fig.1.

Figure 1: PCB photo example, size 4032 × 3850 pixels

To work with images for conversion more effectively, it is possible to use the OpenCV computer vision library, which includes a large number of functions for image conversion, as well as machine learning capabilities. Built-in features also include special functions needed to obtain individual image details or entire objects
using their geometric characteristics.

A set of data with all possible types of PCBs defects was created for the training. To increase the learning speed and reduce the error probability, 1500 images of PCBs with test data were selected, which are part of the scanned images in black and white variant, where white is allocated to the PCB area without copper layer, black is allocated to areas covered with copper layer.

Each image is presented in two formats: with a defect and without a defect, in the size of 640 × 640 pixels, each image contains from 3 to 12 defects. Of the total set of images, 1000 was set aside for training, 500 was set aside for test data.

The maximum number of steps for the neural network learning was determined, which will avoid retraining, it will be equal to 12,000.

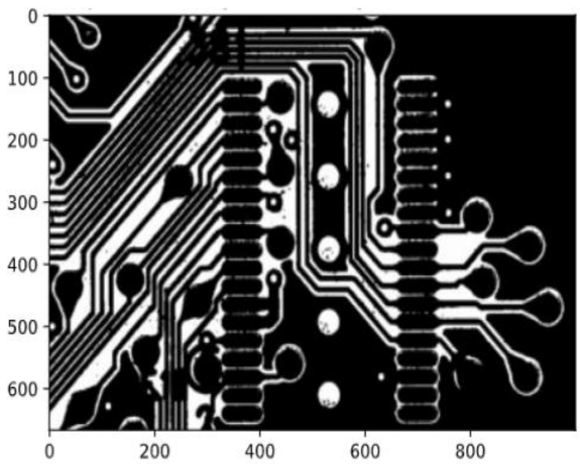
In order to increase the accuracy and speed of the neural network, as well as to get rid of problems that may arise due to uneven lighting of the object under test, it was decided to create a program for image binarization, i.e. convert it into the format for which there are only two colors – black and white.

To create such a program, the Python programming language was chosen due to some advantages, such as:

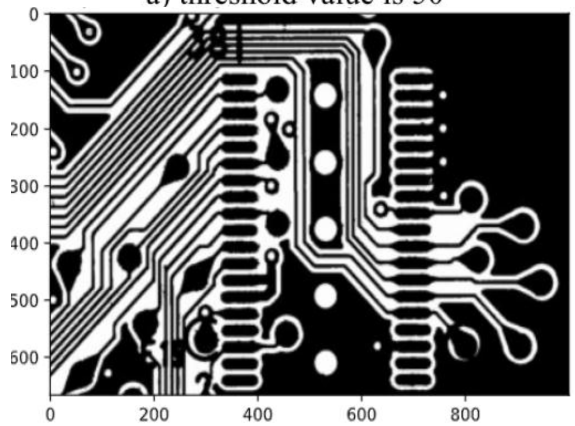
- ease of use;
- large number of libraries for working with images;
- ability to connect a neural network as a third-party library.

For binarization, the most important parameter is the threshold value in the range from 0 to 255. In a black and white image this value indicates the limit at which the image will be divided into black and white. Three threshold values in 50, 100 and 150 were tested and the results were compared, the results of binarization are shown in Fig.2. Thus, we can assume that the best threshold value in this case is 100.

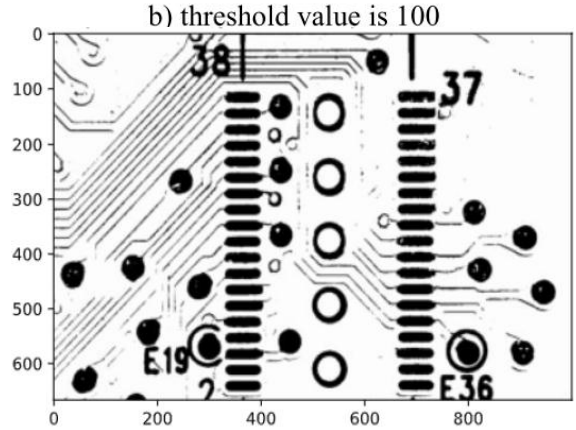
The layers number of neural network is 23 and the stride value is 4.



a) threshold value is 50



b) threshold value is 100



c) threshold value is 150

Figure 2: Examples of binarized images with different threshold values

4. Training and Testing an Optical Inspection Automated System Using Machine Learning

In order to prevent network retraining, with test dataset use the weights with the lowest error value were found. In this case the error can be considered as the found region deviation from the region that was marked manually.

The graph of defective regions recognition. accuracy depending on the training iterations number of developed neural network was constructed (Fig.3). It is experimentally determined that the best result on the test data is in the range from 8300 to 8600 learning iterations of the created neural network.

The neural network speed was tested, the image processing time was from 60 to 120 ms, which allows to process from 10 to 15 images per second. This result is acceptable for use in the PCBs inspection process in real time during their production.

The effect of image size on processing speed was also tested, as the created neural network should be able to be used with high-resolution images taken with a digital camera or with images obtained by scanning. For this check, 10 images with a resolution of 1280 × 1280 pixels were created based on existing images and tested using the same method as the smaller images. The processing time of the enlarged image is in the range from 94 to 143 ms and allows processing the image at a speed of 7-10 frames per second.

It should be noted that the image processing speed depends on its size, but increasing the image area has increased the time non-linearly.

This allows to use the created neural network to work with high-resolution images or images obtained by scanning, which confirms the possibility of its use for inspection of a large PCBs variety, as well as for large volumes of production.

5. Investigation of the Neural Network Accuracy Depending on the Image Inclination Angle

As the test data images with straight lines were used. But in production it is not always possible to obtain such images, and often the obtained PCB images will have an angle relative to the frame boundaries. The factors that may affect the angle image rotation are:

- scanner or camera installed incorrectly;
- lack of a mechanism for products aligning;
- human factor, if testing is performed in an automated mode;
- conductive drawing of non-standard shape.

The inclination angle can also be affected by the geometric features of a particular PCB type, which will be placed in products with complex package geometry.

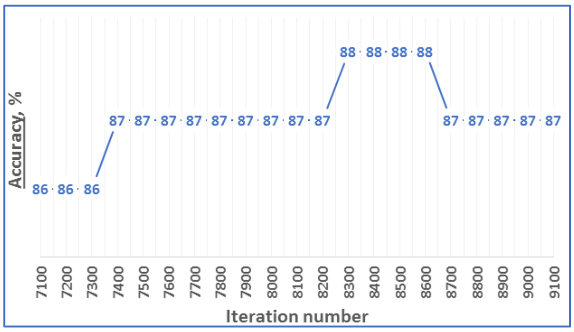


Figure 3: Graph of the trained dataset accuracy

Therefore, in order to ensure that the neural network remains ready of use for PCB inspection and the impact on accuracy depending on the image angle is minimal, it is necessary to perform additional research. Such studies were performed on the original dataset, which was modified, and the following steps were used to perform the study:

- using the program called imagemagick a test images were rotated at 15°, 30° and 45° clockwise, with image size increasing and without original image contours cropping;
- the images were selected in such way, that each of the defects occurred on them exactly 100 times and had different variants of display on the PCB;
- also due to the fact that when the images are rotated, negative areas appear where there is no conductor pattern, the testing program has been modified to take into account the region in which the conductor pattern is located. Images and results obtained outside such a region do not affect the accuracy calculation;
- the size of the region has been reduced by 2 pixels relative to the original image size to prevent errors caused by artifacts of edge smoothing algorithms used in the image conversion process due to the fact that in photos such transformations help to achieve the best result.

One of the examples of rotated images is shown in Fig. 4.

Also, the list of defects was replaced with new ones for each of the transformed images because of the fact that trigonometric transformations to the corresponding lists are impossible due to the peculiarities of defect registration in the yolo neural network and due to the fact that the image was resized.

That is, taking into account such changes in the image, the algorithm for accuracy determining of the neural network was changed.

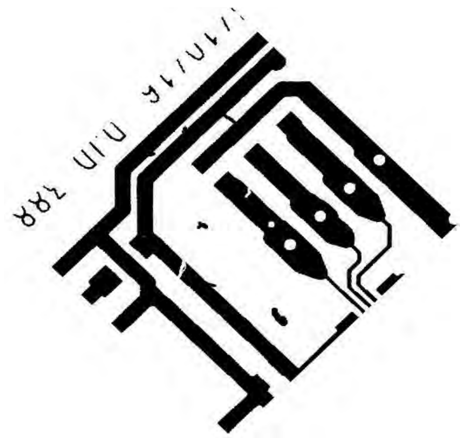


Figure 4: Example of a converted image rotated by 45°

The defect region was reduced, to determine that the defect was classified, instead of calculating the total area between the defined region and the region specified in the text file, an algorithm was chosen according to which the area of intersection of regions relative to a manually defined region will be considered a match. This algorithm allows more accurately determining the defect presence, rather than the accuracy of its location.

The average values of the defects finding accuracy depending on the inclination of the image are given in Table 1.

Table 1
Average accuracy values depending on the angle of the image inclination

Inclination angle	Accuracy, %
0°	88
15°	86
30°	87
45°	88
Average value	87,5

Thus, given the results obtained, it can be assumed that the accuracy of the neural network depends on the angle of the image. But this may be more due to the artifacts or smoothing algorithms used to create the test data than the inaccuracy of the trained neural network.

6. Conclusions

Thus, an automated method of PCBs optical inspection using the created artificial neural network is proposed.

The conducted testing of the neural network confirmed its operability and possibility of use for

PCBs inspection at the stage of production. The software was developed to perform the transformations over images, such as converting an image to a grayscale color space and image binarization, which speeds up the neural network by reducing the size of the input matrix to a binary value per each pixel of the image.

The accuracy of each defects types locating was also investigated, in particular, it was determined that some of the defects are determined more accurately by the neural network than others.

For example, the lack of copper in the conductor pattern has less detection accuracy, and individual areas of the copper layer that are not connected to any of the conductors, the program finds more accurately. The average value of image processing accuracy was 87.5 %.

In addition, it has been found that the image angle affects the accuracy of defects detection and at certain angles the accuracy may decrease by several percent, but not for all types of defects, for example, the accuracy of foreign inclusions does not depend on the angle of the image.

The developed system differs from the existing ones by greater versatility, both in the process of use in production and in the process of implementation due to the fact that although the system requires high-performance hardware, short processing time of each image allows to use hardware resources in shared mode during several production stages or on adjacent production lines.

One example of further improvement of the developed automated inspection system may be the expansion of the dataset with additional images or defects specific to a particular stage of PCBs production. It is also possible to adapt the neural network to equipment used in production, for example, to optical equipment such as scanners or cameras.

7. References

[1] V. Palagin, I. Botsman, I. Razumov-Fryziuk and V. Nevliudova, “Development of Multi-Probe Connecting Devices on Flexible Polyimide Base for MEMS Components Testing,” Proc. of XIV-th International Conference «Perspective technologies and methods in MEMS design» (MEMSTECH’2018), April 18-22, 2018, Polyana (Zakarpattya), Ukraine.

[2] I. Sh. Nevliudov, V. A. Palagin, E. A. Razumov-Frizjuk and I. V. Zharikova, “MEMS Intellect Multiprobes Contacting Devices for Electrical Checking-up of Multilayers Commutative Boards and BGA/CSP Electronic Components,” Proc. of IEEE East-West Design & Test Symposium (EWDTS’2012), September 14-17, 2012. Kharkov, Ukraine.

[3] A. A. R. M. A. Ebayyeh and A. Mousavi, "A Review and Analysis of Automatic Optical Inspection and Quality Monitoring Methods in Electronics Industry," in IEEE Access, vol. 8, pp. 183192-183271, 2020, doi: 10.1109/ACCESS.2020.3029127.

[4] S. S. Zakaria, A. Amir, N. Yaakob, and S. Nazemi, ‘‘Automated detection of printed circuit boards (PCB) defects by using machine learning in electronic manufacturing: Current approaches," IOP Conf. Ser., Mater. Sci. Eng., vol. 767, Mar. 2020, Art. no. 012064.

[5] S. N. D. Chua, S. Mohamaddan, S. J. Tanjong, A. Yassin, and S. F. Lim, ‘‘Detection of bond pad discolorations at outgoing wafer inspections,’’ IEEE Trans. Semicond. Manuf., vol. 31, no. 1, pp. 144–148, Feb. 2018.

[6] S. Fonseka and J. A. K. S. Jayasinghe, ‘‘Feature extraction and template matching algorithms classification for PCB fiducial verification,’’ J. Achievements Mater. Manuf. Eng., vol. 1, no. 86, pp. 14–32, Jan. 2018.

[7] L. Bai, X. Yang, and H. Gao, ‘‘A novel coarse-fine method for ball grid array component positioning and defect inspection,’’ IEEE Trans. Ind. Electron., vol. 65, no. 6, pp. 5023–5031, Jun. 2018.

[8] V. Bortnikova, I. Nevliudov, I. Botsman and O. Chala, “Search Query Classification Using Machine Learning for Information Retrieval Systems in Intelligent Manufacturing,” in CEUR Workshop Proceedings of the 15th International Conference on ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer (ICTERI’2019), June 12-15, 2019, Kherson, Ukraine.

[9] N. Guand, X. Wang, “Computational Design Methods and Technologies: Applications in CAD, CAM and CAE Education”. USA: IGI Global, 2012.

Ternary Wavelets And Their Application In Image Compression

Oleksander Shumeiko[1], Genady Shevchenko [2] and Vladislav Zyba[3]

[1, 3] *Dniprovsk State Technical University, Dniprobudivska st. 2, Kamianske, 51918, Ukraine*
[2] *Scientific Center Noosphere Company, Dnipro, 49000, Ukraine*

Abstract
The filtering of source information plays an important role in the digital data processing problems. The filtering results in dividing the processed information into the "averaged" part, which contains basic information, and "detailing" part. The results of filtering can be used for studying the features of signal as well as for further filtering. The second – high-frequency part – is the most critical for further analysis of signal features. On the other hand, the specific features of human visual perception are such that low-frequency characteristics are more informative, and the high-frequency ones have a large proportion of redundant data which can be neglected without prejudice to the human visual perception.

Keywords
Wavelets, ternary wavelets, image compression.

1. Introduction

In the 1980s and early 1990s, Fourier trigonometric transform [1, 2] was used for the filtration of data. In particular, widely used image compression format JPEG uses Fourier discrete transform for the filtration of each of three color components, by breaking the images into squares with the side of eight pixels. The result is 192 digital filters. With the emergence of wavelet-filters, Fourier trigonometric transform failed to produce the results commensurate with the quality and operating speed of wavelet filters. The 1990s and early 2000s [3] may be considered as the boom of wavelet studies. There were developed Daubechies orthogonal filters, biorthogonal, semi-orthogonal and other wavelets, and there also were obtained effective methods of different information compression based on wavelet filters, DJVU, JPEG-2000, H-264, which have set a new course in the development of the tools for digital information (in particular, multimedia) processing. The aim of this work is to demonstrate the efficiency of image filtering by ternary wavelets in comparison with binary analogs of the same complexity.

2. Haar wavelets

The applied wavelets separate the original signal into two sets of low and high frequencies. Further, using cascade circuit, every set is also separated into two levels, dividing "more averaged" and "more detailing part" [4]. In case of two-dimensional signal (image) processing, the filtering direction alternates – horizontally and vertically. Therefore, initially an image is broken into 4 frequency domains, relatively speaking – "low-low", "high-low", "low-high" and "high-high". At the next step – the same transform is applied to every domain, which results in 16 frequency domains; the next iteration leads to 64 domains, etc.

Let us give an example to illustrate classic (figure 1).

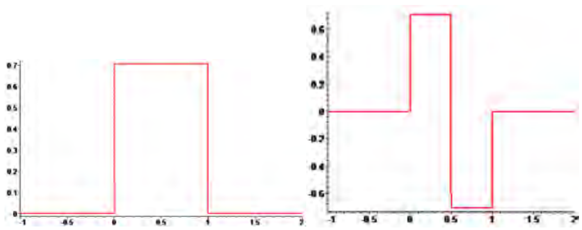


Figure 1: Haar wavelets

EMAIL: shumeiko_a@ukr.net (A. 1); nikk.gena@gmail.com (A. 2); aksel1455@gmail.com (A. 3)
ORCID: 0000-0002-8170-9606 (A. 1); 0000-0003-3984-9266 (A. 2)

$$\varphi(t)=\begin{cases}\frac{1}{\sqrt{2}}, & t\in[0,1]\\ 0, & \text{in other way.}\end{cases}$$

$$\psi(t)=\begin{cases}\frac{1}{\sqrt{2}}, & t\in[0,1/2]\\ -\frac{1}{\sqrt{2}} & t\in(1/2,1]\\ 0, & \text{in other way.}\end{cases}$$

In this case, if the value of original signal is equal to x_i , then the direct wavelet transform is written as

$$s_i=\frac{x_{2i}+x_{2i+1}}{\sqrt{2}}, d_i=\frac{x_{2i}-x_{2i+1}}{\sqrt{2}}.$$

It is easy to see that the inverse transform is, as follows

$$x_{2i}=\frac{s_i+d_i}{\sqrt{2}}, x_{2i+1}=\frac{s_i-d_i}{\sqrt{2}}.$$

Let us give an example of the use of Haar filters for filtration of Lena test image. After three iterations of cascade circuit, we have 64 frequency domains (figure 2).

Figure 2: Domains of Lena image transformation by Haar wavelets

It is clear that the information importance of each domain is different. In order to evaluate it, the norm of the respective frequency domain is used $D^k=\{d_i^k\}$, k=0,1,…,63:

$$Norm(D^k)=\frac{1}{mes(D^k)}\sqrt{\sum_i (d_i^k)^2}.$$

Putting aside the norm of the low-frequency domain, if we arrange the norms in descending order, we will have the following distribution (figure 3).

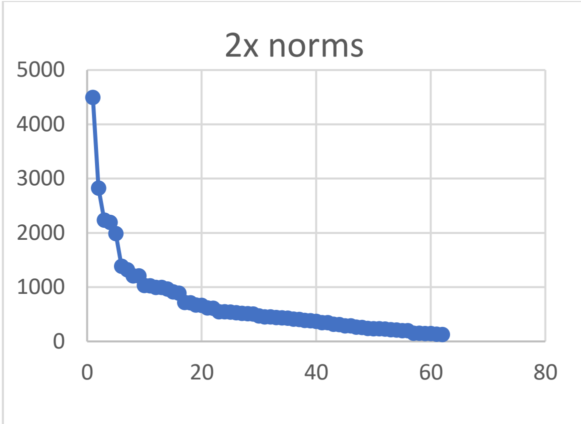


Figure 3: Domains of Lena image transformation by Haar wavelets

The lower is the domain norm, the less is its contribution to image generation; therefore, the domains with the lowest norm can be more roughly processed or even expelled from the process of image restoration. It is clear that image quality will worsen, but the size of the data used for its restoration will also be less.

Haar wavelets are the simplest and are used practically to illustrate or simulate one or another filtering process. In practice, more sophisticated and, respectively, more effective wavelets are used, which nevertheless, like Haar wavelets, separate the signal into low and high component at each step.

3. Symmetric ternary wavelets

However, there is another way to improve the filtering quality, which involves C. Shannon's observation, who has proved that, in the data transfer task, the calculation system in base 3 is more effective than the binary system. In the 1960s, this fact was used to design computing machines "Setun", based on ternary logic.

Concerning the effectiveness of the ternary computer system, it should be noted that, when adding trits in ternary adders, the number of additions is 1.58 times less than in binary adders and, consequently, the performance is 58% higher.

The failure of computing machinery built on ternary logic is primarily connected with element base.

The interest in ternary wavelets has increased recently, which is primarily due to the fact that binary orthogonal wavelets are asymmetric, in contrast to ternary orthogonal wavelets. Professor Petryshyn L.B. and his apprentices have published a series of papers, in particular [5-6], focused on studies in ternary orthogonal wavelets.

The papers [5-6] demonstrate that the use of symmetric ternary functions with length 3 ensures acceleration of the computing process by up to 50%, and reduction in memory capacity requires storing the transform results up to 57%.

In the general case, discrete wavelet transform of continuous function $f(x) \in L_2$ looks like

$$w_{n,m} = \langle f(\cdot) | \psi_{n,m}(\cdot) \rangle,$$

where $\psi_{n,m}(x)$- a system of wavelet functions, which, in the general case, is built based on mother wavelet $\psi(x)$

$$\psi_{n,m}(x) = a_0^{-m/2}\,\psi(a_0^{-m}x - nb_0)$$

where $a_0 \neq 1 -$ compression parameter, b_0-shear parameter, $n, m \in Z$.

In case of wavelet transform based on symmetric ternary functions, the parameter $a_0 = 3, b_0 = 1$ and, respectively (figure 4),

$$\psi_{n,m}(x) = 3^{-m/2}\,\psi(3^{-m}x - n)$$

Let us consider ternary wavelets (figure 4), which are the closest analogs of Haar wavelets.

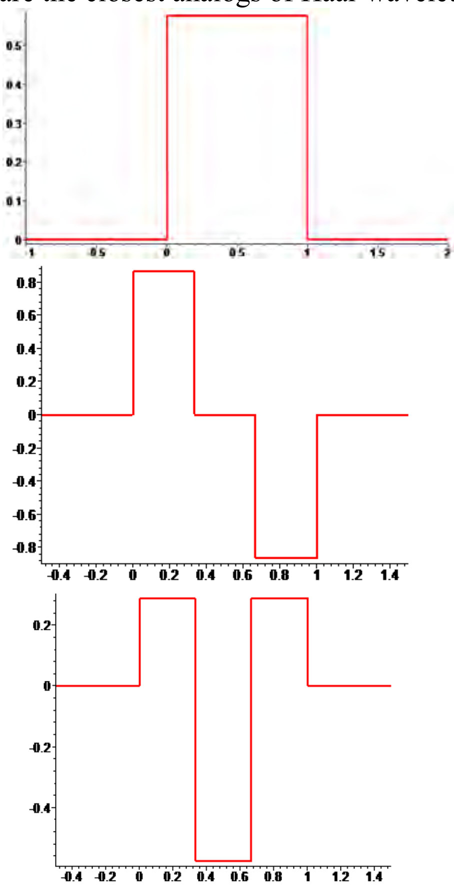


Figure 4: Ternary wavelets

$$\varphi(t) = \begin{cases} \frac{1}{\sqrt{3}}, & t \in [0,1], \\ 0, & \text{in other way.} \end{cases}$$

$$\psi_1 = \begin{cases} \frac{1}{\sqrt{2}}, & t \in \left[0, \frac{1}{3}\right], \\ 0, & t \in \left(\frac{1}{3}, \frac{2}{3}\right), \\ -\frac{1}{\sqrt{2}}, & t \in \left[\frac{2}{3}, 1\right], \\ 0, & \text{in other way.} \end{cases}$$

$$\psi_2 = \begin{cases} \frac{1}{\sqrt{6}}, & t \in \left[0, \frac{1}{3}\right], \\ -\frac{2}{\sqrt{6}}, & t \in \left(\frac{1}{3}, \frac{2}{3}\right), \\ \frac{1}{\sqrt{6}}, & t \in \left[\frac{2}{3}, 1\right], \\ 0, & \text{in other way.} \end{cases}$$

This transform matrix looks like

$$H = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

The orthonormality of the obtained system is evident. In this case, the inverse matrix will coincide with the transposed one $H^{-1} = H^T$

$$H^{-1} = H^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

In this case, the direct wavelet transform is written as

$$s_i = \frac{x_{2i} + x_{2i+1} + x_{2i+2}}{\sqrt{3}}, d_i^1 = \frac{x_{2i} - x_{2i+2}}{\sqrt{2}}, d_i^2 = \frac{x_{2i} - 2x_{2i+1} + x_{2i+2}}{\sqrt{6}}.$$

It is easy to see that the inverse transform is, as follows

$$x_{2i} = \frac{s_i}{\sqrt{3}} + \frac{d_i^1}{\sqrt{2}} + \frac{d_i^2}{\sqrt{6}}, x_{2i+1} = \frac{s_i}{\sqrt{3}} - \frac{2d_i^2}{\sqrt{6}}, x_{2i+2} = \frac{s_i}{\sqrt{3}} - \frac{d_i^1}{\sqrt{2}} + \frac{d_i^2}{\sqrt{6}}.$$

Let us give an example of the use of ternary filters for filtration of Lena test image. After two iterations of cascade circuit, we have 81 frequency domains (figure 5).

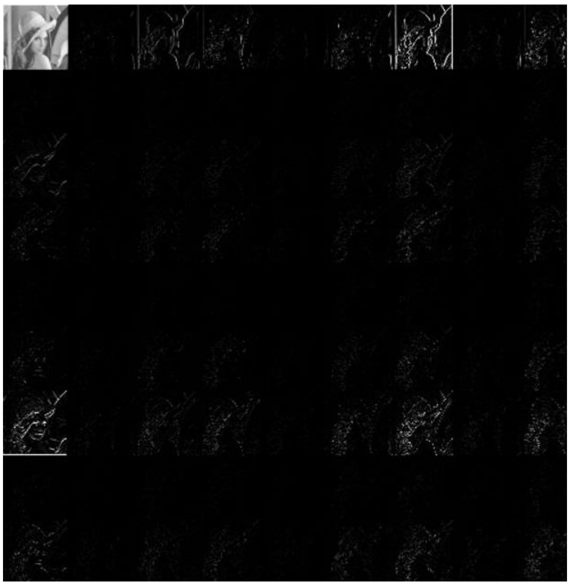


Figure 5: Domains of Lena image transformation by ternary wavelets

Decreasing permutation of the norms of frequency domains (excluding the low-frequency domain) looks like (figure 6)

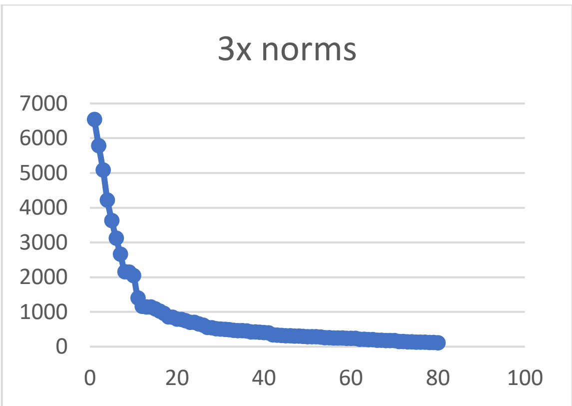


Figure 6: Decreasing permutation of domain norms for ternary wavelet

It can be inferred by the permutation that it is typical for ternary wavelets, in contrast to the binary ones, that the most of the information is grouped into a small number of frequency domains.

4. The results of numerical experiment

Let us provide the results of numerical experiment. Let us take 0.25 of the whole information from Haar wavelets after three iterations of cascade circuit. It makes 16 domains. Let us take 20 from 81 domains of ternary wavelets, which make 0.2469 of all frequency coefficients. Let us select the first domains of the decreasing permutation. Therefore, let us select a quarter (or almost a quarter) of the most significant frequency coefficients. Let us nullify all the remaining domains, i.e., their coefficients will not be taken into consideration when restoring the signal. PSNR will be used as a criterion for the restoration. The acronym PSNR (pealsignal-to-noiseratio) is used for peak signal to noise ratio and it is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

As a rule, PSNR value is in the range from 20 to 40 dB. In case of the coincidence of images, MSE value will be zero, as a result, PSNR takes on a value equal to infinity. In this case, it is generally assumed that PSNR = 100.

For monochrome images I and K of $H \times W$ size, one of which is contaminated with noise caused by the approach of the other one, PSNR is calculated, as follows

$$PSNR = 20\log_{10}\frac{\max\{I\}}{\sqrt{MSE}},\ MSE = \frac{1}{HW}\sum_{i=1}^{W}\sum_{j=1}^{H}\left|I_{i,j} - K_{i,j}\right|^2.$$

Let us compare the original image with the restored one. For the experiment, we used the same TID2008 collection [7] (figure 7), which was created by Eastman Kodak and is used for testing and quality assurance of different image processing methods, to which Lena (the last one in the above-mentioned collection) image was added, which, in fact, is a standard for such kind of tasks.

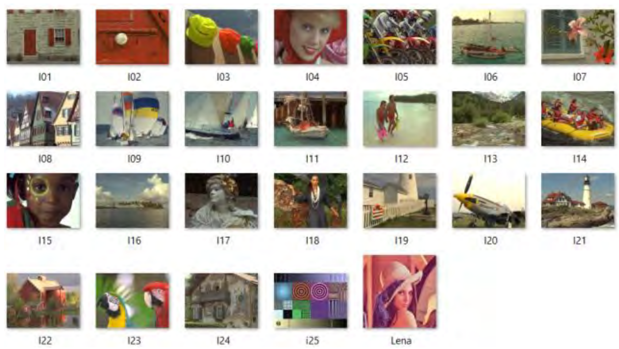


Figure 7: Test collection TID2008

Experimental results are presented in Table 1 and illustrated in figure 8.

Table 1
PSNR values filtering images of the test collection by Haar wavelets and ternary wavelets

	I01	I02	I03	I04	I05
3x	26	31.33	30.86	33.14	22.77
2x	25.17	30.5	30.02	32.21	21.77
	I06	**I07**	**I08**	**I09**	**I10**
3x	25.53	28.98	23.44	29.58	29.21
2x	24.55	27.97	22.51	28.41	28.33
	I11	**I12**	**I13**	**I14**	**I15**
3x	25.89	30.23	22.13	25.59	29.09
2x	25.06	29.28	21.2	24.55	28.7
	I16	**I17**	**I18**	**I19**	**I20**
3x	31.36	28.29	25.15	26.79	27.56
2x	30.3	27.47	24.32	25.72	26.63
	I21	**I22**	**I23**	**I24**	**I25**
3x	25.82	27.94	29.23	26.57	18.57
2x	24.84	27.15	28.16	25.52	17.9
	Lena				
3x	30.93				
2x	29.77				

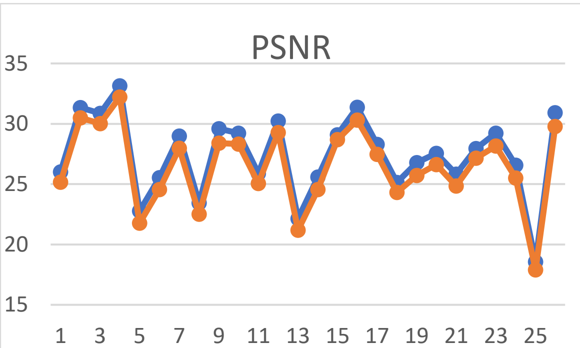


Figure 8: PSNR values filtering images of the test collection by Haar wavelets (brown curve) and ternary wavelets (blue curve)

5. Conclusions

The results of a numerical experiment based on TID2008 test images showed that when filtering with binary and ternary wavelets, followed by image restoration using the same amount of information, the quality of the reconstructed images for ternary wavelets is higher. Based on the results obtained, the following conclusions can be drawn: other things being equal, ternary wavelets, in contrast to binary ones, filter images more efficiently.

6. References

[1] Gonzales R. Digital image processing / R. Gonzales, R. Woods. – M.: Technosphere, 2005 .– 1070 p.

[2] Shumeiko A. Discrete trigonometric transform and its usage in digital image processing / A. Shumeiko, V. Smorodskyi // Econtechmod. An International Quarterly Journal .— 2017 .— No.4(6) .— P.21-26.

[3] Daubechies I. Ten lectures on wavelets, Izhevsk, Scientific and Publishing Center “Regular and Chaotic Dynamics”, 2001.– 464 p.

[4] Mallat S. A Wavelet Tour to Signal Processing / S. Mallat .— M.: Mir, 2005 .— 671 p.

[5] Izmailov A.V. Symmetric ternary functions and their application in digital information processing / A.V. Izmailov, L.B. Petryshyn // Information processing systems. – Kharkiv: HNUPS, 2016. – No.4(141). – P. 41-44.

[6] Izmailov A. Symmetric ternary functions and their application in orthogonal transforms / A. Izmailov, L. Petryshyn // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, 2017. – P. 836-841.

[7] TID2008: Tampere Image Database 2008 [Electronic resource] .— Access mode: https://computervisiononline.com/dataset/110513866

Digital Industry in Ukraine: challenges and perspectives

Tetiana Polozova [1], Iryna Sheiko[2], Irina Kolupaieva[3], Vasyl Otenko[4]

[1, 2, 3] *Kharkiv National University of Radio Electronics, Nauky Avenue, 14, Kharkiv, 61166, Ukraine*
[4] *Simon Kuznets Kharkiv National University of Economics, Nauky Avenue, 9a, Kharkiv, 61166, Ukraine*

Abstract
The article considers the issues of digitalization of Ukrainian industry. The experience of European countries in the formation of digitization policies at the national level is analyzed, as well as the first results of the implementation of such policies are evaluated. Prospects for the development of digital technologies in the industry of Ukraine are considered. It is noted that the main challenges to digital development are the high level of corruption in the country, external threats, political risks, low level of investment by foreign investors. The impact of the pandemic on the growth of the information technology sector in Ukraine is also assessed. Thus, the development of Ukraine's digital economy is a crucial factor for the success, competitiveness of Ukrainian business on the global market, as well as for attracting investment to the country. But also a necessary condition is the state support of digitalization processes without attempts of over-regulation.
Keywords
Digitalization, industry 4.0, productivity, IT sector, COVID-19 pandemic, cybersecurity.

1. Introduction

Digitalization becomes the frontier of innovations in industry and social services. All countries at global level provided actions to improve their digital development. The global economy is rapidly becoming digital. Information and Communications Technology (ICT) is no longer a specific sector but the foundation of all modern innovative economic systems. The Internet and digital technologies are transforming production process worldwide. Industry 4.0 or Manufacturing Internet of Things (IoT) systems connect the components of a production process in a factory. Their purpose is to enable “smart manufacturing”. In smart factories, cyber-physical systems monitor physical processes and make decentralized decisions. Via their IoT connection, these cyber-physical systems can communicate and cooperate with each other and with humans in real time. Connected devices include manufacturing equipment and robots.

That is why it is important to analyze the current perspectives of Ukrainian business sector in digitalization global processes.

In this paper we analyze the digitalization trends in EU members, including digital industry policies and summarize the main perspectives and challenges for Ukraine in global digital production. It is also important to conduct an impact of COVID-19 pandemic in intensity of digitalization worldwide and in Ukraine.

2. Current tendencies in Digital industry policies

Three industrial revolutions have led to changes in the domain of manufacturing – mechanization through water and steam power, mass production in assembly lines, and automation using information technology. However, over the past years, industries together with researchers and policy makers worldwide have increasingly advocated an upcoming fourth industrial revolution.

Scientists pay a lot of attention to the impact of information technology and digitalization on economic development.

C. Foster & S. Azmeh [8] investigated national digital policies, with a focus on China, they proved that these policies often aim at facilitating global integration and linkages.

Author's analysis shows that, under certain conditions, more interventionist approaches can be vital in countering structural challenges, such as power of digital platforms, limitations of domestic digital firms, limited ability to leverage digitalization for broad-based national development.

EMAIL: tetiana.polozova@nure.ua (A. 1); irina.sheiko@nure.ua (A. 2), iryna.kolupaieva@nure.ua (A. 3); ovi@hneu.edu.ua (A. 4)
ORCID: 0000-0001-9956-8816 (A. 1); 0000-0002-5770-3677 (A. 2); 0000-0003-1256-0817 (A. 3); 0000-0002-5979-1084 (A. 4)

Bart Van Ark et al. [1] confirmed that in the 90s of XX century, European economies were not able to utilize the potential of the digital economy so effectively as Canada or United States, mostly due to the restrictive regulations, that made the obstacles to a quick proliferation of the digital economy infrastructure in the sectors that were the users of ICT.

Biagi F. [2] provides literature review concerning ICT and productivity. The author concludes that ICT had a major role in the U.S. productivity acceleration observed in the period 1995-2005. Biagi also pointed out, that ICT is largely responsible for the divergence in productivity paths observed between 1995 and 2005 for the U.S. and the E.U.

T.Polozova et al [22] analyzed digital gaps between EU-members and found out, that Scandinavian countries has greater digital potential than Eastern and Central Europe countries.

World Economic Forum [20] and authors of Network Readiness Index 2020 [17] put attention to the importance of digital development for readiness of future production and post-COVID recovery.

Literature review shows the actuality of digitalization problem and importance of digital progress for economic development. But COVID-19 consequences on digitalization are less investigated due to short period of time for studying effects of pandemic on digital development. Also the results of digitalization in Ukrainian industrial sector did not receive sufficient coverage in economic press. That is why it is actual to analyze digital development in Ukraine business sector and attempts to activate digitalization by different state agencies.

2.1. Digital Industry in European Union

In recent years, the European has paid a lot of attention to the development of digital technology in Industry 4.0 under the slogan "Advanced Production". The formed working group, established in 2013, presented a working document [10], which mainly addressed the problem of reducing the share of industrial production in the GDP of the European Union. European Commission emphasized that digital technologies such as cloud computing, big data, the new industrial Internet, applications, smart factories, robotics and 3D printing had become prerequisites for increased productivity in the industry of the EU. The European Commission has identified three objectives to support the development of digital technologies [10]: faster commercialization of advanced production technologies; reducing the deficit of demand for advanced production technologies; promotion of advanced production skills.

However, the relative contribution of industry to the EU economy is declining. Over the last 40 years, the European economy has lost one third part of its industrial base. By the third quarter of 2014, the value added of production in the EU economy was only 15.3% of total value added, which is 1.2 percentage points less than at the beginning of 2008. This "deindustrialization" is common to all developed economies, in part due to increased production in other parts of the world (including China), the relocation of labor-intensive labor to lower labor costs, and participation in global supply chains with suppliers outside the EU. Moreover, the growing services sector accounts for an increasing share of the overall European economy, leading to a lower relative share of industry [10].

Large investments are needed for companies to move to Industry 4.0. It is projected, that by 2020 Germany alone will need € 40 billion a year. These investments may be unaffordable for small and medium-sized enterprises (SMEs). Therefore, in 2014 a new EU research program "Horizon 2020" was launched - by 2020 the Commission planned to allocate 77 billion Euros of funding, including 24.4 billion Euros for "research excellence" and 17 billion Euros for industrial innovation and so-called key technologies [10].

The concept of i4.0 is based on linkage of virtual and physical parts of business processes along a supply chain. The virtualization is gained by Internet of Things (IoT), Internet of Services (IoS) and Internet of people (IoP). [20].

The European manufacturing industry is responsible for 15% of GDP. Countries with an especially large manufacturing sector include Germany and Ireland, as well as various Eastern European countries. However, for a market to be promising for Industry 4.0 services it also needs to be ready for these techniques. This depends on, for example, production process maturity, degree of automation, degree of innovation, industry openness and internet use [10].

Industry 4.0 readiness is considerably greater in western and northern Europe than in other parts of Europe. As the founder of the movement, Germany scores particularly high.

National Industry 4.0 initiatives in European Union have yielded numerous qualitative and quantitative results at the country level.

For example Germany since 2010 has contributed € 200 million to the Industrie 4.0

initiative (one of ten projects under the German High Technology Strategy 2020 Action Plan) to encourage the development of "smart factories" [12].

The United Kingdom has initiated several policies to make production more sensitive, more sustainable, more open to new markets and more dependent on skilled workers. The most well-known are high-value production centers, called "catapult centers", which help companies to gain access to research and experience in specialized areas such as advanced production and innovation processes. The goal of these centers, which have received more than £200 million in public funding since 2011, is to double the contribution of production to GDP (about 10% in previous decades) [12].

In April 2015, France launched the Industry of the Future plan (Alliance pour l'Industrie du Futur) to set up demonstration centers to demonstrate new products and services. This program has implemented more than 800 loans to companies and drawn up 3,400 business plans, while the Swedish P2030 project has financed 30 projects involving more than 150 companies [12].

In Poland the Future Industry Platform was announced as part of the Responsible Development Plan ('Morawiecki Plan') by the Ministry of Finance and Development in 2016 [7]. Providing industrial financing over a 25-year period, the Morawiecki Plan pursues an agenda of reindustrialisation through new partnerships, export-oriented support measures and comprehensive regional development. With a total planned investment of €235 billion over the next 25 years, the Plan seeks to unleash the potential of the economy to achieve development that improves the quality of life in Poland [21].

Průmysl 4.0 (Industry 4.0) is a national initiative aiming to maintain and enhance the competitiveness of the Czech Republic in the wake of the Fourth Industrial Revolution. The national innovation fund (€1.87 million) brings together financial resources from the European Structural and Investment Funds, and additional €1.2 million from the private sector and it is considered as a possible source of a refundable form of financial support, of which part would be reinvested in the new projects [3].

European experience in the operation of programs to support the development of digitalization in industry can serve as a basis for modernizing the industrial base of the Ukrainian economy and to create an action plan to support the development of innovation and advanced technologies. Analysis of digitalization processes and their regulation in EU-members can be basis for comparing digitalization intensity in Ukraine and European countries.

To investigate impact on digitalization processes on economic development and labor productivity such data were used: cross-section data for European Union countries (EU-28), published by statistical office of the European Union [11]; Digital Economy and Society Index (DESI) [5], KOF Globalization Index [14], the Global Competitiveness ranking [6-7].

On the basis of EU-28 data for 2015-2020 the correlation between DESI index and labour productivity was found (fig. 1)

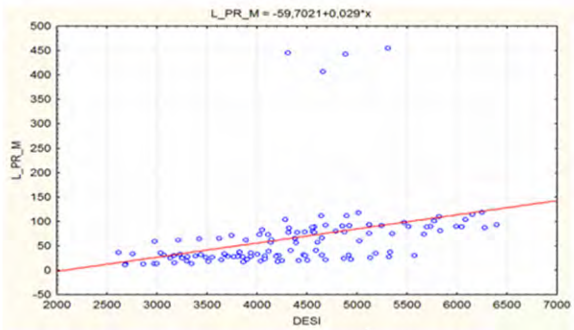


Figure 1: Relation between DESI Index and Labour Productivity level for EU-28 countries

Source: constructed by authors

Also on the base of EU-data with the data of 8 global countries (USA, China, Korea Rep., Japan, Australia, Canada, Norway and Brazil) the influence of digital development on improvement of country's competitiveness ranking was studied (Fig. 2)

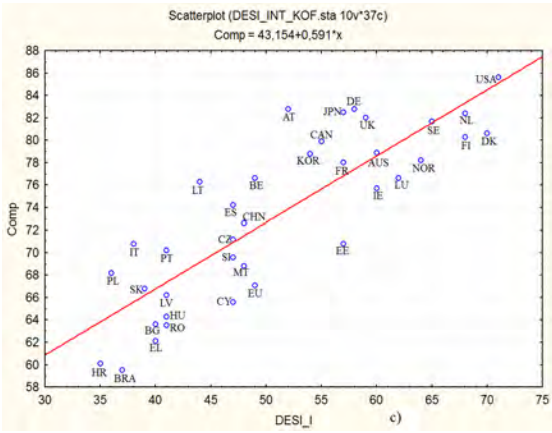


Figure 2: Correlation between DESI International Index and Global Competitiveness Index

Source: constructed by authors

Due to multiple regression modeling on the base of EU-members and 8 global competitors (USA, China, Korea Rep., Japan, Australia, Canada, Norway and Brazil) on the newest available data (KOF Globalization Index, revised in 2020, is based on 2018 data) such equation was received:

$$COMP = 28.34 + 0.213 \cdot KOF + 0.537 \cdot DESI\_I \quad (1)$$

where COMP - Global Competitiveness Index [6-7];

KOF – KOF Globalization Index [14];

DESI_I - DESI International Index [5]

Multiple regression modeling proved that together International DESI and KOF Index provide significant impact on the value of global competitiveness score, due to second multiple regression model. The multiple correlation coefficient regression models is above 0.84 that prove the significant quality of results. Adjusted coefficient of determination (R-square) is about 0,70: we include significant factors, that allows to explain the most part of variability of dependent variables, the rest 30% of variety is explained by residuals.

The further investigations need to find additional factors to increase the quality of regression results. Also additional studies are need to measure the impact of digital technologies on post-pandemic recovery to international business, international trade and transforming of global value chains into some "new normal" form.

Future studies, possibly with longer time series and another methodical approach, should better explore the interaction of digitalization with indicators of labor market (such as education, structural change, labor demand, human capital rating) and allows to explain differences between European countries towards digitalization policies.

2.2. Digital Challenges in Ukraine

In Ukraine, the regulatory framework for the development of the high-tech industry is represented by the following documents: draft order of the Cabinet of Ministers of Ukraine "On approval of the strategy for the development of high-tech industries until 2025 and approval of the Action Plan for its implementation" [19]; Action plan for the implementation of the strategy for the development of high-tech industries until 2020 [18]; orders of the Cabinet of Ministers of Ukraine "On approval of the Concept of development of the digital economy and society of Ukraine for 2018-2020 and approval of the action plan for its implementation" [19], "Digital Agenda of Ukraine 2020" [4], Strategy on integration of Ukraine into the European Union Digital Single Market (“Roadmap”), prepared by Ministry of Digital transformation of Ukraine [16].

The concept of the Strategy for the Development of the High-Tech Industry until 2020 also envisages the development of an export-oriented innovation ecosystem, implementation of a program to involve world high-tech leaders in production and R&D developing in Ukraine and promotion of new technologies among young people (High-Tech Nation) [7].

The “Digital Agenda 2020” project defines the main goals of digitalization in Ukraine as: stimulating the economy and attracting investments; laying the foundations for the transformation of sectors of the economy into competitive and efficient ("digitalization" of business); availability of digital technologies; creating new opportunities for the realization of human capital, development of innovative, creative and "digital" industries and businesses; development and world leadership in the export of "digital" products and services. The document also identifies the necessary steps for the digitalization of Ukraine in the fields of health, infrastructure, ecology, e-commerce, e-government, etc. [4].

According to experts’ opinion from the initiative "Digital Agenda of Ukraine" [3], to reach a GDP of 1 trillion USD in 2030, it will take 3-4 years to actively stimulate the penetration of technology and innovation into the economy of such sectors that could potentially show significant growth, namely: mechanical engineering; military-industrial complex; transport and logistics; agricultural sector; food and processing industry; woodworking; metallurgy.

Also Digital Agenda of Ukraine initiative experts pointed out, that digitalization can brings such effects for Ukrainian business [4]:

– growth of high-tech segments up to 20% per year;

– increase of production capacity - up to 60%;

– an increase in the number of orders executed on time - up to 95%;

– reduction of stocks - up to 20%;

– increase of efficiency of installed equipment - up to 15%;

– reduction of equipment downtime - up to 22%;

– savings in procurement costs - up to 30%;

As for machine-building industry, due to digitalization production lines will reconfigure themselves automatically in order to optimize productivity. Some of that will be driven from above, with production lines responding dynamically to new or amended production orders, tying in seamlessly with logistics and the wider business. Some will be driven from the product itself, communicating with the line to determine the optimal route through the production process. For example, if there is a bottleneck at some point the production line, the product will recognize this and look to see if there are other processes that might be accomplished first, and instruct the line to reroute its progress [4].

High developed traditional sectors can be a good basis for providing elements of digitalization. For Ukraine, machine building is one of the most priority and export oriented industries, where it already has world recognition. At the same time, our enterprises are critically in need of modernization and construction of modern digital models in management. Therefore, the sooner the machine-building enterprises will understand the importance of the development and implementation of: enterprise resource planning systems, supply chain management systems, production process control systems and other enterprise management systems, the sooner Ukraine will take a step towards a strong industry.

Digital Agenda argues why now and in the next 5 years in Ukraine there are no conditions for any positioning among the leaders - the advanced 20-30 post-industrial countries in the world. Instead, Ukraine can be at least a regional leader in the field of complex and science-intensive engineering services as [4]:

- programming in the field of industrial high-tech / creation of new software products, including new technologies 4.0;
- design (electrical, mechanical, electronic, technological, construction, etc.);
- industrial automation and complex engineering (including commissioning of industrial facilities);
- development and production of complex, small-batch or unique products.

Special attention should be paid to the creation of industrial clusters with the presence of high-tech areas, such as: robotics; bioengineering; 3D printing; artificial intelligence with a focus on the world market [4].

In analytical report [15], prepared on request of the Ministry of Digital Transformation of Ukraine, Composite Digitalization Index and Digital Services Trade Restrictiveness Index for Ukraine for 2016-2018 was calculated. The results of the calculation of the Digitalization Index show that the current level of digital development in Ukraine is far below the EU average and behind the new EU Member States.

The dynamics of the Digitalization Index in 2016-2018 indicates a widening in the gap between Ukraine and the EU over this period. In particular, the digitalization of the neighboring EU countries has grown faster than that of Ukraine. As a result, Ukraine has lost traction even compared to the EU countries with the lowest the Digitalization Index (Croatia, Bulgaria, and Romania) (Fig. 3)[15].

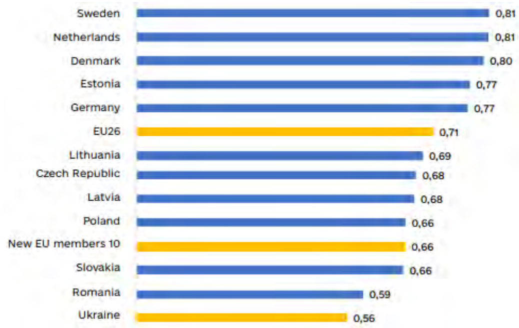


Figure 3: Comparison of the Digitalization Index of Ukraine and EU countries, 2018 [15]

Also authors of the report [15] had found out that the level of digital services trade restrictiveness in Ukraine is much higher than the EU average. For Ukraine, most barriers and regulatory differences in the digital sphere are related to cross border electronic payments and settlements, protection of intellectual property rights on the Internet, lack of practical mechanisms for the application of the electronic digital signature in foreign trade contracts, lack of mutual recognition of electronic identification and electronic trust services between Ukraine and major trading partners.

A case of successful digitalization of domestic machine-building enterprise is represented by state enterprise “FED” (Kharkiv) [13]. In 2011, the company launched a large-scale project on the implementation of an automated system for planning, accounting and analysis of Enterprise Resource Planning (ERP) business processes. And in 2013, FED launched a new project for the implementation of information technologies. Digital transformation of the business using the elements of Industry 4.0 helped the enterprise to increase the capacity of the equipment and improve delivery time. The first results of implementation were obtained during the year. Stability and reliability in terms management already then allowed the enterprise to guarantee the fulfillment of new orders from the world's leading aircraft manufacturers in the amount of more than hundreds of millions of dollars [13].

However, the development of the digital economy depends largely on the regulatory policy of the state and the creation of favorable conditions for all major stakeholders - innovators, investors, corporations. The field where the state can influence is quite wide - from legal protection to direct funding. The main tasks of the state on the way to the digital economy are: standards - significantly increased efficiency and productivity of all staff.

2.3. Impact of COVID-19 Pandemic on Digital Development in Ukraine

The IT sphere in Ukraine is one of the most dynamic and promising. It has long been called the locomotive of the Ukrainian economy. In 2018 alone, IT services brought almost UAH 10 billion in taxes to the Ukrainian budget.

However, the Covid-19 pandemic has made adjustments to the development of information technology around the world. Many clients of IT companies before the start of quarantine in anticipation of the crisis cut all their capital expenditures on IT solutions, leaving only operating. That is, they decided to abandon investment in "hardware", physical IT solutions, such as servers and data warehouses, in favor of services under the model of licenses [23].

But according to expert opinions [15,23], the IT sector in Ukraine has suffered less than other industries. In Ukraine, more than 1,600 companies provide IT services. A feature of this industry is also the presence of a large number of self-employed professionals. Outsourcing has suffered the most in the field of information technology, especially that related to tourism. Companies focused on team optimization, operating costs, and business processes. Most IT professionals worked at home. Exceptions were technical specialists who ensure the operation of the company's infrastructure.

One of the negative factors that significantly affected the industry was that the field of information technology in Ukraine focused on customers in Western countries, actively growing with them. Now that these countries are experiencing a decline in production, the Ukrainian IT industry may also face a drop in orders.

Another response to the Covid-19 pandemic experience in the medium term is to increase the use of machines, robots and other digital technologies in production processes. By replacing human labor, automation and robotics reduce dependence on it. This trend is already in full swing, as digital technologies significantly increase productivity and reduce costs. The use of such technologies to increase resilience to crises affecting production is an additional incentive. However, not all industries and companies have the same opportunities to use digital technology to reduce their vulnerability to crises.

The COVID-19 pandemic, among other things, has raised up cybersecurity issues, as crisis situations have traditionally provoked the intensification of various hacker groups. The main factors that potentially contributed to the increase in destructive cyberactivity were: an increase in the number of potentially vulnerable connections that could compromise information or the organization itself or its employees; intensification of electronic payments, which causes increased attention of cyber criminals to fraudulent activities; increase in the number of phishing attacks - increase in fake emails (with malware attachments) and fake sites (to collect personal and banking information of citizens); further escalation of panic may be one of the goals of influence operations by other states that may wish to take advantage of the situation [23].

The biggest potential danger can be a rapid increase in the number of employees who work remotely and use IT for this purpose - often such people do not have the appropriate skills, and the state of cybersecurity of home devices is not too high. This leads to a potential increase in the number of cyber incidents (including acts of cyber espionage and compromising information).

Some organizations have completely abandoned cybersecurity tools and policies in order to somehow continue their activities.

Although the future is uncertain and no one has a clear idea about what the "new normal" state looks like, there is a high probability that many IT companies will look for an acceptable format of work, combining remote form, online meetings, team meetings in the office at a certain time.

In a changing environment, those who best adapt to new conditions win. This is especially true of the leading sectors of the IT sector, which must meet the needs of business in the organization of online and remote work, and society in the organization of interpersonal communications, education and entertainment.

3. Conclusions

The development of Ukraine's digital economy is a crucial factor for the success, competitiveness of Ukrainian business on the global market, as well as for attracting investment to the country.

An important aspect: thanks to the introduction of the digital economy, small and medium-sized businesses have became global. Ukrainian companies can integrate into international value added networks, some of them are already successfully implemented. This is a way to increase exports and produce more complex and value added products. Such way Ukraine can significantly diversify country's export structure and geography.

Ukraine's integration into the EU's digital single market is a tool to facilitate Ukraine's digitalization in line with European and international standards. Approximation of Ukrainian legislation and standards to EU legislation and standards will reduce regulatory differences between Ukraine and the EU in the digital sphere and accelerate Ukraine's digital development.

As the role of digital technologies in international trade and economy is getting increasingly more significant, Ukraine is interested in integrating into the European digital space, because the development of Ukrainian trade and economy as a whole depends on it in all areas of economic activity.

Digital development in Ukraine should be based on high-developed Ukrainian IT sector. Ukraine has high chance to repeat the success of the domestic IT sector and become at least a regional leader in complex and high-tech engineering services as programming in the field of industrial high-tech, design (electrical, mechanical, electronic, technological, construction, etc.); industrial automation and integrated engineering (including the commissioning of complex industrial sites), development and production of complex, small-scale or unique products.

More than 100 companies on the Fortune 500 list are clients of Ukrainian IT firms Cisco, IBM, Atlassian, Travelport, OpenText, Fluke Corporation, Oracle, Boeing. Jooble, DepositPhotos, Grammarly, GitLab, PetCube, Mobalytics, Preply, Attendify are only a few of the big names that got started in Ukraine. Global market leaders have already shown their interest in Ukrainian startups which lead to a number of big acquisitions. Concentrating on such areas makes the most of the existing strengths of Ukraine and preserves the engineering schools at the corporate level and the state system of technical higher education. The impact of the COVID-19 pandemic on the global economy could provide an additional incentive for companies to make greater use of digital technologies. Digital solutions can compensate the employee's remote work, their communication and team working.

4. Acknowledgements

This paper was prepared in cooperation of scientists from Kharkiv National University of Radioelectronics and Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine

5. References

[1] Van Ark B, Inklaar R, McGuckin R.H., Timmer M.P. The Employment Effects of the `New Economy'. A Comparison of the European Union and the United States. National Institute Economic Review 184(1) (2003) 86-98. doi:10.1177/0027950103184001008

[2] F. Biagi ICT and Productivity: A Review of the Literature. JRC Working Papers on Digital Economy, September, (2013) Joint Research Centre (Seville site)

[3] Czech Republic: "Průmysl 4.0" Digital Transformation Monitor. May, 17 URL:https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_Prumysl%2040_CZ%20v1.pdf

[4] Digital Agenda of Ukraine Hi-Tech Office Ukraine, Ukrainian Institute of Future URL: https://strategy.uifuture.org/kraina-z-rozvinutoyu-cifrovoyu-ekonomikoyu.html#6-2-13

[5] Digital Economy and Society Index (DESI) 2020 – Retrieved from https://ec.europa.eu/digital-single-market/en/digital-economy-and-society-index-desi

[6] The Global Competitiveness Report Special Edition 2020: How Countries are Performing on the Road to Recovery URL: https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2020.pdf

[7] The Global Competitiveness Report 2019 URL:https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf

[8] C.Foster, S. Azmeh Latecomer Economies and National Digital Policy: An Industrial Policy Perspective. The Journal of Development Studies 56:7 (2020) 1247-1262, doi: 10.1080/00220388.2019.1677886

[9] European Commission A Digital Single Market Strategy for Europe – Analysis and

Evidence. Commission Staff Working Document SWD (2015) 100. European Union. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015SC0100&from=EN
[10] European Commission Advancing Manufacturing – Advancing Europe: Report of the Task Force on Advanced Manufacturing for Clean Production, Commission Staff Working Document, SWD (2014) 120, Brussels
[11] Eurostat database (2020). - Retrived from: https://ec.europa.eu/eurostat/web/main/data/database
[12] Industry 4.0 in Europe. CBI. Ministry of Foreign Affairs, 2017, URL:https://www.cbi.eu/market-information/outsourcing-itobpo/industry-40/
[13] IT-Enterprise, SmartFactory Production Management System on FED. Association of Industrial Automation of Ukraine, URL: https://industry4-0-ukraine.com.ua/smart-factory/
[14] KOF Index of Globalization. URL: http://globalization.kof.ethz.ch
[15] Ministry of Digital transformation of Ukraine. Ukraine’s integration into the EU’s digital single market: Analytic report. URL: https://thedigital.gov.ua/projects/yevrointegraciya
[16] Ministry of Digital transformation of Ukraine Strategy on integration of Ukraine into the European Union Digital Single Market (“Roadmap”) URL: https://thedigital.gov.ua/projects/yevrointegraciya
[17] Network Readiness Index 2020:Accelarating Digital transformation in post-COVID Global Economy URL: https://networkreadinessindex.org/wp-content/uploads/2020/11/NRI-2020-V8_28-11-2020.pdf
[18] On approval of the Concept of development of the digital economy and society of Ukraine for 2018-2020 and approval of the action plan for its implementation [Pro skhvalennia Kontseptsii rozvytku tsyfrovoi ekonomiky ta suspilstva Ukrainy na 2018-2020 roky ta zatverdzhennia planu zakhodiv shchodo yii realizatsii]: order of the Cabinet of Ministers of Ukraine URL: https://zakon.rada.gov.ua/laws/show/67-2018-%D1%80#Text
[19] On approval of the strategy for the development of high-tech industries until 2025 and approval of the Action Plan for its implementation [Pro skhvalennia Stratehii rozvytku vysokotekhnolohichnykh haluzei do 2025 roku ta zatverdzhennia planu zakhodiv shchodo yii realizatsii]: draft order of the Cabinet of Ministers of Ukraine URL: https://www.me.gov.ua/Documents/Detail?lang=uk-UA&id=c9b6f0b0-1ed5-4aba-a25e-f824405ccc64&
[20] The Readiness for the Future of Production Report 2018. Insight report. *World Economic Forum (In collaboration with A.T. Kearney)* URL: http://www3.weforum.org/docs/FOP_Readiness_Report_2018.pdf
[21] Poland: Initiative for Polish Industry 4.0 – The Future Industry Platform. *Digital Transformation Monitor.* February, 2018. URL: https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_Poland%20_vf.pdf
[22] T. Polozova, I. Kolupaieva, I. Sheiko Digital Gap in EU Countries and its Impact on Labour Productivity and Global Competitiveness, in J. Maci, P. Maresova, K. Firlej, & I. Soukal (Eds.), Hradec Economic Days, part 1, (2021) 569–570. University of Hradec Králové. https://doi.org/10.36689/uhk/hed/2021-01-065
[23] A. Verba What will happen to the Ukrainian IT market after the pandemic [Shcho bude z ukrainskym IT-rynkom pislia pandemii]. URL: https://nv.ua/biz/experts/rynok-it-karantin-i-pandemiya-koronavirusa-k-chemu-gotovitsya-tehnologicheskim-kompaniyam-novosti-ukrainy-50084249.html

Інтерактивна технологія виділення підмножин ефективних проєктних рішень

Безкоровайний Володимир [1], Семенець Валерій [2], Драз Оксана [3]

[1,2,3] *Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна*

Анотація

Для забезпечення взаємодії особи, що приймає рішення, з автоматичними проєктними процедурами запропоновано інтерактивну технологію виділення вузьких підмножин ефективних варіантів. Технологія об'єднує етапи кількісного й якісного впорядкування ефективних рішень з множини допустимих. Ефективність рішень з підмножини, що виділяється, забезпечується попереднім формування множини Парето-оптимальних варіантів. Запропонована технологія розширює методологічні засади автоматизації процесів підтримки багатокритеріальних проєктних рішень, дозволяє здійснювати коректне скорочення множини ефективних альтернатив для остаточного вибору з урахуванням факторів, що важко піддаються формалізації, знань і досвіду особи, що приймає рішення.

Ключові слова

Автоматизація проєктування, багатокритеріальна оптимізація, ефективне проєктне рішення, функція корисності.

Interactive technology for selecting subsets of effective design solutions

Vladimir Beskorovainyi [1], Valerii Semenets [2] and Oksana Draz [3]

[1,2,3] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract

To ensure the interaction of the decision maker with automatic design procedures, an interactive technology of allocating narrow subsets of effective options is proposed. The technology combines the stages of quantitative and qualitative ordering of effective solutions from a set of acceptable ones. The efficiency of solutions from the allocated subset is provided by the preliminary formation of the set of Pareto-optimal variants. The proposed technology expands the methodological principles of automation of processes to support multi-criteria design decisions, allows to correctly reduce the set of effective alternatives for the final choice, taking into account factors that are difficult to formalize, knowledge and experience of the decision maker.

Keywords

Design automation, multi criteria optimization, effective design solution, utility function.

1. Вступ

Для отримання ефективних і стійких рішень в процесах автоматизованого проєктування об'єктів необхідним є

EMAIL: vladimir.beskorovainyi@nure.ua (A. 1); valerii_semenets@ieee.org (A. 2); oksana.draz@nure.ua (A. 3)
ORCID: 0000-0001-7930-3984 (A. 1); 0000-0001-8969-2143 (A. 2); 0000-0002-9848-8360 (A. 3)

розв’язання цілого комплексу задач системної оптимізації [1]. Переважна більшість таких задач мають комбінаторний характер і розв’язуються за множиною показників якості [2-3]. Задачі багатокритеріального вибору розв’язуються навіть при виборі дизайну кінцевого продукту [4]. Суть проблеми прийняття проектних рішень подається логічним висловлюванням «необхідно знайти найкраще рішення s^o» або формально $< -, s^o >$ (де s^o – найкраще проектне рішення) [5]. При цьому ситуація прийняття проєктного рішення D зазвичай визначена не достатньо чітко. Для переходу до задачі прийняття рішення $< D, s^o >$ здійснюється декомпозиція проблеми на множину допоміжних задач виду: $<< D, - >, < D, s^o >>$ чи $<< -, s^o >, < D, s^o >>$.

У загальному випадку це потребує розв’язання множини задач [6]: визначення мети проєктування об’єкта; формування універсальної множини проєктних рішень S^U; виділення множини допустимих проєктних рішень $S \subseteq S^U$; виділення підмножини Парето-оптимальних та ефективних проєктних рішень $S^E \subset S^P \subset S \subset S^U$; ранжування множини ефективних проєктних рішень S^E; вибір найкращого проєктного рішення $s^o \in S^E$.

Задача визначення мети полягає у встановленні множини і важливості показників (локальних критеріїв) ефективності $k_j(s)$, $j = \overline{1,m}$, які адекватно характеризують проектні рішення [6-7]. Вони визначають співвідношення між функціональними і вартісними показниками проєктних рішень. Задача формування універсальної множини проєктних рішень S^U передбачає генерацію надпотужних множин всіх можливих альтернативних варіантів. Задача визначення множини допустимих рішень $S \subseteq S^U$ полягає у вилученні з універсальної множини підмножини рішень, які не задовольняють встановленим обмеженням $S = S^U \setminus \overline{S}\ :\ k_j(s) \gtrless k_j^*,\ \forall j = \overline{1,m}$ (де k_j^* – граничне значення для j-го показника). При цьому переважна більшість з допустимих рішень, що генеруються з використанням автоматичних комбінаторних процедур, є неефективними [8]. Це потребує попереднього визначення підмножини Парето-оптимальних рішень $S^P \subset S \subset S^U$. Ранжування проєктних рішень і вибір найкращого серед них $s^o \in S^E$ здійснюється з використанням теорії корисності, що передбачає їх якісне або кількісне оцінювання [6]. У процесі якісного оцінювання впорядкування невеликої множини ефективних рішень $s \in S^E$ здійснюється особою, що приймає рішення (ОПР). Для кількісного оцінювання визначають узагальнений критерій ефективності $P(s)$, з використанням якого здійснюється скалярна оцінка і вибір найкращого варіанту: $s^o = arg \max_{s \in S^E} P(s)$.

У переважній більшості випадків підмножина Парето-оптимальних рішень $S^P \subset S \subset S^U$ складається зі значної кількості елементів, а остаточний вибір здійснюється ОПР, яка здатна аналізувати незначну кількість варіантів. ОПР здійснює оцінку варіантів на основі аналізу множини суперечливих локальних критеріїв з використанням методів індивідуального чи колективного експертного оцінювання [9-11]. Вище сказане породжує проблему узгодження взаємодії між процедурами генерації й оцінювання варіантів у технологіях проєктування. Виходом може слугувати звуження множини Парето-оптимальних рішень до підмножини ефективних $S^E \subset S^P$.

Однією з важливих задач цієї проблеми є формування та коректне скорочення множини ефективних альтернатив для остаточного вибору з урахуванням факторів, що важко піддаються формалізації, а також знань і досвіду ОПР.

2. Формування множини Парето-оптимальних рішень

У класичних задачах прийняття рішень множина допустимих рішень вважається заданою [5-6]. Виділення на ній підмножини Парето-оптимальних рішень $S^P \subset S$ полягає у вилученні з множини допустимих неефективних рішень $\overline{S}^P$. У такий спосіб множина допустимих рішень буде розбита на дві непересічних підмножини:

$$S = S^P \bigcup \overline{S}^P, \;\; S^P \bigcap \overline{S}^P = \varnothing . \qquad (1)$$

Варіант проєктного рішення $s^P \in S^P$ називається Парето-оптимальним, якщо на множині допустимих рішень не існує рішення $s \in S$, для якого були б справедливими нерівності:

$$k_i(s) \geq k_i(s^P), \text{ якщо } k_i(s) \to max, \qquad (2)$$

$$k_i(s) \leq k_i(s^P), \text{ якщо } k_i(s) \to min \qquad (3)$$

та хоча б одна з них була строгою.

Зі зростанням потужності множини допустимих рішень $|S|$ та кількості локальних критеріїв $j = \overline{1,m}$ потужність підмножини Парето-оптимальних рішень $|S^P|$ має тенденцію до зростання (рис. 1), а відносна потужність підмножини Парето-оптимальних рішень $|S^P| / |S|$ – до убування (рис. 2).

Для розв’язання задачі виділення підмножини Парето-оптимальних рішень використовують методи парних порівнянь, Карліна, Гермейєра, еволюційного пошуку [8, 12-14].

Метод парних порівнянь дозволяє здійснювати точне розбиття (1) як на опуклих, так і на неопуклих множинах допустимих альтернатив. Його суть полягає в такому [8]. Перший з альтернативних варіантів $s \in S$ включається до множини ефективних. Кожен з наступних варіантів $v \in S$ порівнюється за множиною критеріїв $k_j(s)$, $j = \overline{1,m}$ з кожним з варіантів $s \in S^E$. На першому циклі методу $|S^P| = 1$. Якщо варіант $v \in S$ кращий за будь-який варіант з S^E хоча б за одним з критеріїв $k_j(s)$, $j = \overline{1,m}$, він включається в підмножину S^E. Якщо деякий варіант $s \in S^P$ є гіршим, ніж варіант $v \in S$, він виключається з підмножини S^P, а варіант $v \in S$ включається до неї. Після перегляду всіх альтернативних варіантів $s \in S$ буде виділена повна підмножина Парето-оптимальних варіантів. [2] Цей метод має відносно високу часову складність.

Методи Карліна, Гермейєра й еволюційного пошуку мають регульовану часову складність, проте у загальному випадку не дозволяють здійснювати точне розбиття (1).

Підмножина Парето-оптимальних рішень на опуклій множині S за методом Карліна знаходиться як об'єднання рішень s_j^o, $j = \overline{1,m}$, які оптимізують кожен з локальних критеріїв шляхом розв’язання множини задач [12]:

$$s_j^o = arg\,\max_{s \in S} \{ P(s) = \sum_{j=1}^{m} \lambda_j \xi_j(s) \}, \qquad (4)$$

$$\lambda_j \in \Lambda = \{ \lambda_j : \lambda_j > 0 \;\; \forall j = \overline{1,m}, \;\; \sum_{j=1}^{m} \lambda_j = 1 \}, \qquad (5)$$

де $\xi_j(s)$ – значення функції корисності або нормоване значення j-го локального критерію; λ_i – ваговий коефіцієнт j-го локального критерію.

Підмножина Парето-оптимальних рішень на опуклій чи неопуклій множині S за методом Гермейєра знаходиться як об'єднання рішень s_j^o, $j = \overline{1,m}$, які оптимізують кожен з локальних критеріїв шляхом розв’язання множини задач [12]:

$$s_j^o = arg\,\max_{s \in S} \{ P(s) = \min_{j} \lambda_j \xi_j(s) \}, \qquad (6)$$

$$\lambda_j \in \Lambda = \{ \lambda_j : \lambda_j > 0 \;\; \forall j = \overline{1,m}, \;\; \sum_{j=1}^{m} \lambda_j = 1 \}. \qquad (7)$$

Часова складність методів парних порівнянь, Карліна та Гермейєра може бути знижена при попередньому виділенні підмножини субоптимальних за Парето рішень S', $S^P \subseteq S' \subseteq S$ за методами «сектора» або «сегмента» [8].

Найбільш популярним серед еволюційних вважається метод на основі генетичного алгоритму з недомінантним сортуванням [15]. З його допомогою здійснюється визначення фронту Парето на допустимих множинах S надвеликих розмірів. Він надає збіжність до фронту і хороший розподіл рішень по всьому фронту.

Часова й ємнісна складності формування підмножини Парето-оптимальних рішень можуть бути суттєво знижені, якщо застосувати метод парних порівнянь уже на етапі генерації допустимих варіантів.

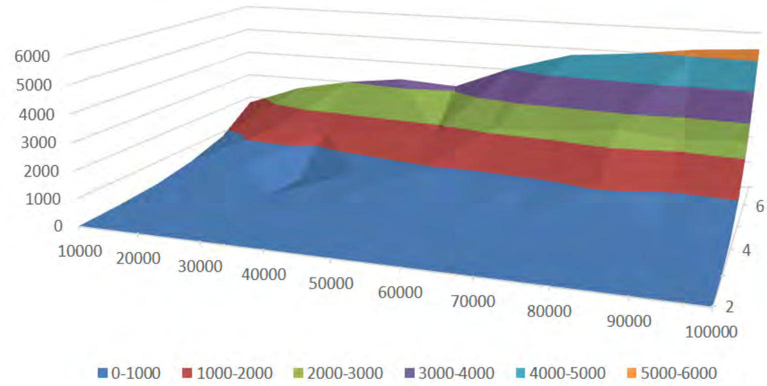


Рисунок 1: Залежність абсолютної потужності S^P від потужності S та кількості локальних критеріїв m

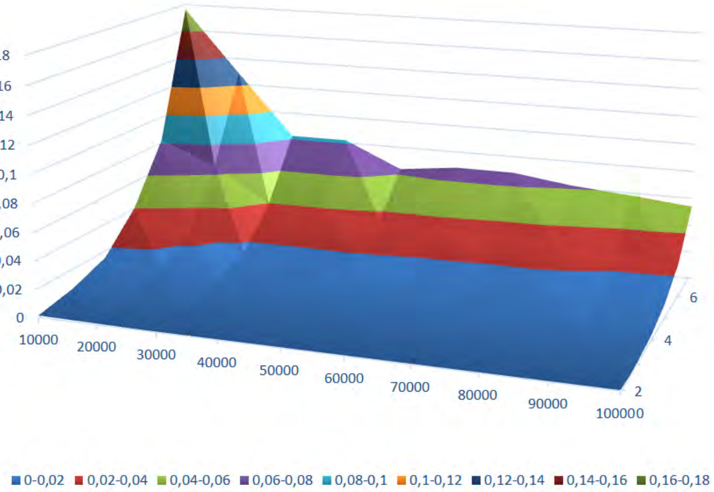


Рисунок 2: Залежність відносної потужності S^P від потужності S та кількості локальних критеріїв m

3. Звуження підмножини ефективних проєктних рішень

В задачах проєктування до підмножини Парето-оптимальних рішень може входити досить велика кількість варіантів (рис. 1). Для остаточного ж вибору ОПР необхідно надавати лише найкращі з них. Здійснимо оцінку кожного з варіантів $s \in S^P$ з використанням скалярної функції загальної корисності $P(s)$, яка й визначатиме їхнє впорядкування за якістю [6]:

$$\forall s,v \in S^P : s \sim v \leftrightarrow P(s) = P(v)\,;$$

$$s \succ v \leftrightarrow P(s) > P(v)\,;\; s \succsim v \leftrightarrow P(s) \geq P(v)\,. \quad (8)$$

Для оцінювання можна скористатися класичною адитивною згорткою зважених значень функцій корисності локальних критеріїв [6]:

$$P(s) = \sum_{j=1}^{m} \lambda_j \xi_j(s)\,, \quad (9)$$

$$\xi_j(s) = \left[\frac{k_i(s) - k_i^-(s)}{k_i^+(s) - k_i^-(s)} \right]^{\alpha_i}, \; j = \overline{1,m}\,, \quad (10)$$

де λ_i – коефіцієнт важливості i-го критерію; $\xi_i(s)$ – значення функції корисності i-го критерію; $k_i(s),\, k_i^+,\, k_i^-$ – значення i-го локального критерію для рішення s, його найкраще та найгірше значення; α_i – параметр, який визначає вид залежності (10): $\alpha_i = 1$ – лінійна; $0 < \alpha_i < 1$ – випукла догори; $\alpha_i > 1$ – опукла донизу.

Для параметричного синтезу функції (9) можна скористатись експертними методами або методом компараторної ідентифікації [3, 6, 7, 9]. Значення функції (9) для варіантів з множини Парето S^P дозволяють здійснити їх ранжування та виділити підмножину найкращих серед них $S^E \subseteq S^P$. Після цього ОПР здійснює остаточний вибір найкращого варіанта s^o серед незначної кількості найкращих з підмножини S^E.

4. Висновки

Встановлено, що більшість задач проєктування є багатокритеріальними та мають комбінаторний характер. Процеси прийняття остаточних проєктних рішень здійснюється ОПР з використанням методів експертного оцінювання незначної кількості варіантів. Для забезпечення взаємодії ОПР з автоматичними процедурами проєктування запропоновано інтерактивну технологію виділення вузьких підмножин ефективних рішень, яка об'єднує етапи їх кількісного й якісного впорядкування. Ефективність рішень з підмножини, що виділяється, забезпечується попереднім формування множини Парето-оптимальних рішень. Запропонована технологія розширює методологічні засади автоматизації процесів підтримки багатокритеріальних проєктних рішень, дозволяє здійснювати коректне скорочення множини ефективних альтернатив для остаточного вибору з урахуванням факторів, що важко піддаються формалізації, знань і досвіду ОПР. Напрямком подальших досліджень може розробка моделей задач прийняття рішень для умов невизначеності цілей і даних, які присутні в технологіях проєктування.

5. Література References

[1] A. Kossiakoff, W. N. Sweet, S. J. Seymour, S. M. Biemer, Systems Engineering Principles and Practice. Hoboken, New Jersey: A John Wiley & Sons (2011).

[2] I. Kaliszewski, T. Kiczkowiak, Miroforidis J. Mechanical design, Multiple Criteria Decision Making and Pareto optimality gap, Engineering Computations 33(3) (2016).

[3] S. Greco, M. Ehrgott, J.R. Figueira Multiple Criteria Decision Analysis – State of the Art Surveys. New York: USA, Springer, (2016).

[4] Leyva López, J.C., J.J. Solano Noriega, O. Ahumada Valenzuela et al., A preference choice model for the new product design problem, Operation Research (2021).

[5] Э.Й. Вилкас, Е.З. Майминас, Решение: теория, информация, моделирование, М.: Радио и связь, (1981).

[6] Э.Г. Петров, Н.А. Брынза, Л.В. Колесник, О.А. Писклакова, Методы и модели принятия решений в условиях многокритериальности и неопределенности, Херсон: Гринь Д.С. (2014).

[7] V. Beskorovainyi, G. Berezovskyi, Estimating the properties of technological systems based on fuzzy sets, Innovative Technologies and Scientific Solutions for Industries 1 (1) (2017).

[8] Vladimir V. Beskorovainyi, Lubomyr B. Petryshyn, Olha Yu. Shevchenko, Specific subset effective option in technology design decisions, Applied Aspects of Information Technology 3 (1) (2020).

[9] К. Э. Петров, А. А. Дейнеко, О. В. Чалая, И. Ю. Панферова, Метод ранжирования вариантов при проведении процедуры коллективного экспертного оценивания, Радіоелектроніка, інформатика, управління 2 (2020).

[10] M. Bernasconi, C. Choirat, R. Seri, Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence, European Journal of Operational Research 232 (2014).

[11] О.А. Подоляка, А.Н. Подоляка, Применение порядковой нормализации и скремблирования критериев для решения многокритериальных задач, Автомобіль і електроніка. Сучасні технології, 8 (2015).

[12] В.С. Михалевич, В.Л. Волкович, Вычислительные методы исследования и проектирования сложных систем, М.: Наука (1982).

[13] Deb Kalyanmoy, Deb Debayan, Analysing mutation schemes for real-parameter genetic algorithms. International Journal of Artificial Intelligence and Soft Computing 4(1) (2014).

[14] О. Шадура, Модифікація генетичних алгоритмів на основі методу нецетрованих головних компонент та стандартні тести, World Science № 4(44) (2019).

[15] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE transactions on evolutionary computation, 6(2) (2002).

Математична модель оптимального розподілу внутрішніх інвестицій банка на основі метода послідовного аналізу варіантів

Давид Дабагян[1], Михайло Годлевський[2]

[1, 2] *НТУ «ХПІ»,вул.Кирпичова, 2, Харків, 61002, Україна*

Анотація

Внаслідок низки факторів, таких як політика регулятора, пандемія, на банківському ринку України є високий рівень конкуренції. Банки зацікавлені в залученні нових клієнтів та утриманні існуючих, та готові проводити внутрішнє інвестування. Стаття пропонує математичну модель оцінки банка клієнтами, та модель оптимального розподілу внутрішніх інвестицій, збудовану на основі модернізації методу послідовного аналіза варіантів (ПАВ).

Ключові слова

Послідовний аналіз варіантів, оптимальний розподіл інвестицій, метод аналіза ієрархій, метод парних порівнянь

Mathematical model of optimal distribution of internal investments of the bank based on the method of sequential analysis of options

Davyd Dabahian [1], Mykhailo Godlevskyi [2]

[1, 2] *NTU "KhPI", Kyrpychova str,2 , Kharkiv, 61002, Ukraine*

Abstract

Due to a number of factors, such as the regulator's policy, the pandemic, there is a high level of competition in the banking market of Ukraine. Banks are interested in attracting new customers and retaining existing ones, and are ready to make domestic investments. The article proposes a model for assessing the bank by clients, and proposes a method for the optimal distribution of internal investments. The method is based on a modification of the method of sequential analysis of options

Keywords

Sequential analysis of options, optimal distribution of investments, analytic hierarchy process, paired comparisons method

1. Вступ

З урахуванням ситуації, що склалася в банківському секторі України, а саме - високим рівнем конкуренції, специфічної політикою регулятора, пандемією коронавируса і т.д. - банки максимально зацікавлені в залученні нових клієнтів, і утриманні існуючих. Для цього, банки готові інвестувати в оптимізацію власної роботи (це і

EMAIL: d.dabagyan@gmail.com (A. 1);
mgodlevskyi@gmail.com (A.2);
ORCID: 0000-0003-2872-0598 (A.2);

є внутрішні інвестиції). Раніше, в роботі 1 були сформовані основні критерії оцінки комерційного банку клієнтом. Таким чином, перед банком стоїть завдання розподілу кінцевої суми інвестицій за цими критеріями з максимальним ефектом - тобто максимальне збільшення власної привабливості з точки зору клієнтів. Це завдання може вирішуватися в рамках побудови довгострокової стратегії банку, як частина стратегічного планування.

Мета даної роботи - запропонувати математичну модель, яка дозволить оптимальним чином розподілити конкретну суму внутрішніх інвестицій Банку за критеріями оцінки. В основі математичної моделі, з урахуванням специфіці предметної області, лежить нечітка модифікація методу послідовного аналізу варіантів, описана в [2].

2. Модель оптимального розподілу інвестицій

2.1. Критерії оцінювання Банка

Першим кроком для побудови ієрархічної моделі оцінки банка, стало визначення загальної ієрархії оцінки банка. Така ієрархія враховує ті факти, що на українському ринку представлені також банки з іноземним капіталом, які діють на міжнаціональному рівні. На ієрархії (на верхньому рівні) наведено міжнаціональний рівень – це рівень міжнародних банків, які діють на території декількох країн. Наступний – рівень окремих країн. Далі – регіональний рівень, тобто, адміністративні одиниці країн. Під ним йде рівень кластерів клієнтів – вікових, соціальних, та інших груп населення, які є діючими або потенційними клієнтами банка. Далі йдуть власне критерії – рівень груп критеріїв, та конретних критеріїв. І на найнижчому рівні йдуть альтернативи – банки. Загальна ієрархія оцінки банків наведена на рисунку 1.

Рисунок 1: Загальна ієрархія оцінки банка

Слід більш детально розглянути рівні груп критеріїв, та критеріїв. Їх було взято з роботи [1]. Сама ієрархічна модель оцінки банку клієнтом з цієї роботи була змінена, і доопрацьована - були додані нові критерії (Рисунок 2).

Критерії розбиті на 4 групи.

Слід окремо зазначити, що використання нечіткої логіки та методу аналізу ієрархій (його модифікацій) для вирішення задач прийняття рішень суттєво освітлене в багатьох існуючих працях. Наприклад, робота [2] пропонує (та наводить приклад) використання нечіткого МАІ для вирішення задачі визначення ступеню надійності нафтогазового обладнання. В рамках цієї роботи, нечітка модифікація МАІ буде використана при побудові моделі оцінки банка.

Рисунок 2: Ієрархія критеріїв оцінювання

2.2. Опис моделі оцінки банка

Для визначення конкретних рекомендацій щодо оптимізації роботи банка за рахунок підвищення його привабливості, планується виконати наступну послідовність дій (Рисунок 3).

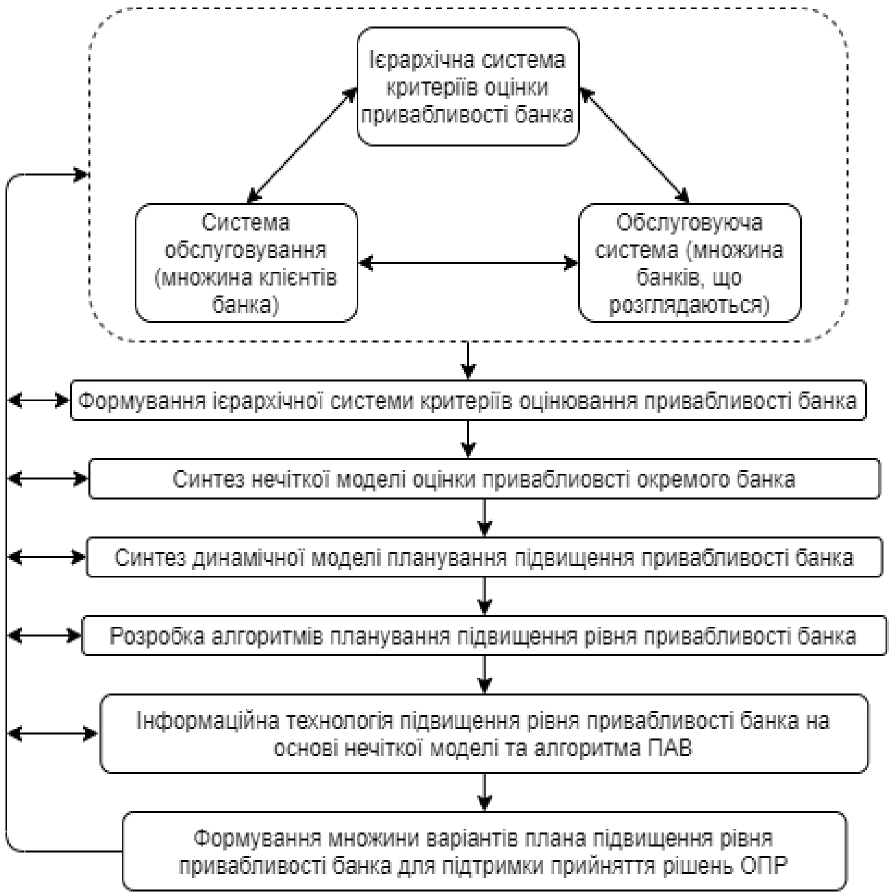


Рисунок 3: Послідовність дій щодо підвищення рівня привабливості банка

2.3. Нечітка модель оцінки банка

Для побудови нечіткої моделі оцінки окремого банка, потрібно виконати наступні дії:

- Визначення території, на якій представлено банк;
- Визначення основних кластерів клієнтів на цій території;
- Побудова ієрархії критеріїв на основі існуючих ієрархій;
- Застосування нечіткої модифікації МАІ [2] для побудови моделі оцінки.

На цьому етапі, частину критеріїв буде відсіяно як несуттєві – це буде зроблено експертним шляхом, тобто, співробітники відділів аналізу, методології Банка винесуть рішення щодо суттєвості критеріїв.

Результатом виконання цих кроків, буде модель оцінки банка, яка на вхід приймає альтернативи та дані матриць парних порівнянь альтернатив, застосовує до них процедуру фазифікації, та на виході дає упорядковану за ступенем зменшення привабливості множину альтернатив. За основу взято підхід, наведений у роботі [2].

2.4. Динамічна модель планування підвищення привабливості банка

Для такої моделі, першим етапом буде визначення планового періоду, на якому планується розподіляти інвестування. З урахуванням того факту, що конкуренти теж будуть працювати над оптимізацією свого функціонування, слід враховувати зміни їхньої оцінки. Для цього, слід побудувати прогноз їхнього стану на основних інтервалах планового періоду. Такий прогноз можливо здійснити на основі аналізу звітності банків, показників їхньої діяльності [3], рекламних кампаній банків, та комерційної розвідки.

Результатом цього етапу мають стати оцінки банків-конкурентів на основних інтервалах планового періоду.

Розглянемо динамічну модель.

Планування проводиться на деякому плановому періоді, який розбито на підперіоди $t \in [0, T]$, де T – кількість підперіодів, t_0 - початковий стан процесу. Будемо вважати, що на t-му підперіоді, стан Банка характеризується деяким значенням Z_t. Перехід у наступний стан здійснюється за рахунок керуючого впливу x_t, та може бути записаний як $Z_t = F_t(Z_{t-1}, x_t), t \in [1, T]$. Внаслідок того, що процес здійснюється в рамках обмежених ресурсів, вважатимемо, що $x_t \in X_t$, де X_t – множина можливих варіантів інвестування на t-му підперіоді планування. Оцінка привабливості Банка здійснюється описаним вище способом (за допомогою побудованої ієрархії оцінювання).

На основі таких даних, можна формувати множину варіантів розподілу інвестицій на плановому періоді.

В якості основного математичного апарату, пропонується одна з модифікації методу послідовного аналізу варіантів – нечітка модифікація, яку було описано в статті [4]. Пропонується її застосування до розв'язуваної задачі розподілу інвестицій.

Метод послідовного аналізу варіантів широко застосовується у задачах підтримки прийняття рішень. Наприклад, робота [5] наводить приклад використання метода для вирішення задачі розподілу віртуальних машин по серверам центру обробки даних – тобто, вирішує задачу прийняття рішень з протирічними критеріями, і пропонує особі, що приймає рішення, перелік варіантів, які є оптимальними.

В даній роботі використовується модифікація метода послідовного аналізу варіантів.

Вихідними даними для завдання розподілу інвестицій, є:

- Множина критеріїв оцінювання банка $q_i, i = \overline{1..m}$;
- Початкова сума інвестицій P.

Пропонується наступний алгоритм вирішення задачі.

Першим кроком необхідно сформувати множину варіантів. Для цього, розіб'ємо вихідну суму інвестицій P на n шматочків, тобто візьмемо якийсь крок. Тоді у нас виходить $k = P/n$.

Результатом буде множина варіантів, загальною кількістю

$$C_k^m = \frac{k!}{m! * (k-m)!}.$$

Тобто, всі можливі комбінації сум інвестувань в критерії.

Подальші дії - відповідно до описаної в статті [4] процедурі. Тобто, задається якийсь допуск δ, з яким порівнюється ефект від інвестування в кожен варіант. Далі проводиться розрахунок матриць парних порівнянь на кожному інтервалі планового періоду, з урахуванням варіанта інвестування, що розглядається. Тобто, наприклад, перший варіант – направлення усієї суми інвестувань у критерій «Дистанційні сервіси». Результатом буде суттєве зміцнення позиції по дистанційному обслуговуванню, але втрата позицій по іншим критеріям (наприклад, кредитування). І якщо результат інвестування (позиція банка в упорядкованій множині альтернатив, або його відносна оцінка у процентах) менша за величину допуску, то варіант відсіюється.

Процедура повторюється для всіх інтервалів планового періоду.

Результатом проходу всіх кроків буде оптимальна множина варіантів, яку буде аналізувати особа, яка приймає рішення.

У разі, якщо результуюча множина буде містити занадто багато значень, можна повторити процедуру з меншим допуском.

3. Висновки

У роботі запропоновано модель вирішення задачі розподілу внутрішніх інвестицій банку, на основі модифікації методу послідовного аналізу варіантів. Запропонована модель піддається автоматизації, і не вимагає багаторазового ручного введення інформації, і відповідно - застосовна на практиці. Подальшою роботою буде проектування і розробка інформаційної системи, яка реалізує описану вище модель.

4. Література References

[1] Dabahian D., Godlevskyi M. “Bank attractiveness evaluation method based on soft computing in the analytic hierarchy process”. ICT in Education, Research, and Industrial Applications (ICTERI 2018) : proc. of the PhD Symposium at 14th Intern. Conf., Kyiv, Ukraine, 16.05 (2018): 8-15.

[2] Mukhametzianov I.Z. — Fuzzy logical inference and fuzzy analytic hierarchy process in decision support systems: an application to assessing the reliability of technical systems // Cybernetics and programming – №2 (2017): 59-77.

[3] Bank.Gov.ua, Bank financial performance statistics, 2021. URL: https://bank.gov.ua/ua/statistic/supervision-statist/data-supervision#1.

[4] Malyar N., Shwalagin O. "Fuzzy procedures of the sequential analysis of variants in combinatory optimization problems and their application”. International journal “Information technologies and knowledge” (2012): 81-87.

[5] Rolik A., Bodaniuk M. “Algorithm of sequential analysis f options in the problem of distributing virtual machines in data processing center”. Adaptive automatic control systems. №21 (2012): 62-69.

Комп'ютерний контроль якості дизельного палива

Шпак О.І., вул. С. Бандери, 12, м. Львів, 79013, Україна

Анотація

Дана робота представляє розроблення давача. Він буде вимірювати цетанове число дизельного палива. Пізніше порівнювати його із запрограмованим стандартним значенням. Та видавати інформацію про стан палива на автозаправній станції. Ця інформація буде надсилатися по BLE (Bluetooth Low Energy) на мобільний телефон споживача.

Ключові слова

Дизельне паливо, давач, BLE (Bluetooth Low Energy), якість, контроль.

Computer control of diesel fuel quality

Shpak O.I. S.Bandery, 12 St., Lviv, 79013, Ukraine

Abstract

This work represents the development of the sensor. It will measure the cetane number of diesel fuel. Later compare it with the programmed default value. And publish information about the state of fuel at the gas station. This information will be sent via BLE (Bluetooth Low Energy) to the consumer's mobile phone.

Keywords

Diesel fuel, sensor, BLE (Bluetooth Low Energy), quality, control.

1. Вступ

Висока конкуренція на ринку роздрібних продажів дизельного палива і боротьба за клієнта вимагає від паливних компаній активних маркетингових дій. Одним з найголовніших методів перевірки є контроль якості палива, що реалізується на автозаправних станціях (АЗС) і донесення інформації до клієнта про інструменти цього контролю різними способами.

Наукові дослідження з метою удосконалення оперативних методів контролю дизельного палива є своєчасними та актуальними, оскільки стандартні лабораторні методи призначені для дослідження фізико-хімічних показників є досить складними в реалізації та вимагають громіздкого обладнання. Вони не придатні для здійснення оперативного контролю в місцях розповсюдження, тому пропонується комп'ютерний метод контролю якості дизельного палива, який дозволить визначити якість палива за коротший термін і більш оперативно. Застосування комп'ютерного методу має на меті застосування давача, який дозволить встановити зв'язок між рівнем якості дизельного палива та телефоном споживача за допомогою ІЧ спектру. Дослідження нових методів контрою дозволять встановити залежності електричних параметрів дизельного палива у широких діапазонах частот електромагнітного поля від хімічної природи ДП. Це дозволить розширити оперативність контролю якості дизельного палива даним методом.

2. Визначення рівня якості дизельного палива.

Для визначення рівня якості дизельного палива у паливному баці автомобіля пропонується новий метод контролю якості ДП. Він полягає у передачі даних від продавця до покупця. Тобто, інформуванні клієнтів про якість палива, що реалізується на АЗС. Результатом роботи даного методу є

Oksana.i.shpak@lpnu.ua
ORCID: 0000-0002-2093-9029

інформація про рівень якості дизельного палива на екрані телефону споживача.

Рівень якості поділяється на три складові: задовільна якість, добра якість і відмінна якість палива [1].

2.1. Застосування давача

При застосуванні даного методу використовується давач, який вимірює показники якості палива і порівнює їх із стандартизованими. Як результат роботи видає звіт про рівень якості палива. Давач представляє собою пристрій (рисунок 1), який дозволяє виміряти цетанове число палива. Ця інформація передається у буфер, де виміряні показники порівнюються зі стандартними. На основі цих значень і формується рівень якості ДП, який передається по Bluetooth на мобільний телефон споживача.

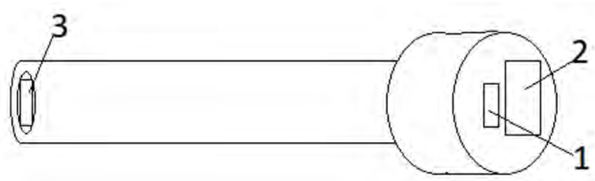


Рисунок 1: Вигляд пристрою.
1 - Bluetooth модуль nRF24L01.
2 – мікроконтролер Arduino nano.
3 - давач вимірювання якості ДП.

2.2. Цетанове число дизельного палива та методи його підключення

На багатьох АЗС можна зустріти дизельне паливо низької якості, яке може нашкодити двигуну. Тому важливо враховувати цетанове число дизпалива. Це один з перших критеріїв, який говорить про його якість. Є й інші, але саме цей найважливіший. Оптимальне значення цетанового числа знаходиться в діапазоні 50-60. При цих значеннях досягається максимальне "насичення" палива.

Непрямі методи визначення цетанового числа в основному спираються на аналіз компонентного складу [2]. В основному це різного роду інфрачервоні спектрометри, проте є приклади і успішного використання для цієї мети хроматографії і мас-спектрометрії. Інфрачервоні експрес-аналізатори необхідно перекалібровувати при будь-якій зміні складу дизельного палива - наприклад, при його змішуванні з біодизелем. Крім того, вони не здатні працювати з паливом, у яке додані поліпшувачі цетанового числа. Існують невеликі ІЧ-спектрометри які видають аналоговий сигнал. Для цього теоретично потрібно знати, які дані з конкретним спектрометром буде давати еталонне дизельне паливо. Його необхідно зафіксувати у буфері даних. Сам ІЧ-спектрометр запрограмовується на цетанове число в межах 50-60. Тобто, по Bluetooth буде приходити аналоговий сигнал в межах від 50 до 60, який буде означати рівень якості ДП. Зафіксоване стандартне значення цетанового числа порівнюється з виміряним значенням цетанового числа. Чим більше значення, тим показник буде кращий.

2.3. Підключення давача контролю якості у бак автомобіля.

Підключення пристрою в бак автомобіля буде відбуватися за схожим принципом, що і давач розходу палива [3]. Спочатку йде підготовка та проводиться ретельне обстеження паливного бака. Пізніше монтаж самого давача. Завершальним етапом є тарування паливної ємності, яке полягає в поступовому (порціями) наповнені баку паливом з паралельної фіксацією показників кожного заливу.

Схематичне підключення давача вимірювання рівня якості з розташуванням плати керування та Bluetooth модуля спереду і зверху показано на рисунку 2 та рисунку 3 відповідно.

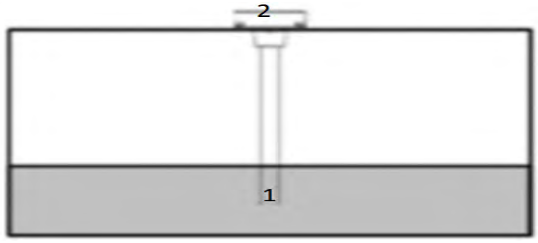


Рисунок 2: Вигляд встановленого пристрою спереду.
1 - Розташування плати керування та Bluetooth модуля.
2 - Розташування давача.

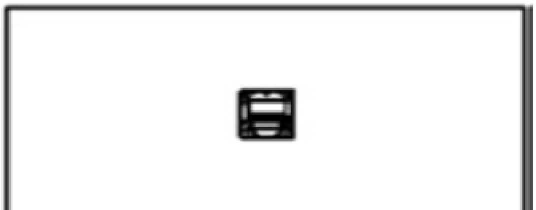


Рисунок 3: Вигляд встановленого пристрою зверху.

Таке розташування дає декілька переваг:
1. 1. До “мозгу” схеми можна добратися дуже просто.
2. 2. Полегшує заміну елементів в разі браку.
3. 3. Полегшує встановлення датчика в бак.

2.4. Під’єднання давача до мобільного телефону споживача.

Щоб під’єднати давач до мобільного телефону використовується технологія BLE (Bluetooth Low Energy) [4]. BLE це протокол передавання даних “повітрям”, який широко використовується в різних сферах. Найкраще ця технологія підходить для сценарію: дані, які не часто передаються і не швидко обробляються. Великою перевагою BLE є його стійкість в роботі і низьке енергоспоживання. Протокол BLE міститься в модулях з вбудованим програмним забезпеченням. Модулі під’єднуються до кінцевого пристрою. Для під’єднання телефону до пристрою буде використовуватися додаток nRF Connect for Mobile.

Для встановлення BLE необхідно виконати декілька кроків:

1. Сканувати всі доступні пристрої;
2. Підключитися до потрібного пристрою за допомогою кнопки “connect”.

2.5. Зчитування (передача даних) інформації з давача на телефон.

Дані про рівень якості ДП на мобільному телефоні споживача можна побачити після підключення в сервісах пристрою [5, 6]. Передати виміряний сигнал датчика в реальному часі з Arduino-Mikrocontroller на Android-смартфон можна через Bluetooth.

При вимірювання цетанового числа у баці автомобіля і порівнянні його із стандартизованим, споживач на своєму гаджеті має побачити повідомлення про якість палива. Наприклад: «відмінна якість палива», якщо цетанове число знаходиться в межах від 55-60.

3. Розроблення структурної схеми (системи) контролю якості ДП.

Структурна схема розроблювального пристрою буде мати вигляд, представлений на рисунку 4. Вона включає в себе такі елементи: еталон якості ДП, керуюча плата, датчик якості ДП, блютуз модуль, телефон.

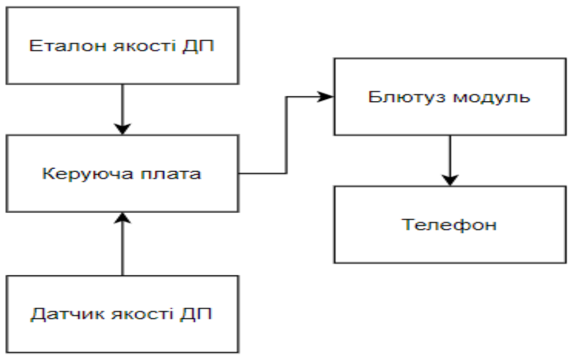


Рисунок 4: Структурна схема пристрою, що розробляється.

4. Висновки

Розроблюваний пристрій, для визначення рівня якості ДП, дозволяє виміряти цетанове число палива. Отримана інформація передається у буфер, де виміряні показники порівнюються зі стандартними. На основі цих значень і формується рівень якості ДП, який передається по Bluetooth на мобільний телефон споживача. На екрані телефону покупець бачить повідомлення, наприклад: «відмінна якість палива». Таким чином, споживач оцінює рівень якості ДП на АЗС за повідомленням на екрані свого телефону.

5.Література

[1] Шпак О.І., Столярчук П.Г., Юзевич В.М. Дослідження якості біологічного дизельного палива. //Вимірювальна та обчислювальна техніка в технологічних процесах. Науково-технічний журнал. Хмельницький – 2012. - № 3. – С.136-143.

[2] Цетанове число дизельного палива, 2010. URL: https://koneks-oil.ru/blog/cetanovoe-chislo-dizelnogo-topliva.html.

[3] Встановлення давача в бак автомобіля, 2019. URL: https://itob.ru/blog/ustanovka-datchika-urovnya-topliva//

[4] Огляд і робота з блютуз модулями, 2015. URL: https://arduinomaster.ru/datchiki-arduino/arduino-bluetooth-hc05-hc06/.

[5] Пристрої передачі даних, 2010. URL: http://pps.vtc.vn.ua/-5

[6] Передача даних з пристрою на телефон, 2016. URL: https://www.instructables.com/How-to-Receive-Arduino-Sensor-Data-on-Your-Android/

Доцільність використання геоінформаційних підсистем у складі Автоматизованої системи управління Збройних Сил України

Олександр Мясіщев[1], Наталія Литвиненко[2], Олексій Федченко[3]

[1] Хмельницький національний університет, вул. Інститутська, 11, Хмельницький, 29016, Україна
[2, 3] Військовий інститут Київського національного університету імені Тараса Шевченка, вул. Ломоносова, 81, Київ, 03189, Україна

Анотація
Геоінформаційне забезпечення військ є важливою складовою при прийняті рішень на всіх ланках військового управління та при вирішенні військово-прикладних задач.
Широке впровадження засобів електронно-обчислювальної техніки в системи управління Збройних Сил країни зумовлює необхідність розроблення нових підходів щодо автоматизації процесів обробки геопросторової інформації, вдосконалення організації, технологій і технічних засобів отримання інформації про місцевість, її опрацювання та передачі користувачам.
Аналіз напрямків використання геоінформаційних систем для військової сфери свідчить про необхідність застосування геоінформаційних підсистем в складі автоматизованої системи управління (АСУ) Збройних Сил України.
Актуальність та необхідність створення і впровадження геоінформаційних підсистем військового призначення як складової інформаційного забезпечення АСУ обумовлена загальним підвищенням вимог до оперативності, повноти і якості інформаційного забезпечення процесу управління військами (силами).

Ключові слова
геоінформаційне забезпечення, геоінофрмаційні технології, геоінформаційна підсистема, геопросторові дані, автоматизована система управління, Збройні Сили України.

Expediency of Using Geoinformation Subsystems as a Part of the Automated Control System of the Armed Forces of Ukraine

Olexander Myasischev[1], Nataliia Lytvynenko[2] and Olexiy Fedchenko[3]

[1] Khmelnitsky National University, Institutska Str., 11, Khmelnytsky, 29016, Ukraine
[2, 3] Military Institute of Taras Shevchenko National University of Kyiv, Lomonosova Str., Kyiv, 03189, Ukraine

Abstract
The geoinformation support of troops is an important component in decision-making at all levels of military management and in solving military-applied tasks.
The widespread introduction of electronic computing in the control systems of the Armed Forces of the country necessitates the development of new approaches to automate the processing of geospatial information, improve the organization, technology and technical means of obtaining information about the area of its processing and transmission to users.

EMAIL: alex56ma@gmail.com (A.1); n123n@ukr.net (A.2); a_fedchenko@ecomm.kiev.ua (A.3)
ORCID: 0000-0003-1269-425X (A.1); 0000-0002-2203-2746 (A.2); 0000-0003-1343-3828 (A.3)

The analysis of using directions of geoinformation systems for the military sphere testifies to necessity of geoinformation subsystems application as a part of the Automated control system (ACS) of Armed Forces of Ukraine.
The urgency and necessity of creating and implementing the geographic information subsystems for military purposes as a component of information support of the ACS is due to the general increase in requirements for efficiency, completeness and quality of information support of the command and control process.

Keywords
geoinformation support, geoinformation technologies, geoinformation subsystem, geospatial data, automated control system, Armed Forces of Ukraine.

1. Вступ

Сучасні тенденції розвитку та реформування Збройних Сил України зумовлюють обов'язкове залучення геоінформаційних технологій у військову сферу. Геоінформаційне забезпечення військ є вважливою складовою при прийняті рішень на всіх ланках військового управління та при вирішенні військово-прикладних задач.

Широке впровадження засобів електронно-обчислювальної техніки в системи управління Збройних Сил країни зумовлює необхідність розроблення нових підходів щодо автоматизації процесів обробки геопросторової інформації, вдосконалення організації, технологій і технічних засобів отримання інформації про місцевість її опрацювання й передачі користувачам.

Аналіз напрямків використання геоінформаційних систем для військової сфери свідчить про необхідність розробки єдиної технології підготовки геопросторових даних для застосування в геоінформаційних підсистемах АСУ Збройних Сил України.

Актуальність та необхідність створення і впровадження геоінформаційних систем військового призначення як складової інформаційного забезпечення АСУ обумовлена загальним підвищенням вимог до оперативності, повноти і якості інформаційного забезпечення процесу управління військами (силами).

2. Основна частина

Застосування інформаційних технологій у військовій сфері має свою специфіку, внаслідок чого технології, що використовуються, повинні відповідати ряду вимог, серед яких особливо важливою є підтримка безпечної роботи у повсякденній діяльності та при виконанні бойових завдань.

Сучасні інформаційні та мережеві технології у військовій справі є основою для інтеграції органів управління, засобів розвідки, спостереження, ураження, цілевказання, а також угруповань військ у високоадаптивну глобальну систему [1, 2].

Геоінформаційні технології дозволяють сформувати єдину картину ситуаційної обізнаності, розробити сучасні методи контролю і управління засобами збройної боротьби, включаючи безпілотні і роботизовані комплекси, підвищити прозорість і оперативність роботи штабів та інших органів управління.

На сучасному етапі розбудови Збройних Сил важливою особливістю розвитку систем управління є надбання цими системами властивості наскрізного масштабування – від систем стратегічного рівня до систем управління безпосередньо окремими підрозділами (та навпаки). В той же час, склад та зміст інформаційного забезпечення для вирішення завдань управління діями військ (сил) на різних рівнях ієрархії органів військового управління має відмінності.

В цьому сенсі не становить виключення і геоінформаційне забезпечення.

Для усвідомлення вимог до геоінформаційного забезпечення розглянемо деякі аспекти роботи систем управління діями військ (сил) на різних рівнях [3, 4].

На стратегічному рівні система управління повинна збирати, обробляти і поширювати інформацію, тим самим дозволяючи військам (силам) оперативно і адекватно реагувати на постійні зміни оперативної обстановки.

Забезпечення постійної актуальності відображення загальної оперативної обстановки підвищує ситуаційну обізнаність командирів всіх рівнів, дозволяючи приймати зважені рішення. Основні можливості, необхідні при відтворені оперативної обстановки у реальному часі, включають в себе можливість інтегрувати координати і зображення від датчиків різного типу, проводити аналіз співвідношення цих даних з метою усунення дублювань, обумовлених перекриттям поля зору датчиків. Датчики різного типу мають різні формати передачі інформації, періодичність та типи помилок, а також інші параметри. Крім того, важливу роль відіграє координатна прив'язка місця знаходження датчиків. Для забезпечення однозначного та точного отримання координат необхідно, щоб усі виміри базувалися на одному і тому ж картографічному матеріалі, або мали однозначні методи перетворення координат. Іншими словами, система управління повинна використовувати єдиний координатний простір. Реалізація цієї вимоги дозволяє створювати якомога точніше представлення оперативної обстановки, яка інтегрується з інформацією із різних джерел.

Автоматизований аналіз ситуації передбачає здатність використовувати можливості визначення напрямку розвитку подій безпосередньо в системі управління, в тому числі на основі просторового аналізу поведінки об'єктів оперативної обстановки. Такі системи надають додатковий рівень аналізу для прийняття рішень органами управління. Існує кілька підходів аналізу ситуації, включаючи аналіз кінематичних характеристик об'єктів оперативної обстановки на географічних мережах, просторово-часовий аналіз змін території в районі проведення операцій військами (силами), аналізу інших просторово пов'язаних даних. Приклади таких підходів включають ідентифікацію загроз на основі автоматичного відстеження характеристик та інших кінематичних даних таких, як наближення або проникнення на обмежені території, перевищення конкретної висоти / швидкості для певних типів траєкторій руху об'єктів оперативної обстановки, виявлення маневрів об'єктів, перешкод, що унеможливлюють рух тощо.

Властивість генералізації просторово розподіленої інформації для відображення на інформаційних екранах і табло в органах управління різного рівня в межах єдиного геоінформаційного простору дає змогу надавати інформацію про об'єкти оперативної обстановки в обсязі, необхідному для певного рівня. Наприклад, данні розвідки про нові кулеметні вогневі точки внесені в систему одразу відображаються на операційних табло (екранах портативних комп'ютерів) командирів тактичних підрозділів і в той же час приховані (невідображені) на операційних екранах стратегічного рівня.

Якщо органам управління вищого рівня знадобиться провести додатковий аналіз дій тактичних підрозділів, то повинна бути можливість отримати всю інформацію, яку використовує командир тактичної ланки.

При проведені планування застосування військ (сил) та в ході ведення бойових дій важливу роль має оперативне затвердження плануючих та бойових документів у вигляді рішень командирів всіх рівнів. Однією з особливостей геоінформаційного забезпечення повинна бути функція видання та збереження бойових графічних документів в електронній формі з можливістю колективної роботи з ними. В той же час система управління повинна підтримувати процес розгляду та затвердження електронних документів, у тому числі графічних на основі електронних карт.

Така автоматизація скорочує час відгуку для виконання завдання, гарантує, що завданням було надано офіційне затвердження, і знімає навантаження з координації процесу прийняття рішень, надаючи більше часу для оперативної діяльності персоналу органів управління [5,6].

В процесі планування застосування військ (сил) бере участь певна кількість осіб, відповідальних за окремі напрямки оперативного забезпечення. Зрозуміло, що вимоги до представлення інформації для вирішення завдань окремого виду оперативного забезпечення різні. В даній ситуації від геоінформаційного забезпечення вимагається можливість створювати потрібне інформаційне оточення для кожної окремої посадової особи, не втративши при цьому властивостей системи управління щодо узгодженості та синхронізації всіх її елементів.

Реалізація такої вимоги досягається використанням єдиного ядра геоінформаційного забезпечення, яке надає розробникам програмного забезпечення і користувачам базові геоінформаційні сервіси та інструментарій для їх спільного використання при вирішенні спеціалізованих завдань. При цьому ядро геоінформаційного забезпечення має діяти в просторово розподіленому середовищі, бути захищеним від зовнішнього втручання і бути максимально доступним для внутрішніх користувачів згідно встановленого розподілу повноважень.

Використання такого підходу дає змогу забезпечити єдиний (уніфікований) підхід до використання просторово розподіленої інформації, гнучкого масштабування

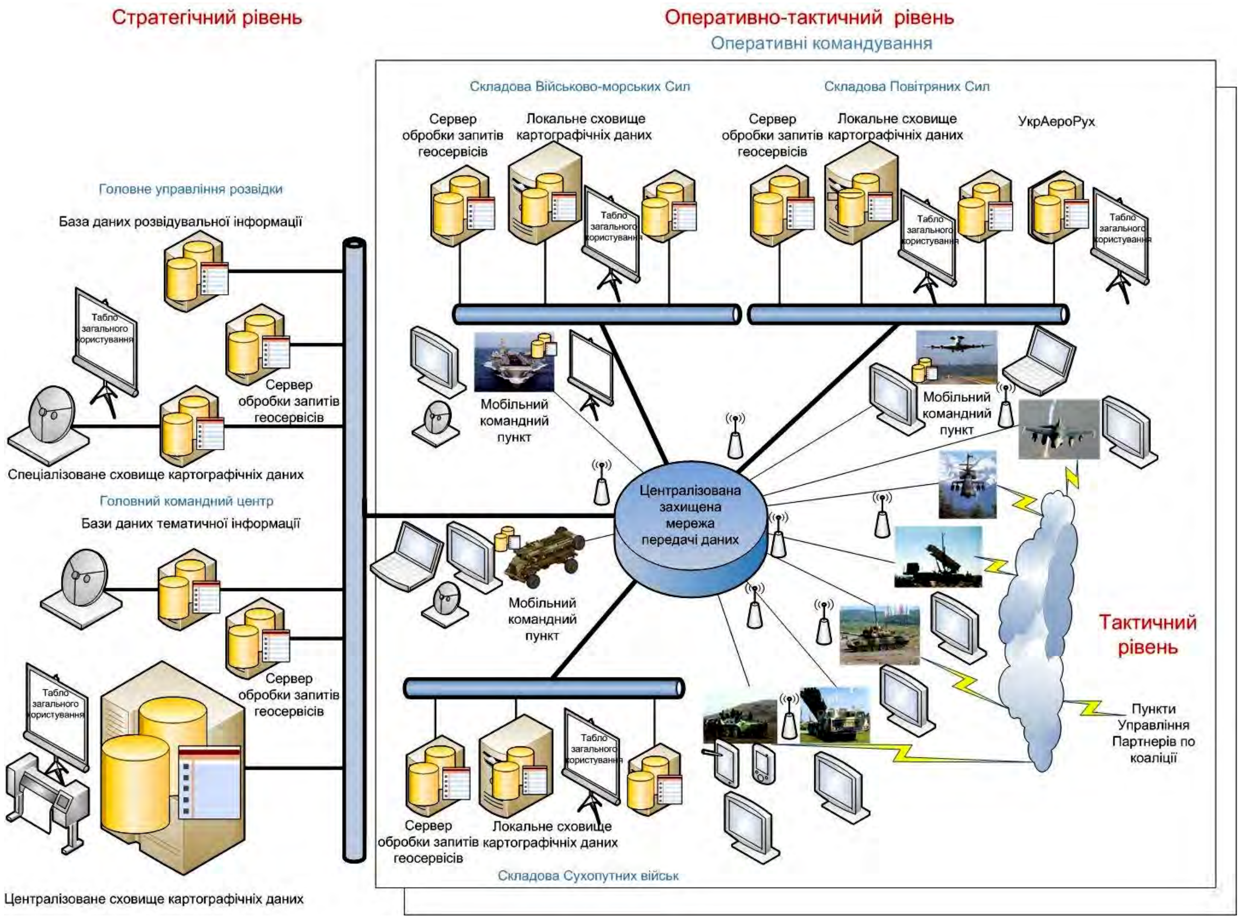


Рисунок 1: Геоінформаційне забезпечення системи управління

системи від тактичного рівня до стратегічного рівня (і навпаки), та дозволяє зосередитися на вирішенні прикладних завдань .

Геоінформаційне забезпечення системи управління повинне складатися з окремих предметно-орієнтованих модулів в рамках загальної сервісно-орієнтованої архітектури (рис. 1).

Геоінформаційне забезпечення системи управління повинне використовувати функціональні компоненти, які вбудовуються у відкриту, розширювану програмно-апаратну платформу для досягнення інтероперабільності використання просторово розподіленої інформації, використовуючи стандартні правила в рамках базових послуг, які підтримують обмін повідомленнями, перетворення даних задля подальшого їх поширення (розповсюдження) шляхом публікації та підписки на стандартні і спеціально створені шаблони проектування.

Процес планування підготовки та застосування військ (сил) включає оперативне планування в рамках одного виду оперативного забезпечення для різних рівнів, яке повинне бути узгоджене з іншими видами оперативного забезпечення за місцем та послідовністю в часі. Отже, геоінформаційне забезпечення цього процесу повинне забезпечити сумісну роботу між різними органами управління (ланками та службами) Збройних Сил України, не втративши при цьому єдність підходів до використання просторово розподіленої та іншої інформації [7].

При цьому за актуальність та достовірність тематичної інформації в межах окремого виду оперативного забезпечення відповідає її власник (ініціатор). Це інформація про стан систем зв'язку, озброєння та військової техніки, обладнання інженерних позицій, стан логістичного забезпечення, наявність та стан особового складу, дані розвідки тощо, тобто вся інформація, яка може бути введена в систему управління та використана в процесі підготовки та застосування військ (сил).

Система відображення оперативної обстановки та супутніх даних повинна забезпечувати графічне представлення інформації у робочій зоні перегляду, дозволяючи включати відео, графіки, таблиці та інші нетипові для геоінформаційного забезпечення види зображення інформації, які відносяться до просторово-розподілених об'єктів [8].

Будь-які зміни оперативної обстановки або рішень щодо виконання завдань військами (силами) одразу повинні з'являтися у відповідних місцях та у відповідних посадових осіб.

Для більш детального ознайомлення з місцевістю та всебічного вивчення її властивостей застосовуються тривимірні моделі. Ці моделі місцевості генеруються на основі даних про рельєф, закладених в базі даних електронних карт, та уточнюються за даними аеро-, фото- та космічної зйомки [9].

Тривимірні моделі місцевості використовують при аналізі оперативної обстановки, в задачах моделювання бойових дій військ (сил), в комплексних розрахункових задачах за напрямками оперативного забезпечення в ході проведення навчань особового складу та посадових осіб органів військового управління.

Як відомо, операції військ (сил) за типом поділяються на операції сухопутних військ (сил), повітряних військ (сил), ППО, військово-морських сил та міжвидових угрупувань військ (сил). Отже, геоінформаційне забезпечення повинно враховувати особливості процесу планування та проведення операцій кожного з наведених типів.

В реальних умовах підготовка операцій, як правило, здійснюється органами військового управління відповідних видів збройних сил та узгоджується з силами забезпечення і органами військового управління вищого рівня. Бойове застосування військ (сил) здійснюється виключно органами військового управління за належністю [10]. В такій ситуації геоінформаційне забезпечення повинне надати заздалегідь налаштовані шаблони сервісів та технологій взаємодії між ними [8].

Останнім часом паралельно з фазами планування та проведення операцій військ (сил) проводяться інформаційні операції. Для підвищення ефективності такого роду дій, як правило, використовується просторовий аналіз на основі геоінформаційного забезпечення. Для цього використовується інформація з баз картографічних даних та інших джерел про райони проживання етнічних груп населення, умови їх життя, наявність корисних копалин, промислових підприємств, об'єктів культурної та соціальної сфери.

Для досягнення тактичної переваги на полі бою необхідно забезпечити постійну ситуаційну обізнаність командирів тактичної ланки управління. В цьому сенсі завданнями геоінформаційного забезпечення є забезпечення командирів детальними картами, планами та знімками району ведення бойових дій, відображеннями об'єктів оперативної обстановки на портативних пристроях в умовах суттєвих обмежень пропускної здатності ліній передачі даних.

З цією метою для успішного проведення операцій завчасно створюються і доводяться до військ топографічні карти масштабу

1:50 000, 1:100 000, 1:200 000, 1:500 000, 1:1 000 000, а також карти масштабу 1:10 000 та 1:25 000. У зв'язку з тим, що тактичні підрозділи діють на невеликих територіях, дані карти є більш зручними, вони надають більш детальну інформацію про місцевість. На них більш точно відображаються промислові, господарські та військові об'єкти, дорожні шляхи, лінії електропередач, ріки та озера, гірські перевали тощо [11]. Досвід застосування інженерних топографічних підрозділів під час проведення миротворчих операцій продемонстрував, що найбільш ефективним є виготовлення топографічних карт, карт району цілей, різних видів спеціальних карт, фотокарт, різних видів довідок про місцевість безпосередньо в районі проведення операцій військ (сил) [12].

ГІС дає можливість створювати інформаційні продукти, точно відповідають вимогам користувача, дають можливість візуалізації картографічної інформації, недоступної для паперових карт. Тривимірний показ місцевості з конкретної точки місцезнаходження спостерігача або віртуальний огляд місцевості дають повнішу картину командиру, ніж паперова карта з нанесеною на ній обстановкою.

3. Висновки

Таким чином, зростаючі можливості використання перспективних інформаційних технологій у військовій сфері змушують переглянути традиційні підходи до використання засобів та методів сумісного аналізу наявної інформації щодо дій військ (сил). Зростаюча динаміка ведення збройної боротьби та великий обсяг інформації, яка підлягає аналізу, призводить до того, що людина вже не в змозі за дуже короткий час провести аналіз необхідного обсягу інформації та прийняти зважене рішення. За таких обставин виникає необхідність застосування нових технологій управління силами та засобами збройних сил для аналізу та прийняття рішень, відображення ситуацій, що склалися. Вони дають можливість прогнозувати можливі варіанти розвитку подій та пропонувати командирам будь-якого рівня різні шляхи досягнення оптимального результату.

Для вирішення таких завдань необхідно використовувати технології, які ефективно поєднують простір та час зі значними за обсягом супутніми даними у вигляді атрибутивної інформації щодо об'єктів оперативної обстановки, довідкової інформації про театр (район) ведення бойових дій, кліматичні умови, дані розвідки тощо. Саме для вирішення подібних завдань в арміях багатьох держав світу використовуються геоінформаційні технології.

Геоінформаційні підсистеми в складі АСУ Збройних Сил України забезпечить наступне:

- формування єдиного геоінформаційного простору в межах театру воєнних дій військ (сил);
- відображення оперативної обстановки на картографічному фоні з використанням загальноприйнятої символіки;
- генералізацію (масштабування) картографічної інформації в залежності від завдань, які вирішуються;
- постачання картографічних даних для забезпечення постановки завдань щодо реалізації рішень командирів всіх рівнів ієрархії підпорядкованості;
- обробку координатної та растрової інформації від всіх видів розвідки (аеро-, фото- космічні знімки);
- поєднання просторово розподіленої інформації з інформацією з тематичних баз даних, довідковою та іншою інформацією;
- надання технологій для автоматизованого аналізу оперативної обстановки з метою отримання вихідних даних для виявлення загроз та прогнозування їх розвитку, проведення моделювання і надання рекомендацій командирам відповідних рівнів;
- геоінформаційну підтримку автоматизованого вирішення завдань стосовно видів оперативного забезпечення Збройних Сил України (комплекси розрахункових задач);

- автоматизоване видання бойових графічних документів;
- сумісність автоматизованих систем та засобів автоматизації АСУ ЗС України з іншими подібними системами, в частині просторово розподіленої інформації, за умов виконання завдань у складі коаліції військ (сил).

4. Література References

[1] Цифровий вимір ЗСУ. За яких умов це можливо? 2019. URL: http://opk.com.ua/%D1%86%D0%B8%D1%84%D1%80%D0%BE%D0%B2%D0%B8%D0%B9-%D0%B2%D0%B8%D0%BC%D1%96%D1%80-%D0%B7%D1%81%D1%83-%D0%B7%D0%B0-%D1%8F%D0%BA%D0%B8%D1%85-%D1%83%D0%BC%D0%BE%D0%B2-%D1%86%D0%B5-%D0%BC%D0%BE%D0%B6%D0%BB/.

[2] Організація автоматизованих систем управління військами (силами) Збройних Сил України: навч. посіб. / Київ: НУОУ, 2010. 192 с.

[3] Фролов В.С. Структурно-логічна схема Єдиної автоматизованої системи управління Збройних Сил України. Наука і оборона. Київ, 2012. № 1. С. 15-23.

[4] Федченко О., Литвиненко Н. Підвищення ефективності управління військами за рахунок автоматизації геоінофрмаційної складової. Зб. наук. праць ВІКНУ. Київ, 2019. № 63. С. 47-53.

[5] Присяжнюк С.П., Филатов В.Н., Федоненков С.П. Геоиформационные системы военного назначения: учебник / СПб.: БГТУ, 2009. 210 с.

[6] Пермяков О.Ю., Солонніков В.Г., Прібилєв Ю.Б. та ін. Використання інформаційних технологій та застосування космічних систем в інтересах військ (сил). Київ: НУОУ ім. Івана Черняховського, 2014. 208 с.

[7] GIS in Defense and Intelligence Communities, 2011. URL: https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/brochures/pdfs/gis-in-defense-vol3.pdf.

[8] Geospatial Service-Oriented Architecture (SOA), 2011. URL: https://www.semanticscholar.org/paper/SOA-Approach-to-Integration%3A-XML%2C-Web-services%2C-and-Juric/a879c98143cb10adbfb6161011244f9706fb0cf6.

[9] Військовий стандарт 01.110.013 Видання 1. «Топогеодезичне і навігаційне забезпечення. Стандарти масштабів для сухопутних карт та аеронавігаційних морських карт (STANAG 3677) IGEO (Edition 3), IDT)».

[10] Військовий стандарт 01.110.020 Видання 1. «Топогеодезичне і навігаційне забезпечення. Формат обміну цифрової картографічної інформації».

[11] Класифікатор інформації, яка відображається на топографічних картах масштабів 1:10 000, 1:25 000, 1:50 000, 1:100 000, 1:200 000, 1:500 000, 1:1 000 000. Київ, 2002.

[12] Петлюк І.В., Власенко С.Г. Використання геоінформаційних систем військового призначення в миротворчих цілях: зб. матер. наук.-пр. конференції (Львів, 29 січня 2015 р.). Львів: 2015. С. 253-255.

Визначення Ізотопного Складу Радіоактивних Джерел за Допомогою Нейронної Мережі

Анна Скрипник [1], Манап Хажмурадов [2]

[1, 2] Національний науковий центр "Харківський фізико-технічний інститут", Інститут фізики високих енергій та ядерної фізики, вул. Академічна, 1, Харків, 61108, Україна

Анотація

Побудовано штучну нейронну мережу. За допомогою нейронної мережі визначено ізотопний склад джерел радіоактивного випромінювання за гамма-спектрами. Гамма-спектри отримано моделюванням методом Монте-Карло на основі набора бібліотек Geant4.

Ключові слова

Нейронна мережа, радіоактивне джерело, гамма-спектр

Determination of Isotopic Composition of Radioactive Sources with Using Neural Network

Anna Skrypnyk [1], Manap Khazhmuradov [2]

[1, 2] National Science Center "Kharkiv Institute of Physics and Technology", Institute of High-Energy Physics and Nuclear Physics , Akademichna St., 1, Kharkiv, 61108, Ukraine

Abstract

An artificial neural network is constructed. The isotopic composition of radioactive radiation sources is determined in gamma-spectra with using a neural network. Gamma-spectra are obtained by Monte-Carlo simulation based on a set of Geant4 libraries.

Keywords

Neural network, radioactive source, gamma-spectrum

1. Вступ

В більшості випадків для визначення ізотопного складу радіоактивних джерел за гамма-спектрами використовуються методики, в основі яких лежить пошук піків повного поглинання. Останнім часом все більше дослідників займаються розробкою методик ідентифікації нуклідів за допомогою штучних нейронних мереж. В роботі Olmos та ін. [1] було застосовано алгоритм, в основі якого лежить асоціативна пам’ять. Yoshida та ін. в своїй роботі [2] використали персептрон для ідентифікації нуклідів.

В даній роботі для вирішення задачі ідентифікації нуклідів використовується штучна нейронна мережа з одним прихованим шаром.

2. Побудова нейронної мережі

Побудовано штучну нейронну мережу. В якості вхідних даних було використано гамма-спектри. 5 гамма-спектрів ^{57}Co, ^{137}Cs, ^{131}I, ^{133}Ba та ^{241}Am було отримано моделюванням методом Монте-Карло. За допомогою них розраховано 500 гамма-спектрів змішаних радіоактивних джерел. Кількість вхідних нейронів складає 60. Вхідні

EMAIL: annyskryp@gmail.com (A. 1); khazhm@kipt.kharkov.ua (A. 2)
ORCID: 0000-0002-4920-5125 (A. 1); 0000-0003-4458-3669 (A. 2)

спектри було розбито на 60 частин. Встановлено, що для даної задачі достатньо одного прихованого шару та 40 прихованих нейронів в ньому. Кількість прихованих нейронів було вибрано експериментально. Нейронна мережа включає в себе 5 вихідних нейронів. Кожен з вихідних нейронів відповідає за один із нуклідів (^{57}Co, ^{137}Cs, ^{131}I, ^{133}Ba та ^{241}Am), вказує на його наявність або відсутність в радіоактивному джерелі.

Для побудови нейронної мережі було використано пакет Neuroph [3], що являється програмним комплексом для розробки нейронних мереж поширених архітектур.

3. Моделювання гамма-спектрів

Гамма-спектри були отримані за допомогою методів Монте-Карло на основі набора бібліотек Geant4, що призначені для моделювання проходження заряджених частинок, нейтронів та гамма-квантів скрізь речовину. У комп'ютерних експериментах в якості детектора було використано напівпровідниковий CdZnTe детектор розміром 6×6×3 мм. Отримано функції відгуку CdZnTe детектора на випромінювання ^{57}Co, ^{137}Cs, ^{131}I (рисунок 1), ^{133}Ba та ^{241}Am.

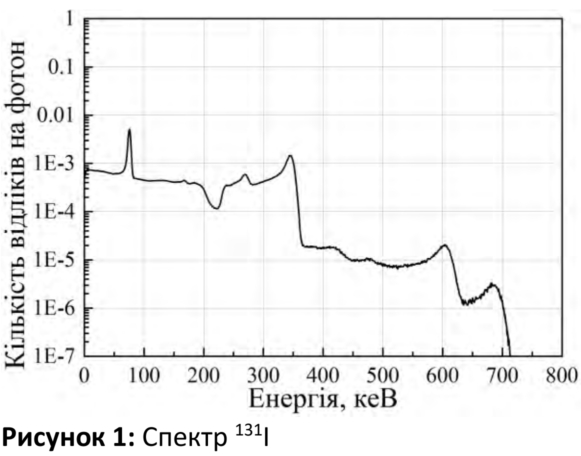


Рисунок 1: Спектр ^{131}I

Розраховано функції відгуку CdZnTe детектора на випромінювання змішаних джерел, до складу яких можуть входити ^{57}Co, ^{137}Cs, ^{131}I, ^{133}Ba та ^{241}Am.

4. Результати

На рисунку 2 представлено вихід нейронної мережі для змішаного джерела.

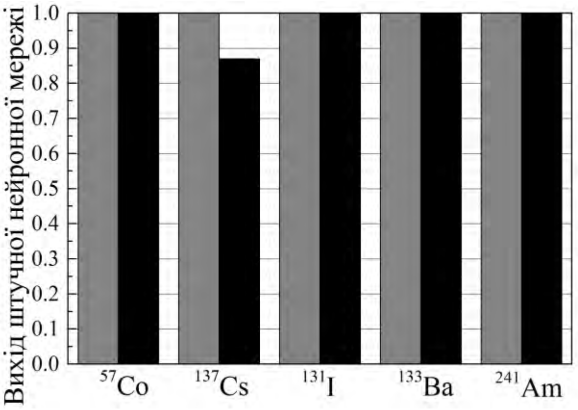


Рисунок 2: Вихід нейронної мережі для джерела ^{57}Co + 0.55 ^{137}Cs + 0.7 ^{131}I + ^{133}Ba + ^{241}Am

Сірим кольором позначено бажаний вихід, чорним – отриманий.

Результати можна інтерпретувати наступним чином: значення більше 0.5 на виході вказує на наявність нукліда в радіоактивному джерелі, яке досліджується, значення менше 0.5 – на його відсутність.

5. Висновки

За допомогою штучної нейронної мережі було ідентифіковано ^{57}Co, ^{137}Cs, ^{131}I, ^{133}Ba та ^{241}Am в змішаних радіоактивних джерелах. Результати вказують на успішну ідентифікацію.

6. Література References

[1] P. Olmos, J. C. Diaz, J. M. Perez, G. Garcia-Belmonte, P. Gomez, V. Rodellar, Application of neural network techniques in gamma spectroscopy, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 312 (1992) 167-173. doi.org/10.1016/0168-9002(92)90148-W.

[2] E. Yoshida, K. Shizuma, S. Endo, T. Oka, Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 484 (2002) 557-563. doi.org/10.1016/S0168-9002(01)01962-3.

[3] Java Neural Network Framework Neuroph. URL: http://neuroph.sourceforge.net.

Побудова підсистеми управління та автономного руху спеціального автомобіля

Станіслав Абрамов [1], Олег Лупаленко [2]

[1] *Військова академія, вул. Фонтанська дорога, 10, м. Одеса, 65009, Україна*
[2] *Військова академія, вул. Фонтанська дорога, 10, м. Одеса, 65009, Україна*

Анотація
Акцентовано увагу на створенні підсистеми управління та автономного руху спеціального автомобіля та її особливостях. Визначено завдання та вимоги до субсистем, які входять до підсистеми управління та автономного руху зазначеного типу автомобілів. Наводиться характеристика сенсорів системи LiDAR як складових частин субсистем багатоканального спостереження та комплексування інформації, автономного руху. Вказано на необхідність комплексування різних датчиків для виявлення та розпізнавання об’єктів, перешкод і забезпечення автономності руху автомобілів.

Ключові слова
автономний рух, підсистема управління та автономного руху, субсистема, сенсор, субсистема багатоканального спостереження та комплексування інформації, розпізнавання об’єктів, обробка інформації, 2D-LiDAR, сканер, 3D-LiDAR, тривимірна модель, навігація, субсистема автономного руху

Construction of the Subsystem of Control and Autonomous Movement of the Special Vehicle

Stanislav Abramov [1], Oleh Lupalenko [2]

[1] *Military Academy, Fontanska street, 10, Odessa, 65009, Ukraine*
[2] *Military Academy, Fontanska street, 10, Odessa, 65009, Ukraine*

Abstract
Emphasis is placed on the creation of a control subsystem and autonomous movement of a special vehicle and its features. Tasks and requirements to subsystems which are a part of a subsystem of management and autonomous movement of the specified type of cars are defined. The characteristics of LiDAR system sensors as components of subsystems of multichannel monitoring and information complexation, autonomous motion are given. The need to integrate different sensors to detect and recognize objects, obstacles and ensure the autonomy of traffic is indicated.

Keywords
autonomous motion, control and autonomous motion subsystem, subsystem, sensor, subsystem of multichannel observation and information complexation, object recognition, information processing, 2D-LiDAR, scanner, 3D-LiDAR, three-dimensional model, navigation, autonomous motion subsystem

EMAIL: xamamupa@ukr.net (A. 1); bessarabia_od@ukr.net (A. 2)
ORCID: 0000-0002-1522-4248 (A. 1); 0000-0002-9457-8437 (A. 2);

1. Вступ

Досвід застосування підрозділів Державної служби України з надзвичайних ситуацій свідчить про те, що автомобільна техніка (АТ) та її можливості суттєво впливають на результати виконання службових завдань. Поряд з тим тактико-технічні характеристики наявної АТ не відповідають сучасним вимогам, більшість наявних зразків автомобілів фізично та морально застаріли [1-4]. Один з напрямів рішення проблеми – створення спеціального автомобіля (СА) для виконання завдань в умовах надзвичайних ситуацій.

2. Автономні технології в автомобілях спеціального призначення

Аналіз останніх публікацій [5-12] свідчить про те, що в пріоритеті є створення машин, які здатні рухатися дистанційно-керовано або автономно. Так, великі виробники автомобілів і багато інших компаній, працюють над створенням автономних систем, в яких об'єктом дослідження стають підсистеми управління та автономного руху (ПСУ та АР) автомобіля. Така підсистема поєднує інформацію від різнотипних датчиків навколишнього середовища, аналізує її з використанням штучного інтелекту та видає управлінські команди органам управління [13].

За класифікацією SAE International серед систем допомоги водієві (англ. ADAS - Advanced Driver Assistance System) існує шість класів автономності: від рівня 0 – повністю ручне управління з можливістю попередження про небезпечні ситуації на дорозі, до рівня 5 – повністю безпілотний автомобіль [14].

Силові відомства різних країн є замовниками різних типів автомобілів, в першу чергу вантажних, з функцією автономного руху [5-7,10-12].

Так, компанія Oshkosh Defense вдосконалює технологію безпілотного транспортного засобу на базі вантажного автомобіля Oshkosh PLS. Зазначена технологія забезпечує перевезення небезпечних вантажів в режимах: «Leader-Follower», автономний рух і дистанційно керований.

Система автономного руху Oshkosh PLS складається з трьох комплектів, що встановлюються до неї. Перший – комплект активної безпеки з дротовим підключенням, якій з'єднує органі управління рухом машини з іншим інструментарієм. Функції автономності надходять з комплекту Robotics Research, який складається з датчиків та програмного забезпечення, які повідомляють транспортному засобу, що робити. Третій – це власне інтерфейс машини-лідера, який включає в себе сенсорний екран і дозволяє водію керувати іншими безпілотними автомобілями [5,6].

У 2016 році, в Україні на базі КрАЗ «Спартан» створено СА із дистанційним керуванням. Вітчизняна система Pilotdrive, яка встановлена на СА, оснащена комплексом спеціальних датчиків, що дозволяють автомобілю легко орієнтуватися в дорозі. Управління автомобілем здійснюється за допомогою планшета або операторської станції [9]. Систему можна «вчити»: режим Teach-inDrive дозволяє запам'ятати і відтворити певний маршрут [10].

Згідно звіту 2020 Autonomous Vehicle Technology Report [15], в світі не існує працюючих технологій 5-го рівня автономності.

Метою даної роботи є побудова архітектури підсистеми управління та автономного руху спеціального автомобіля.

2.1.1. Структура та завдання підсистеми управління та автономного руху спеціального автомобіля

Підсистема управління та автономного руху СА призначена для забезпечення управління рухом машини, режимами роботи її складових та функціонуванням спеціального обладнання [16].

Вона включає сукупність датчиків, які реагують на стан (параметри, характеристики) навколишнього середовища, забезпечують навігацію на основі пристроїв системи технічного зору (СТЗ) і програмне забезпечення СА.

Підсистема управління та автономного руху, в свою чергу є системою та складається з субсистем:

- *багатоканального спостереження та комплексування інформації* (БКС та КІ), що забезпечує отримання та обробку інформації від датчиків з метою визначення характеристик маршруту (зони) руху та його картографування;
- *навігації*, що забезпечує автоматичну орієнтацію та визначення місця розташування СА з прив’язкою до електронної карти місцевості або тривимірних і цифрових моделей рельєфу (об’єктів);
- *автономного руху*, що забезпечує автоматичне управління рухом машини;
- *обміну інформацією* між СА і оператором (пунктом управління) [17].

Головним завданням ПСУ та АР є забезпечення вирішення низки взаємопов'язаних завдань:

- комплексна обробка інформації від бортових датчиків, СТЗ та систем навігації з прив'язкою до карти місцевості;
- спостереження в звичайних і складних умовах видимості;
- розпізнавання об'єктів (перешкод на маршруті руху);
- створення тривимірних і цифрових моделей рельєфу (об’єктів);
- автоматичне планування маршрутів руху в реальному масштабі часу.

Для забезпечення ефективної роботи СА ПСУ та АР комплектують наступними датчиками: акустичними, оптичними, оптико-електронними, інфрачервоними, системи LiDAR тощо.

Основними завданнями субсистеми БКС та КІ (в частині підвищення можливостей з управління СА та інформативності водія) є:

- забезпечення відеоспостереження в складних умовах видимості шляхом поліпшення якості зображення за рахунок фільтрації завад, обумовлених опадами, туманом, димовою завісою, темною порою доби, а також його відновлення та доповнення за рахунок комплексування відеоданих СТЗ та різних сенсорів);
- відновлення геометрії робочої зони, створення тривимірних і цифрових моделей рельєфу, об’єктів, їх проекцій та перетинів з маршрутом руху;
- розпізнавання об'єктів (перешкод) з виділенням їх на зображеннях електронної карти та сформованої інформаційно-аналітичної моделі.

В якості бортових СТЗ пропонується застосовувати датчики 2D-LiDAR (таблиця 1) [19-21] і 3D-LiDAR на опорно-поворотних пристроях, які в тому числі закомплексовані з відеокамерами.

Таблиця 1

Характеристики датчиків 2D-LiDAR

Параметри	**LMS291-S05**	**UTM-30LX-EW**	**NAV350-3232**
Джерело світла	ІЧ (довжина хвилі λ = 905 nm)	Лазер (довжина хвилі λ = 905 nm)	Лазер (довжина хвилі λ = 905 nm)
Частота сканування	75 Гц	-	8 Гц ± 5 %
Поле зору, (горизонт.)	180°	270°	360°
Робоча область, м	0-80	0,1-60	0,5-250
Робоча напруга, В	24	12	24
Габаритні розміри, (Д х В х Ш), мм	156×155×210	62×62×87	115×121×222
Вага, кг	4,5	0,37	2,4

Для розпізнавання ґрунту пропонується застосовувати закомплексовані СТЗ, які об'єднують сканери 3D-LiDAR (таблиця 2) [22-24], кольорові відеокамери, тепловізор, блок навігації та комп'ютер з програмним забезпеченням.

Таблиця 2

Характеристики датчиків 3D-LiDAR

Параметри	**HDL-32E**	**RS-LiDAR-32**	**HDL-64E**
Джерело світла	Лазер (довжина хвилі λ = 903 нм)	Лазер (довжина хвилі λ = 905 нм)	Лазер (довжина хвилі λ = 903 нм)
Вертикальний кут огляду	40° (від +10° до -30°)	40° (від +15° до -25°)	27° (від +2° до -25°)
LiDAR канали	32	32	64
Поле зору, (горизонт.)	360°	360°	360°
Робоча область	0,05-100 м	0,2-200 м	1-120 м
Похибка вимірювання	±2 см *	±3 см	±2 см *
Робоча напруга, В	24	9-32	12-32
Габар. розміри, (Діаметр х Висота), мм	144 × 85	115 × 96	215 × 283
Вага, кг	1,36	1	12,7

* на відстані 25 м

Сканери системи LiDAR використовують різні принципи вимірювання: час прольоту світла, фазову модуляцію, інтерферометрію та тріангуляцію. У багатьох випадках перші три принципи об'єднуються в метод, якій відомий як вимірювання часу прольоту (англ. TOF – Time-of-Flight). Датчики тріангуляції вимірюють короткі відстані (максимум кілька метрів) з високою точністю, тоді як датчики TOF – для коротких, так і для великих відстаней [25,26].

Субсистема автономного руху (ССАР) являє собою комплекс програмно-апаратних засобів, призначений для безпосереднього управління рухом СА в автоматичному та напівавтоматичному режимах.

Інформація, яка потрібна для прийняття рішення ССАР, отримується від датчиків субсистеми БКС та КІ, за допомогою яких вона вирішує гальмувати чи збільшити швидкість.

Автоматичний режим передбачає рух СА (тільки з 4-м рівнем автономності ССАР) в автономному режимі по заданому маршруту в певну точку або в складі колони (за алгоритмом «Leader-Follower») або повернення машини у вихідне положення (точку старту) за умови наявності високо деталізованої тривимірної карти [27].

Напівавтоматичний режим (ступінь автономності ССАР: рівень 1-3) дозволяє здійснювати рух об'єкта під керуванням людини-оператора, як з місця водія, так і з місця старшого машини з передачею ряду функцій, які виконуються в штатному режимі водієм, на систему управління (наприклад, оптимальне автоматичне перемикання передач) [28].

3. Висновки

Сьогодні, в галузі машинобудування, основним перспективним напрямом є розробка безпілотних автомобілів.

Хоча концепції автономних автомобілів різняться, всі ці транспортні засоби покладаються на використання набору датчиків для сприйняття навколишнього середовища, розширеного програмного забезпечення для обробки вхідних даних і визначення маршруту руху транспортного засобу, а також набору виконавчих систем і механізмів для прийняття рішень.

Застосування субсистеми багатоканального спостереження та комплексування інформації в підсистемі управління та автономного руху спеціального автомобіля покращить його існуючі експлуатаційні характеристики та розширить область застосування даного типу машин і дозволить отримати наступні результати: підвищення живучості машини та екіпажу; автоматичне виявлення та розпізнання об’єктів (перешкод) з високою точністю та якістю їх відображення; точність навігації (положення машини); можливість створення тривимірних і візуально-подібних моделей місцевості (об’єктів).

Архітектура підсистеми управління та автономного руху спеціальних автомобілів складається з субсистем: багатоканального спостереження та комплексування інформації; навігації; автономного руху; обміну інформацією.

4. Література References

[1] Лупаленко О. Автомобільне базове шасі Високомобільних десантних військ [Електронний ресурс] / О. Лупаленко, В. Биковець // Наукова конференція «Новітні шляхи створення, технічної експлуатації, ремонту і сервісу автомобілів» – 2015.

[2] Крайник Л. Основні напрями розвитку середньотонажних військових автомобілів / Л. Крайник, М. Грубель // Автомобильный транспорт, № 40. – 2017.

[3] Ukrainian Military Pages – 2020. – https://www.ukrmilitary.com/2020/10/bars.html.

[4] Defense Express. – 2021. – https://defenceua.com/news/zvatimut_jogo_sajgak_viznacheno_vimogi_do_novogo_legkovogo_avtomobilja_pidvischenoji_prohidnosti_dlja_zs_ukrajini-3681.html.

[5] Kelsey D. Atherton Oshkosh robot trucks could roll out to the Army by 2020 / Kelsey D. Atherton // Defense News. – 2018.

[6] Oshkosh Autonomous Technology // Oshkosh Defense – https://oshkoshdefense.com/technology/oshkosh-autonomous-technology/.

[7] Szatkowski D. Military working to make its autonomous technology smarter / Danielle Szatkowski // Automotive News Europe. – 2019.

[8] Гусляков О. Концептуальні засади створення вітчизняних бойових роботизованих систем важкого класу для потреб ЗСУ / О. Гусляков, А. Довгополий, І. Чепков. // Збірник тез доповідей «Перспективи розвитку озброєння та військової техніки Сухопутних військ». – 2019. – С. 11–12.
[9] Перший безпілотний КрАЗ проходить випробування // Ukrainian Military Pages. https://www.ukrmilitary.com/2016/10/unmanned-kraz.html.
[10] Ященко В. Інтелектуальні роботизовані системи в забезпеченні безпеки життєдіяльності в надзвичайних ситуаціях мирного і воєнного часу / В. Ященко // Математичні машини і системи, № 1. – 2020.
[11] Type-X RCV // Milrem Robotics. – 2021. – https://milremrobotics.com/type-x/.
[12] Lockheed Martin's Autonomous Mobility Applique System logs more than 55,000 testing miles // AUVSI News. – 2017.
[13] Автопилот. Беспилотный автомобиль // TAdviser – 2020. –https://www.tadviser.ru.
[14] Shuttlewoth J. SAE Standards News: J3016 automated-driving graphic update / Jennifer Shuttlewoth // SAE International. – 2019. – https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic.
[15] 2020 Autonomous Vehicle Technology Report // Wevolver. – 2020. – https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report.
[16] Григорьев А. П. Построения системы противодействия наземным боевым роботам / А. П. Григорьев, В. К. Набок. // Збірник наукових праць Військова академія (м. Одеса). – 2014. – С. 47.
[17] Парамонов Н. Б. Доверенная программно-аппаратная среда «Эльбрус» бортовых вычислительных средств робототехнических комплексов / Н. Б. Парамонов, Д. А. Ржевский, В. И. Перекатов. // Вопросы радиоэлектроники. – 2015. – С. 159–169.
[18] Григорьев А. П. Модель обоснования оперативно-тактических требований и тактико-технических характеристик к наземним боевым робототехническим комплексам / А. П. Григорьев, О. В. Кравчук, В. К. Набок. // Збірник наукових праць Військова академія (м. Одеса). – 2014. – С. 128–134.
[19] Лазерный 2D-сканер LMS2xx / LMS291 / Outdoor // Sick. Sensor Intelligence – https://www.sick.com/ru/ru/detection-and-ranging-solutions/-2d-lidar/lms2xx/lms291-s05/p/p109849.
[20] Distance Data Output/UTM-30LX-EW // Hokuyo. Optical data transmission device, Photo sensor, Auto counter, and Automatic door – https://www.hokuyo-aut.jp/search/single.php?serial=170.
[21] Датчики 2D-LIDAR NAV350-3232 / NAV3xx // Sick. Sensor Intelligence – https://cdn.sick.com/media/pdf/1/41/041/dataSheet_NAV350-3232_1052928_ru.pdf.
[22] HDL-32E. High Resolution Real-Time 3D Lidar Sensor // Velodyne Lidar – https://velodynelidar.com/products/hdl-32e/.
[23] RoboSense RS-LiDAR-32 3D Laser Range Finder // Generationrobots – https://www.generationrobots.com/en/403307-robosense-rs-lidar-32-laser-range-finder.html..
[24] HDL-64E High Definition Real-Time 3D Lidar // Velodyne Lidar – https://velodynelidar.com/products/hdl-64e/.
[25] Білецький В. Р. Датчики LiDAR у сільському господарстві. / В. Р. Білецький, М. О. Бондарчук. // Матеріали XXI Міжнародної наукової конференції „Сучасні проблеми землеробської механіки”. – 2020. – С. 30–32.
[26] Курпас Д. В. Системи діагностики земної поверхні для наземних безпілотних транспортних засобів / Д. В. Курпас, О. В. Тищенко, О. М. Бевза. // Матеріали 13-ї науково-практичної конференції «Перспективні напрямки сучасної електроніки». – 2019. – С. 132–140.
[27] Формирование моделей виртуальной реальности и информационно-навигационных полей для обеспечения автономного функционирования РТК специального назначения / [В. С. Лапшов, В. П. Носков, И. В. Рубцов та ін.]. // Известия ЮФУ. Технические науки. – 2017. – №2 (187) – С. 248–265.
[28] Калинин А. В. Средства, обеспечивающие автономное движение наземных РТК / А. В. Калинин, В. П. Носков, И. В. Рубцов. // Известия ЮФУ. Технические науки. – 2012. – № 11 (136). – С. 71–81.

Рефлекторна інтелектуальна система УМ та її застосування для прогнозування Чемпіонату Європи з футболу 2020

Юрій Тесля[1], Юлія Хлевна[2], Олексій Єгорченков[2], Григорій Заспа[1], Євгенія Катаєва[1], Андрій Хлевний[2]

[1] *Черкаський державний технологічний університет, бульвар Шевченка,460, 18000, Україна*
[2] *Київський національний університет імені Тараса Шевченка, вул.Володимирська, 64, 0103, Україна*

Анотація

Продемонстрована можливість застосування методу несилової (інформаційної) взаємодії для вирішення задач прогнозування. Показано, що розроблена з використанням цього методу система УМ дозволила більш точно спрогнозувати результати Чемпіонату Європи з футболу 2020 року, ніж штучний інтелект, аналітичний центр та фахівці букмекерських контор.

Ключові слова

Метод несилової взаємодії, прогнозування, футбол, рефлекторні системи

Reflective intelligent system UM and its application for forecasting the 2020 European Football Championship Results

Iurii Teslia [1], Iulia Khlevna[2], Oleksiy Yegorchenkov[2], Hryhoriy Zaspa[1], Kataieva Yevheniia[1], Andrii Khlevnyi[2]

1 Cherkasy State Technological University, 460 Shevchenko blvd., 18006, Cherkasy, Ukraine, 18000
[2] *Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Kyiv, Ukraine, 01033*

Abstract

The possibility of applying the method of nonviolent (information) interaction to solve forecasting problems is demonstrated. It is shown that the UM system developed using this method allowed to predict the results of 2020 European Football Championship more accurately than artificial intelligence, analytical center and bookmakers.

Keywords

Method of nonviolent interaction, forecasting, football, reflective systems

1. Вступ

Важливим питанням діяльності людини є прогнозування – в бізнесі, політиці, економіці, спортивних заходах, природніх явищах і т.д. Для вирішення цих та інших задач прогнозування розроблено багато методів та створено чимало різноманітних аналітичних систем [1 – 4]. Але поки що їх результати не задовольняють вимоги працівників багатьох сфер.

Тому є необхідність в подальших наукових дослідженнях, які спрямовані на побудову ефективних прогнозуючих систем.

EMAIL: teslyas@ukr.net (A. 1); yuliya.khlevna@gmail.com (A. 2); alexee@ukr.net (A. 3); gzaspa@ukr.net (A. 4); kataevae@ukr.net (A. 5) andlev@ukr.net (A. 6);
ORCID: 0000-0002-5185-6947 (A. 1); 0000-0002-1874-1961 (A. 2); 0000-0003-1390-5311 (A. 3) ; 0000-0001-8726-7100 (A. 4); 0000-0002-9668-4739 (A. 5); 0000-0001-8942-6670 (A. 6)

2. Аналіз літературних даних та постановка проблеми

Поки що немає універсальних систем прогнозування, які б краще за людину прогнозували в різних предметних областях. Найпоширеніші системи прогнозування створюються на нейромережевих технологіях, навчання яких проводиться в межах конкретної предметної області діяльності [1 – 4]. Інший принцип закладено в системи, що використовують результати, отримані в теорії несилової взаємодії [5-7]. На сьогодні вже створено системи розпізнавання змісту природномовного тексту, оцінки інвестиційних пропозицій, оцінки впливу шкідливих речовин у водних ресурсах регіону на здоров'я населення, і т.д. [5-7].

Тому зрозумілим є бажання спробувати використати цю теорію для побудови систем прогнозування. Зокрема, результатів футбольних матчів, як сфери, яка є дуже популярною, і в якій є достатньо даних для оцінки ефективності систем прогнозування.

3. Мета і задачі дослідження

Метою роботи є перевірка можливості застосування методу несилової взаємодії [5] для прогнозування результатів футбольних матчів. Для досягнення мети необхідно розв'язати наступні завдання: розробити систему прогнозування; провести експериментальну перевірку розробленої системи.

4. Виклад основного матеріалу

4.1. Рефлекторна система прогнозування результатів футбольних матчів УМ

З використанням методу несилової взаємодії була створена рефлекторна система прогнозування результатів футбольних матчів – УМ. В якості чинників, що враховуються для прогнозування були використані: поле, на якому проходить матч (1-своє, 0-нейтральне, -1 – чуже); кількість очок, набраних в попередньому матчі; кількість очок, набраних в 3-х попередніх матчах; кількість очок, набраних в 5-ти попередніх матчах; різниця рейтингу команд (можна використовувати різницю місць в турнірній таблиці); як зіграли між собою попередній матч.

Були обрані саме ці чинники, які, з одного боку можуть бути отримані автоматично з спеціалізованих сайтів для всіх матчів, з іншого боку по ним з 2010 року велось накопичення статистичного матеріалу (близько 158000) матчів. І це дозволило застосувати цю статистику для прогнозування Чемпіонату Європи з футболу 2020.

Для кожного чинника, виходячи з статистики попередніх матчів, розраховується величина його впливу на прогноз [5]:

Ймовірність результату (виграш, нічия, програш) оцінюється по сумарному несиловому (інформаційному) впливу:

$$p(A_k/X) = \frac{n(|d(A_k/X) - \overline{d(A_k)}| \le d_0)}{N}, k = \overline{1,3}, \quad (1)$$

де X – чинники; A_k- результат матчу (A_1 – виграш; A_2 – нічия; A_3 – програш); $n(|d(A_k/X) - d(A/X)| \le d_0)$ – кількість попередніх матчів, у яких величина впливу відхилялась від розрахованої не більше ніж на d_0; $d(A_k/X)$ - величина впливу на результат A_k всіх чинників прогнозованого матчу; $\overline{d(A_k)}$ - величина впливу на результат A_k в попередніх матчах; N –кількість матчів, які є в пам'яті рефлекторної прогнозуючої системи.

Значення $d(A_k/X)$ і $d(A_k)$ отримуються з використанням методу несилової взаємодії [5].

На відміну від систем, що будуються з використанням нейромереж в системі УМ відсутні навчальна і контрольні вибірки, відсутній сам процес навчання. В прогнозуванні результатів матчів, які ще не відбулися використовується статистична інформація по всім матчам, які відбулися Це значно прискорює і спрощує роботу такої системи, порівняно з системами, що використовують штучні нейромережі. Розглянемо результати прогнозування.

4.2. Використання системи УМ для прогнозування Чемпіонату Європи з футболу - 2020

Система була перевірена на матчах Чемпіонату Європи з футболу 2020 (в подальшому ЧЄ-2020). 26 червня 2021 року, до початку матчів 1/8 фіналу, було опубліковано прогноз штучного інтелекту Stats Perform [8], бельгійського аналітичного сервісу KU Leuven [9], та прогноз системи УМ [10].

Для оцінки відхилення прогнозу від фактичних результатів було взято критерій середньоквадратичного відхилення розрахованої ймовірності від фактичного результату:

$$K_i^s = \sqrt{\frac{\sum_{j=1}^{M}\left(p_{ji} - r_j\ \right)^2}{M}}, \qquad (2)$$

де K_i^s – оцінка ефективності i-ої системи прогнозування (s=1 – розглядаються ймовірності проходження всіх команд; s=2 – розглядаються ймовірності тільки для тих команд, які пройшли до наступної стадії); p_{ji} – прогнозована ймовірність проходження до наступної стадії ЧЄ-2020, чи стати чемпіоном j-ої команди, в прогнозах i-ої системи; r_j – фактична ймовірність проходження до наступної стадії ЧЄ-2020, чи виграшу j-ою командою; M – кількість прогнозів.

По закінченню ЧЄ-2020, для всіх прогнозів було отримано наступні результати:

$$K_{KU\ Leuven}^1 = 0{,}388569;$$
$$K_{\text{ШІ}}^1 = 0{,}36942;$$
$$K_{\text{УМ}}^1 = 0{,}340982.$$

Рефлекторна прогнозуюча система дала прогноз на 7,7% кращий, ніж система штучного інтелекту!

Крім того, було порівняно результати штучного інтелекту, системи УМ з прогнозами букмекерів [11]. Але оскільки букмекери прогнозували лише проходження в наступну стадію, то було використано критерій (2) в розрізі тих команд, що проходили до наступної стадії (M=30 в формулі 2).

Результати прогнозування з використанням критерію (2) по відношенню до команд, що вийшли в наступну стадію наступні:

$$K_{\text{KU Leuven}}^2 = 0{,}5278;$$
$$K_{\text{Експерти}}^2 = 0{,}489042;$$
$$K_{\text{ШІ}}^2 = 0{,}498294;$$
$$K_{\text{УМ}}^2 = 0{,}464156.$$

У цьому випадку, рефлекторна прогнозуюча система УМ на 7,4% краще склала прогноз, ніж система штучного інтелекту. І на 4% краще, ніж експерти букмекерських контор.

Використавши розподіл Фішера до отриманих значень, виконано оцінку статистичних гіпотез про перевищення дисперсії інших систем над системою УМ. В зв'язку з невеликою вибіркою прийнято 70% довірчий інтервал. В таблиці 1 наведено розрахунок для всіх команд (критерій K_i^1). Для обох систем гіпотеза підтверджена.

Таблиця 1

Перевірка гіпотези про перевищення середньоквадратичного відхилення, отриманого різними системами над середньоквадратичним відхиленням, отриманим системою УМ з використанням критерію Фішера для всіх команд по всім стадіям ($F_{0,70;63;63}$ =1,141919)

Система	σ_i^2	$F_{n;n} = \frac{\sigma_i^2}{\sigma_{\text{УМ}}^2}$	Підтверджена гіпотеза $\sigma_i^2 > \sigma_{\text{УМ}}^2$
KU Leuven	0,150986	1,298594	Так (0,85)
ШІ	0,136471	1,173756	Так (0,74)
УМ	0,116269	1	-

В таблиці 2 наведено розрахунок для команд, що пройшли до наступної стадії (критерій K_i^2). Гіпотеза була підтверджена лише для бельгійського аналітичного сервісу KU Leuven. А для інших систем вона попадає в 60% довірчий інтервал.

Таблиця 2

Перевірка гіпотези про перевищення середньоквадратичного відхилення, отриманого різними системами над середньоквадратичним відхиленням, отриманим системою УМ з використанням критерію Фішера для команд, що вийшли до наступної стадії ($F_{0,70;29;29}$ = 1,217257)

Система	σ_i^2	$F_{n;n} = \frac{\sigma_i^2}{\sigma_{\text{УМ}}^2}$	Підтверджена гіпотеза $\sigma_i^2 > \sigma_{\text{УМ}}^2$
KU Leuven	0,278573	1,293037	Так (0,75)
ШІ	0,248297	1,152506	Ні (0,65)
Експерти	0,239162	1,110106	Ні (0,61)
УМ	0,215441	1	-

5. Висновки

Потреба в системах прогнозування була і є досить високою. Як показало експериментальне застосування рефлекторної інтелектуальної системи для прогнозування результатів футбольних матчів, теорія несилової взаємодії, яка лежить в основі цієї системи, має значну перспективу в цьому напрямку. Показано, рефлекторна прогнозуюча система по точності прогнозу ЧЄ-2020 перевершує традиційну систему штучного інтелекту, одну з аналітичних систем і ні в чому не поступається експертам у цій області.

Але цей результат вимагає підтвердження на більш значних за обсягом вибірках при проведенні інших турнірів. Автори планують продовжити розпочату роботу.

6. Література

[1] А.Г. Гранберга (и другие.), Статистическое моделирование и прогнозирование. Финансы и статистика, М, 1990.

[2] П. І. Бідюк, Ймовірнісно-статистичні методи моделювання і прогнозування: монографія. В Миколаїв, 2014.

[3] В. М. Синеглазов, Е. И. Чумаченко, В. С. Горбатюк, Интеллектуальные методы прогнозирования. Освіта України, 2013.

[4] С.Д. Штовба, В.В. Вивдюк, Прогнозирование результатов футбольных матчей на основе нечетких правил.URL: http://shtovba.vk.vntu.edu.ua/file/ bc407aa9b5afed2e9e9a6c8e5a6c6ab7.pdf

[5] Ю.Н. Тесля, Введение в информатику Природы, Черкассы,: McLaut, 2010.

[6] Teslia, I., Pylypenko, V., Popovych, N., Chornyy, O. The Non-Force Interaction Theory for Reflex System Creation with Application to TV Voice Control. 6th International Conference on Agents and Artificial Intelligence (ICAART 2014), LERIA, France, pp. 288-296.

[7] I. Teslia, Experimental Verification of General Validity of Non-force Interaction Formula, URL: https://assets.researchsquare.com/files/rs-138538/v2_stamped.pdf (06/06/2021 last accessed).

[8] Штучний інтелект назвав шанси збірної України подолати Швецію у 1/8 фіналу. URL: https://zbirna.com/post/iskusstvennyi-intellekt-nazval-shansy-sbornoi-ukrainy-proiti-shveciyu-v-1-8-i-finala

[9] Максим Фомичев. Кто выиграет Евро-2020? Франция – фаворит перед началом плей-офф. Режим доступу: https://superbets.guru/kto-vyigraet-evro-2020-franciya-favorit-pered-nachalom-plej-off/

[10] Повертаємось до футболу. URL: https://www.facebook.com/photo?fbid=268928181669588&set=a.118474746714933

[11] Ставки и коэффициенты на Евро-2020 (2021). Режим доступу: https://legalbet.ru/match-center/tournaments/chempionat-evropi-po-futbolu-2021/)

Оцінювання Рівня Вивченості Суб’єкта Навчання

Валентин Козлов [1], Юрій Козлов [2], Володимир Кобзєв [3], Інна Мощенко [4]

[1] *Національна академія Національної гвардії України, пл. Захисників України 3, Харків, 61166, Україна*
[2, 3, 4] *Харківський національний університет радіоелектроніки, пр. Науки 14, Харків, 61166, Україна*

Анотація
Розглянуто метод, придатний для вирішення завдань експертного оцінювання рівня вивченості суб’єкта навчання.

Ключові слова
Експертний метод, шкали, результати навчання

Assessing the Level of Learning Subject of Training

Valentin Kozlov[1], Yuri Kozlov[2], Volodymyr Kobziev[3] and Inna Moschenko[4]

[1] National Academy National Guard of *Ukraine*, Zakhysnykiv Ukrainy Square 3, Kharkiv, 61166, Ukraine
[2, 3, 4] *Kharkiv National University of Radio Electronics, Ave. 14, Kharkiv, 61166, Ukraine*

Abstract
The method of expert assessment, given to determine the level of study of the subject is considered.

Keywords
Expert method, scales, learning outcomes

1. Вступ

Прийнятим і пропонованим до вжитку методам і системам оцінювання рівня вивченості суб’єктам навчання (СН) притаманні недоліки, пов’язані з використанням різних шкал і збільшенням їх довжини з метою поліпшення якості розрізнення суб’єктів або об’єктів порівняння, відсутністю чітких критеріїв застосування шкал, що обумовлює суб’єктивність оцінок. Доречно відмітити також різноманіття та необґрунтованість процедур оброблення результатів, їх надмірну ”заматематизованість”, що робить більшість із них неприйнятними для застосування у повсякденній діяльності освітян. Це обумовлює актуальність проблеми усунення перелічених вище недоліків і розробити придатний до застосування широким колом працівників системи освіти метод кількісного оцінювання рівня вивченості суб’єкта навчання, як логічне продовження попередніх робіт авторів.

Під винесеним у заголовок терміном “вивченість” будемо розуміти цивільний аналог словникового терміна “виучка (військова)” – показник (бажано інтегральний), що характеризує проміжний або кінцевий результат навчання як рівень засвоєння СН знань, умінь та навичок у певній галузі людської діяльності.

2. Виклад основного матеріалу

Серед більш ніж 80,4 млн англомовних Google-публікацій, вибраних за ключовими словами: learning+pedagogical measurements +scales+knowledge assessment system+expert

EMAIL: kozlov1945ve@gmail.com (A. 1); yurii.kozlov@nure.ua (A. 2); volodymyr.kobziev@nure.ua (A. 3); my1485@ukr.net (A. 4)
ORCID: 0000-0003-4452-3009 (A. 1); 0000-0002-6165-4978 (A. 2); 0000-0002-8303-1595 (A. 3); 0000-0002-2738-0037 (A. 4)

method+are ranging, з декількох десятків перших із них за релевантністю та цитуємостю, заслуговує увагу стаття [1]. В ній розглянуто питання вимірювання якості послуг. Відзначено, що складність вирішення цього питання обумовлена впливом людського фактора, залежністю процесу надання послуг від об'єктивних, суб'єктивних, реляційних та організаційних впливів. Незважаючи на це, існують інструменти, загальними рисами яких є використання анкет, які досліджують різні детермінанти якості обслуговування, та надання "числової" інтерпретації зібраної інформації. Майже всі анкети для збору даних використовують оцінки від 1 до 7 або від 1 до 5.

Зібрані дані потім перетворюються і інтерпретуються як числа. Така нумерація (або скаляризація) спрощує обробку даних для аналітиків, але виключає з розгляду логіку призначення оцінок експертами (оцінювачами), що зумовлює так званий ефект "спотворення" інформації. Ця проблема стимулювала розробку відповідного інструменту під назвою Qualitornetro II, орієнтованого на обробку даних лінгвістичних шкал.

Використання лінгвістичних (вербальних або мовних) шкал накладає багато обмежень в процес опрацювання даних, наприклад, у розрахунок відстані між двома елементами шкали, яку повністю визначено в числовій шкалі. У цьому випадку, щоб "агрегувати" інформацію, що надходить від різних оцінювачів, припускають гіпотезу про однорідне тлумачення мовних термінів шкали кожним експертом.

Метод Qualitornetro II для оцінювання рівня наданої (або передбачуваної) послуги при опитуванні передбачає використання вербальної шкали важливості із запереченнями: ніякий N = Neg(P); дуже низький VL = Neg(VH); низький L = Neg(H); середній M = Neg(M); високий H = Neg(L); дуже високий VH = Neg(VL); ідеальний P = .Neg(N). Алгоритм процедури визначення глобального індикатора очікуваної та наданої якості отримано за допомогою методів та інструментів формальної логіки і нечіткої математики, прийнятних для обробки мовної шкали. Сам алгоритм у статті не розкритий.

Схожий за сутністю підхід [2] передбачає використання чотирибальних шкал порядку для оцінювання рівня вивченості суб'єкта навчання (СН) і нескладного алгоритму обробки результатів. Підкреслимо, що значення оцінок традиційної чотирибальної шкали (ЧШ) виражають словами "незадовільно", "задовільно", "добре" і "відмінно", що свідчить про їх нечисловий характер. Сама шкала є аналогічною наведеній у абзаці вище. У процесі оцінювання експерти (викладач або група викладачів) виставляють бали ЧШ як числову інтерпретацію відповіді (судження) на запитання (твердження) типу "Відповідає чи ні зміст судження смислу твердження?": Так – 5; Скоріше так, чим ні – 4; Скоріше ні, чим так – 3; Ні – 2. У разі утруднень з визначенням відповідного бала (назвемо це "ефектом буріданова віслюка") експерт виставляє оцінку, що дорівнює середині інтервалу між оцінками або середині шкали. Такий підхід забезпечує однорідне тлумачення мовних термінів шкали кожним із експертів. Відмітимо також, що ведення до застосування удосконаленої чотирибальної шкали (УЧШ), яка передбачає використання так званого ноніуса, робить застосування десяти- і дванадцятибальної шкал недоцільним.

Особі, що приймає рішення, рівень вивченості СН доцільно подавати інтегральним показником вивченості (ІПВ) – скаляризованим ("згорнутим") значенням деякого вектора оцінок ознак – часткових показників вивченості (ЧПВ). При цьому, ЧПВ можуть бути як кількісними, так і якісними. Оскільки для випадкових подій (отримання оцінки) та величин (самих оцінок) неможливо встановити закони розподілу, існує так звана нестохастична невизначеність [3]. Це обумовлює необхідність використання методів обробки даних загальної природи [4] і застосування загальної процедури нечіткого виводу з теорії нечітких множин [29], [30]: фаззіфікація, композиція як етап отримання результуючої функції належності (РФН); дефаззіфікація як етап вибору шуканого точного числового значення характеристики положення описаної РФН змінної. Зазвичай, це результат оцінювання знання, уміння, навички за декількома запитаннями (завданнями, задачами тощо) у балах чотирибальної або/та удосконаленої чотирибальної шкали – середнє арифметичне значення в УЧШ. Для вузьких шкал це не суперечить теорії [5].

3. Висновки

Викладене дозволяє запропонувати до застосування метод оцінювання рівня вивченості суб'єкта навчання як послідовність таких процедур: оцінюють значення ЧПВ СН у будь-який спосіб за будь-якою з кількісних шкал або шкал порядку; приводять до УЧШ та усереднюють отримані значення часткових показників; визначають (при необхідності) вагу ЧПВ експертним методом; розраховують інтегральний показник вивченості (ІПВ) суб'єкта навчання ваговим або іншим методом.

Кінцевий результат експертного оцінювання – ранжируваний список СН, який подають відповідній і відповідальній особі для прийняття рішення.

Література References

[1] F. Franceschini, S. Rossetto, “Service qualimetrics: The Qualitometro II method”, Quality Engineering, 12:1, 13-20. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1080/08982119908962552.

[2] В. Є. Козлов, Ю. В. Козлов, В. Г. Кобзєв, І. О. Мощенко, “Метод оцінювання рівня вивченості суб'єкта навчання”, у Інформаційні системи та технології ІСТ-2020: матеріали 9-ї Міжнародної науково-технічної конференції 17-20 листопада 2020 р., Харків: Друкарня Мадрид, 2020. – с. 160-162.

[3] Е. С. Вентцель, Теория вероятностей, Москва,: КноРус, 2010.

[4] А. Пегат, Нечеткое моделирование и управление, Москва, БИНОМ, 2013.

[5] А.И. Орлов, Эконометрика, Ростов-на-Дону: Феникс, 2009.

Использование автоматизированной тестирующей системы при изучении алгоритмов и структур данных

Александр Щербаков [1], Юрий Скорин [2], Ирина Ушакова [3]

[1] *Харьковский национальный экономический университет им. С. Кузнеца, пр. Науки, 9-А, Харьков, 61166, Украина*
[2] *Харьковский национальный экономический университет им. С. Кузнеца, пр. Науки, 9-А, Харьков, 61166, Украина*
[3] *Харьковский национальный экономический университет им. С. Кузнеца, пр. Науки, 9-А, Харьков, 61166, Украина*

Аннотация
С целью повышения эффективности изучения студентами IT-специальностей алгоритмов и структур данных в статье предлагается использование автоматизированных тестирующих систем. Были проанализированы существующие подходы к решению этой задачи, выделено два основных варианта организации учебного процесса с использованием автоматизированных систем тестирования, выявлены преимущества и недостатки каждого из них. Предложено использование системы DOTS, которая, с одной стороны, позволяет быстро и эффективно организовать процесс автоматической проверки решений студентами задач на базовые структуры данных и алгоритмы, а с другой, оставляет преподавателю возможность максимально контролировать это процесс и управлять ним.

Ключевые слова
Алгоритмы и структуры данных, автоматизированная тестирующая система, эффективность учебного процесса

Using an automated testing system in the study of algorithms and data structures

Oleksandr Shcherbakov [1], Yuri Skorin [2], Iryna Ushakova [3]

[1] *Simon Kuznets Kharkiv National University of Economics, Nauki Ave., 9-A, Kharkiv, 61166, Ukraine*
[2] *Simon Kuznets Kharkiv National University of Economics, Nauki Ave., 9-A, Kharkiv, 61166, Ukraine*
[3] *Simon Kuznets Kharkiv National University of Economics, Nauki Ave., 9-A, Kharkiv, 61166, Ukraine*

Abstract
To increase the efficiency of students of IT specialties studying algorithms and data structures, the article proposes the use of automated testing systems. The existing approaches to solving this problem were analyzed, two main options for organizing the educational process using automated testing systems were identified, the advantages and disadvantages of each of them were identified. The use of the DOTS system is proposed, which, on the one hand, allows you to organize the process of automatic verification of student solutions quickly and efficiently to problems for basic data structures and algorithms, and on the other hand, leaves the teacher the opportunity to maximally control this process and manage it.

Keywords
Algorithms and data structures, automated testing system, educational process efficiency

EMAIL: oleksandr.shcherbakov@hneu.net;
skorin.yuriy@gmail.com; Varavina.ira@gmail.com
ORCID: 0000-0001-7555-2545; 0000-0002-4613-3154; XXXX-0000-0001-8315-0917

1. Введение

Стремительный рост и развитие информационных технологий во всем мире, повышение уровня сложности решаемых задач, влекут за собой повышение требований к уровню подготовки специалистов для IT-отрасли. Выпускник IT-специальности университета должен обладать профессиональными компетентностями в сфере информационных технологий, совмещать глубокую и всестороннюю теоретическую подготовку в области компьютерных наук с умением выполнять широкий спектр практических задач в данной области. Согласно международному образовательному стандарту в области компьютерных наук, известному как Computing Curricula, одной из основных фундаментальных дисциплин при подготовке будущих IT-специалистов, является дисциплина «Алгоритмы и структуры данных» [1]. Целью изучения дисциплины «Алгоритмы и структуры данных» является изучение эффективных структур и алгоритмов обработки данных в информационных системах, формирование у студентов компетентностей в проектировании, разработке и анализе алгоритмов, оценке их сложности для адекватного моделирования предметных областей, а также реализации алгоритмов в виде программных и информационных систем. В большинстве университетов в рамках дисциплины «Алгоритмы и структуры данных» изучаются такие темы как алгоритмы сортировки и поиска, алгоритмы на графах, алгоритмы вычислительной геометрии, алгоритмы поиска в строках, комбинаторные и эвристические алгоритмы. Основными структурами данных, рассматриваемыми в этой дисциплине, являются линейные списки, очереди, стеки, деревья и графы [2–4]. Реализация всех вышеперечисленных алгоритмов и структур данных в виде программ с одной стороны позволяет студентам более глубоко понять теоретический материал и получить практические навыки его использования, а с другой – дает преподавателям возможность убедиться, что у студентов сформировались необходимые компетентности. Но при этом актуальным становится вопрос тестирования тех программ, которые разрабатывают студенты. Очевидно, что, так называемое, ручное тестирование является слишком трудоемким процессом и малоэффективно, поэтому все больше преподавателей используют автоматизированные тестирующие системы. Целью данной статьи является анализ существующих подходов и поиск эффективного решения проблемы автоматизированного тестирования программ студентов при изучении алгоритмов и структур данных.

2. Обзор литературы

Идея использования в учебном процессе автоматизированных тестирующих систем не является новой. Во многих учебных заведениях при изучении дисциплин, связанных с алгоритмизацией и программированием, используются различные автоматизированные средства обучения и проверки, о чем свидетельствует большое количество публикаций по данной теме [5–8]. Проведенный анализ показал, что можно выделить два основных подхода к организации автоматизированного тестирования студенческих программ.

Первый – это использование специальных информационных систем, выполняющих автоматическую проверку программ. Одно из определений понятия автоматизированная тестирующая система приведено в [7] и звучит так – это комплекс программных (или программно-аппаратных) средств, которые позволяют в автоматическом режиме проводить тестирование решений пользователей в рамках объекта тестирования. Существует достаточно большое количество таких систем, например, такие как PCMS2, Contester, DOMjudge dudge и другие, но одной из наиболее распространённых является Ejudge [9]. Ejudge - это система для проведения различных мероприятий, в которых необходима автоматическая проверка программ. Несмотря на то, что основное назначение системы Ejudge — это проведение олимпиад и других соревнований по программированию, система нашла также широкое применение в учебном процессе для обучения программированию в различных университетах [8]. Но у такого подхода есть один существенный недостаток. Система

Ejudge не располагает базой задач с тестами, преподаватель должен их разработать самостоятельно. Это обстоятельство существенно усложняет процесс и требует достаточно много времени.

Вторым возможным вариантом применения автоматизированных тестирующих систем для изучения алгоритмов и структур данных, лишенным указанного недостатка, является использование Интернет-ресурсов, рассчитанных на использование в процессе обучения или тренировки уже готовых задач. Типичными представителями таких систем является сайт «Школа программиста» [10] или сайт “E-olymp” [11]. В базе данных каждого из этих двух сайтов находится более тысячи задач самого различного уровня сложности. Методика использования этих ресурсов для подготовки учащихся к олимпиадам по программированию изложена работах [12-14]. Например, на сайте “E-olymp” преподаватель может создавать группы учеников (студентов), приглашать в эти группы участников, которые предварительно уже зарегистрировались на этом сайте, создавать на базе существующих задач различные соревнования, просматривать общий рейтинг участников (по результатам всех соревнований) и рейтинг по результатам отдельных соревнований. Хотя основным назначением сайта является подготовка к олимпиадам, о чем говорит само название сайта, его можно использовать и на занятиях по программированию [14].

Безусловно, наличие готовой баз задач существенно сокращает время на подготовку турнира. Но, в то же время, отсутствие возможности увидеть код каждого решения студента, отсутствие возможности аннулировать решение, если его автор нарушает установленные преподавателем требования, другие ограничения снижает эффективность такого подхода именно в процессе изучения алгоритмизации.

3. Решение проблемы

В Харьковском национальном экономическом университете имени Семена Кузнеца на кафедре информационных систем вот уже несколько лет при проведении лабораторных и практических занятий по дисциплинам «Алгоритмы и структуры данных», «Программирование», «Основы алгоритмизации» используется система DOTS, созданная группой разработчиков под руководством Заслуженного учителя Украины Арзубова Н.А. DOTS (Distributed Olympiad Test System) — это онлайн-платформа для обучения программированию и развития алгоритмического мышления. [15] Система позволяет студентам более эффективно изучать алгоритмизацию и программирование, дает возможность в любой момент отправить свое решение в виде программы на проверку и практически мгновенно получить вердикт проверяющей системы на свое решение.

Широкие возможности система предоставляет и преподавателям. Преподаватель может для каждой дисциплины для каждой группы создать лабораторную или практическую работу (в системе они называются турниры), наполнить ее задачами из базы готовых задач, установить время начала и время окончания и получает максимальный контроль над процессом выполнения студентами заданий. Система поддерживает 28 различных компиляторов для всех широко распространенных языков программирования, таких как C, C++, Java, C#, Python, Kotlin, Haskell, Ruby, JavaScript и др. Интерфейс преподавателя в системе DOTS приведен на рис. 1.

Система имеет большую базу готовых задач, разделенных на три большие группы – задачи для новичков на элементарные алгоритмические конструкции, такие как разветвление или циклы, задачи для изучающих алгоритмы и структуры данных, а также задачи олимпиадного характера. В рамках каждой группы все задачи также разделены по темам. Например, для изучения алгоритмов и структур данных выделены следующие темы – алгоритмы целочисленной арифметики, алгоритмы сортировки, алгоритмы на графах, алгоритмы на деревьях, задачи на динамическое программирование, алгоритмы работы со строками, алгоритмы вычислительной геометрии, комбинаторные алгоритмы и т. д. Каждая задача имеет свой уровень сложности и преподаватель легко и быстро может сформировать комплект заданий для каждой учебной группы и даже для отдельных студентов, с учетом их уровня подготовки в теории алгоритмов. Пример экрана преподавателя с комплектом задач для лабораторной работы по теме «Алгоритмы на графах» приведен на рис.2. При необходимости преподаватель может в любой момент изменить набор задач, добавив в него или дополнительные задачи, или, наоборот, заменить сложную задачу на более легкую.

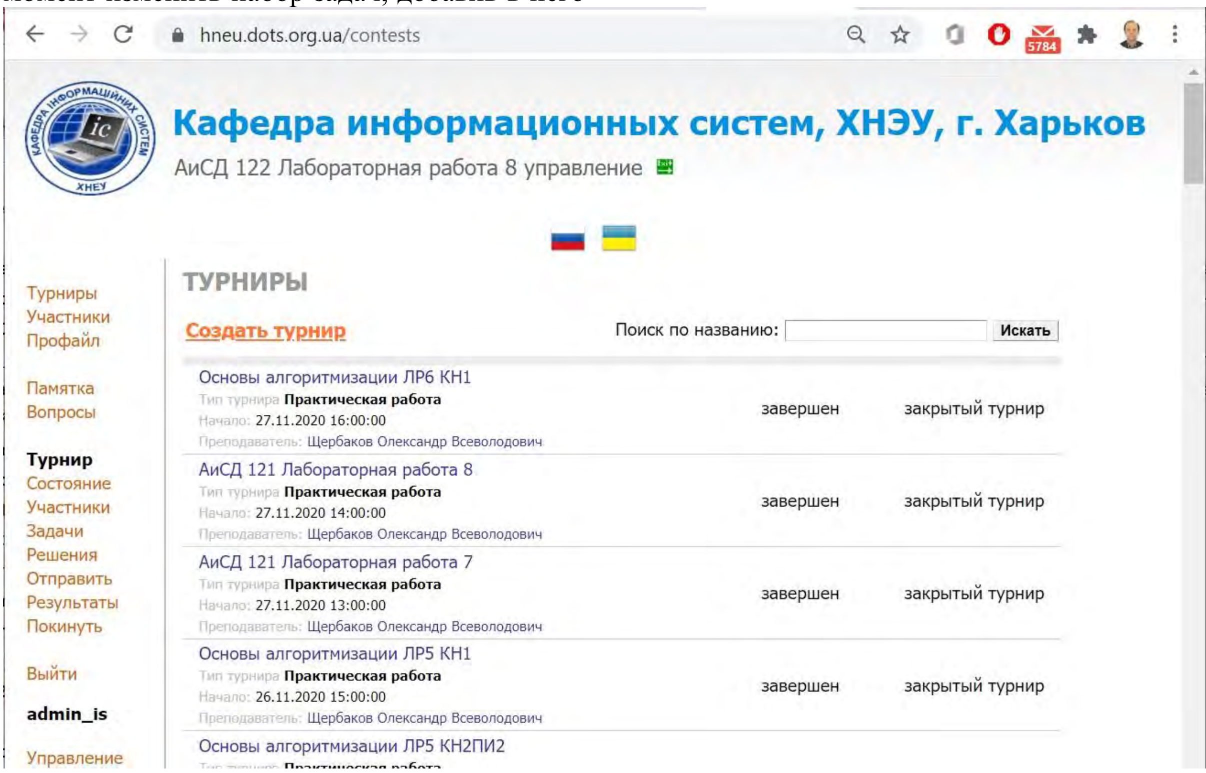


Рисунок 1. Интерфейс преподавателя в системе DOTS

.

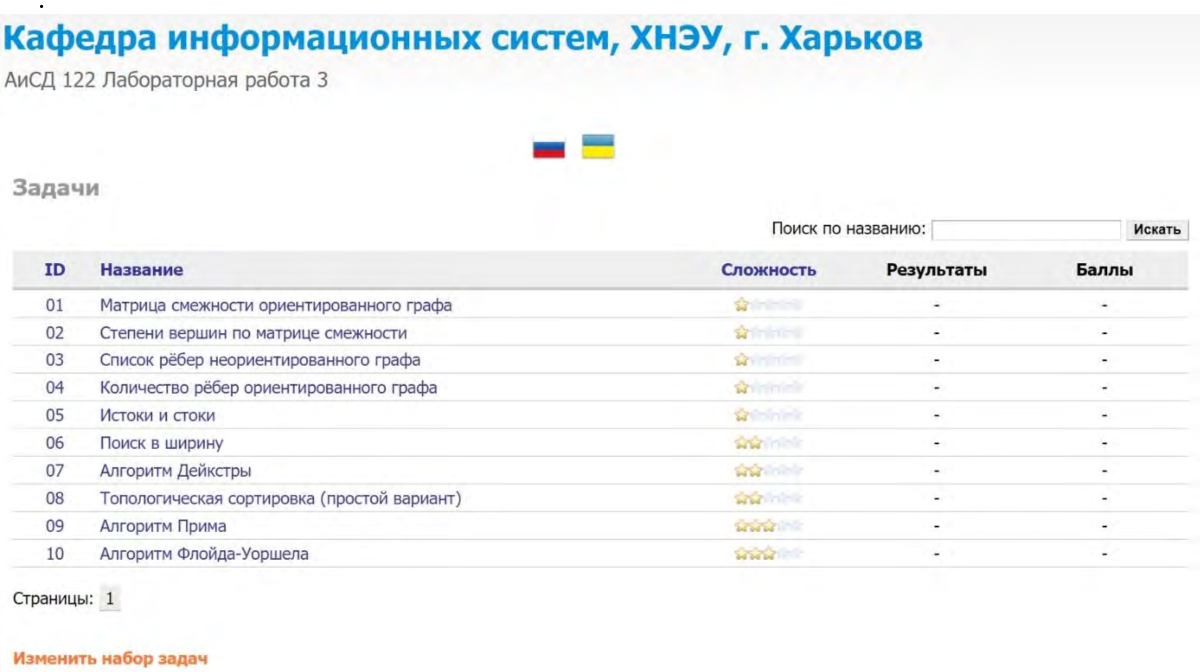


Рисунок 2. Пример экрана преподавателя с комплектом задач

Все задачи по алгоритмам и структурам данных в системе имеют одинаковый формат. Есть четкая формулировка задачи, указан формат входных данных, в котором прописаны все ограничения на входные данные, а также указан порядок ввода исходных данных. Обязательно указывается формат выходных данных, чтобы студент понимал в какой последовательности выводить результаты. Это очень важный момент, так как при автоматической проверке система посимвольно сравнивает результат работы программы студента с результатом эталонного решения и неправильный формат вывода результатов может привести к вердикту WA – Wrong Answer (Неправильный ответ). Кроме того, обязательно указывается ограничение на время выполнения программы и ограничения на количество выделяемой памяти. Таким образом автоматически проверяется насколько эффективным в плане вычислительной сложности является решение студента. Если решение студента выполняется дольше установленного ограничения по времени, система автоматически прекращает дальнейшее тестирование и выдает вердикт TL – Time Limit Exceeded (Превышено время выполнения). А если решение требует больше памяти, чем предусмотрено ограничениями задачи, то система выдает сообщение ML – Memory Limit Exceeded (Превышено ограничение по памяти).

Пример формулировки задачи в системе DOTS приведен на рис. 3.

Задача 06

06: Поиск в ширину

Сложность:

Ограничение по времени:	0.2 секунды
Ограничение по памяти:	64 мегабайта

Поиск в ширину (англ. *breadth-first search, BFS*) – один из методов обхода графа. Пусть задан граф $G = (V, E)$ и выделена исходная вершина s. Алгоритм поиска в ширину систематически обходит все рёбра G для «открытия» всех вершин, достижимых из s, вычисляя при этом расстояние (минимальное количество рёбер) от s до каждой достижимой из s вершины. Алгоритм работает как для ориентированных, так и для неориентированных графов.

Дан ориентированный граф. Найдите расстояние от одной заданной вершины до другой. Расстояние – это количество ориентированных рёбер заданного графа, по которым надо «пройти», чтобы попасть из одной заданной вершины в другую.

Формат входных данных

В первой строке входных данных содержатся три натуральных числа N, S и F ($1 \leqslant N \leqslant 100$, $1 \leqslant S \leqslant N$, $1 \leqslant F \leqslant N$) – количество вершин в графе и номера начальной и конечной вершин соответственно. Далее в N строках задана матрица смежности графа. Если значение в j-м элементе i-й строки равно 1, то в графе есть направленное ребро из вершины i в вершину j.

Формат выходных данных

Выведите одно число – минимальное расстояние от начальной вершины до конечной. Если пути не существует, выведите 0.

Пример

тест	ответ
5 5 3 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 0 0 1 0 0 1 1 0	1

Рисунок 3. Пример условия задачи

Для каждого турнира система автоматически формирует рейтинговую таблицу, где указывается количество баллов, набранных каждым студентом группы за каждую задачу, общее количество баллов, набранное студентом за все задачи, а также полная статистика по группе в целом. Пример итоговой таблицы с результатами выполнения студентами одной группы лабораторной работы с помощью автоматизированной тестирующей системы приведен на рис. 4. Преподаватель видит, какой средний балл в группе по каждой задаче, сколько человек пыталось решать каждую задачи и сколько из них решили эту задачу на полный балл.

Количество максимальных баллов за задачу можно устанавливать для каждой задачи отдельно и это количество может быть произвольным. Таким образом можно настроить систему оценивания под любую шкалу оценок.

Большим преимуществом данной системы для преподавателя является возможность просмотреть исходный код всех решений, которые студенты сдавали на проверку. При этом, даже если студент отправил повторно свое решение по одной и той же задаче, в системе хранятся все его предыдущие решения. В системе предусмотрен режим, в котором оценка выставляется не автоматически, а только после того, как преподаватель посмотрел исходный код программы и принял это решение. Таким образом можно контролировать не только соответствие результата эталонному, но и проверять, насколько студент выполнил дополнительные требования к этой задаче, например, реализовал именно тот алгоритм или способ решения задачи, который определил преподаватель. Кроме того, если преподаватель пришел к выводу, что студент выполнял задание не самостоятельно или серьезно нарушил требования к решению, то такое решение преподаватель может аннулировать.

Кафедра информационных систем, ХНЭУ, г. Харьков

Основы алгоритмизации ЛР5 КН1

Результаты

Место	Участник	01	02	03	04	05	06	07	08	09	10	Всего
1-5.	Авраменко Леон	100	100	100	100	100	100	100	100	100	100	1 000
1-5.	Баєв Вадим	100	100	100	100	100	100	100	100	100	100	1 000
1-5.	Клос Максим	100	100	100	100	100	100	100	100	100	100	1 000
1-5.	Петрушина Діана	100	100	100	100	100	100	100	100	100	100	1 000
1-5.	Пирог Данило	100	100	100	100	100	100	100	100	100	100	1 000
6.	Копитін Алім	100	100	100	100	100	100	100	100	64	100	964
7-9.	Нестеренко Данило	100	100	100	100	100	100	57,60	100	80	100	937,60
7-9.	Фомічов Максим	100	100	100	100	100	100	57,60	100	80	100	937,60
7-9.	Цатурян Ігор	100	100	100	100	100	100	57,60	100	80	100	937,60
10.	Цупін Григорій	100	100	100	100	100	100	100	100	80	32	912
11.	Лактіонов Артем	100	100	100	100	100	22,40	57,60	100	73,60	100	853,60
12.	Почтарьова Юлія	100	100	86,40	100	100	16	57,60	100	80	100	840
13.	Пилипенко Дмитро	100	100	100	100	100	35,20	57,60	51,20	73,60	100	817,60
14.	Польовик Ілля	100	72,20	86,40	100	100	16	100	100	80	32	786,60
15.	Щербаков Сергій	100	100	100	100	100	16	57,60	100	80	32	785,60
16.	Дерев`янко Євгеній	100	100	76,80	100	100	100		100		32	708,80
17.	Нечепуренко Вероніка	100	72,20	100	0	67,20	100	57,60	100	80	25,60	702,60
18.	Половніков Ярослав	100	100	100	100	100			100		100	700
19.	Онопрієнко Алік	100	72,20	0	100	81,60	22,40	67,20	100	80	0	623,40
20.	Кушнарьов Даніель	100	100	86,40		100		100	100			586,40
21.	Кондратов Максим	100		100		67,20	100		100	80		547,20
22.	Діденко Дмитро	100	100	100	0	100			76,80			476,80
	Итоги	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**Всего**
	Средний балл	100	96,03	92,55	90	96,18	75,16	79,33	96,73	83,96	76,51	823,52
	Пытались решить	22	21	22	20	22	19	18	22	18	19	203
	Решили полностью	22 (100%)	18 (85%)	17 (77%)	18 (90%)	19 (86%)	13 (68%)	9 (50%)	20 (90%)	5 (27%)	13 (68%)	154 (75%)

Рисунок 4. Итоговая таблица с результатами студентов

4. Выводы

В статье было обосновано использование автоматизированных тестирующих систем при изучении алгоритмов и структур данных. Были проанализированы существующие подходы и выявлены их преимущества и недостатки. В качестве автоматизированной тестирующей системы для организации проверки решений студентов при изучении дисциплины «Алгоритмы и структуры данных» было предложено использовать систему DOTS, которая, по мнению авторов, является наиболее эффективным инструментом для достижения поставленной цели. К преимуществам данной системы нужно отнести наличие большой базы задач самого различного уровня сложности, начиная от самых простых и до задач олимпиадного характера, широту охвата тем по алгоритмам и структурам данных, лёгкость и удобство создания новых турниров, наполнения их задачами и добавления в эти турниры студентов, возможность автоматической проверки всех задач и формирования итоговых таблиц с результатами этой проверки, возможность для преподавателя контролировать все этапы выполнения студентами задания, просматривать исходный код программ, отправленных на проверку и управлять всем процессом. Опыт применения системы DOTS на кафедре информационных систем ХНЭУ им. С.Кузнеца позволяет со всей уверенностью утверждать, что использование автоматизированной тестирующей системы при изучении алгоритмов и структур данных в итоге дает возможность эффективно формировать компетентности выпускников в проектировании, разработке и анализе алгоритмов, а также реализации алгоритмов в виде программных и информационных систем.

5. Литература

[1] Computing Curricula 2020: Paradigms for Global Computing Education // URL: https://dl.acm.org/doi/book/10.1145/3467967.

[2] Meleshko Ye.V., Yakymenko M.S., Polishchuk L.I. Alhorytmy ta struktury danykh: Navchalnyi posibnyk dlia studentiv tekhnichnykh spetsialnostei dennoi ta zaochnoi formy navchannia. – Kropyvnytskyi: Vydavets – Lysenko V.F., 2019. – 156 s.

[3] Tkachuk V.M. Alhorytmy i struktura danykh: Navchalnyi posibnyk / V.M.Tkachuk. - Ivano-Frankivsk: Vydavnytstvo Prykarpatskoho natsionalnoho universytetu imeni Vasylia Stefanyka, 2016.- 286 s.

[4] Korotieieva T.O. Alhorytmy ta struktury danykh: navch. posibnyk / T.O. Korotieieva. – Lviv: Vydavnytstvo Lvivskoi politekhniky, 2014. – 280 s.

[5] Zubkov O. V., Semicheva N. L. Rol' avtomatizirovannyh testiruyushchih sistem v processe formirovaniya navykov algoritmicheskogo myshleniya // Pedagogicheskij IMIDZH. 2019. № 4 (45). S. 550–565.

[6] Ajdarov K. A. Balakaeva G.T. Primenenie testiruyushchej onlajn sistemy dlya razvitiya prakticheskih navykov obuchayushchihsya v programmirovanii // Materialy 48-oj nauchno-metodicheskoj konferencii «Rol' vysshih uchebnyh zavedenij v modernizacii obshchestvennogo soznaniya: perekhod k modeli «Universitet 4.0»». 18-19 yanvarya 2018 g. – Almaty: Kaz. un-t, 2018. – s. 17–20.

[7] Gasparyan A. V., Timoshina N. V. Osobennosti avtomatizacii proverki zadach po programmirovaniyu // ITportal. 2018. № 2 (18). URL: http://itportal.ru/science/tech/osobennosti-avtomatizatsii-proverki/.

[8] Gorchakov L. V., Stas' A. N., Kartashov D. V. Obuchenie programmirovaniyu s ispol'zovaniem sistemy Ejudge // Vestnik Tomskogo gosudarstvennogo pedagogicheskogo universiteta. 2017. Vyp.9 (186). S. 109–112.

[9] eJudge – sistema avtomaticheskoj proverki programm. URL: https://ejudge.ru/.

[10] SHkola programmista. URL: https://acmp.ru/

[11] Internet-portal orhanizatsiino-metodychnoho zabezpechennia dystantsiinykh olimpiad z prohramuvannia dlia obdarovanoi molodi. - URL: https://www.e-olymp.com.

[12] Alekseev A.V. Podgotovka shkol'nikov k olimpiadam po informatike s ispol'zovaniem vebsajta: uchebno-metodicheskoe posobie dlya uchashchihsya 7-11 klassov. / A.V. Alekseev, S.N. Belyaev. – Hanty-Mansijsk: RIO IRO, 2008. – 284 s.

[13] Zhukovskyi S.S. “E-olimp” – systema avtomatychnoi perevirky zadach ta provedennia olimpiad z informatyky v interneti /Kompiuter u shkoli ta simi. – №1 (65). – 2008. – S.48-50.

[14] Zhukovskyi S. S. Vykorystannia Internet-portalu orhanizatsiino-metodychnoho zabezpechennia "E-OLIMP" dlia pidhotovky obdarovanykh shkoliariv do olimpiady z informatyky / S. S. Zhukovskyi // Kompiuter u shkoli ta simi - 2010. - № 8. - S. 47-48.

[15] DOTS — onlajn-platforma dlya obucheniya programmirovaniyu. URL: https://dots.org.ua/

Нові підходи до створення систем запобігання витоку мовної інформації за межі контрольованої зони

Нужний Сергій Миколайович

Національний університет кораблебудування імені адмірала Макарова, пр. Героїв України 9, м. Миколаїв, 54007, Україна

Анотація

В Україні та багатьох інших країнах білий шум або його клони використовуються для активних систем постановки завад в каналах витоку мовної інформації, що дозволяє зловмисникам перехоплювати конфіденційну інформацію. Запропоновано структуру системи захисту мовної інформації на базі генератора завад скремблерного типу. Перехід в системах захисту мовної інформації до такої структури дозволяє відмовитися від застарілого, малоефективного в сучасних умовах, енергетичного зашумлення мовної інформації і перейти до більш продуктивного комплексного методу – енерго-інформаційного маскування. Аналіз такого типу завад показує їх високу стійкість до сучасних методів математичної обробки цифрових фонограм, фільтрації завад, реінжинірингу та виділення голосів дикторів. На основі використання методів системного аналізу (абстрагування, експертної оцінки, формалізації та ін.) показано високу ефективність такої системи захисту. Уточнено вимоги до процесу формування основних критеріїв для систем постановки активних завад і їх основних компонентів, а також обмежень у організації і технології їх використання.

Ключові слова

Захист мовної інформації, системи постановки активної завади, мовоподібний сигнал

New approaches to development of systems which prevent speech information from leaking outside the controlled area

Sergey Nuzhniy

Admiral Makarov National University of Shipbuilding, 9 Heroiv Ukrainy ave., Mykolayiv, 54007, Ukraine

Abstract

In Ukraine and many other countries, white noise or its clones are used for active jamming systems for speech information leakage channels, which allows attackers to intercept confidential information. The structure of the system of speech information protection on the basis of the noise generator of scrambler type is offered. The transition in the systems of protection of speech information to such a structure allows to abandon the outdated, inefficient in modern conditions, energy noise of speech information and move to a more productive comprehensive method – energoinformational masking. Analysis of this type of masking shows their high resistance to modern methods of mathematical processing of digital phonograms, interference filtering, reengineering and voice selection of speakers. Based on the use of system analysis methods (abstraction, expert evaluation, formalization, etc.), the indifference of the active interference signal of such a protection system to the considered filtration methods is shown. The requirements to the process of formation of the main criteria for systems of active interference and their main components, as well as limitations in the organization and technology of their use are specified.

Keywords

Protection of speech information, active interference systems, speech-like signal

EMAIL: s.nuzhniy@gmail.com
ORCID: 0000-0002-7706-0453

1. Вступ

Мова є найбільш уживаним засобом обміну інформацією. Особливе значення вона набуває при створенні та обговоренні нових та діючих проектів. При цьому, в багатьох випадках, озвучується інформація обмеженого доступу – конфіденційна (конструкторська, економічна, організаційна, know-how та інша), а в певних випадках і таємна інформація. Таким чином, особливого значення набуває її захист.

Існуючі в Україні, та й в більшості країн світу, методи та технології проектування систем захисту мовної інформації засновані на використанні систем постановки активних завад в акустичному та вібраційному каналах витоку інформації. Основою таких систем є генератор сигналу завади. Згідно нормативних документів, в Україні це виключно генератори «білого» шуму або його «кольорові» клони. Однак, на даний час, таке рішення є застарілим і не забезпечує необхідного рівня захищеності інформації.

В роботі розглянуто особливості побудови та функціонування систем захисту мовної інформації (СЗМІ) та запропоновано рішення, яке дозволяє розробляти більш ефективні системи захисту мовної інформації.

2. Аналіз літературних джерел та постановка проблеми

Технології виявлення та розпізнавання мови диктора в останні десятиріччя отримала новий імпульс. Це зумовлено декількома факторами:

- Широким використанням цифрових технологій для запису та обробки фонограм
- Використання цифрових фонограм в системах ідентифікації та контролю доступу
- Перехід на цифрові технології в радіозв'язку та телефонії
- Використання цифрових методів та технологій фільтрації завад та покращення якості цифрових фонограм

Найбільш актуальними загрозами є:

- Можливість виявлення зловмисником наявності мови в фонограмі
- Фільтрація зловмисником сигналу завади та розпізнавання мови диктора (отримання несанкціонованого доступу до лінгвістичної складової фонограми)
- Ідентифікація зловмисником дикторів за особливостями їх мовлення по фонограмі/перехопленій розмові

Аналіз літературних джерел показує значний прогрес в розвитку методів та технологій виявлення та розпізнавання мови диктора при наявності фонових шумів.

В [1] наведені результати порівняльного аналізу можливості виявлення та відновлення мови диктора на фоні мовних завад, що створені іншими дикторами (завади типу «мовний хор» / «cocktail party»). В роботі показана різниця в спектральному складі англійської та китайської (Mandarin Chinese) мов. Зазначається, що в англійській мові превалюють голосні приголосні, а в китайській – глухі приголосні. В роботі зроблено висновок, що для китайської мови доцільніше використовувати енергетичне маскування. В роботі використано метод просторового розділення джерел сигналу, що дав позитивний результат. Однак метод має ряд суттєвих обмежень і тому потребує подальших досліджень.

Метод просторового розділення в [2] використано для дослідження величини внеску двох базових складових, які змінюються в залежності від перерозподілу джерел звуку – бінаурального конверту (ENV) та тимчасової тонкої структури (TFS). Встановлено, що сигнали ENV є важливими для розбірливості мови, однак бінауральні сигнали TFS мають вирішальне значення для перцептивно відокремлення різних співрозмовників і вирішення проблеми «cocktail party». Обмеженнями такого дослідження є розміщення основного та маскувального джерел сигналу. Основне джерело завжди розміщувалось навпроти приймача сигналу, а джерело-маска – переміщувалось по азимуту. Це дало змогу використати ефект парцептивності. Для систем захисту мовної інформації від витоку акустичними та вібраційними каналами таке розділення можливе тільки за умови використання багатоканальних систем несанкціонованого перехоплення мовної інформації.

Вплив адитивного білого гаусового шуму та високочастотної фільтрації на розбірливість мови при співвідношеннях сигнал/шум (signal-to-noise ratios – SNR) від – 26 до 0 дБ розглянуто в [3]. Оцінено вісім об'єктивних показників:

- короткочасна об'єктивна розбірливість

(short-time objective intelligibility – STOI),

- запропонований варіант, названий STOI+,
- розширена короткочасна об'єктивна розбірливість (extended short-time objective intelligibility – ESTOI),
- нормована метрика коваріації (normalised covariance metric – NCM),
- нормована метрика кореляції обвідної піддіапазону (normalised subband envelope correlation metric – NSEC),
- дві похідні метрики з індексу розбірливості зв'язаної мови (CSII)
- індекс передачі мови з методом регресії на основі конверта (STI).

В [3] показано, що нормалізація та вирізання збільшують похибку прогнозування STOI та зменшують надійність метрики, коли мова змішується з білим гаусовим шумом при низьких глобальних SNR

В [4] порівнюються два алгоритми на основі масок, які застосовуються до мови, змішаної з білим гаусовим шумом (white Gaussian noise – WGN) при низьких показниках SNR (від -29 до -5 дБ). Вони включають ідеальну двійкову маску (Ideal Binary Mask – IBM) з локальним критерієм, встановленим на 0 дБ, та маску ідеального співвідношення (Ideal Ratio Mask – IRM). Продуктивність короткочасної об'єктивної розбірливості (STOI) та варіанта STOI (званої STOI+) порівнюють з показниками інших монометричних показників розбірливості (зрозумілості), які можна використовувати до та після обробки на основі маски. Результати показують, що IRM можуть бути використані для отримання майже максимальної розбірливості мови (> 90 для матеріалу речення) навіть при дуже низьких значеннях SNR. Показано, що широко використовуваний IBM з LC = 0 дБ (local SNR criterion – LC) працює відносно погано для WGN та SNR при – 17 дБ і нижче. Результати IRM показали значне покращення розбірливості порівняно з IBM з LC = 0 навіть при надзвичайно низьких рівнях SNR (до – 29 дБ).

Таким чином, як видно з наведених джерел, розробка методів фільтрації шумової завади та відновлення мовного сигналу, в останні роки, набули поширення та застосовуються для вирішення різних задач. Однак це створює додаткові можливості для зловмисників та робить надзвичайно актуальною задачу розробки СЗМІ, стійких до таких методів фільтрації шумової завади та реінжинірингу.

3. Мета та задачі дослідження

Метою дослідження є розробка нової концепції створення СЗМІ від витоку акустичними та вібраційними каналами за межі контрольованої зони з урахуванням об'єктивних та суб'єктивних факторів.

Для досягнення мети в роботі були поставлені наступні завдання:

1. Сформувати перелік об'єктивних та суб'єктивних факторів, що впливають на ефективність роботи СЗМІ
2. Запропонувати нові принципи побудови системи постановки активних завад та її вузлів

4. Аналіз факторів, що впливають на ефективність роботи системи захисту мовної інформації

Ефективність роботи СЗМІ залежить від багатьох факторів, які можна розділити на дві групи – об'єктивні та суб'єктивні.

До об'єктивних факторів віднесемо ті, що не залежать від проектувальника і власника СЗМІ. Відповідно, суб'єктивними будуть фактори, які пов'язати тим чи іншим чином з бажаннями, кваліфікацією та можливостями проектувальника і власника СЗМІ.

4.1. Об'єктивні фактори

Проектування та експлуатація СМЗІ у всіх розвинених країнах відбувається в відповідності до нормативного законодавства даної країни, методів, методик та технологій, затверджених відповідними керуючими органами. Відхилення від вимог є суттєвим порушенням і має «жорсткі» реакції та наслідки – тимчасова заборона на роботу об'єкту, заборона на виконання певних функцій, адміністративна та кримінальна відповідальність керівництва об'єкту.

До об'єктивних факторів відносяться:

- Мета та задачі СЗМІ
- Структура СЗМІ
- Тактико-технічні характеристики СЗМІ

4.2. Суб'єктивні фактори

Контроль за працездатністю СЗМІ в виділених приміщеннях (ВП) є задачею спеціально виділеного підрозділу або окремого співробітника. При цьому, в більшості випадків, їх доступ в ці приміщення є суттєво регламентованим, а співробітники, що працюють в ВП не мають відповідної кваліфікації.

Таким чином можна виділити наступні суб'єктивні фактори, які мають суттєвий вплив на працездатність СЗМІ:

- Проектувальники системи:

1. Вибір обладнання, яке встановлюється (безпосередньо тип та марка обладнання, його колір, схема включення та інше).

Аналіз показує, що в більшості випадків є можливість вибору із встановленого переліку обладнання – тобто проявляється певна суб'єктивність проектувальника

2. Кваліфікованість та компетентність проектувальника, які проявляється при оптимізації структури СЗМІ та виконання певних вимог замовника робіт.

Досвіт показує, що при створенні комплексних систем захисту інформації, зазвичай є певний надлишок кількості та можливостей певних елементів обладнання. Замовник може вимагати змінити певні можливості обладнання (наприклад, може вважати, що система акустоконтролю рівня звукового тиску в ВП є засобом підслуховування та інше)

3. Забезпечення паритету між ергономікою ВП і вимогами нормативних документів до СЗМІ.

- Співробітники, що працюють в ВП:

1. Порушення вимог персоналом ВМ організаційних заходів при виконанні технологічного процесу

В [5] наведено приклад проведення спецоперації АНБ (США) по перехопленні мовної інформації з офісних приміщень іншої спецслужби. Основним каналом витоку інформації стало дистанційне знімання інформації з відкритих вікон. За результатами проведеної спецоперації в США прийнято ряд нормативних документів, що суттєво підвищують вимоги до систем безпеки в державних установах.

2. Частина персоналу ВП, зазвичай, має низький рівень знань та компетентностей з інформаційної безпеки:

2.1. Відсутність систематизованих знань (рівень підготовки – «любитель», тобто знання отримані з кінофільмів та белетристичної літератури) призводить до стану «я все знаю!» та «це все маячня!» (особливо керівництво організацій).

Так, наявність мікрофону в системі автоматичного коригування рівня вихідного сигналу віброакустичної завади (пристрій „Топаз ГША-4МК" / Топаз – Комфорт) був сприйнятий керівником однієї з державних установ як пристрій підслуховування. Він наказав його видалити, а так як рівень фону при роботі системи після цього заважав йому, то систему просто вимкнули.

2.2. Залучення до робіт в ВП некваліфікованого персоналу

Термін «інсайдерська не навмисна загроза» в більшості випадків означає залучення до виконання низько кваліфікованих робіт в ВП некваліфікованого персоналу (прибиральники, електрики, санітари тощо) та сторонніх осіб. Це може призвести як до «випадкового» вимкнення окремих блоків системи безпеки (вимкнення живлення, пошкодження ліній підключення випромінювачів та інше) так і розбалансування / пошкодження пристроїв. Тобто СЗМІ не зможе виконувати свої функції.

- Співробітники спеціально виділеного підрозділу (наприклад, відділу захисту інформації – ЗІ) або відповідальний співробітник (група ЗІ):

1. Недостатня кваліфікація співробітників

В відповідності до вимог ДССЗЗІ України [6], кадрове комплектування відділів/груп ЗІ повинно відбуватись виключно спеціалістами, що мають профільну освіту. Однак, в багатьох організаціях та державних установах, посилаючись на недофінансування та відсутність бажаючих, відповідні посади займають співробітники без відповідної фахової освіти

2. Неналежне виконання співробітниками своїх обов'язків

Аналіз показу, шо в багатьох організаціях та державних установах основним навантаженням співробітників відділів/груп ЗІ є супровід роботи комп'ютерної та оргтехніки, а не проведення заходів по забезпеченню захисту інформації на об'єкті

5. Варіант побудови системи постановки активних завад та її вузлів

Аналіз сучасних методів та технологій виявлення і відновлення мови диктора при значних рівнях шумових завад (як типу білий (гаусовий) шум [3,4,7] та і при використанні завад типу «мовний хор» / «cocktail party» [1,2]) показує їх високу ефективність. Одночасно, це показує низьку ефективність методів та технологій, які є нормативними на даний час, а також недоліки методів визначення рівня захищеності мовної інформації [7–9].

Таким чином, для створення дійсно ефективної системи необхідно використати обмеження, які є в розглянутих метода виявлення, фільтрації завади та відновлення мовного сигналу.

5.1. Альтернатива для білого (гаусового) шуму

Перевагами білого шуму є рівномірність спектральної щільності на всьому діапазоні частот. При цьому самі частоти спектру є випадковими, тобто постійно змінюються. Це стало основною перевагою білого шуму та сприяло його використанню в системах постановки завад.

Однак, при використанні вейвлет-фільтрації [7–9,11], двійкових масок (IBM, IRM) [3,4] та інших методів, заснованих на видаленні випадкових імпульсних сигналів, це і є його головною вадою. Для аналізу використовують віконне перетворення (по октавним смугам) на середньотермінових відрізках часу, співставних з тривалістю фонем/алофонів в людській мові – 0,05…0,25 с.

Таким чином, використання сигналу, який по спектральній щільності був би близьким до білого шуму та створеного на базі модифікованих блоків мови диктора, дає змогу отримати заваду, до якої вказані методи фільтрації є індиферентними.

В загальному вигляді, це є перехід від енергетичного маскування мовного сигналу до інформаційно-енергетичного маскування.

Прикладом такого сигналу маскування є завада «скремблерного» типу, синтез та використання якої розглянуто в [7–10].

5.2. Альтернатива для «мовного хору» / «cocktail party»

Головним принципом, який використовується при синтезі сигналу завади типу «мовний хор» / «cocktail party», є інформаційне маскування. Його принцип заснований на властивостях людського слуху виділяти основного диктора із «мовного хору» при SNR ≥ 0 дБ. Якщо співвідношення SNR знаходиться в діапазоні -5 дБ $\leq$ SNR ≤ 0 дБ, то відбувається різке погіршення розбірливості мови (сприйняття аудитором мовної інформації). Подальше зменшення SNR призводить до повного маскування мови диктора [12,13].

На практиці розрізняють два типи таких систем [7.8]:

1. «Мовний хор» на основі мови диктора (дикторозалежні системи)
2. Залучення фонограм сторонніх дикторів (дикторонезалежні системи)

В першому випадку в системах використовують різні методи маніпуляції з мовним сигналом диктора [7,8]. В більшості випадків використовують тимчасову та/або частотну реверберацію [14]. Такий підхід суттєво підвищує рівень маскування та стійкість сигналу завади до фільтрації. Це є основною перевагою цього методу маскування. Однак, як видно з [14] (див. рис. 6 «Спектри потужності сигналу x(t), шуму $n_1(t)$ та суміші y(t) для SNR_0=12 дБ»), використання таких методів має ряд суттєвих недоліків, основним з яких є можливість проведення реінжинірингу (наприклад, процедури сліпої дереверберації [15]).

Для другого типу систем використовують:

- Набори фонограм, заготовлених заздалегідь.

При виготовленні фонограм можуть бути задіяні співробітники організаціє, втому числі й ті, що працюють в ВП. В цьому випадку зловмисник повинен визначитись з інформаційною цінність перехоплених сигналів. Недоліком методу є обмеженість кількості фонограм, що призводить до їх повторного використання, а отже й зниження рівня захищеності інформації.

- Використовують сигнали з радіо- та телеканалів.

Недоліком даного методу є доступ зловмисника до джерел сигналу.

Головними методами реінжинірингу для другого типу систем є експертні системи [1,2], в тому числі й з використанням нейронних мереж [12,16].

Отже, можна констатувати, що синтез шумової завади за методами «мовний хор» / «cocktail party» створює значні труднощі для зловмисника. На даний час існує декілька методів фільтрації та відновлення мовного повідомлення, втім всі вони потребують значної обчислювальної потужності. Цей показник зазвичай був вирішальним при визначенні можливості доступу зловмисника до інформації, однак поява швидкісного інтернету та майнінгових ферм фактично знівелювало його вагомість.

Вирішенням проблеми є створення для експертної системи конкурентної невизначеності в наслідок наявності великої кількості середньотривалих фонемних блоків створених з мови диктора за деякий попередній проміжок часу.

Такі характеристики має запропонована в [9] система захисту мовної інформації на основі генератора мовоподібного сигналу завади скремблерного типу.

5.3. Особливості побудови системи постановки активних завад та її вузлів

При створенні системи постановки активних завад необхідно врахували наступні чинники:

- Наявність в ВП системи озвучування приміщення

Такі системи встановлюють в кімнатах проведення перемовин, актових залах, тощо. Наявність такої системи дозволяє інтегрувати її з системою постановки активних завад. Структура такої системи представлена в [9].
В разі її відсутності постає задача узгодження структури та місць розміщення її
елементів в приміщенні. При цьому набуває значення наступний чинник.

- Психологічний фактор

Постійне акцентування уваги на питаннях кіберзлочинності, а саме на можливості використання мікрофонів та камер смартфонів для шпіонажу за власниками, створили ефект необґрунтованої підозрілості у багатьох людей. Це призводить до неадекватних реакцій на малогабаритні мікрофони в системах захисту мовної інформації (наприклад, в [9] блок контролю рівнів небезпечних сигналів).

Вирішенням цієї проблеми може бути:

1. Використання прихованих мікрофонів, які встановлені безпосередньо в корпусі СЗМІ
2. Використання мікрофону (спікеру), який вбудований в систему внутрішнього зв'язку
3. Використання акустичних і вібраційних вимірювачів небезпечних сигналів на межі ВП (контрольованої зони) [9]

При використанні першого та другого варіантів можливе виникнення інцидентів, пов'язаних з можливістю доступу до таких приладів сторонніх людей. Третій варіант надає більший рівень захищеності, однак має суттєво більшу вартість.

6. Висновки

На основі використання методів системного аналізу (абстрагування, експертної оцінки, формалізації та ін.) набули подальшого розвитку методи побудови перспективних систем протидії витоку мовної інформації, яка озвучується на об'єктах інформаційної діяльності, акустичними та вібраційними каналами, що дозволило сформувати перелік об'єктивних та суб'єктивних факторів, що впливають на ефективність роботи СЗМІ на основі систем постановки активної завади.

Запропонувати нові принципи побудови системи постановки активних завад та її вузлів шляхом використання в системі постановки активних акустичних та вібраційних завад генераторів мовоподібного сигналу скремблерного типу, що дає змогу розробляти системи захисту мовної інформації, стійких до її відновлення методами фільтрації та реінжинірингу.

7. Література References

[1] Xianhui, Wang, and Li Xu «Speech perception in noise: Masking anunmasking». Journal of Otology 16 (2021), 109 – 119. Doi:10.1016/j.joto.2020.12.001

[2] Jayaganesh Swaminathan, Christine R. Mason, Timothy M. Streeter, Virginia Best, Elin Roverud and Gerald Kidd Jr. «Role of Binaural Temporal Fine Structure and Envelope Cues in Cocktail-Party Listening». Journal of Neuroscience 36 (31) (2016), 8250 – 8257 Doi:10.1523/JNEUROSCI.4421-15.2016

[3] Simone Graetzer and Carl Hopkins, «Intelligibility prediction for speech mixed with white Gaussian noise at low signal-to-noise ratios», The Journal of the Acoustical Society of America 149 (2021), 1346 – 1362 Doi:10.1121/10.0003557

[4] Graetzer, Simone & Hopkins, Carl. «Comparison of ideal mask-based speech enhancement algorithms for white noise and low mixture signal-to-noise ratios», Proceedings of the 23rd International Congress on Acoustics 9 to 13 September 2019 in Aachen, Germany. Doi: 10.18154/RWTH-CONV-239141

[5] Snouden, E. «Lichnoe delo». Moskva: Eksmo, 410. (2020). Available at: https://readli.net/edvard-snouden-lichnoe-delo/

[6] Nakaz Administratsiyi Derzhspetszv'yazku vid 19.06.2015 roku № 023

[7] Blintsov, V., Nuzhniy, S., Parkhuts, L., Kasianov, Y. «The objectified procedure and a technology for assessing the state of complex noise speech information protection». Eastern-European Journal of Enterprise Technologies, 5 (9 (95)), (2018). 26 – 34. Doi: 10.15587/1729-4061.2018.144146

[8] Nuzhnyy, S. M. «Improved technology of evaluation of stage of protection of license information». Modern Information Security, 1 (33), (2018). 66–73. Available at: http://journals.dut.edu.ua/index.php/ dataprotect/article/view/1796

[9] Blintsov, V., Nuzhniy, S., Kasianov, Y., Korytskyi, V. «Development of a mathematical model of scrambler-type speech-like interference generator for system of prevent speech information from leaking via acoustic and vibration channels». Technology Audit and Production Reserves, 5 (2 (49)), (2019). 19–26. Doi: 10.15587/2312-8372.2019.185133

[10] Blintsov, V., Nuzhniy, S., Kasianov, Y., & Korytskyi, V. «Mathematical model of the system of active protection against eavesdropping of speech information on the scrambler generator». Eureka: Physics and Engineering, (3), (2020). 11-22. Doi: 10.21303/2461-4262.2020.001241

[11] Renevey, P., Drygajlo, A. «Entropy based voice activity detection in very noisy conditions». In EUROSPEECH – 2001, (2001). 1887–1890. Available at: https://www.isca-speech.org/archive/archive_papers/eurospeech_2001/e01_1887.pdf

[12] Brungart D. S. «Informational and energetic masking effects in the perception of two simultaneous talkers». The Journal of the Acoustical Society of America 109 (3), (2001). 1101–1109. Doi:10.1121/1.1345696

[13] Brungart, D. S., Simpson, B. D. «The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal». The Journal of the Acoustical Society of America. 112 (2), (2002). 664-676. Doi: 10.1121/1.1490592

[14] Prodeus, A. M., Vityk, A. V., & Didenko, D. Y. Sub'ektivne otsínyuvannya yakostí i rozbírlivostí movnikh signalov, spotvorenikh sintezovanimi shumami. Mikrosistemi, Yelektronika i Akustika, 22 (6), (2017). 56–63. Doi: 10.20535/2523-4455.2017.22.6.101929

[15] Zverev V.A. «Slepaya dereverberatsiya rechevogo signala». Akusticheskiy zhurnal, 54, 2, (2008). 307-314. Available at: http://www.akzh.ru/pdf/2008_2_307-314.pdf

[16] Wang, Quan; Muckenhirn, Hannah; Wilson, Kevin; Sridhar, Prashant; Wu, Zelin; Hershey, John; Saurous, Rif; Weiss, Ron; Jia, Ye and Moreno, Ignacio. «VoiceFilter: Targeted Voice Separation by Speaker-Conditioned Spectrogram Masking». (2018). Available at: https://arxiv.org/pdf/1810.04826.pd

Методика киберзащиты компьютерных сетей

Хорошко В.О. [1], Зыбин С.В. [1], Хохлачова Ю.Е. [1], Аясрах А. [1], Аль-Далваш А. [1]

[1] *Национальный авиационный университет, пр-т Любомира Гузара 1, Киев, 03058, Украина*

Аннотация

В данной работе разработано методику многокритериальной маршрутизации в компьютерных сетях, которая учитывает группу критериев качества, характеризует кибербезопасность пользователей компьютерных сетей, группу критериев качества обслуживания в компьютерной сети и группу критериев качества, характеризующих надёжность и живучесть технических средств передачи информации в условиях воздействия различных помех и возмущений. Отличительной особенностью предложенной методики является сведение задачи многокритериальной маршрутизации в компьютерной сети к простой задаче о кратчайшем пути, для решения которой существует множество алгоритмических приёмных и аппаратурных средств решения.

Ключевые слова

кибербезопасность, киберзащита, компьютерные сети.

Methodology for cyber protection of computer networks

Khoroshko V. [1], Zybin S. [1], Khokhlachova Y. [1], Ayasrah A. [1], Al-Dalvash A. [1]

[1] *National aviation university, 1, Liubomyra Huzara ave., Kyiv, 03058, Ukraine*

Abstract

This paper has developed a methodology for multicriteria routing in computer networks, which takes into account a group of quality criteria, characterizes the cybersecurity of users of computer networks, a group of quality criteria for service in a computer network and a group of quality criteria characterizing the reliability and survivability of technical means of information transmission under the influence of various interference and indignation.

Keywords

cybersecurity, cybersecurity, computer networks.

1. Введение и цель исследования

Непрерывный рост сложности компьютерных сетей и резкое возрастание объёмов передаваемой информации ставят актуальную задачу повышения информационной безопасности (кибербезопасности) пользователей сетей.

Цель работы заключается в разработке методики многокритериальной маршрутизации в компьютерных сетях.

2. Методика маршрутизации

Первый этап этой методики реализуется путём формирования частных критериев качества. Особенности этапа заключаются в том, что среди множества частных критериев

EMAIL: professor_va@ukr.net (A.1); zysv@ukr.net (A.2); hohlachova@gmail.com (A.3); ahmadaesr@gmail.com (A.4); abduiiah.dalosh@gmail.com (A.5)
ORCID: 0000-0001-6213-7086 (A.1); 0000-0002-2670-2823 (A.2); 0000-0002-1883-8704 (A.3); 0000-0003-4392-1806 (A.4); 0000-0002-1003-9182 (A.5)

качества, характеризующих кибербезопасность, качество обслуживания сети, надёжность функционирования технических средств в условиях внутренних и внешних помех и возможностей необходимо выбрать наиболее существенные, которые в значительной степени влияют на эффективность работы компьютерной сети. Ограничивать количество частных критериев качества необходимо потому, что вычислительная сложность алгоритмов многокритериальной оптимизации растёт пропорционально количеству учитывающих критериев качества [1]. Кроме того, в каждой проблеме многокритериальной оптимизации существует определённое число частных критериев качества, которое является необходимым и достаточным для решения

Задачи многокритериальной оптимизации [2,3]. В этом случае использование любых дополнительных критериев или их сочетаний не изменяет результатов решения задачи. С изменяет результаты решения задачи. Для формирования такого количества частных критериев качества используем метод экспертных оценок. Как известно, в основе метода экспертных оценок лежит индивидуальное мнение эксперта об каждом критерии качества и его вкладе в обобщённый критерий качеств. В [2] предлагается уменьшить субъективизм и воспользоваться экспертными оценками нескольких экспертов, что позволит учитывать их «коллективное» мнение. На основе «коллективного» мнения специалистов в области кибербезопасности в компьютерных сетях установлено, что в систему частных критериев следует включить следующие группы критериев.

В первую группу входят частные критерии, которые характеризуют качество и уровень кибербезопасности пользователей компьютерных сетей.

Вторую группу частных критериев качества следует формировать на основе требований и на основу обслуживания в коллективных сетях.

В третью группу частных критериев качества входят критерии качества, оценивающие направленность функционирования технических средств м каналов передачи шифрования в условиях действий внешних, внутренних помех и возможностей.

В результате реализации первого этапа методики сформировано n_1 критериев первой группы K_{11} , K_{21} , …, $K_{n_1 1}$, n_2 критериев второй группы K_{12}, K_{22}, …, $K_{n_2 2}$ и n_3 критериев качества третьей группы K_{13}, K_{23}, …, $K_{n_3 3}$.

На втором этапе методики переводим максимальные критерии качества в минимальные и формируем систему ограничений:

$$0 \leq К_{i1} \leq К_{i1m}, i = \overline{1, n_1};$$
$$0 \leq К_{j2} \leq К_{j2m}, j = \overline{1, n_2};$$
$$0 \leq К_{s3} \leq К_{s3m}, S = \overline{1, n_3},$$

где $K_{i1m}, K_{j2m}, K_{s3m}$ – предельно решительные изменения критериев качества.

Третий этап методики состоит в формировании обобщённого критерия качества. Необходимость в этом этапе возможностей потому, что частные критерии конфликтуют между собой и улучшением одного критерия приводит к одновременному ухудшению других частных критериев качества [2]. Задачи многокритериальной относятся к классу непростых задач и трудностей между различными решениями, которые характеризуются несравнимыми наборами знаний частных критерий качества.

Поэтому реализация третьего этапа методики является сложной проблемой, не имеющей единственного подкода или метода решения. Реализация возможностей применима для задачи свёртки частных критериев по нелинейной схеме компромиссов.

$$К_0 = \sum_{i=1}^{n_1} \frac{\alpha_i}{1 - \frac{К_{i1}}{К_{i1m}}} + \sum_{j=1}^{n_2} \frac{\beta_j}{1 - \frac{К_{j2}}{К_{j2m}}} + $$
$$+ \sum_{s=1}^{n_3} \frac{\gamma_s}{1 - \frac{К_{s3}}{К_{s3m}}},$$

где весовые коэффициенты $\alpha_i \geq 0, \beta_j \geq 0, \gamma_s \geq 0$;

$$\sum_{i=1}^{n_1} \alpha_i + \sum_{j=1}^{n_2} \beta_j + \sum_{s=1}^{n_3} \gamma_s = 1.$$

Свёртки по нелинейной схеме компонентов дают строго модальную функцию обобщённого критерия качества $К_0$ при условии непрерывности и строгой выпуклости частных критерий качества [3]. Свойство унимодальности функции $К_0$ проявляется в области ограничений в виде единственной точки минимума функции $К_0$. Это обстоятельство существенно упрощает осуществлять минимум функции $К_0$ и является аргументом в пользу выбора нелинейной схемы компромиссов для решения задач маршрутизации.

Четвёртый этап этой методики состоит в формировании математической модели компьютерной сети в виде графа G (V, U), где $V = \{V_1, V_2, V_3, \dots, V_p\}$ – множество вершин, а $U = \{U_1, U_2, U_3, \dots, U_q\}$ – множество ребер. Вершины графа моделируют узлы – источники и узлы приёмники шифрования. Направление ребра графа сопоставляет каналы передачи шифрования между узлом источником и узлом приёмником. Рёбрам графа присваивается вес, предназначенный значениям обобщённого критерия качества для данного канала передачи шифрования. Пусть в графе G (V, U) выделено несколько чередуемых последовательностей вершин и рёбер $(V_0, U_1, V_1, U_2, \dots, V_{K-1}, U_K, V_K) = \mu$. В теории графов последовательность μ называется маршрутом из вершины V_0 в вершину V_K, если любая пара соседних элементов из μ инициирована в G. Во взвешенном графе G (V, U) ряд длинной пути поменяется суммарно все входных в него рёбер. Расстоянием $\rho(V_i, V_j)$ между вершинами V_i и V_j в графе G называется длина кратчайшего пути. Согласно эти определениям задача маршрутизации компьютерной сети может быть представлена как задача о кратчайшем пути между вершиной графа моделирующей узел-источник информации и вершиной графа моделирующей узел-приёмник информации:

$$\min_M L = \sum_{r=}^{R} K_{or},$$

где K_{or} – иерархический обобщённый критерий качества для второго ребра графа; L – длина кратчайшего пути; R – количество рёбер графа вдоль маршрута от вершины моделирующей узел – источник до вершины моделирующей узел – приёмник информации; M – множество всех допустимых маршрутов M между узлом – источником и узлом – приёмником информации. Достоинством предлагаемой методики является сведение задачи многокритериальной маршрутизации компьютерной сети к задаче о кратчайшем на графе, которая решается современными средствами маршрутизации компьютерных сетях.

Пятый этап методики заключается в решении задачи о кратчайшем пути на взвешенном графе G, вне рёбер которого пропорциональны иерархическому обобщённому критерию качества. Для реализации этого этапа методики могут применяться любые известные программные и аппаратные средства решения задачи о кратчайшем пути, на пример алгоритм Дейкстры или параллельные средства маршрутизации.

3. Выводы

Предложенная методика многокритериальной маршрутизации в компьютерных сетях, которая учитывает группу критериев качества, характеризуют кибербезопасность пользователей компьютерных сетей, группу критериев качества обслуживания в компьютерной сети и группу критериев качества, характеризующих надёжность и живучесть технических средств передачи информации в условиях воздействия различных помех и возмущений. Отличительная особенность предложенной методики является в сведении задачи многокритериальной маршрутизации в компьютерной сети к простой задаче о кратчайшем пути, для решения которой существует множество алгоритмических приемных и аппаратурных средств решения.

4. Литература

[1] Kobozeva. A.A. Analysis of security information / A.A. Kobozeva, V.A. Khoroshko – K: DCNKT, 2009.

[2] Voronin A.N. Multicriteria synthesis of dynamical systems / A.N. Voronin – K: 1992.

[3] Voronin A.N. Vector optimization of dynamic systems. 2nd edition / A.N. Voronin Yu.K. Zlatdinov and others. – К: Technics, 2003.

Метод анализа инструментальных погрешностей измерения параметров траекторий движения летательных аппаратов оптико-электронными станциями

Андрей Тевяшев [1], Олег Земляный [2], Игорь Шостко [3], Антон Колядин [4]

[1] *Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, 61166, Украина*
[2] *Институт радиофизики и электроники имени А. Я. Усикова НАН Украины, ул. Ак. Просфоры 12, Харьков, 61085, Украина*
[3] *Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, 61166, Украина*
[4] *Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, 61166, Украина*

Аннотация

В работе предложен метод анализе инструментальных погрешностей измерения параметров траекторий движения летательных аппаратов оптико-электронными станциями вызванных остаточными ошибками процедур горизонтирования и привязки платформы оптико-электронных станций к местному меридиану. Метод позволяет оценить влияние этих ошибок на статистические свойства оценок параметров траекторий движения летательных аппаратов, повысить их эффективность и снизить их смещённость за счёт временной и пространственной избыточности. В [1] рассматривается влияние погрешностей юстировки РЛС на точность оценивания координат объекта ЛА в декартовой системе координат РЛС, способ определения данных погрешностей и возможность их учёта в математической обработке

Ключові слова

Оптико-электронные станции, инструментальные погрешности, летательные аппараты, траектории движения, смещённые и эффективные оценки

Method for analyzing instrumental errors in measuring the parameters of trajectories of movement of aircraft by optoelectronic stations

Andrei Tevyashev [1], Oleg Zemlyaniy [2], Igor Shostko [3] and Anton Koliadin [4]

[1] *Kharkiv National University of Radio Electronics, Nauki Ave., 14, Kharkiv, 61166, Ukraine*
[2] *O. Ya. Usikov Institute for Radiophysics and Electronics National Academy of Sciences of Ukraine, st. Ak. Prosphora 12, Kharkiv, 61085, Ukraine*
[3] *Kharkiv National University of Radio Electronics, Nauki Ave., 14, Kharkiv, 61166, Ukraine*
[4] *Kharkiv National University of Radio Electronics, Nauki Ave., 14, Kharkiv, 61166, Ukraine*

Abstract

The paper proposes a method for analyzing instrumental errors in measuring the parameters of the trajectories of movement of aircraft by optoelectronic stations caused by residual errors in the procedures for leveling and binding the platform of optoelectronic stations to the local meridian. The method makes it possible to assess the influence of these errors on the statistical properties of estimates of the parameters of the trajectories of aircraft movement, to increase their efficiency and to reduce their bias due to temporal and spatial redundancy.

Keywords

Optoelectronic stations, instrumental errors, aircraft, motion trajectories, biased and effective estimates

EMAIL: andrew.teviashev@nure.ua (A. 1); zolvas@ukr.net (A. 2); iss010166@gmail.com (A. 3); antonkoliadin@gmail.com (A. 4)
ORCID: 0000-0001-5261-9874 (A. 1); 0000-0002-0271-722X (A. 2); 0000-0002-5612-3080 (A. 3); XXXX-XXXX-XXXX-XXXX (A. 4)

1. Введение

Современные оптико-электронные станции (ОЭС) всё чаще используются в качестве основных инструментальных средств для высокоточного определения параметров траекторий движения летательных аппаратов (ЛА), при полигонных испытаниях беспилотных летательных аппаратов (БПЛА), боеприпасов, при фиксировании траектории падения ракет, определении промаха при стрельбе по воздушным целям, обнаружении стартовых позиций и запусков ракет дальнего действия и др. Проблемам использования ОЭС для мониторинга воздушного пространства, внешнетраекторных измерений и метрологической аттестации ОЭС посвящено достаточно большое количество работ [1–9]. Однако проблема высокоточных траекторных измерений остаётся чрезвычайно актуальной. Решение этой проблемы требует детального анализа причин возникновения, степени влияния и способов подавления (нейтрализации) всех типов погрешностей измерения ОЭС траекторных параметров ЛА [10].

2. Оптико-электронные станции

Современные ОЭС для траекторных измерений реализуются в модульном исполнении и включают в себя следующие модули:

1. Опорно-поворотный с двухосевой платформой с вращением по азимуту и углу места, на которой размешен оптико-электронный модуль с системой глобального позиционирования и горизонтирования платформы.
2. Оптико-электронный – с телевизионной камерой, тепловизором и лазерным дальномером.

ОЭС обеспечивает круговой обзор воздушного пространства в оптическом и инфракрасном диапазонах частот, обнаружение, идентификацию, распознавание и автоматическое сопровождение ЛА.

Для определения параметров траекторий движения ЛА необходимо выбрать систему координат, в которой будут осуществляться траекторные измерения и определение (оценивание) параметров траекторий движения ЛА.

3. Системы координат

В зависимости от выбора начала отсчета, системы координат классифицируются на: геоцентрические (с началом отсчета в центре масс Земли *A*) и топоцентрические (с началом отсчета в точке *O* на поверхности Земли). Гринвичская система координат (ГСК) является геоцентрической системой координат и определяется следующим образом:

- ось AZг направлена по оси вращения Земли к Северному полюсу;
- ось AXг проходит через точку пересечения Гринвичского меридиана с плоскостью экватора;
- ось AYг дополняет прямоугольную систему до правой.

Местная декартовая система координат (МДСК) является топоцентрической системой координат и определяется следующим образом:

- ось OYм направлена по отвесной линии к земной поверхности;
- ось OXм расположена в плоскости, перпендикулярной отвесной линии в точке O и составляет с плоскостью меридиана начала системы координат угол Am – геодезический азимут, отсчитываемый по направлению движения часовой стрелки от направления на север;
- ось OZм проведена так, чтобы образовалась правая прямоугольная система координат.

Местная сферическая система координат (МССК) является топоцентрической системой координат, оси которой совпадают с осями МДСК.

В общеземных системах координат положения точек в пространстве могут быть заданы пространственными прямоугольными координатами X, Y, Z и геодезическими координатами B, L, H.

Геодезическая широта B – это угол между нормалью и плоскостью экватора. Геодезическая долгота L – это угол между плоскостью меридиана данной точки и плоскостью начального меридиана. Геодезической высотой H является отрезок нормали от точки до поверхности эллипсоида [11].

Для определения параметров траекторий движения ЛА используется местная декартовая система координат (МДСК) с началом в точке расположения ОЭС, координаты которой (L,B,H) определяются средствами GPS навигации и две инструментальные системы координат ОЭС (ИСК ОЭС): сферическая и декартовая системах координат, которые имеют общее начало с МДСК и, в идеале, совпадающие оси. ОЭС обеспечивает прямые измерения местоположения ЛА $P(\alpha_p, \beta_p, D_p)$ в инструментальной сферической системе координат ОЭС: наклонную дальность от начала координат до ЛА, азимут и угол места. Однако любые измерения обладают различными типами погрешности: погрешность измерения – отклонение результата измерения от истинного значения измеряемой величины; инструментальная погрешность обусловлена несовершенством применяемых средств измерений или их калибровки (юстировки); динамическая погрешность измерений – свойственна измерениям быстро меняющихся величин (погрешность сопровождения); статическая погрешность измерений – погрешность результатов измерений координат местоположения ОЭС (L,B,H) средствами GHS навигации; систематическая погрешность измерения – составляющая общей погрешности, остающаяся постоянной, или закономерно изменяющаяся при повторных измерениях; случайные погрешности, которые определяются совместным воздействием ряда случайных факторов.

Для минимизации влияния возникающих погрешностей необходимо использовать методы анализа причин их возникновения и применения специальных средств обработки результатов измерений. Данная статья посвящена анализу инструментальных погрешностей измерения параметров траекторий движения ЛА ОЭС, связанных с фактическим несовпадением направлений осей МДСК и ИДСК ОЭС, т.е. погрешностей юстировки (горизонтирования) платформы ОЭС и погрешностью привязки местной системы координат ОЭС к местному меридиану – погрешностью направления оси OX с вектором ON – направлением на север (NORD) из точки O МДСК.

4. Местные системы координат

На рис. 1 представлены местные инструментальные системы координат ОЭС и координаты точки P (местоположения ЛА) в момент времени t в местной инструментальной сферической системе координат ОЭС (МИССК) $P(\alpha_p, \beta_p, D_p)$ и местной инструментальной декартовой системе координат (МИДСК) $P(x_p, y_p, z_p)$.

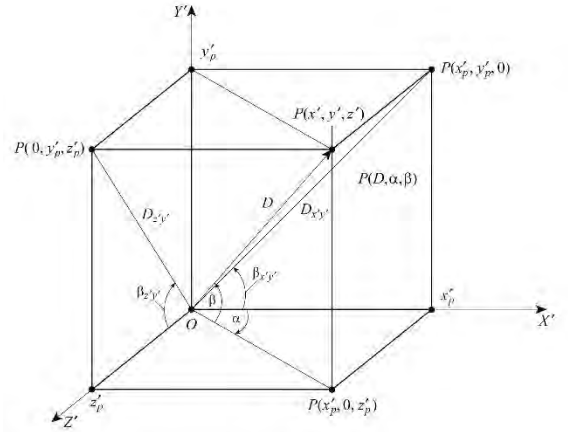


Рисунок 1: Координаты точки P в МИССК – $P(D_p, \alpha_p, \beta_p)$ и МИДСК – $P(x'_p, y'_p, z'_p)$ системах координат

Переход от сферической к декартовой системе координат точки P осуществляется по формулам:

$$y'_p = D\sin\beta, \tag{1}$$

$$x'_p = D\cos\beta\cos\alpha, \tag{2}$$

$$z'_p = D\cos\beta\sin\alpha. \tag{3}$$

5. Инструментальные системы координат ОЭС

Инструментальные системы координат ОЭС жёстко связаны с конструктивными особенностями реализации ОЭС, её пространственным расположением на поверхности земли и направлением оси OX, относительно которой осуществляется отсчёт (измерение) азимута.

Основной причиной возникновения инструментальных погрешностей измерения азимута и угла места ЛА являются фактические ошибки между направлениями осей МДСК и МИДСК ОЭС. Для их устранения или, по крайней мере, минимизации используются различные технологии и оборудование, включая лазерное, оптическое и навигационное. Однако полностью устранить инструментальные погрешности измерения

ОЭС азимута, угла места и наклонной дальности практически не удаётся. Как уже отмечалось ранее, ОЭС обеспечивает прямые измерения местоположения ЛА D_p – наклонной дальности, α_p – азимута и β_p – угла места в МИССК ОЭС. Далее эти координаты, в соответствии с выражениями (1)–(3), переводятся в координаты (X′,Y′,Z′) МИДСК. Так как дальнейшая обработка результатов траекторных измерений осуществляется в МДСК, то координаты местоположения ЛА (X′,Y′,Z′) должны быть переведены в координаты (X,Y,Z) МДСК. Если МИДСК полностью совпадает с МДСК, то такой перевод осуществляется путём приравнивания соответствующих координат, т.е. X=X′, Y=Y′, Z=Z′. В реальности эти системы координат имеют общее начало, но, как правило, отличаются по направлению на соответствующие углы по каждой из осей координат. Для проведения анализа инструментальных погрешностей измерения местоположения ЛА, связанных с погрешностями юстировки (горизонтирования) платформы ОЭС и погрешностью привязки местных систем координат ОЭС к местному меридиану, рассмотрим модели преобразования декартовых систем координат, связанных с поворотами координатных плоскостей вокруг координатных осей.

6. Повороты координатных плоскостей

Поворот плоскости X′OY′ на положительный угол вокруг оси OZ′. Система координат Z′X′Y′ правая, поэтому положительному углу α_z вокруг оси OZ′ соответствует поворот осей координат на угол α_z против часовой стрелки. На рис. 2 представлены координаты проекции точки P на плоскости X′OY′ и XOY.

При повороте плоскости X′OY′ вокруг оси OZ′ на $\alpha_z > 0$ координаты точки P относительно оси OZ не изменяются, не изменяются D и угол $\beta_{x'y'}$.

В этом случае координаты $x'_p = D_{x'y'} \cos\beta_{x'y'}$, $y'_p = D_{x'y'} \sin\beta_{x'y'}$, соответственно равны:

$$x'_p = D_{x'y'} \cos\beta_{x'y'}, \tag{4}$$

$$y'_p = D_{x'y'} \sin\beta_{x'y'}. \tag{5}$$

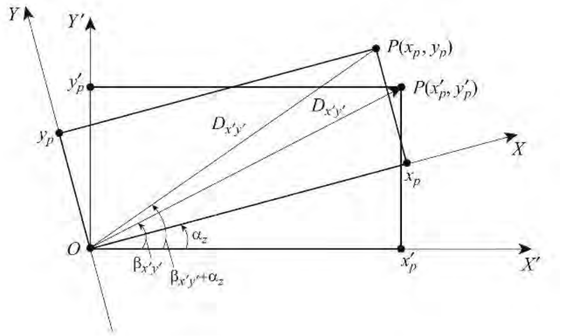


Рисунок 2: Координаты проекции точки P на плоскости X′OY′ и XOY, повёрнутой на положительный угол $\alpha_z > 0$ вокруг оси OZ′

Для вычисления координат точки P в МДСК (x_p, y_p), воспользуемся выражением:

$$\begin{aligned} x_p &= D_{x'y'} \cos(\beta_{x'y'} + \alpha_z) = \\ &= D_{x'y'}\left[\cos\beta_{x'y'} \cos\alpha_z - \sin\beta_{x'y'} \sin\alpha_z\right] = \\ &= x'_p \cos\alpha_z - y'_p \sin\alpha_z, \end{aligned} \tag{6}$$

$$\begin{aligned} y_p &= D_{x'y'} \sin(\beta_{x'y'} + \alpha_z) = \\ &= D_{x'y'}\left[\sin\beta_{x'y'} \cos\alpha_z + \cos\beta_{x'y'} \sin\alpha_z\right] = \\ &= x'_p \sin\alpha_z + y'_p \cos\alpha_z. \end{aligned} \tag{7}$$

В матричной форме:

$$P_z(\alpha_z) = \begin{pmatrix} \cos\alpha_z & -\sin\alpha_z & 0 \\ \sin\alpha_z & \cos\alpha_z & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{8}$$

Тогда

$$x_p = P_z(\alpha_z) x'_p. \tag{9}$$

При повороте плоскости X′OY′ вокруг оси OZ′ на отрицательный угол $\alpha_z < 0$ (по часовой стрелки) имеем:

$$\begin{aligned} x_p &= D_{x'y'} \cos(\beta_{x'y'} - \alpha_z) = \\ &= D_{x'y'}\left[\cos\beta_{x'y'} \cos\alpha_z + \sin\beta_{x'y'} \sin\alpha_z\right] = \\ &= x'_p \cos\alpha_z + y'_p \sin\alpha_z, \end{aligned} \tag{10}$$

$$\begin{aligned} y_p &= D_{x'y'} \sin(\beta_{x'y'} - \alpha_z) = \\ &= D_{x'y'}\left[\sin\beta_{x'y'} \cos\alpha_z - \sin\alpha_z \sin\beta_{x'y'}\right] = \\ &= y'_p \cos\alpha_z - x'_p \sin\alpha_z. \end{aligned} \tag{11}$$

В этом случае матрица поворота $P_z(-\alpha_z)$ принимает вид:

$$P_z(-\alpha_z) = \begin{pmatrix} \cos\alpha_z & \sin\alpha_z & 0 \\ -\sin\alpha_z & \cos\alpha_z & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12}$$

В этом случае:

$$x_p = P_z(-\alpha_z) x'_p \tag{13}$$

Поворот плоскости Y′OZ′ на положительный угол $\alpha_x > 0$ вокруг оси OX′. Система координат X′Y′Z′ – правая, поэтому положительному углу $\alpha_x > 0$ вокруг оси OX′ соответствует поворот осей координат на угол $\alpha_x > 0$ против часовой стрелки. В этом случае матрица поворота принимает вид:

$$P_x(\alpha_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_x & -\sin\alpha_x \\ 0 & \sin\alpha_x & \cos\alpha_x \end{pmatrix}. \quad (14)$$

При $\alpha_x < 0$:

$$P_x(-\alpha_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_x & \sin\alpha_x \\ 0 & -\sin\alpha_x & \cos\alpha_x \end{pmatrix}. \quad (15)$$

Поворот плоскости Z′OX′ на положительный угол $\alpha_y > 0$ вокруг оси OY′. Система координат Y′Z′X′ – правая, поэтому положительному углу $\alpha_y > 0$ вокруг оси OY′ соответствует поворот осей координат на угол $\alpha_y > 0$ против часовой стрелке. В этом случае матрица поворота принимает вид:

$$P_y(\alpha_y) = \begin{pmatrix} \cos\alpha_y & 0 & \sin\alpha_y \\ 0 & 1 & 0 \\ -\sin\alpha_y & 0 & \cos\alpha_y \end{pmatrix}. \quad (16)$$

При $\alpha_y < 0$:

$$P_y(-\alpha_y) = \begin{pmatrix} \cos\alpha_y & 0 & -\sin\alpha_y \\ 0 & 1 & 0 \\ \sin\alpha_y & 0 & \cos\alpha_y \end{pmatrix}. \quad (17)$$

Определители матриц $P_x(\cdot), P_y(\cdot), P_z(\cdot)$ равны 1 и обладают свойством ортогональности:

$$P^{\mathrm{T}}(\cdot) \cdot P(\cdot) = P(\cdot) P^{\mathrm{T}}(\cdot) = E,$$

Откуда следует, что обращение матрицы поворота соответствует ее транспонированию:

$$P^{-1}(\cdot) = P^{\mathrm{T}}(\cdot). \quad (18)$$

Матрица повротов на соответствующие углы получается в виде произведения матриц поворотов на те же углы:

$$P_{xyz}(\cdot) = P_x(\cdot) P_y(\cdot) P_z(\cdot), \quad (19)$$

при этом её определитель будет также равен 1.

Не нарушая общности, рассмотрим матрицу поворотов плоскости X′OZ′ (платформы опорно-поворотного модуля ОЭС) на положительные углы α_x – вокруг оси OX′ ($\alpha_x > 0$) и α_z – вокруг оси OZ′ ($\alpha_z > 0$) имеет вид:

$$P_{xz}(\alpha_x, \alpha_z) = P_x(\alpha_x) \cdot P_z(\alpha_z) =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_x & -\sin\alpha_x \\ 0 & \sin\alpha_x & \cos\alpha_x \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha_z & -\sin\alpha_z & 0 \\ \sin\alpha_z & \cos\alpha_z & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \cos\alpha_z & -\sin\alpha_z & 0 \\ \cos\alpha_x \sin\alpha_z & \cos\alpha_x \cos\alpha_z & -\sin\alpha_x \\ \sin\alpha_x \sin\alpha_z & \sin\alpha_x \cos\alpha_z & \cos\alpha_x \end{pmatrix}. \quad (20)$$

В этом случае координаты точки P в МДСК x равны:

$$x = P_{xz}(\alpha_x, \alpha_z) \cdot x'_p,$$

где x'_p – координаты точки P в МИДСК.

Матрица поворотов системы координат (x', y', z') относительно начала координат на положительные углы α_x, α_z, α_y имеет вид:

$$P_{xzy}(\alpha_x, \alpha_z, \alpha_y) = P_{xz}(\alpha_x, \alpha_z) \cdot P_y(\alpha_y) =$$

$$= \begin{pmatrix} \cos\alpha_z & -\sin\alpha_z & 0 \\ \cos\alpha_x \sin\alpha_z & \cos\alpha_x \cos\alpha_z & -\sin\alpha_x \\ \sin\alpha_x \sin\alpha_z & \sin\alpha_x \cos\alpha_z & \cos\alpha_x \end{pmatrix} \times$$

$$\times \begin{pmatrix} \cos\alpha_y & 0 & \sin\alpha_y \\ 0 & 1 & 0 \\ -\sin\alpha_y & 0 & \cos\alpha_y \end{pmatrix}.$$

Элементы матрицы $P_{xzy}(\alpha_x, \alpha_z, \alpha_y)$ имеют вид:

$$P_{11} = \cos\alpha_z \cos\alpha_y;$$

$$P_{12} = -\sin\alpha_z;$$

$$P_{13} = \cos\alpha_z \sin\alpha_y;$$

$$P_{21} = \cos\alpha_x \sin\alpha_z \cos\alpha_y + \sin\alpha_x \sin\alpha_y;$$

$$P_{22} = \cos\alpha_x \cos\alpha_z;$$

$$P_{23} = \cos\alpha_x \sin\alpha_z \sin\alpha_y - \sin\alpha_x \cos\alpha_y;$$

$$P_{31} = \sin\alpha_x \sin\alpha_z \cos\alpha_y - \cos\alpha_x \sin\alpha_y;$$

$$P_{32} = \sin\alpha_x \cos\alpha_z;$$

$$P_{33} = \sin\alpha_x \sin\alpha_z \sin\alpha_y + \cos\alpha_x \cos\alpha_y. \quad (21)$$

В этом случае координаты точки P МДСК равны:

$$x_p = P_{xzy}(\alpha_x, \alpha_z, \alpha_y) \cdot x'_p, \quad (22)$$

где $x_p = (x_p, y_p, z_p)^{\mathrm{T}}$, $x'_p = (x'_p, y'_p, z'_p)^{\mathrm{T}}$ – координаты точки P в МДСК и МИДСК.

7. Модели измерений

ОЭС обеспечивает прямые измерения местоположения ЛА в МИССК: наклонную дальность – D, азимут – α и угол места – β.

Введём следующие обозначения: $\xi(\omega)$ – случайная величина, где $\omega \in \Omega$, (Ω, B, P) – декартово произведение вероятностных пространств (Ω_i, B_i, P_i), $i = 1, 2, ..., n$ – декартово произведение вероятностных пространств; $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$, $P = P_1 \times P_2 \times ... \times P_n$, $B = B_1 \times B_2 \times ... \times B_n$; Ω_i – пространство элементарных событий; B_i – σ-алгебры событий; P_i – вероятностные меры на B_i); $\xi(\tilde{\omega})$ – конкретная реализация случайной величины; $\underset{\omega}{M}\{\xi(\omega)\} = \bar{\xi}$ – математическое ожидание случайной величины; σ_{ξ}^2 – дисперсия случайной величины $\xi(\omega)$. Не нарушая общности, будем полагать, что ошибки измерений распределены по нормальному закону с известными параметрами:

$$D = D(\omega) \cong N(\bar{D}, \sigma_D^2), \tag{23}$$

$$\alpha = \alpha(\omega) \cong N(\bar{\alpha}, \sigma_\alpha^2), \tag{24}$$

$$\beta = \beta(\omega) \cong N(\bar{\beta}, \sigma_\beta^2), \tag{25}$$

где $D(\omega), \alpha(\omega), \beta(\omega)$ – случайные величины, характеризующие ошибки (погрешности) измерений: $D(\omega)$ – дальности, $\alpha(\omega)$ – азимута и угла места $\beta(\omega)$; $\bar{D}, \bar{\alpha}, \bar{\beta}$ – математические ожидания и $\sigma_D^2, \sigma_\alpha^2, \sigma_\beta^2$ – дисперсии измеряемых величин; $D(\tilde{\omega}), \alpha(\tilde{\omega}), \beta(\tilde{\omega})$ – результаты измерений (реализаций) случайных величин.

Предполагается, что $D(\omega), \alpha(\omega), \beta(\omega)$ – независимые случайные величины.

Для привязки местоположения ОЭС к топооснове (координат точки O – начала местных систем координат) используются средства GPS позиционирования, обеспечивающих косвенные измерения геодезических координат (B, L, H).

Предполагается, что ошибки измерения геодезических координат также имеют нормальное распределение:

$$B = B(\omega) \cong N(\bar{B}, \sigma_B^2), \tag{26}$$

$$L = L(\omega) \cong N(\bar{L}, \sigma_L^2), \tag{27}$$

$$H = H(\omega) \cong N(\bar{H}, \sigma_H^2), \tag{28}$$

где $\bar{B}, \bar{L}, \bar{H}$; $\sigma_B^2, \sigma_L^2, \sigma_H^2$ – соответственно, математические ожидания и дисперсии измеряемых величин.

8. Модели юстировки платформы ОЭС

Юстировка платформы ОЭС заключается в решении двух задач: горизонтирования и привязки оси OX МИСК к местному меридиану.

Измерение значений азимута $\alpha(\omega)$ и угла места $\beta(\omega)$ осуществляются в МИСК. Необходимо провести горизонтирование платформы ОЭС таким образом, чтобы ее поверхность совпадала (как можно точнее) с плоскостью XOZ МДСК. Задача привязки оси OX МИСК к местному меридиану заключается в определении точного направления на север из нулевой точки O МИСК ОЭС. Решение задачи юстировки платформы ОЭС заключается, в том, чтобы обеспечить выполнение равенства нулю математических ожиданий результатов измерений углов поворотов $\alpha_x, \alpha_z, \alpha_y$. В реальных условиях, даже при использовании современных технологий, добиться этого не удаётся ни по одной из осей. Поэтому в качестве модели юстировки платформы будем использовать модель с отличными от нуля математическими ожиданиями углов поворотов α_x, α_z, α_y осей платформы вида:

$$\alpha_x = \alpha_x(\omega) \cong N(\bar{\alpha}_x, \sigma_{\alpha_x}^2), \tag{29}$$

$$\alpha_z = \alpha_z(\omega) \cong N(\bar{\alpha}_z, \sigma_{\alpha_z}^2), \tag{30}$$

$$\alpha_y = \alpha_y(\omega) \cong N(\bar{\alpha}_y, \sigma_{\alpha_y}^2), \tag{31}$$

где, как и ранее, $\bar{\alpha}_x$, $\bar{\alpha}_z$, $\bar{\alpha}_y$ – математических ожиданий а $\sigma_{\alpha_x}^2, \sigma_{\alpha_z}^2, \sigma_{\alpha_y}^2$ – дисперсии измеряемых углов поворотов платформы.

Для оценки метрологических характеристик ОЭС траекторных наблюдений был проведен вычислительный эксперимент.

9. Вычислительный эксперимент

Для исследования влияния инструментальных ошибок юстировки платформы ОЭС на значение систематических ошибок оценок местоположения ЛА был проведен вычислительный эксперимент, суть которого заключалась в следующем.

1. Задавались значения математического ожидания углов поворота платформы ОЭС вокруг соответствующих осей из следующего диапазона:

$$\bar{\sigma}_x, \sigma_y, \sigma_z \in [0°;\ 0{,}5°;\ 1{,}0°].$$

2. Задавались наборы значений измеренных величин $D(\tilde{\omega})$, $\beta(\tilde{\omega})$, $\alpha(\tilde{\omega})$ в МИССК из следующих диапазонов:

$$D(\tilde{\omega}) \in [1000\text{ м};\ 5000\text{ м};\ 10000\text{ м};\ 15000\text{ м}];$$

$$\beta(\tilde{\omega}) \in [0°;\ 30°;\ 45°;\ 60°;\ 75°];$$

$$\alpha(\tilde{\omega}) \in [0° \div 360°] \text{ с шагом } 1{,}0°.$$

3. Для каждого набора измеренных величин $D(\tilde{\omega})$, $\alpha(\tilde{\omega})$, $\beta(\tilde{\omega})$ вычислялись координаты ЛА в МДСК:

$$x' = x'(\tilde{\omega}) = D(\tilde{\omega})\sin\beta(\tilde{\omega});$$

$$y' = y'(\tilde{\omega}) = D(\tilde{\omega})\cos\beta(\tilde{\omega})\cos\alpha(\tilde{\omega});$$

$$z' = z'(\tilde{\omega}) = D(\tilde{\omega})\cos\beta(\tilde{\omega})\sin\alpha(\tilde{\omega}).$$

4. Далее, используя матрицу поворотов (21), вычислялись значения координат ЛА в МДСК:

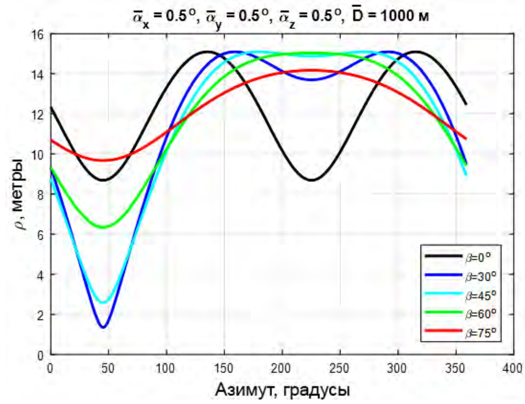
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P_{xyz}\left(\bar{\alpha}_x, \bar{\alpha}_z, \bar{\alpha}_y\right)\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

5. Вычислялось значение систематической ошибки оценки координат ЛА:

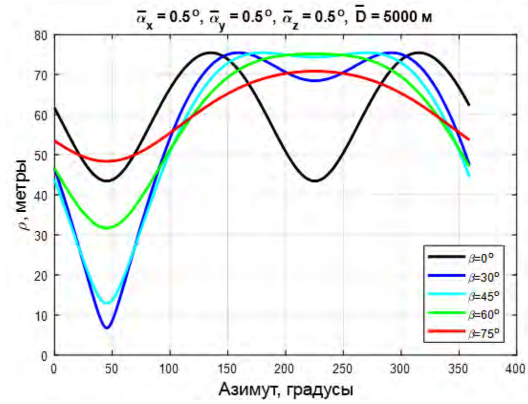
$$\rho = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}.$$

На рис. 3 (a–l) приведены зависимости изменения значений систематической ошибки ρ оценки координат местоположения ЛА от изменения азимута при различных значениях математических ожиданий ошибок юстировки платформы ОЭС, различных значениях наклонной дальности до ЛА и угла места.

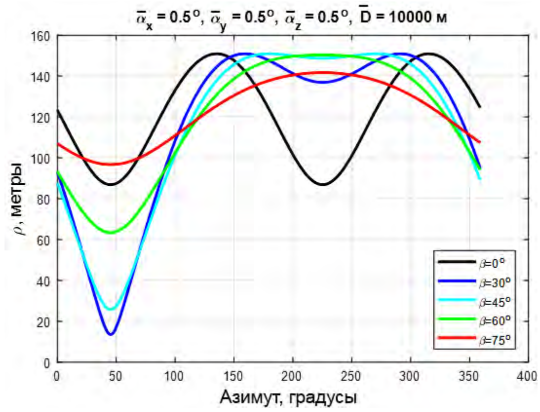
На рис. 4 (a–h) приведены аналогичные зависимости при изменении азимута с шагом 1,0°.



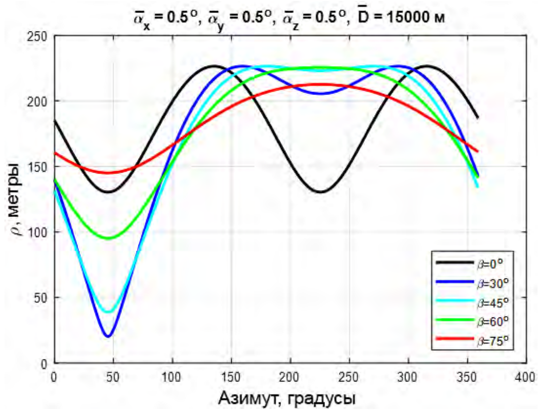
a



b

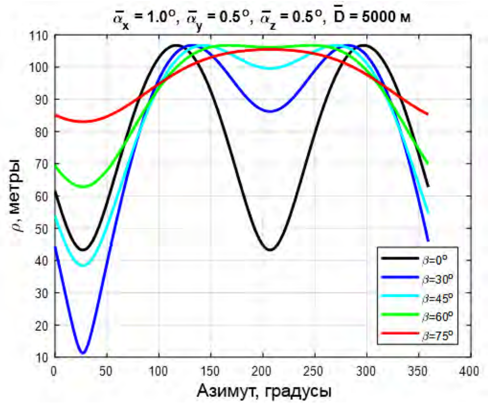


c

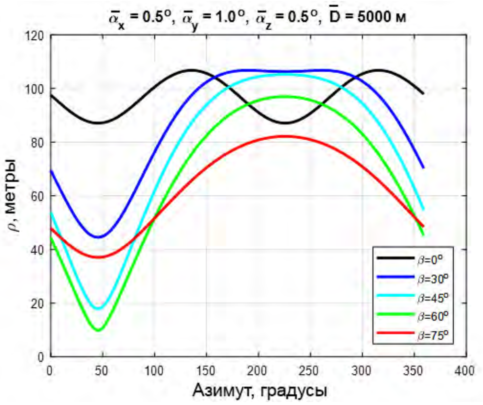


d

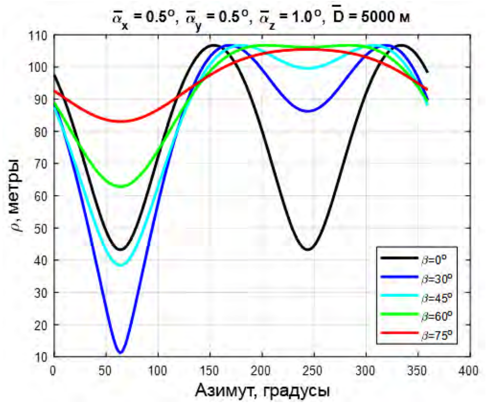
Рисунок 3 (a, b, c, d): Зависимости инструментальной погрешности от изменения азимута



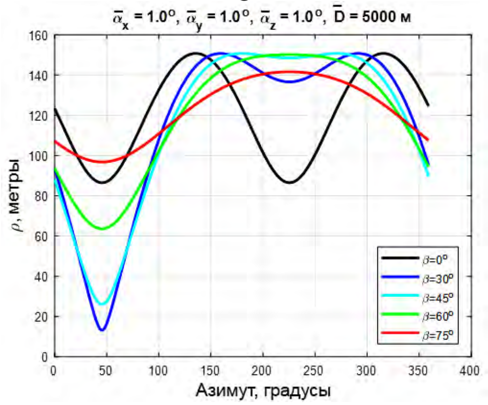
e



f

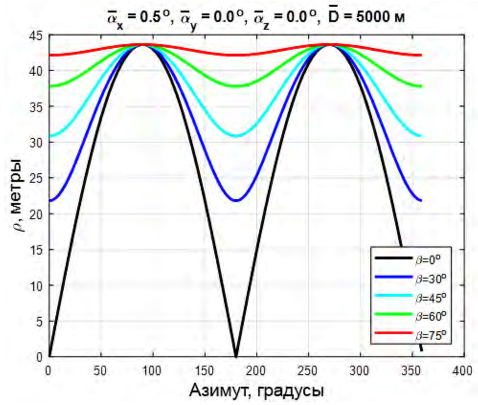


g

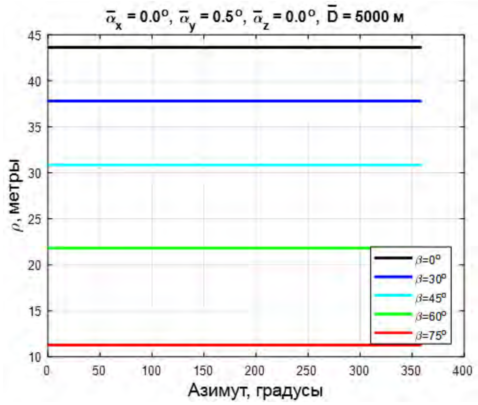


h

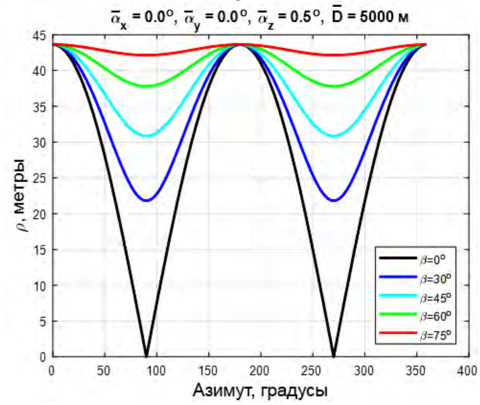
Рисунок 3 (e, f, g, h): Зависимости инструментальной погрешности от изменения азимута



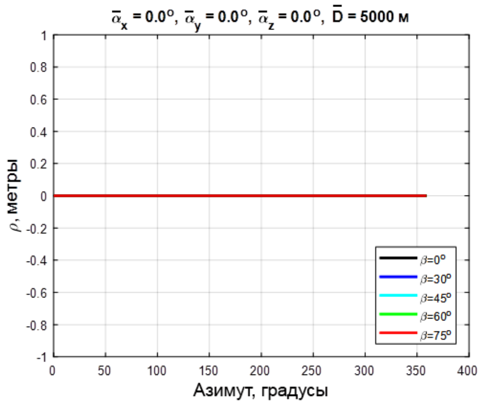
i



j

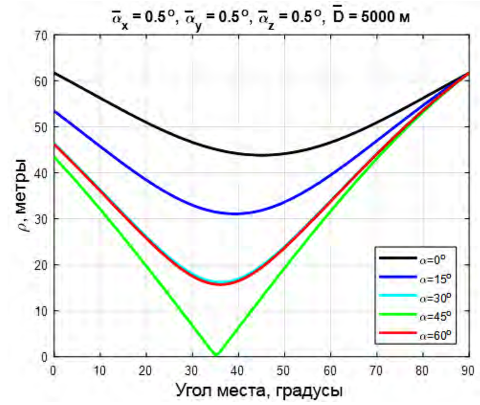


k

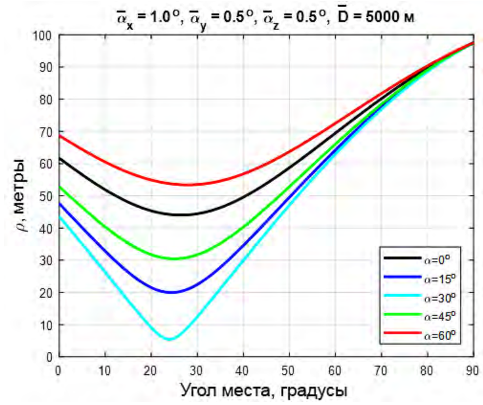


l

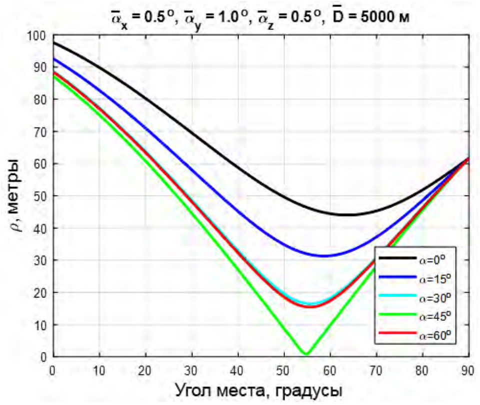
Рисунок 3 (i, j, k, l): Зависимости инструментальной погрешности от изменения азимута



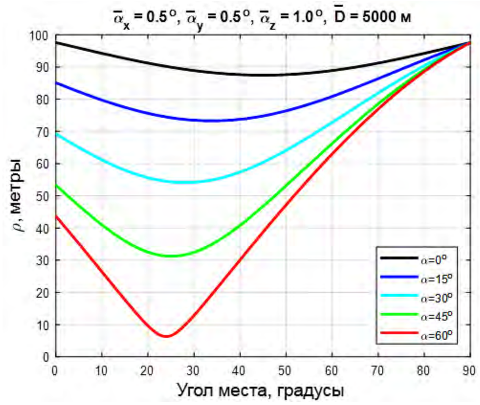
a



b

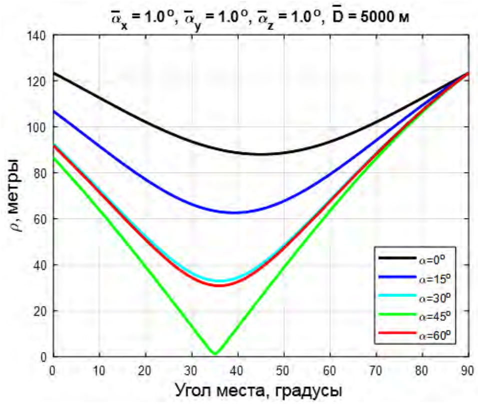


c

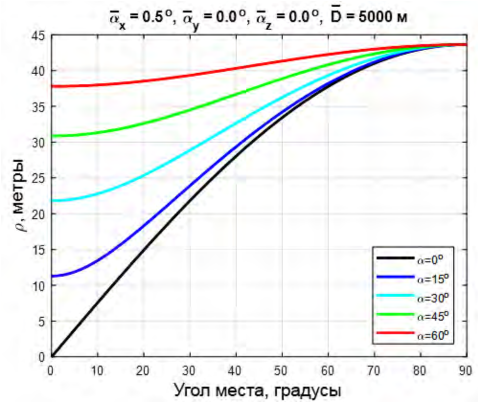


d

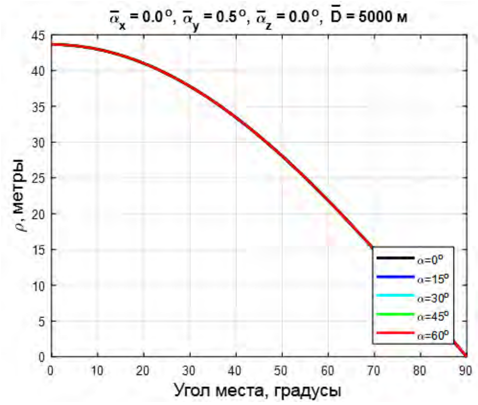
Рисунок 4 (a, b, c, d): Зависимости инструментальной погрешности от изменения угла места



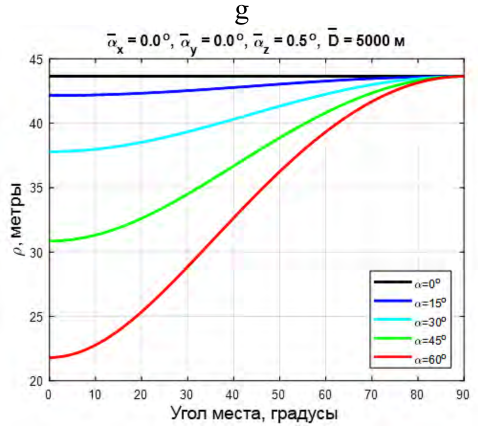
e



f



g



h

Рисунок 4 (e, f, g, h): Зависимости инструментальной погрешности от изменения угла места

10.Заключение

Анализ зависимостей инструментальной погрешности измерения ОЭС координат местоположения ЛА (рис. 3, 4) от исследуемых параметров, позволяет сделать следующие выводы:

1. Зависимость численного значения инструментальной погрешности ρ (систематической ошибки) определения координат ЛА, значительно и существенно нелинейная от исследуемых параметров.
2. Для снижения инструментальной погрешности, необходимо разработать и реализовать технологичность юстировки платформы ОЭС с максимальной погрешностью не более ±0,2 мрад. по каждой из осей.
3. Для снижения зависимости инструментальной погрешности от влияния внешних факторов (температуры окружающей среды), необходимо использовать гиростабилизируемую платформу ОЭС на основе лазерных гироскопов.

11.Литература References

[1] Додонов А. Г., Путятин В. Г., Валетчик В. А. Обработка оптических измерений траектории летательных объектов. Реєстрація, зберігання і оброб. даних, 2004, т. 6, № 1, 38–52.

[2] Путятин В. Г., Додонов В. А. Об одной задаче высокоточных траекторных измерений оптическими средствами. Реєстрація, зберігання і обробка даних, 2017, т. 19, № 2, 36–54.

[3] Шапиро Л., Стокман Д. Компьютерное зрение. М.: Издательство «БИНОМ. Лаборатория знаний», 765 с.

[4] Szeliski Richard. Computer Vision: Algorithms and Applications. Springer, 2010, 933 с.

[5] Shostko I., Tevyashev A., Kulia, Y., Koliadin A. Optical-electronic system of automatic detection and high-precise tracking of aerial objects in real-time. The Third International Workshop on Computer Modeling and Intelligent Systems, CMIS, 2020, 784–803.

[6] Tevyashev Andrey, Shostko Igor, Neofitniy Mihail, Koliadin Anton. Laser Opto-Electronic Airspace Monitoring System in The Visible and Infrared Ranges. 2019 IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments, APUAVD 2019 – Proceedings, 2019, 170–173. DOI: 10.1109/APUAVD47061.2019.8943887.

[7] Tevjashev A. D., Shostko I. S., Neofinyi M. V., Kolomiyets S. V., Kyrychenko I. Yu., Pryimachov Yu. D. Mathematical model and method of optimal placement of optical-electronic systems for trajectory measurements of air objects at test. Odessa Astronomical Publications, Vol. 32 (2019), 171–175.

[8] Shostko I., Tevyashev A., Neofitnyi M., Kulia Y. Information-measuring system of polygon based on wireless sensor infocommunication network. Chapter in the book Lecture Notes on Data Engineering and Communications Technologies, 48. Publisher: Springer Nature, 2021, 649–674.

[9] Shostko I., Tevyashev A., Neofitnyi M., Ageyev D., Gulak S. Information and Measurement System Based on Wireless Sensory Infocommunication Network for Polygon Testing of Guided and Unguided Rockets and Missiles. 2018 International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, PIC S and T 2018 – Proceedings, 2019, 705–710.

[10] Лысенко И. В. Способ автоматизированного оперативного контроля горизонтирования радиотехнических средств траекторных измерений, TRANSPORT BUSINESS IN RUSSIA, № 6, 2016, 45–50.

[11] Жданюк В. Ф. Основы статистической обработки траекторных измеркний. М.: Советское радио, 1978, 384 с.

Accelerated learning of the neural network ADALINE in the presence of stationary correlated noise

Oleksandr Bezsonov[1], Oleg Rudenko[1], Natalia Serdiuk[1], Oleg Lebediev[1] and Valentyn Lebediev[1]

[1]*Kharkiv National University of Radio Electronic, Kharkiv, 61166, Ukraine*

Abstract

The problem of constructing a recurrent form of a multi-step ADALINA learning algorithm in the presence of stationary correlated interferences is considered. The basic relations describing the processes of accumulation of new and dumping of obsolete information are obtained. It is shown that the algorithms under consideration use the transformation of a random vector with correlated components into a random vector with uncorrelated components.

Keywords

Correlation, learning, recurrent algorithm, correlation matrix, centering, interference, inclusion of information, reset of information

1. Introduction

ADALINA (adaptive linear element) became the first linear neural network proposed by Widrow B. and Hoff M., which was an alternative to the perceptron [1]. Subsequently, this element and the algorithm of its training have found wide application in the tasks of identification, management, filtering, etc. The Widrow-Hoff learning algorithm is Kachmazh's algorithm for solving systems of linear algebraic equations. The properties of this algorithm in solving the identification problem are quite fully described in [2]. In [3], the regularized Kachmazh algorithm (Widrow-Hoff) was used to train ADALINA in the task of estimating nonstationary parameters and the corresponding estimates of convergence rate and accuracy were obtained.

It should be noted that the learning process of ADALINA can be significantly accelerated if instead of a one-step Widrow-Hoff (Kachmazh) algorithm to utilize a multi-step algorithm that uses a limited number of measurements, i.e. has limited memory, and is based on the current regression analysis algorithm [4-6].

Usually when constructing neural network models and their training algorithms, it is assumed that the noise in the measurements is not correlated. The presence of correlated noise complicates the evaluation procedure. Therefore, the correlation of noise is often neglected, while obtaining deliberately suboptimal results.

Recurrent estimation with correlated disturbances was considered in [7-9]. In [7] a recurrent form of MLS was proposed, in [8] a procedure of the stochastic approximation type was studied, the coefficients of which at each step are chosen to be optimal in the sense of minimizing the amount of a posteriori variance of estimates. In [9] it was shown that both of these methods have much in common and their experimental comparison was performed, which showed that in the case of a very strong correlation RMLS has greater computational stability.

The aim of this work is to obtain a recurrent form of a multi-step learning algorithm under conditions of stationary correlated noise.

2. Construction of a recurrent form of the current regression analysis algorithm

The problem of object identification described by the following equation is considered

$$Y_n = X_n c^* + \Xi_n, \tag{1}$$

EMAIL: oleksandr.bezsonov@nure.ua, (A.1); oleh.rudenko@nure.ua (A.2); nataliya.serdyuk@nure.ua (A.3); oleh.lebediev@nure.ua (A.4); lebedevvalen@gmail.com (A.5)

ORCID: 0000-0001-6104-4275 (A.1); 0000-0003-0859-2015 (A.2); 0000-0002-0107-4365 (A.3); 0000-0001-5998-0136 (A.4); 0000-0002-0095-7481 (A.5)

where $Y_n = (y_1, y_2, ... y_n)^T$ – vector of output signals; $X_n^T = (x_1, x_2, ..., x_n)^T = \left(X_{n-1}^T x_n\right)$ – matrix of input signals; $c^* = (c_1^*, c_2^*, .. c_N^*)^T$ – vector of parameters to be evaluated; $\Xi_n = (\xi_1, \xi_2, ..., \xi_n)^T$ – disturbance vector.

Covariance matrix D_n order n disturb ξ_{n+1} has the appearance

$$D_n = M\left\{\Xi_n \Xi_n^T\right\} = \begin{bmatrix} d_{1,1} & d_{1,2} \dots & d_{1,n-1} & d_{1,n} \\ d_{2,1} & d_{2,2} \dots & d_{2,n-1} & d_{2,n} \\ \dots & \dots & \dots & \dots \\ d_{n,1} & d_{n,2} \dots & d_{n,n-1} & d_{n,n} \end{bmatrix} = \begin{bmatrix} D_{n-1} & d_{n-1} \\ d_{n-1}^T & d_{nn} \end{bmatrix}, \qquad (2)$$

where $d_{ij} = M\left\{\xi_i \xi_j\right\}$;

$$d_{n-1}^T = \left(d_{n,1}, d_{n,2}, ..., d_{n,n-1}\right) = M\left\{\xi_n \Xi_{n-1}^T\right\}.$$

As is known [10], the application of evaluation

$$c_n = \left(X_n^T X_n\right)^{-1} X_n^T Y_n \qquad (3)$$

to the model with correlated noise gives estimates whose variances will be underestimated.

Gaussian - Markov estimate (MLS), obtained by minimizing the quadratic functional, has the form

$$c_n = \left(X_n^T D_n^{-1} X_n\right)^{-1} X_n^T D_n^{-1} Y_n. \qquad (4)$$

An algorithm of current regression analysis that has the form

$$c_{n|L} = (X_{n|L}^T X_{n|L})^{-1} X_{n|L}^T Y_{n|L}, \qquad (5)$$

where

$$Y_{n|L} = \begin{pmatrix} Y_{n-1|L-1} \\ --- \\ y_n \end{pmatrix} = \begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L-1} \end{pmatrix} \qquad (6)$$

– vector $L \times 1$;

$$X_{n|L} = \begin{pmatrix} X_{n-1|L-1} \\ --- \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{n-L+1}^T \\ --- \\ X_{n|L-1} \end{pmatrix} \qquad (7)$$

– matrix $(L \times 1) \times N$;

was proposed in [4], and in [5] a modification of this algorithm is considered, which uses the mechanism of forgetting past information (smoothing). Here $L = const (L \ge N)$ – algorithm's memory.

Feature of algorithms with $L = const$ is that the matrices and observation vectors used in the construction of estimates at each step of the assessment are formed as follows: they include information about the newly arrived measurements and exclude information about the oldest. Depending on how these matrices and vectors are formed (new information is added first, and then obsolete information is excluded, or obsolete information is first excluded, and then new information is added), two forms of evaluation are possible.

Let's dwell on this in more detail.

By analogy with the Gaussian-Markov estimate (4), following estimate can be proposed

$$c_{n|L} = (X_{n|L}^T D_{n|L}^{-1} X_{n|L})^{-1} X_{n|L}^T D_{n|L}^{-1} Y_{n|L}, \qquad (8)$$

where

$$D_{n|L} = \begin{bmatrix} D_{n-1,L-1+1} & \vdots & d_{n-1} \\ - & - & - \\ d_{n-1}^T & \vdots & d_{n,n} \end{bmatrix}, \qquad (9)$$

$$d_{n-1}^T = \left(d_{n,n-L+1}, d_{n,n-L+2}, ..., d_{n,n-1}\right) = M\left\{\xi_n \Xi_{n-1|L-1}^T\right\}$$

Due to the fact that the matrix $Dn|L$ has a block representation, it can be written

$$D_{n|L}^{-1} = \begin{bmatrix} D_{n-1,L-1}^{-1} + \frac{D_{n-1,L-1}^{-1} d_{n-1} d_{n-1}^T D_{n-1,L}^{-1}}{\alpha_n} & \vdots & -\frac{D_{n-1,L-1}^{-1} d_{n-1}}{\alpha_n} \\ \dots & \dots & \dots \\ -\frac{d_{n-1}^T D_{n-1,L-1}^{-1}}{\alpha_n} & \vdots & \frac{1}{\alpha_n} \end{bmatrix}, \qquad (10)$$

where $\alpha_n = d_{nn} - d_{n-1}^T D_{n-1,L-1}^{-1} d_{n-1}$.

Similar relations are used in [7] to obtain a recurrent form of MLS.

Consider the case of stationary noise ξ.

Normalized covariance matrix for stationary noise with dispersion σ_ξ^2 looks like

$$R_{n|L} = D_{n|L} / \sigma_\xi^2 = \begin{bmatrix} 1 & \rho_1 \dots & \rho_{n-L} & \rho_{n-L+1} \\ \rho_1 & 1 \dots & \rho_{n-3} & \rho_{n-L} \\ \dots & \dots & \dots & \dots \\ \rho_{n-L+1} & \rho_{n-L} \dots & \rho_1 & 1 \end{bmatrix}, \qquad (11)$$

where ρ_i – the value of the normalized correlation function of the process ξ.

With this in mind, the estimate (8) takes the form

$$c_{n|L} = (X_{n|L}^T R_{n|L}^{-1} X_{n|L})^{-1} X_{n|L}^T R_{n|L}^{-1} Y_{n|L}. \qquad (12)$$

Let's introduce $(n-m+1)$ – measurable vector

$$r_{m,n}^T = (\rho_n, \rho_{n-1}, ...\rho_m), \quad (m \le n).$$

Then the matrix $R_{n|L}$ can be represented in block form

$$R_{n|L} = D_{n|L} / \sigma_\xi^2 = \begin{bmatrix} R_{n-1|L-1} & r_{n-1,1} \\ r_{n-1,1}^T & 1 \end{bmatrix}, \qquad (13)$$

where $r_{n-1,1}^T = (\rho_{n-1}, \rho_{n-2}, ...\rho_1)$, and the inverse matrix $R_{n|L}^{-1}$ will be calculated as follows:

$$R_{n|L}^{-1} = \\ = \begin{bmatrix} R_{n-1|L-1}^{-1} + \\ + \dfrac{R_{n-1|L-1}^{-1} r_{n-1,1} r_{n-1,1}^T R_{n-1|L-1}^{-1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} & -\dfrac{R_{n-1|L-1}^{-1} r_{n-1,1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} \\ -\dfrac{r_{n-1,1}^T R_{n-1}^{-1}}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} & \dfrac{1}{1 - r_{n-1,1}^T R_{n-1|L-1}^{-1} r_{n-1,1}} \end{bmatrix}. \qquad (14)$$

3. The procedure for adding new information

Suppose that on $(n-1)$-th tact an estimate is received

$$c_{n-1|L} = \\ = \left(X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L}\right)^{-1} X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} \qquad (15)$$

The arrival of new information (adding a new dimension) leads to the calculation of the estimate, which by analogy with (15) can be written as follows:

$$c_{n|L+1} = \\ = (X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1})^{-1} X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} \qquad (16)$$

where

$$Y_{n|L+1} = \begin{pmatrix} Y_{n-1|L} \\ --- \\ y_n \end{pmatrix} = \begin{pmatrix} y_{n-L+1} \\ --- \\ Y_{n|L} \end{pmatrix} \qquad (17)$$

– vector $(L+1)\times 1$;

$$X_{n|L+1} = \begin{pmatrix} X_{n-1|L} \\ --- \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{n-L+1}^T \\ --- \\ X_{n|L} \end{pmatrix} \qquad (18)$$

– matrix $(L+1)\times N$.

We introduce following notation

$$K_{n|L+1}^{-1} = \left(X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1}\right);$$
$$K_{n-1|L}^{-1} = \left(X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L}\right); \qquad (19)$$
$$K_{n|L}^{-1} = \left(X_{n|L}^T R_{n|L}^{-1} X_{n|L}\right)$$

$$\alpha_{j,n} = \rho_j - r_{j+1,n+j-1} R_{n-1|L-1}^{-1} r_{1,n-1} \\ (j = 0,1,...) \qquad (20)$$

and calculate $K_{n|L+1}^{-1}$

$$K_{n|L+1}^{-1} = X_{n-1|L}^T R_{n-1|L}^{-1} X_{n-1|L} + \\ + \frac{X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1} r_{n-1,1}^T R_{n-1|L}^{-1} X_{n-1|L}}{\alpha_{n,0}} - \\ - \frac{x_n r_{n-1,1}^T R_{n-1|L}^{-1} X_{n-1|L}}{\alpha_{n,0}} - \qquad (21) \\ - \frac{X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1} x_n^T}{\alpha_{n,0}} + \frac{x_n x_n^T}{\alpha_{n,0}} = \\ = K_{n-1|L}^{-1} + x_n^* x_n^{*T},$$

where $x_n^* = \dfrac{x_n - X_{n-1|L}^T R_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}$

Similarly, calculate

$$X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} = X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} + x_n^* y_n^*, \qquad (22)$$

where $y_n^* = \dfrac{y_n - Y_{n-1|L} R_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}.$

Adding $x_n^* x_n^{*T} c_{n-1|L}$ to both parts (15)

$$K_{n-1|L}^{-1} c_{n-1|L} + x_n^* x_n^{*T} c_{n-1|L} = \\ = X_{n-1|L}^T R_{n-1|L}^{-1} Y_{n-1|L} + x_n^* x_n^{*T} c_{n-1|L}$$

and subtracting (15) from (16), taking into account the properties $K_{n-1|L}^{-1}$ and $X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1}$), we obtain

$$K_{n|L+1}^{-1}\left(c_{n|L+1} - c_{n-1|L}\right) = x_n^*\left(y_n^* - c_{n-1|L}^T x_n^*\right)$$

or

$$c_{n|L+1} = c_{n-1|L} + K_{n|L+1} x_n^* \left(y_n^* - c_{n-1|L}^T x_n^*\right), \qquad (23)$$

where

$$K_{n|L+1} = K_{n-1|L} - \frac{K_{n-1|L} x_n^* x_n^{*T} K_{n-1|L}}{1 + x_n^{*T} K_{n-1|L} x_n^*}. \qquad (24)$$

4. Procedure for discarding obsolete information

When discarding outdated information obtained on $n-L+1$-th step, we come from the assessment $c_{n|L+1}$ to evaluation $c_{n|L}$. To obtain the appropriate rules for the correction of the assessment, proceed as follows.

We use the block representation of the covariance matrix $R_{n|L+1}$

$$R_{n|L+1} = \\ = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_{n-L+1} \\ \rho_1 & 1 & \rho_1 & \rho_{n-L+1,n} \\ \dots & \dots & \dots & \dots \\ \rho_{n-L+1} & \rho_{n-L} & \rho_{n-L-1} & 1 \end{bmatrix} = \\ = \begin{bmatrix} 1 & r_{n-1,1}^T \\ r_{n-1,1} & R_{n|L} \end{bmatrix}, \tag{25}$$

where $r_{n-1,1}^T = (\rho_{n-L+1}, \rho_{n-L}, \dots, \rho_{n-1})$, and the representation of the inverse matrix $R_{n|L+1}^{-1}$ as

$$R_{n|L+1}^{-1} = \\ = \begin{bmatrix} \dfrac{1}{\alpha_{n-L+1,0}} & -\dfrac{r_{n-L+1}^T R_{n,L}^{-1}}{\alpha_{n-L+1,0}} \\ -\dfrac{R_{n|L}^{-1} r_{n-L+1}}{\alpha_{n-L+1,0}} & \begin{matrix} R_{n,L}^{-1} + \\ + \dfrac{R_{n,L}^{-1} r_{n-L+1} r_{n-L+1}^T R_{n,L}^{-1}}{\alpha_{n-L+1,0}} \end{matrix} \end{bmatrix}, \tag{26}$$

where $\alpha_{n-L+1,0} = 1 - r_{n-L+1,1}^T R_{n,L}^{-1} r_{n-L+1,1}$,

In this case

$$K_{n|L+1}^{-1} = \left(X_{n|L+1}^T R_{n|L+1}^{-1} X_{n|L+1} \right) = \\ = \frac{x_{n-L+1} x_{n-L+1}^T}{\alpha_{n-L}} + \\ + \frac{X_{n|L}^T R_{n|L}^{-1} d_{n-L+1} d_{n-L+1}^T R_{n|L}^{-1} X_{n|L}}{\alpha_{n-L}} - \\ - \frac{x_{n-L+1} d_{n-L+1}^T R_{n|L}^{-1} X_{n|L}}{\alpha_{n-L+1}} - \\ - \frac{X_{n|L}^T R_{n|L}^{-1} d_{n-L+1} x_{n-L+1}^T}{\alpha_{n-L+1}} + \\ + X_{n|L}^T R_{n|L}^{-1} X_{n|L} = K_{n|L}^{-1} + x_{n-L+1}^* x_{n-L+1}^{*T}, \tag{27}$$

where

$$x_{n-L+1}^* = \frac{x_{n-L+1} - X_{n|L}^T D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}.$$

We have the same

$$X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} = \\ = X_{n|L}^T R_{n|L}^{-1} Y_{n|L} + x_{n-L+1}^* y_{n-L+1}^*, \tag{28}$$

where $y_{n-L+1}^* = \dfrac{y_{n-L+1} - Y_{n|L} D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}$.

Subtract from both parts (16) $x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1}$. Then

$$\left(\left(X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} \right)^{-1} - x_{n-L+1}^* x_{n-L+1}^{*T} \right) c_{n|L+1} = \\ = X_{n|L+1}^T R_{n|L+1}^{-1} Y_{n|L+1} - x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1}. \tag{29}$$

Taking into account that

$$\left(X_{n|L}^T R_{n|L}^{-1} X_{n|L} \right) c_{n|L} = X_{n|L}^T R_{n|L}^{-1} Y_{n|L}, \tag{30}$$

subtract from (30) the relation (16) taking into account the expressions for $K_{n|L}^{-1}$ and $X_{n|L}^T R_{n|L}^{-1} Y_{n|L}$ we obtain:

$$K_{n|L}^{-1} \left(c_{n|L} - c_{n|L+1} \right) = \\ = x_{n-L+1}^* x_{n-L+1}^{*T} c_{n|L+1} - x_{n-L+1}^* y_{n-L+1}^*,$$

where

$$c_{n|L} = c_{n|L+1} - \\ - K_{n|L} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n|L+1}^T x_{n-L+1}^* \right), \tag{31}$$

but

$$K_{n|L}^{-1} = K_{n|L+1}^{-1} - x_{n-L+1}^* x_{n-L+1}^{*T}.$$

Then

$$K_{n|L} = K_{n|L+1} + \frac{K_{n|L+1} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n|L+1}}{1 - x_{n-L+1}^{*T} K_{n|L+1} x_{n-L+1}^*}. \tag{32}$$

Thus, the algorithm will look like (the first two relations describe the inclusion of new information, and the next two – the rejection of obsolete)

$$c_{n|L+1} = c_{n-1|L} + K_{n|L+1} x_n^* \left(y_n^* - c_{n-1|L}^T x_n^* \right); \tag{33}$$

$$K_{n|L+1} = K_{n-1|L} - \frac{K_{n-1|L} x_n^* x_n^{*T} K_{n-1|L}}{1 + x_n^{*T} K_{n-1|L} x_n^*}. \tag{34}$$

$$c_{n|L} = c_{n|L+1} - \\ - K_{n|L} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n|L+1}^T x_{n-L+1}^* \right), \tag{35}$$

$$K_{n|L+1} = K_{n|L} + \frac{K_{n|L+1} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n|L+1}}{1 - x_{n-L+1}^{*T} K_{n|L+1} x_{n-L+1}^*}; \tag{36}$$

If at first the outdated information is rejected, and then new is included, the algorithm takes the form

$$c_{n-1|L-1} = c_{n-1|L} - \\ - K_{n-1|L-1} x_{n-L+1}^* \left(y_{n-L+1}^* - c_{n-1|L}^T x_{n-L+1}^* \right); \tag{37}$$

$$K_{n-1|L-1} = K_{n-1|L} + \frac{K_{n-1|L} x_{n-L+1}^* x_{n-L+1}^{*T} K_{n-1|L}}{1 - x_{n-L+1}^{*T} K_{n-1|L} x_{n-L+1}^*}; \tag{38}$$

$$c_{n|L} = c_{n-1|L-1} + K_{n|L} x_n^* \left(y_n^* - c_{n-1|L-1}^T x_n^* \right); \tag{39}$$

$$K_{n|L} = K_{n-1|L-1} - \frac{K_{n-1|L-1} x_n^* x_n^{*T} K_{n-1|L-1}}{1 + x_n^{*T} K_{n-1|L-1} x_n^*}; \qquad (40)$$

Here

$$x_{n-L+1}^* = \frac{x_{n-L+1} - X_{n|L}^T D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \qquad (41)$$

$$x_n^* = \frac{x_n - X_{n-1|L}^T D_{n-1|L}^{-1} r_{n-1,1}}{\sqrt{\alpha_{n,0}}}; \qquad (42)$$

$$y_{n-L+1}^* = \frac{y_{n-L+1} - Y_{n|L} D_{n|L}^{-1} r_{n-L+1,1}}{\sqrt{\alpha_{n-L+1,0}}}; \qquad (43)$$

$$y_n^* = \frac{y_n - Y_{n-1|L} D_{n-1|L}^{-1} d_{n-1}}{\sqrt{\alpha_n}}. \qquad (44)$$

5. Features of assessment for the case of stationary interference

In this case the output signal is converted taking into account the entered notations y_n^* can be written as follows

$$y_n^* = y_n - r_{1,n-1}^T R_{n-1|L-1}^{-1} Y_{n-1|L-1}. \qquad (45)$$

Calculating $M\left\{\overline{y}_n^* \overline{y}_{n-i}^*\right\}, i = 1, 2, \ldots,$ it is not difficult to show that all these mathematical expectations of the products are zero. Therefore, the centered component of the transformed process, which is determined by relation (45) is a discrete white noise, i.e. relation (44) describes the transformation of a random vector with correlated components into a random vector with uncorrelated components.

6. Conclusions

Thus, the problem of constructing a recurrent form of the current regression analysis algorithm is considered, which allows estimating unknown parameters in the presence of stationary correlated interferences.

The basic relations describing the processes of accumulation of new and dumping of obsolete information are obtained. It is shown that the considered algorithms use the transformation of a random vector with correlated components into a random vector with uncorrelated components.

7. References

[1] Widrow B. Adaptive switching circuits, / B. Widrow, M. Hoff // IRE WESCON Convention Record. Part 4. New York: Institute of Radio Engineers, 1960. – P.96–104.

[2] Либероль Б.Д. Исследование сходимости одношаговых адаптивных алгоритмов идентификации. / Б.Д. Либероль, О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2018. – 5. – С.19 – 32.

[3] Руденко О.Г. Регуляризованный алгоритм обучения адалины в задаче оценивания нестационарных параметров / О.Г. Руденко, А.А. Бессонов // Управляющие системы и машины. – 2019. –1. – С.22 – 30.

[4] Перельман И.И. Оперативная идентификация объектов управления./ И.И. Перельман // М.: Энергоиздат,1982. – 272 с.

[5] Руденко О.Г. Модифицированный алгоритм текущего регрессионного анализа в задачах идентификации и прогнозирования. / О.Г. Руденко, И.Д. Теренковский, А. Штефан, Г.А. Ода // Радиоэлектроника и информатика. – 1998. –№ 4(05). – С.58–61.

[6] Rudenko O.G. Adaline robust multistep training algorithm / O.G. Rudenko, O.O. Bezsonov // Control Systems and Computers. – 2020. – № 3. – P.15–27.

[7] Аведьян Э.Д. Рекуррентный метод наименьших квадратов при коррелированных помехах / Э.Д. Аведьян //Автоматика и телемеханика . –1975. –5. – С.67–75.

[8] Медведев Г.А. Рекуррентное оценивание при помощи коррелированных наблюдений / Г.А. Медведев // Автоматика и телемеханика . 1974. – 5. – С.110–116.

[9] Медведев Г.А. О рекуррентных оценках по коррелированным наблюдениям / Г.А. Медведев, Г.А. Хацкевич // Автоматика и телемеханика . – 1979. – 8. – С.69–75.

[10] Джонстон Дж. Эконометрические методы./ Дж. Джонстон – М.: Статистика,1980. – 444 с.

Використання еволюційних алгоритмів для навчання нейронних мереж

Олійник О.О. [1], Дудар З.В. [2] Зибіна К.В. [3]

[1] *ХНУРЕ, проспект Науки, 14, Харків, Харківська область, 61000, Україна*
[2] *ХНУРЕ,проспект Науки, 14, Харків, Харківська область, 61000, Україна*
[3] *ХНУРЕ,проспект Науки, 14, Харків, Харківська область, 61000, Україна*

Анотація

Нейронні мережі та генетичні (еволюційні) алгоритми, а також деякі інші, відносяться до загальної сфери знання - системам штучного інтелекту. Вони мають спільні риси: є необхідність у навчанні системи, для початку роботи потрібна деяка база фатів (вона ж навчальна вибірка), результатом роботи є, зазвичай, вибір рішення з заздалегідь встановленого переліку. Таким чином, метою даного дослідження є вивчення нейронних мереж, еволюційних алгоритмів і можливість застосування других для навчання перших. Застосування генетичного алгоритму з одного боку дозволяє значно спростити розробку структури нейронної мережі, а з іншого боку дозволяє значно скоротити витрати на навчання цієї мережі, в порівнянні з методом зворотного поширення помилки.

Ключові слова

Нейронні мережі, еволюційний алгоритм, навчання нейронної мережи, ІСТ-2021.

Using evolutionary algorithms to train neural networks

Olyynik O.O.[1], Dudar Z.V. [2] and Zibina K.V.[3]

[1] *KHNURE, Prospect Science, 14, Kharkiv, Kharkiv region, 61000, Ukraine*
[2] *KHNURE, Prospect Science, 14, Kharkiv, Kharkiv region, 61000, Ukraine*
[3] *KHNURE, Prospect Science, 14, Kharkiv, Kharkiv region, 61000, Ukraine*

Abstract

And neural networks and genetic (evolutionary) algorithms, as well as some others, belong to the general field of knowledge - artificial intelligence systems. They have common features: there is a need to train the system, to get started requires a certain base of veils (it is also a training sample), the result of the work is, usually, the choice of a solution from a predetermined list. Thus, the purpose of this study is to study neural networks, evolutionary algorithms and the possibility of using the latter to train the former. The use of a genetic algorithm on the one hand can greatly simplify the development of the structure of a neural network, and on the other hand can significantly reduce the cost of training this network, compared with the method of error backpropagation.

Keywords

Neural networks, evolutionary algorithm, neural network training., IST-2021

EMAIL: oleksandr.oliinyk1@nure.ua (A. 1); zoia.dudar@nure.ua (A. 2); kateryna.zybina@nure.ua (A. 3)
ORCID: 0000-0002-7734-3963 (A. 1); 0000-0001-5728-9253 (A. 2); 0000-0002-9955-4710 (A. 3)

1. Вступ

Нейронна мережа - це послідовність нейронів, з'єднаних між собою синапсами. Завдяки такій структурі, наприклад, машина отримує можливість аналізувати і навіть запам'ятовувати різну інформацію. Нейронні мережі також здатні не тільки аналізувати вхідну інформацію, а й відтворювати її зі своєї пам'яті. [1]

Найчастіше нейронні мережі вирішують завдання класифікації, тому загальноприйнятим в їх конфігурації є наступне: заздалегідь створюється база еталонних значень, кожному з яких відповідає єдиний нейрон, розташований на виході з мережі.

Нейронна мережа складається з штучних нейронів пов'язаних між собою. Кожен нейрон має кілька вхідних і один вихідний зв'язок. В процесі роботи значення вхідних змінних передаються по міжнейронних зв'язках і множаться на вагові коефіцієнти, отримані значення виважено підсумовуються в нейроні. Така структура дозволяє нейронним мережам, на відміну від лінійних методів порівняння, будувати нелінійні залежності, які більш точно описують набори даних.

Сучасні нейронні мережі - багатошарові, тобто нейрони в них об'єднані в шари, при цьому виходи нейронів попереднього шару є входами нейронів наступного. У кожному шарі обробка даних виконується паралельно.

Крім завдань класифікації нейронні мережі можуть бути використані для вирішення завдань прогнозування, пошуку прихованих закономірностей, стиснення даних, сегментування, передбачення тощо.

Як би не основною проблемою даного методу є зашумленість профілів-матриць збуджених нейронів вхідного шару. Тому вкрай необхідна наявність якісної навчальної вибірки. Якщо вона буде низької якості, то сильна зашумленість призведе до великої кількості помилок. Втім до того ж може призвести і надмірно великий розмір навчальної вибірки.

Хотілося б так само більш докладно розглянути завдання, що вирішуються нейронними мережами. В [1] їм дають такі визначення:

«Класифікація - розподіл даних за параметрами. Наприклад, на вхід надається список клієнтів і потрібно вирішити, кому з них надавати кредит, а кому ні. Цю роботу може зробити нейронна мережа, аналізуючи таку інформацію як: вік, платоспроможність, кредитна історія тощо.

Передбачення - можливість прогнозувати наступний крок. Наприклад, зростання або падіння акцій, ґрунтуючись на ситуації фондового ринку.

Розпізнавання - в даний час, саме широке застосування нейронних мереж. Використовується в Google, коли ви шукаєте фото або в камерах телефонів, коли він визначає положення вашого обличчя і виділяє його і багато іншого»

Основною перевагою нейронних мереж перед класичними алгоритмами є можливість не тільки слідувати заздалегідь закладеним правилами поведінки, а й, спираючись на власний досвід, виробляти нові складні знання і здійснювати власні відкриття.

Базовий принцип роботи нейронних мереж заснований на підсумовуванні отриманого досвіду, безпосередньо не відноситься до вирішення конкретного завдання. Після навчання на прикладах, наприклад написаних цифр, нейромережа може визначити значення написаних рукою чисел, яких до цього не бачила. Також нейромережі можуть виробляти знання для використання людиною і переносити їх в інші програмні продукти. Прикладами таких завдань є медична діагностика і пошук найкращих варіантів лікування з урахуванням персональних особливостей пацієнта.

При вирішенні завдань з використанням нейронних мереж існує дві основні проблеми: вибір оптимальної структури нейронної мережі і побудова ефективного алгоритму навчання нейронної мережі.

Говорячи про ці проблеми нейронних мереж, хотілося б процитувати [2]:

«Оптимізація нейронної мережі спрямована на зменшення обсягу обчислень за умови збереження точності рішення задачі на необхідному рівні. Параметрами оптимізації в нейронної мережі можуть бути:

1. Розмірність і структура вхідного сигналу нейромережі;
2. Синапси нейронів мережі. Вони спрощуються за допомогою видалення з мережі або завданням "потрібної" або "оптимальної" величини ваги синапсу;

3. Кількість нейронів кожного шару мережі: нейрон цілком видаляється з мережі, з автоматичним видаленням тих синапсів нейронів наступного шару, за якими проходив його вихідний сигнал;
4. Кількість шарів мережі.

Друга проблема полягає в розробці алгоритмів навчання нейромереж, що дозволяють за прийнятний час налаштувати їх на розпізнавання заданого набору вхідних образів ».

Обидві ці проблеми можуть бути вирішені з використанням генетичних алгоритмів. Розглянемо генетичний алгоритм більш докладно.

2. Генетичний алгоритм

Генетичний алгоритм - комп'ютерна програма, яка намагається емулювати реальний еволюційний процес. Або інакше, це алгоритм знаходження глобального екстремуму в багатоекстремальній функції, який полягає в паралельній обробці безлічі альтернативних рішень. У нього є декілька ключових змінних: розмір популяції, число поколінь (кількість ітерацій алгоритму), функція пристосованості і кількість мутацій.

У роботі даного алгоритму є так само декілька важливих нюансів. Його робота мало пов'язана з часом. При спробі знайти конкретну особину з певною пристосованістю може змінюватися різна кількість поколінь. У зв'язку з цим неможливо передбачити час вирішення конкретного завдання. Це пов'язано з використанням всередині алгоритму механізму випадкових чисел. Існує три класичних способу визначити точку виходу з алгоритму:

1. Якщо різноманітність в популяції відсутня або дуже мала.
2. Якщо досягнута задана пристосованість.
3. Якщо досягнута задана заздалегідь ітерація (покоління).

Для генерації нових поколінь використовуються прості методи: селекція, схрещування, мутація [2].

Оператор селекції відбирає особин за значенням їх функції пристосованості. Для цього може бути використаний метод рулетки або турнірного відбору. Метод рулетки - при такому відборі члени популяції з більш високою пристосованістю будуть вибиратися частіше, ніж особини з низькою. Метод турніру - з популяції вибираються, випадковим чином, кілька рядків і між ними проводиться турнір, тобто з них вибирається найкращий, який потім і потрапляє в підсумковий масив. Існує ще кілька методів селекції, але в цій статті розглядати їх ми не будемо.

Оператор схрещування, він же оператор кросовера. Може бути однокрапковим і багатокрапковим. У однокрапкового, геном особини розрізається в одній точці і нащадку дістаються, наприклад, ліва половина одного з батьків і права іншого. У багатокрапковому часто застосовують метод гребінця. Це виглядає так: якщо довжина геному особини n, то число точок розриву дорівнює (n-1), при цьому, нащадку гени дістаються наступним чином: парні від першого батька, непарні від другого. Цей оператор виконується з найбільшою ймовірністю серед інших. Зазвичай ймовірність його виконання ставлять > 60%. Як мені здається, найкраще візуально змогли його показати в статті [2]. Дивіться на рисунку 1.

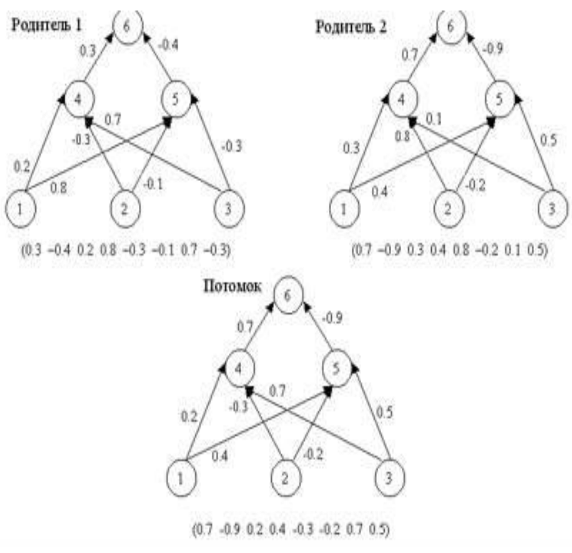


Рисунок 1: приклад оператора схрещування

Оператор мутації - стохастична зміна частини хромосом. Він необхідний для захисту від передчасної збіжності і зациклення в локальному екстремуму. Це досягається за рахунок того, що кожен ген особини, що зазнала мутації змінюється на інший ген. Правда, з досить малою вірогідністю. Приклад показаний на малюнку 2.

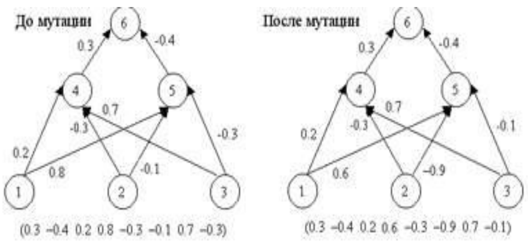


Рисунок 2: Приклад виконання оператору мутації

3. Використання генетичного алгоритму для навчання нейронних мереж

Генетичний алгоритм для навчання нейронної мережі можна використовувати двома способами [4]:

1. Шукати безпосередньо комбінацію ваг синаптичних зв'язків нейронної мережі;
2. Шукати необхідні параметри нейронної мережі: кількість шарів і нейронів у них, функцію активації тощо.

Перший спосіб дозволяє виключити процес навчання, оскільки сам його замінює. А в другому випадку генеруються різні мережі і з використанням обраного методу навчання - навчаються. Природно, другий спосіб набагато більш ресурсовитратний, однак, він дозволяє відносно швидко знайти нейронну мережу меншого розміру, яка все ще вирішує поставлену задачу.

Таким чином, резюмуємо. У застосуванні генетичного алгоритму для навчання нейронних мереж є наступні істотні плюси:

1. Прискорене вивчення простору пошуку, за рахунок використання безлічі паралельних рішень, а не одного послідовно застосованого;
2. малочутлива до зростання розмірності безлічі оптимізації.

А також цей підхід має наступні недоліки:

1. При використанні цього алгоритму існує необхідність в істотних обчисленнях, що означає і великі часові витрати;
2. Недоліком є приналежність до класу евристичних алгоритмів.

Метод зворотного поширення помилки - класичний метод навчання нейронної мережі. Якраз його альтернативою і є генетичні алгоритми. Навчання цим методом зводиться до підбору значень ваг прямоспрямованої нейронної мережі, ґрунтуючись на методах найшвидшого спуску. Основна проблема цього методу - «залипання» в точках локального мінімуму.

На рисунку 3 надано приклад нейронної мережі, що виконує операцію XOR.

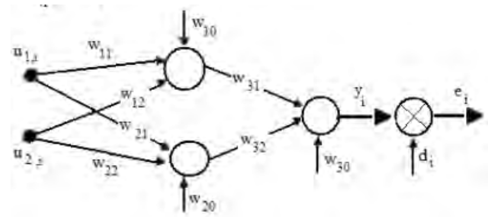


Рисунок 3: Нейронна мережа, що реалізує операцію XOR[5]

Для неї потрібно підібрати оптимальні ваги w_{10}, w_{11}, w_{12}, w_{21}, w_{22}, w_{31}, w_{33}, w_{10}, w_{20}, w_{30}, що мінімізують значення цільової функції, яка визначається середньоквадратичною похибкою (формула 1):

$$E = \frac{1}{2}\sum_{i=1}^{4}(d_i - y_i)^2 \qquad (1)$$

Параметрами завдання є ваги, які і являють собою можливе рішення. Відповідно, при використанні генетичного алгоритму значення цих ваг і будуть генами кожної конкретної особини.

Так само хотілося б розглянути експеримент проведений для статті [2].

В цьому експерименті, автори порівнювали ефективність навчання нейронної мережі з використанням метода зворотного поширення помилки і за допомогою генетичних алгоритмів.

Тепер трохи про суть експерименту. Автори взяли повнозв'язну двошарову нейронну мережу з наступними характеристиками: 4 вхідних сигнали, 3 нейрона прихованого шару, 1 вихідний сигнал. Навчання вважалося завершеним, коли на всьому протязі навчальної вибірки реальний вихідний сигнал дорівнював потрібному значенню.

Що стосується результатів експерименту, то тут хотілося б процитувати самих авторів: «Для навчання нейронної мережі досить було виконати 500 генерацій за допомогою генетичних алгоритмів. Алгоритмам зворотного поширення потрібно у багато разів більше число ітерацій - близько 500000 ітерацій, де одна ітерація - це повний перерахунок всіх навчальних даних (вагові коефіцієнти, похибка, значення виходів нейронної мережі). Зауважимо, що при цьому дві генерації в генетичному алгоритмі еквівалентні одній ітерації алгоритму зворотного поширення, так як зворотне поширення на даному навчальному прикладі складається з двох частин - прямий прохід (обчислення виходу мережі і помилки) і зворотний прохід (налагодження ваг). Генетичні алгоритми виконує тільки першу частину. Таким чином, дві генерації займають менше часу, ніж обчислення єдиної ітерації алгоритму зворотного поширення »[2]. Результат також для наочності представлений на рисунку 4.

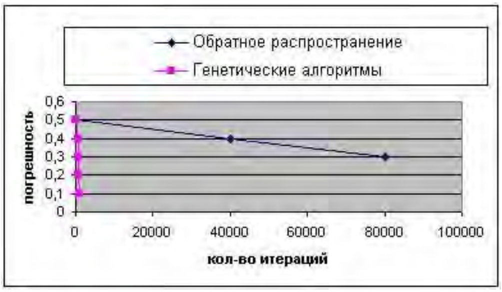


Рисунок 4: Результат порівняння швидкості навчання нейронної мережі за допомогою зворотного розповсюдження та генетичного алгоритму [2].

4. Висновки

На сьогоднішній день нейронні мережі є одним з пріоритетних напрямків в області штучного інтелекту. У нейронних мереж багато важливих властивостей. І ключовою є здатність до навчання. Навчання нейронної мережі в першу чергу це зміна «значущості» синаптичних зв'язків між нейронами.

Генетичні алгоритми застосовуються в багатьох сферах комп'ютерних технологій. І що особливо важливо для нас, вони використовуються при навчанні нейронних мереж.

В даний час нейронні мережі та генетичні алгоритми перейшли з перспективних методів вирішення завдань в доступні і широко застосовуються. Основний напрямок при розробці нових методів і вирішенні завдань вже існуючими пов'язано зі скороченням витрат часу на вибір структури і навчання нейронної мережі. Як показують приклади наведені вище, застосування генетичного алгоритму з одного боку дозволяє значно спростити розробку структури нейронної мережі, а з іншого боку дозволяє значно скоротити витрати на навчання цієї мережі, в порівнянні з методом зворотного поширення помилки. Ні, це не має на увазі що генетичні алгоритми завжди краще методу зворотного поширення чи інших методів використовуваних на етапах проектування і навчання нейронних мереж. Однак при вирішенні таких завдань, як нейроуправління, де нейронні мережі використовуються для управління складними системами - йому дуже складно підшукати конкурентів.

5. Література References

[1] TN Kondrat'ev Evolutionary Computing: Neural Networks and Genetic Algorithms // XXII International Conference on Soft Computing and Measurements. St. Petersburg: LETI, 2019

[2] Yu.I. Solomka Application of genetic algorithms for training neural networks. 15-17 April 2004 р. Materials of the conference. Part 1. - Vinnytsia: BI MAUP, 2004. - 85-90.

[3] Melanie Mitchell. An Introduction to Genetic Algorithms. Massachusetts Institute of Technology, 1998. - 280.

[4] Shumkov Evgeniy Aleksandrovich, Chistik Igor Konstantinovich Using genetic algorithms for training neural networks // Scientific journal of KubGAU. 2013. №91. URL: https://cyberleninka.ru/article/n/ispolzovanie-geneticheskih-algoritmov-dlya-obucheniya-neyronnyh-setey (дата обращения: 10.08.2021).

[5] Application of genetic algorithms for training neural networks // dspace.susu.ru URL: http://dspace.susu.ac.ru/handle/0001.74/1448 (дата обращения: 10.08.2021).

КОРЕКЦІЯ СИСТЕМИ ПІДГОТОВКИ В СПОРТИВНИХ ЄДИНОБОРСТВАХ ЗАСОБАМИ СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Андрій Литвиненко[1], Ганна Грохова[2], Юлія Дорофєєва[3]

[1,2,3] *Харківський національний університет радіоелектроніки, пр. Науки. 14, Харків, 61166, Україна*

Анотація

Представлено дослідження ефективності застосування сучасних інформаційних технологій для прогнозу та корекції стану нелінійних, надскладних, ієрархічних, відкритих систем на прикладі українських національних спортивних єдиноборств. Показано особливості функціонування синергетичних систем, які знаходяться в режимі із загостренням.

Ключові слова

Інформаційні технології, режим із загостренням, крапка біфуркації, корекція системи спортивної підготовки.

CORRECTION OF TRAINING SYSTEM IN SPORTS MARTIAL ARTS BY MEANS OF MODTRN INFORMATION TECHNOLOGIES

Andrii Lytvynenko [1], Grokhova Ganna [2] and Yuliia Dorofieieva [3]

[1, 2, 3] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract

The research of efficacy of application of modern information technologies for prediction and correction of nonlinear, superset, hierarchical, open systems on the application of Ukrainian national sports single combats is presented. The peculiarities of functioning of synergetic systems that are in the regime with overgrowth are shown.

Keywords

Informational technologies, mode with a head start, bifurcation edge, correction of the system of sports training

1. Вступ

В останні десятиріччя інформаційні технології розвиваються швидким темпом і суттєво впливають на ефективність наукових досліджень. Значно зростають можливості отримання інформації, її аналізу, перетворенню та архівуванню. Особливо цінним є застосування інформаційних технологій в процесі дослідження надскладних систем різної природи [2]. Однією з важливих тенденцій розвитку сучасної науки є суттєве зростання кількості міждисциплінарних досліджень. Не залишилася осторонь цієї тенденції й наука о спорті [5].

EMAIL: andrii.lytvynenko@nure.ua (A. 1); anna0282gr@gmail.com (A. 2);
yuliia.hubnytska@nure.ua (A.3)
ORCID: 0000-0002-0352-9511 (A. 1); 0000-0003-3562-1272 (A. 2);
0000-0002-6246-042X (A. 3)

Спортивні єдиноборства [4], а зокрема різновиди українських національних єдиноборств застосовують методику досліджень тренувальної і змагальної діяльності побудовану на основі сучасних інформаційних технологій та методології синергетичного міждисциплінарного дослідного підходу [7].

2. Ефективність змагальної діяльності в українських національних видах спортивних єдиноборств

Спортивний зміст змагальних поєдинків в українських національних видах єдиноборств полягає в тому, що бійці наносять по супротивнику удари руками і ногами та захищаються від них. Також правилами змагань дозволена борцівська техніка ведення поєдинку: кидки, больові і задушливі прийоми, а також утримання. Ефективність застосованих технічних дій визначається суддями за стандартизованими критеріями [1]. Залікові бали (1 бал за 1 удар) нараховуються спортсмену боковими суддями та рефері за удари, які досягли цілі і мали великі показники потужності. Тобто відбувається кількісне оцінювання бойових дій. Уразі якщо удар привів до нокдауну супротивника нараховується якісна оцінка (половина перемоги) з кількісним еквівалентом 6 балів). Перемога присуджується спортсмену у разі, якщо він набрав більшу кількість балів ніж супротивник, або завершив поєдинок нокаутом чи больовим та задушливим прийомом.

2.1. Періоди різкого загострення загальної боротьби

Дослідження, проведені нами з 2012 по 2021 роки показують наявність в спортивних поєдинках окремих періодів, в яких різко загострюється конфліктна взаємодія між єдиноборцями. Спортсменів ми при цьому розглядаємо у якості нелінійних, надскладних, ієрархічних, відкритих систем, які увійшли в режим із загостренням. При цьому суттєво змінюється управління, енергообмін, спортивна тактика та арсенал атакуючих та захисних дій. Експертний аналіз бойових дій спортсменів в ці періоди та їх оцінювання суддями показує, що бойова ефективність в періоди різкого загострення змагальної боротьби суттєво впливає на результат поєдинку в цілому .

2.1.1. Корекція тренувального процесу із застосуванням сучасних інформаційних технологій

Нами розроблено методичний підхід до організації тренувального процесу при якому акцентується увага на зростання ефективності бойових дій спортсменів в періоди різкого загострення змагальної боротьби[6]. Для його реалізації засобами математичної статистики та сучасних інформаційних технологій визначаються коефіцієнти ефективності атакуючих і захисних дій в цілому за поєдинок та в періоди різкого загострення спортивної боротьби. Окремо в періодах різкого загострення спортивної боротьби визначається коефіцієнт ефективності бойових дій в крапці біфуркації. Нами запропоновано для визначення ефективності бойових дій методом експертних оцінок визначати удари, які суттєво вплинули на хід та кінцевий результат поєдинок в періоди різкого загострення змагальної боротьби.

Окремим напрямком є дослідження взаємозв’язку показників серцевих циклів за пульсом та показників артеріального тиску для діагностики та корекції психічного стану спортсменів [2]. На початку, в середині та на закінченні підготовчого періоду підготовки за допомогою тонометра визначаються показники пульсу та артеріального тиску до та після навантаження різної спрямованості. Отримані дані, за допомогою спеціально розробленого членами студентського наукового гуртка ХНУРЕ програмного забезпечення співставляються із результатами спортивної діяльності на змаганнях різного рівня.

Встановлено, що в експериментальній групі, в періоди різкого загострення змагальної боротьби, 35.5 % прямих ударів руками доходять до цілі. З них 62 % отримували оцінку суддів, з яких 7 % приводили до повної або значної втрати бойових якостей супротивника.

Отримані дані дають змогу для корекції ключових параметрів різних сторін спортивної підготовки єдиноборців: фізичної, технічної, тактичної, психологічної та інтегральної [7].

3. Висновки

Застосування методології сучасного міждисциплінарного дослідного підходу до вивчення нелінійних, надскладних, ієрархічних, відкритих систем дає змогу отримати дані о ключових параметрах боротьби в спортивних єдиноборствах ударного та змішаного типів.

Аналіз даних об ефективності бойових дій в періоди різкого загострення змагальної боротьби отриманих за допомогою сучасною апаратурою відеофіксації з наступною обробкою із застосуванням спеціально розробленого програмного забезпечення дає змогу робити результативну корекцію тренувальної та змагальної діяльності.

4. Подяки

Автори дослідження висловлюють подяку організаціям, які сприяли отриманню даних о параметрах спортивної боротьби в ударних та змішаних видах єдиноборств: громадській організації «Українська федерація хортингу», громадській організації «Харківська обласна федерація хортингу», громадській організації «Федерація Окінавького Годзю-рю карате-до і Рюкю кобудо Джундокан України», громадській організації «Харківська обласна федерація годзю-рю карате».

5. Література References

[1] Губницька Ю. С., Оцінювання технічних дій в національному виді єдиноборств – хортингу у змагальному розділі «Двобій» / Ю.С. Губницька, А.М. Литвиненко // Теорія і методика хортингу : зб. Наукових праць. – К.: Палівода А.В. 2018. – Вип. 8. С. 132-138.

[2] Губницька Ю. С., Інформаційні технології в ідентифікації психологічного стану єдиноборців в процесі змагальної діяльності / Ю.С. Губницька, А.М. Литвиненко // Матеріали 8-й Міжнародної науково-технічної конференції «Інформаційні системи та технології», Коблеве - Харків 9-14 вересня 2019 р. – Х.: ХНУРЕ, С. 179-180.

[3] Литвиненко А.М., Співвідношення тренувальних засобів загальної та спеціальної спрямованості в річному циклі спортивної підготовки кваліфікованих єдиноборців в хортингу. Теорія і методика хортингу : зб. Наукових праць. – К.: Палівода А.В. 2015. – Вип. 3. С. 205-210.

[4] Литвиненко А.М., Общая теория единоборств. Сб. Спортивные единоборства: теория, практика и перспективы развития. – Х.: 2004. – С. 39 – 41.

[5] Griban, G., Influence of sambo wrestling training on students' physical fitness. Sport Mont, 19 (1), 89-95. doi: 10.26773/smj.210219. (2021).

[6] Griban, G., Formation of the students' volitional qualities in the process of physical education. International Journal of Human Movement and Sports Sciences, 8 (6), 505-517. doi: 10.13189/saj.2020.080625. http//www.hrpub.org

[7] Malynskyi, I., Horting as a Means of Physical Education for Pupils and Students Problems and prospects of physical culture and sports development and healthy lifestyle formation of different population groups. Monograph. Opole: The Academy of Management and Administration in Opole; pp.286, illus., tabs., bibls. Part 5. 2020. Horting as a Means of Physical Education for Pupils and Students (pp. 226–273).

Мультиагентные системы поддержки принятия решений в проектах по инжинирингу бизнеса

Борис Железко [1]

[1] *Белорусский национальный технический университет, пр-т Независимости, 65, Минск, 220013, Беларусь*

Аннотация
В данной работе проведено обобщение методов построения мультиагентных систем поддержки принятия решений применительно к задачам маркетингового инжиниринга и реинжиниринга бизнеса.

Ключевые слова
Инжиниринг, реинжиниринг, мультиагентная система, система поддержки принятия решений, обобщенный объект, модель предметной области

Multi-agent decision support systems for business engineering projects

Boris Zhalezka [1]

[1] *Belarusian National Technical University, Nezavisimosti ave., 65, Minsk, 220013, Belarus*

Abstract
This paper summarizes the methods for building multi-agent decision support systems in relation to the tasks of marketing engineering and business reengineering.

Keywords
Engineering, reengineering, multi-agent system, decision support system, generalized object, subject area

1. Введение

Маркетинговый инжиниринг и реинжиниринг бизнеса (МРБП) – разновидность реинжиниринга бизнес-процессов (РБП), ориентированная на достижение основных целей маркетинговой деятельности (расширение объема продаж и рынков сбыта; увеличение занимаемой роли на рынке; рост прибыли и обеспечение обоснованности принимаемых руководством фирмы решений в области производственно-сбытовой и научно-технической деятельности). Этим он отличается от, например, стратегического корпоративного РБП, целью которого является поиск стратегического инвестора.

Исследования в сфере мультиагентных систем поддержки принятия решений (МА СППР) и их применения для цифровизации процессов управления проектами по МРБП ведутся достаточно давною. Однако, в связи с исключительно высокой динамикой развития информационных технологий, разработки в области МА СППР требуют постоянной адаптации к уровню достижений научно-технического прогресса и изменениям во внешней бизнес-среде.

Несмотря на имеющиеся разработки в данной области, существующие технологии поддержки принятия решений предполагают адаптацию только отдельных компонент СППР и не обеспечивают адаптацию модели предметной области. Это приводит к использованию неактуальных и

EMAIL: boriszh@bntu.by (A.1); ORCID: 0000-0003-4166-1035 (A.1)

недостоверных данных в СППР, что отрицательно сказывается на эффективности принятия решения в условиях быстро меняющейся внешней среды.

В цикле проведенных исследований под руководством автора данные проблемы в значительной мере решены благодаря разработке механизмов адаптации моделей предметной области к условиям задач принятия решений и своевременного обновления необходимых для этого баз моделей, данных, знаний [1, 2], а также коллекций прецедентов (предметных коллекций [3]).

Цель данной статьи – обобщение результатов цикла работ по развитию методологии проектирования информационно-аналитических СППР. Сравнительный анализ различных подходов в этой области выполнен в работах [4-5].

2. Концепция обобщенного объекта

Обычно объект О может представляться в виде тройки

O = <Data, Met, Mes>, где Data, Met, Mes, соответственно, множество внутренней информации объекта (данных), множество его собственных процедур для манипулирования этими данными (методов) и внешний интерфейс для взаимодействия с другими объектами предметной области (ПрО), например, допустимое множество сообщений о событиях вне и внутри ПрО. Однако необходимость учета развития МА СППР требует более общего и гибкого механизма их описания и моделирования. Такой механизм может быть построен на основе дальнейшего обобщения объектно-ориентированного подхода и, в частности, собственно понятия объекта в концептуальной модели ПрО [6–9].

Введено понятие обобщенного объекта GO = <Data, Met, Model, Knowl, Mes, Link>, в котором, кроме данных Data, методов и свойств Met дополнительно инкапсулируются модели Model, формализованные знания Knowl, а также возможные сообщения Mes и связи Link для взаимодействия данного объекта с другими объектами обобщенной ПрО. Такую модель обобщенной ПрО можно рассматривать как разновидность мультиобъектой нейронной сети, если к обычному способу взаимодействия объектов, путем обмена сообщениями, добавить возможность активизации объектов (передачи возбуждения) по связям, снабженным весами (приоритетами) с указанием для каждого экземпляра объекта значения уровня порога срабатывания.

Последние могут изменяться в зависимости от этапа развития системы, решаемой задачи, накопленного статистического опыта решения подобных задач, наиболее часто решаемых задач и т. п. Взаимодействие между GO может осуществляться посредством: сообщений; изменения структуры связей системы GO, которая отделена от функциональной их части и представлена динамическим списком связей, а также изменения весов (фильтрации) связей и значений порогов срабатывания.

Каждый GO может иметь несколько состояний и переходить из одного состояния в другое в зависимости от поступающих сообщений, которые являются результатом деятельности других GO. При этом GO изменяет свое состояние, когда значение его возбуждения превышает некоторый ненулевой порог срабатывания.

В целом такую модель ПрО можно рассматривать как иерархию абстракций, представленную классами обобщенных системных, проблемных и пользовательских объектов. Состояние ПрО фактически зависит от состояния каждого GO и очередей сообщений на входе и выходе этих GO. Последнее можно рассматривать как базу фактов о событиях, на основании которых можно определить машину вывода для интерпретации существующих и порождения новых фактов в процессе имитации функционирования моделируемой системы.

3. Концепция динамического управления требованиями к МА СППР

Основные группы специалистов-экспертов различной квалификации (например, потребители, маркетологи, изготовители и разработчики) формулируют требования к МА СППР в своих профессиональных терминах ПрО путем их объектно-ориентированной классификации (перечисления участвующих в решении задач классов объектов, их свойств, взаимосвязей и поведения).

Основным принципом классификации является спецификация классов объектов на основе множества внутренних свойств, присущих объектам класса. Затем эти требования последовательно детализируются до тех пор, пока проект МА СППР не будет полностью описан в терминах базовых объектов используемых инструментальных средств. К настоящему времени концепция динамического управления требованиями к МА СППР без ограничения общности уже внедрена в практику проектирования ряда мультиагентых систем поддержки принятия решений, где она получила наибольшее теоретическое развитие и детализацию [2, 7–9]. При этом выявлено, что развитие МА СППР может рассматриваться как процесс взаимного развития и адаптации двух типов систем. Первый из них – это формализованные системы требований UPD(t), которые представляют собой систему баз знаний группы специалистов различного профиля и разной квалификации о ПрО. Второй тип – слабо формализованные системы трансформации требований F(UPD(t), t) по мере развития знаний вышеназванных групп специалистов и научно-технического прогресса, в частности, новых информационных технологий.

4. Выводы

На основе концепций обобщенного объекта и динамического управления требованиями к МА СППР разработан комплекс инструментальных методов и методик поддержки принятия эффективных управленческих стратегических и инвестиционных решений, который дополнен методами автоматизации рутинных и творческих операций интерактивного построения моделей многокритериального выбора наилучшей альтернативы из заданного множества альтернатив (объектов, стратегий), оцениваемых по ряду критериев (показателей эффективности, качества).

Применение данных результатов позволило на порядок уменьшить сроки создания моделей (проводить интерактивное моделирование и прогнозирование), снизить требования к квалификации пользователей любого ранга в области информационных технологий и моделирования, обеспечить возможность непосредственного личного участия в этой процедуре первых лиц организаций, сделать ее более «прозрачной», а результаты более обоснованными и объяснимыми.

Кроме того, существенно развиты методики оценки эффективности [10-11] и качества [12] МА СППР, основанные на динамическом анализе степени удовлетворения требований различных целевых групп специалистов-экспертов, а также разработаны методические рекомендации по внедрению МА СППР в проекты по стратегическому корпоративному [9, 13] и маркетинговому реинжинирингу [8, 14-15], которые позволяют спланировать и реализовать технические и организационные процессы внедрения МА СППР в задачи реинжиниринга, а также осуществить обоснованный выбор МА СППР, необходимых для повышения эффективности результатов проведения проектов.

Разработано математическое и программное обеспечение семейства информационно-аналитических (и мультиагентных) СППР, которые по основным функциональным характеристикам соответствуют международным аналогам, а по некоторым параметрам (простота освоения, используемое математическое обеспечение и др.) превосходят их, а также мета-технология построения методик анализа и рейтингования экономических объектов, которые могут быть использованы в проектах по маркетинговому инжинирингу и реинжинирингу бизнеса, а также в научной деятельности для создания методик и методов поддержки принятия решений на основе рейтингования различных объектов и субъектов экономической деятельности.

Предлагаемые разработки соответствует уровню развития методического и технологического обеспечения, требуемого в проектах по МРБП в современных условиях, когда рынок консалтинговых услуг в этой сфере еще только формируется.

Дальнейшим этапом работ предполагается проведение экспериментальной количественной оценки эффективности использования мультиагентных систем.

5. Литература

[1] Б.А. Железко, Системы поддержки принятия решений: вопросы создания и

примеры использования, под ред. А.Н. Морозевича, Минск, КИВТ НАН Беларуси, 1998.
[2] K. Navitskaya, B. Zhalezka, Information and Analytical Support of Decision-Making Procedures in Strategic Corporate Reengineering, Eastern European Journal of Regional Studies (2016), Volume 2, Issue 2, December 2016, pp. 41-49.
[3] Х.Э.Р.М. Виссия, В.В. Краснопрошин, А.Н. Вальвачев, Принятие решений в информационном обществе: учебное пособие, СПб., Лань, 2019.
[4] Ю.Ф. Тельнов, Компонентная методология реинжиниринга бизнес-процессов на основе управления знаниями, диссертация на соискание ученой степени доктора экономических наук по специальности 08.00.13 – математические и инструментальные методы экономики, Москва, МЭСИ, 2003, 339 с.
[5] Т.А. Ермакова, Б.А. Железко, Е.А. Корначенко, Стратегический корпоративный реинжиниринг процессов управления группой компаний (2010), Т.А. Ермакова, Научные труды БГЭУ (2010), с. 128 – 134.
[6] К.А. Аксенов, Ван Кай, Разработка и применение метода реинжиниринга бизнес-процессов на основе мультиагентного моделирования, монография, Ульяновск: Зебра, 2016.
[7] H. Karkanitsa, Adaptive Decision Support Systems, In: Proceedings of the Internernational Conference Pattern Recognition and Information Processing (PRIP'2019), Minsk, 21–23 May 2019, pp. 342–345.
[8] Б.А. Железко, Мультиагентные системы поддержки принятия решений в проектах по маркетинговому инжинирингу бизнеса, Міжнародна науково-практична конференція «Застосування інформаційних технологій у підготовці та діяльності сил охорони правопорядку», Збірник тез доповідей (м. Харків, 17 березня 2020 р.), Харків, 2020, с. 168 – 169.
[9] Б.А. Железко, Г.Н. Подгорная, Методическое и инструментальное обеспечение стратегического корпоративного реинжиниринга, Научные труды БГЭУ (2018), Вып. 11, с. 171 – 178.
[10] Б.А. Железко, О.А. Синявская, Оценка эффективности баз правил и моделей в системах поддержки принятия решений, 17 міжнародна конференція з автоматичного управління «Автоматика-2010», тези доповідей, том 2, Харків: ХНУРЭ, 2010, с. 102-104.
[11] Б.А. Железко, О.А. Синявская, Методы оценки эффективности интеллектуальных компонент систем поддержки принятия решений, Научные труды БГЭУ (2010), с. 139 – 144.
[12] O.A. Siniavskaya, B.A. Zhelezko, Creation of the rating of stock market analytical systems on the base of expert qualitative estimations, in: Proceedings of The Second International Conference on Fuzzy Sets and Soft Computing in Economics and Finance (FSSCEF 2006), Saint-Petersburg, 2006, pp. 208—215.
[13] Б.А. Железко, Теоретические основы и технологии построения систем информационно-аналитической поддержки принятия решений в стратегическом корпоративном реинжиниринге, Информационные системы и технологии: материалы 4-й Международной научно-технической конференции, Харьков, 21-27 сентября 2015 г.: тезисы докладов, редкол.: А.Д. Тевяшев (отв. ред.) и др., НТМТ, 2015, с. 163–164.
[14] Б.А. Железко, Инструментальное и методическое обеспечение маркетингового реинжиниринга бизнеса, в сборнике материалов 15-го Международного научного семинара «Мировая экономика и бизнес-администрирование малых и средних предприятий», проводимого в рамках 17-й Международной научно-технической конференции «Наука – образованию, производству, экономике», 24-25 января 2019 года, Минск: Право и экономика, 2019, с. 117-118.
[15] V. Siniauskaya, N. Viarbitsky, B. Zhalezka, Using of the concept of algorithmic marketing and machine learning in the marketing data processing, in: Proceedings of the 14th International Conference on Pattern Recognition and Information Processing (PRIP'2019), 21-23 May 2019, BSUIR, Minsk, "Bestprint", 2019, pp. 380 – 383.

Алгоритмы работы мультиагентной системы поддержки принятия решений в сетевой торговле

Борис Железко [1], Ольга Синявская [2]

[1] *Белорусский национальный технический университет, пр-т Независимости, 65, Минск, 220013, Беларусь*
[2] *Белорусский государственный экономический университет, Партизанский пр-т, 26, Минск, 220070, Беларусь*

Аннотация
В данной работе рассмотрен пример проектирования мультиагентной системы поддержки принятия решений в сетевой торговле на базе платформы Java Agent Development Framework (JADE). Разработаны алгоритмы поведения 4 интеллектуальных агентов, участвующих в процессе торгов. Приведены примеры результатов работы системы в виде протоколов коммуникации агентов. Представленные в статье алгоритмы могут быть использованы при компонентном проектировании и разработке различных торговых мультиагентных систем независимо от используемой агентной платформы.

Ключевые слова
Агент, мультиагентная система, система поддержки принятия решений, Java Agent Development Framework

Algorithms of Multi-Agent Decision Support Systems Work in the Net Trading

Boris Zhalezka [1], Volha Siniauskaya [2]

[1] *Belarusian National Technical University, Nezavisimosti ave., 65, Minsk, 220013, Belarus*
[2] *Belarusian State Economic University, Partisan ave., 26, Minsk, 220070, Belarus*

Abstract
This paper considers an example of application of multi-agent decision support system in the sphere of net trading. Algorithms of 4 intellectual trading agents behavior are developed. The examples of results of trading multi-agent system work are shown as the protocols of agents' communication. The algorithms represented in the paper can be used for component design and developing of different trading multi-agent decision support system independently of the used agent platform.

Keywords
Agent, multi-agent system, decision support system, Java Agent Development Framework

1. Введение

Исследования, связанные с использованием агентов и мультиагентных систем поддержки принятия решений (МА СППР), являются одними из интенсивно развивающихся в области искусственного интеллекта. Мультиагентные системы представляют собой совокупности интеллектуальных агентов, осуществляющих поиск данных и процедур, подходящих для решения поставленных пользователем задач, путем кооперации в процессе выработки этих решений. МА СППР характеризуются многообразием компонент, сложностью архитектуры, множеством вариантов взаимодействия агентов и разнообразием факторов внешней среды [1]. Цель данной статьи – разработать алгоритмы работы МА СППР в сфере сетевой торговли.

EMAIL: boriszh@bntu.by (A. 1); volha.siniauskaya@outlook.com (A. 2) ORCID: 0000-0003-4166-1035 (A. 1); 0000-0002-5635-1796 (A. 2)

2. Пример мультиагентной системы в сетевой торговле

Мультиагентные системы имеют широкую сферу применения, практически они могут использоваться в любой предметной области, где присутствуют и взаимодействуют между собой участники с определёнными ролями и функциями. Одной из востребованных сфер применения является электронная торговля [2]. В рассматриваемом примере представлен один из случаев сетевой торговли, когда необходимо реализовать одинаковые товары, предлагаемые разными продавцами, не связанными между собой. При этом требуется организовать централизованную доставку покупателям приобретенного товара. Частный пример такой ситуации – разовая или сезонная покупка скоропортящейся продукции.

МА СППР для сферы торговли реализована на базе агентной платформы JADE (Java Agent DEvelopment Framework) [3-4]. Она включает в себя 4 агента (TraderAgent, EmployerAgent, MajorBuyerAgent, BuyerAgent), назначение и алгоритмы их работы будут описаны далее. Интерфейс одного из агентов приведен на рисунке 1.

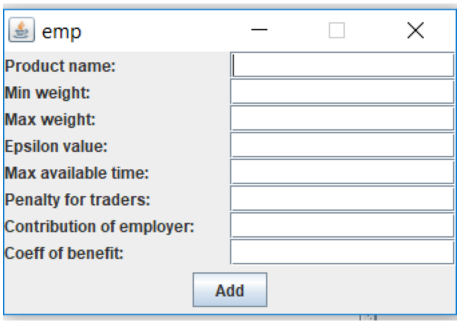


Рисунок 1: Пример интерфейса интеллектуального агента EmployerAgent

2.1. Алгоритм поведения агента TraderAgent

Агент TraderAgent имеет определенный объем некоторого продукта, который он хочет продать за максимальную цену агенту EmployerAgent. Алгоритм поведения TraderAgent включает следующие шаги.

Шаг 0. Найти всех доступных EmployerAgent агентов, занимающихся покупкой такого же товара. Если таковых нет, то повтор шага. Если такие есть, то всем найденным агентам послать сообщение о запросе, какой у них коэффициент прибыли от сделки. Перейти к шагу 1.

Шаг 1. Получить ответы от всех агентов EmployerAgent и выбрать среди них лучший. Если лучший до сих пор null, то перейти к шагу 0. Если лучший не равен null, то перейти к шагу 2.

Шаг 2. Отправить лучшему EmployerAgent предложение о сделке. Перейти к шагу 3.

Шаг 3. Принять ответ от EmployerAgent. Если сообщение положительное, проверить, удовлетворяет ли окончательное предложение начальным параметрам (минимальному проценту прибыли). Если да, то послать утвердительное сообщение о завершении сделки со своей стороны и перейти к шагу 4, если нет – перейти к шагу 0. Если сообщение отрицательное – перейти к шагу 0.

Шаг 4. Принять окончательный ответ от EmployerAgent. Если сообщение положительное, то процесс завершен, иначе – перейти к шагу 0.

2.2. Алгоритм поведения агента EmployerAgent

Агент EmployerAgent покупает у агентов TraderAgent некоторый объем товара по выгодному для себя тарифу. Затем пытается договориться о его продаже с MajorBuyerAgent, также по выгодной себе цене. В случае неудачных переговоров пересматривает свою начальную цену. Алгоритм поведения EmployerAgent включает следующие шаги.

Шаг 0. В течение времени maxAvailabilityTime отправлять ответ о своем тарифе агентам TraderAgent и принимать предложения о сделке от агентов TraderAgent с учетом перегрузки транспортного средства. Если время принятия предложений истекло или текущий объем товара равен максимально возможному, то перейти к шагу 1.

Шаг 1. Если текущий объем продукта больше либо равен минимально установленному, то выслать свое предложение агентам TraderAgent, которые участвовали в сделке, с учетом своего тарифа. Перейти к шагу 2. Если текущий объем товара находится в промежутке [minWeight–epsilon, minWeight), то выслать свое предложение агентам TraderAgent, которые участвовали в сделке, с учетом своего тарифа и штрафа за недобор до минимального объема. Перейти к шагу 2. Если текущий объем товара меньше, чем minWeight–epsilon, то выслать участвующим в сделке агентам отказы. Перейти к шагу 0.

Шаг 2. Принять от всех TraderAgent, участвующих в сделке, согласия либо отказы. Пересчитать текущий объем товара. Перейти к шагу 3.

Шаг 3. Если текущий объем продукта все ещё находится в приемлемых пределах, оповестить всех TraderAgent, участвующих в сделке, о её успешном завершении, перейти к шагу 4. Иначе оповестить об отмене сделки и перейти к шагу 0.

Шаг 4. В течение установленного времени maxAvailabilityTime отвечать на запросы о текущем объеме товара и его стартовой цене агентам MajorBuyerAgent. Перейти к шагу 5.

Шаг 5. Принимать предложения либо отказы от агентов MajorBuyerAgent. Если количество предложений больше нуля, перейти к шагу 6, иначе – к шагу 0.

Шаг 6. Принять финальные предложения от всех MajorBuyerAgent, которые сделали предложение на предыдущем шаге. Перейти к шагу 7.

Шаг 7. Выбрать из всех предложений, полученных на предыдущем шаге, лучшее. Выслать агенту MajorBuyerAgent, сделавшему лучшее предложение, сообщение о согласии на сделку. Всем остальным выслать отказы. Перейти к шагу 8.

Шаг 8. Принять финальное согласие о сделке от агента MajorBuyerAgent, пересчитать свой объем товара. Если объем товара равен нулю, то процесс завершен, иначе перейти к шагу 4.

2.3. Алгоритм поведения агента MajorBuyerAgent

Агент MajorBuyerAgent является покупателем, который сам лично общается с EmployerAgent о возможной сделке. Также он может выборочно договариваться о сделке с BuyerAgent, чтобы выдвинуть совместное предложение агенту EmployerAgent. Алгоритм поведения агента MajorBuyerAgent включает следующие шаги.

Шаг 0. В течение времени maxAvailabilityTime отправлять ответ о своем тарифе агентам TraderAgent и принимать предложения о сделке от агентов TraderAgent с учетом перегрузки транспортного средства. Если время принятия предложений истекло или текущий объем товара равен максимально возможному, то перейти к шагу 1.

Шаг 1. Получить сообщения от всех агентов MajorBuyerAgent. Выбрать агента, предлагающего наиболее выгодные условия. Перейти к шагу 2.

Шаг 2. Отправить сообщение о подписке агенту, сделавший лучшее предложение, всем остальным – сообщение об отмене сделки. Перейти к шагу 3.

Шаг 3. В течение некоторого времени отвечать на запросы о сделке от BuyerAgent. Если ни одного запроса не было обнаружено, перейти к шагу 6, иначе – перейти к шагу 4.

Шаг 4. Принять от всех BuyerAgent предложения либо отказы о сделке. Если предложений было 0, перейти к шагу 6.

Шаг 5. Выбрать для сделки агентов BuyerAgent, которые сделали предложения раньше, с учетом лимита на количество товара.

Шаг 6. Сделать финальное предложение по сделке EmployerAgent. Перейти к шагу 7.

Шаг 7. Дождаться ответа от агента EmployerAgent. В случае положительного ответа отослать всем агентам BuyerAgent информацию об успешном завершении сделки и перейти шагу 8, иначе – отослать отказы и перейти к шагу 0.

Шаг 8. Получить от всех агентов BuyerAgent утвердительные письма и отправить агенту EmployerAgent информацию об успешном завершении сделки.

2.4. Алгоритм поведения агента BuyerAgent

Агент BuyerAgent общается с EmployerAgent с помощью MajorBuyerAgent и не является определяющим в сделке. Алгоритм его поведения включает следующие шаги.

Шаг 0. Найти всех подходящих агентов MajorBuyerAgent для сделки и отправить запрос на сделку. Перейти к шагу 1.

Шаг 1. Получить информационные сообщения от агентов MajorBuyerAgent и выбрать среди них того, который предлагает лучший вариант. Если такой агент был найден, то перейти к шагу 2, иначе – к шагу 0.

Шаг 2. Отправить предложение о сделке лучшему агенту MajorBuyerAgent, всем остальным – отказы.

Шаг 3. Получить согласие либо отказ от MajorBuyerAgent. В случае согласия алгоритм завершает работу, иначе – перейти к шагу 0.

2.5. Пример работы мультиагентной системы

Для того, чтобы запустить приложение, необходимо ввести в командной строке JADE следующие параметры (рисунок 2).

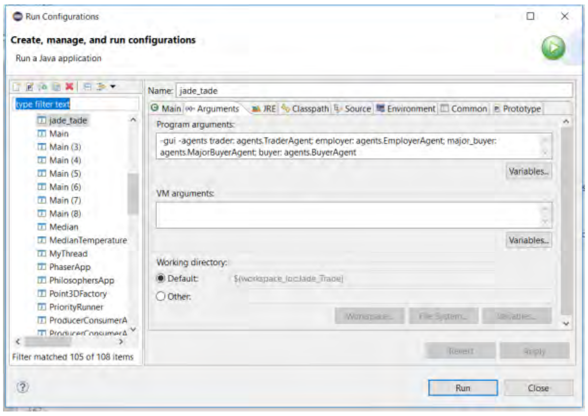


Рисунок 2: Аргументы командной строки для запуска мультиагентной системы на платформе JADE

Примеры протоколов работы данной мультиагентой системы принятия решений представлены на рисунках 3 и 4.

```
tr1@192.168.0.23:1099/JADE: I confirm offer
hello! step = 3?
tr2@192.168.0.23:1099/JADE: I confirm offer
tr3@192.168.0.23:1099/JADE: I cancel offer
emp@192.168.0.23:1099/JADE: Ok, the deal is confirmed
```

Рисунок 3: Протокол работы МА СППР с отказом одного из агентов

В первом случае (рисунок 3) у агентов TraderAgent1-3 в сумме оказалось массы продукта меньше, чем минимально установленное значение у EmployerAgent. Поэтому EmployerAgent дополнительно взимает с агентов TraderAgent штраф за недобор. Как видно из протокола, двух агентов это устроило, а третьего – нет, так как в таком случае он терял для себя неприемлемое количество прибыли.

Во втором случае (рисунок 4) у агентов MajorBuyerAgent и BuyerAgent1-2 получилось набрать массы продукта в заказе ровно столько, сколько было у EmployerAgent. Поэтому сделка успешно завершилась.

```
mj@192.168.0.23:1099/JADE: Receive an accept proposal from emp@192.168.0.23:1099/JADE
mj@192.168.0.23:1099/JADE: Send accept deal info to buyers
( agent-identifier :name emp@192.168.0.23:1099/JADE  :addresses (sequence http://DESKTOP-DNV5
mj@192.168.0.23:1099/JADE: Receive accept deal message from buyer emp@192.168.0.23:1099/JADE
b2: Receive accept deal msg from mj
b1: Receive accept deal msg from mj
b2: Send confirm message to mj
( agent-identifier :name emp@192.168.0.23:1099/JADE  :addresses (sequence http://DESKTOP-DNV5
```

Рисунок 4: Протокол работы МА СППР с успешным завершением сделки

3. Выводы

В статье представлены результаты эксперимента по реализации подхода, предложенного в работе [1], на базе библиотек мультиагентных систем. В результате проведенного исследования спроектирована МА СППР для автоматизации торговой площадки, учитывающая предпочтения всех участников процесса, обеспечивающая гибкость, устойчивость и согласованность их взаимодействий. Разработаны алгоритм поведения каждого из агентов-участников торгового процесса. Продемонстрированы результаты работы МА СППР. Представленные алгоритмы могут использоваться при компонентном проектировании различных МА СППР в сфере торговли независимо от среды разработки. В качестве развития данного исследования предполагается проведение экспериментов по комплексной оценке эффективности применения библиотеки JADE для построения мультиагентных систем в сфере электронной коммерции.

4. Литература

[1] A.N.Morozevich, B.A. Zhalezka, Personal multi-agent decision support systems, in: Proceedings of the 6th International Conference on Pattern Recognition and Information Processing, Minsk-Szczecin, 2001, pp. 233–236.

[2] T.V. Solodukha, B.A. Zhelezko, Developing a multi-agent system for e-commerce, in: Knowledge Acquisition and Management, Research papers of Wroclaw University of Economics, Publishing House of Wroclaw University of Economics, Wroclaw, 2011, pp. 117 – 125.

[3] JADE site, JAVA Agent DEvelopment Framework, 2021. URL: https://jade.tilab.com/.

[4] F. Bellifemine, G. Caire, D. Greenwood, Developing multi-agent systems with JADE, Wiley, 2007.

Композиційні Особливості Побудови Системи Обліку Споживання Природного Газу Об'єктами Малої і Середньої Продуктивності

Віктор Луценко [1], Сергій Бондарев [2], Юрій Пономарьов [3]

[1, 2, 3] *Інститут транспорту газу АТ «Укртрансгаз», вул. Конєва, 16, Харків, 61108, Україна*

Анотація
Розроблення автоматичної системи збирання та обліку природного газу з використанням композиційного методу. Наведено організаційну структуру та компонентний склад системи збору даних. Розроблення програмної платформи для побудови автоматизованих систем моніторингу процесів енергообліку.

Ключові слова
Комерційний облік газу, композиційний метод, програмна платформа, автоматизовані системи.

Composite Features of Construction of the System of the Accounting of Consumption of Natural Gas by Objects of Small and Medium Productivity

Композиційні особливості побудови систем обліку для малих та середніх споживачів природного газу

Viktor Lutsenko [1], Serhii Bondarev [2], Yurii Ponomariov [3]

[1, 2, 3] *Institute of Gas Transportation JSC "Ukrtransgaz", Koneva St., 16, Kharkiv, 61108, Ukraine*

Abstract
Development of automatic gas collection and metering system using a composite method. Provided are the organizational structure and the component composition of the data acquisition system. Development of a software platform for building automated energy monitoring and metering systems.

Keywords
Commercial gas accounting, composite method, software platform, automated systems.

Принцип композиції (інтеграції) при побудові автоматизованих систем полягає в можливості об'єднання за певними правилами і різними способами безлічі вихідних елементів (модулів, підсистем) за допомогою безлічі зв'язків в єдину систему та у виявленні загальносистемних властивостей і функцій новоствореної системи.

Цей принцип поширюється не тільки на способи формування структури складної системи з її окремих компонентів (елементів), але і на способи формування характеристик системи з характеристик підсистеми, глобальної мети системи з локальних цілей підсистем тощо.

До даного принципу за своїм змістом близький відомий в теорії складних систем принцип агрегування, на основі якого побудована теорія агрегативних систем.

EMAIL: lutcenko-va@utg.ua (A, 1); bondarev-sa@utg.ua (A, 2); ponomarev-yv@utg.ua (A, 3)
ORCID:

До цього принципу близькі і еквівалентні за змістом принцип модульності і принцип комплексування, реалізація яких полягає в об'єднанні в єдину систему первинних елементів - окремих модулів або комплексів.

Наразі основним методом реалізації інформаційних систем слугує підхід, який засновано на застосуванні інтегрованого системного програмного забезпечення з найважливішим компонентом - системою керування базами даних (СКБД).

Завдяки концепції бази даних (БД) забезпечується незалежність опису предметної області і завдань додатків від структур даних і методів їх оброблення, програм - від логічної структури баз даних, логічної структури даних - від методів їх фізичної організації. Незалежність даних розуміється як здатність СКБД створювати різні уявлення про одних і тих же даних, що зберігаються, залишаються інваріантними до змін середовища функціонування БД. Необхідна ступінь незалежності даних зазвичай досягається в результаті реалізації стандартів щодо запровадження зовнішнього, концептуального і внутрішнього рівнів визначення та маніпулювання даними.

Спираючись на цей підхід, в АТ «Укртрансгаз» розроблялася і в даний час реалізована система збирання газовимірювальної інформації та обліку споживання природного газу на технологічних об'єктах магістральних газопроводах і підземних сховищ газу, що отримало назву Ask-2.

У 2021 році розпочато розроблення автоматичної системи збирання та обліку споживання природного газу об'єктами малої і середньої продуктивності (топкові котли і комплексні автоматизовані котельні).

В основу розроблення нового програмного продукту (ПП), який отримав назву ПП ObliKot, покладено композиційний метод використання раніше розроблених служб Ask-2 [1]. Як інтегруючі інструменти розробляються відповідні вебмодулі. Це дозволило скоротити час розроблення і оптимізувати людські та технічні ресурси.

ПП ObliKot призначено для автоматичного збирання та зберігання в центральній базі даних (ЦБД) первинної газовимірювальної інформації, зокрема, витрату та об'єм газу, обсяги газу в енергетичних показниках, режими функціонування (аварії та втручання) автоматичних обчислювачів і коректорів (ОіК) витрати газу, які входять до складу вузлів обліку газу на котельнях, формування комерційних та технологічних звітів про об'єм та енерговміст газу, а також для автоматизування документообігу.

Об'єктом автоматизації є система комерційного обліку газу, що використовується для теплових котелень на балансі АТ "Укртрансгаз".

Наразі в ручному режимі від структурних підрозділів надходить інформація про спожитий обсяг газу та для диспетчерської служби готується підсумковий документ, який передається електронною поштою для складання балансу спожитого природного газу по АТ "Укртрансгаз".

Окремо інформація про спожитий газ надається постачальникам газу – представникам операторів газорозподільної мережі. Організаційну структуру системи збору даних побудовано за ієрархічним типом управління (див. рис. 1).

Програмний продукт ObliKot забезпечує автоматичний і ручний процес опитування приладів обліку кількості газу, передачі і зберігання отриманих даних в ЦБД, а також перегляд газовимірювальної інформації в ЦБД.

Структура ПП ObliKot має сервіс-орієнтовану архітектуру, яка базується на використанні розподілу, слабо пов'язаних, замінних компонентів, оснащених стандартизованими інтерфейсами для взаємодії за стандартизованими протоколами.

Локальна база даних (ЛБД) ПП ObliKot побудована у вигляді двох баз даних, основної і резервної, інформаційно синхронізованих і територіально розподілених. Резервна БД використовується в разі неможливості використання основної БД.

Центральна база даних ObliKot розгорнута і функціонує в центрі оброблення даних АТ "Укртрансгаз" і забезпечує збереження актуальної інформації для всього функціоналу програмного продукту ObliKot. Оскільки новітній композиційний продукт частково засновано на використанні програмних компонентів, які наразі ще знаходяться в експлуатації, з точки зору спадкоємності необхідно деякий час зберігати наявність застосованого формату даних у вигляді бінарної бази, так званого hostlib-формату.

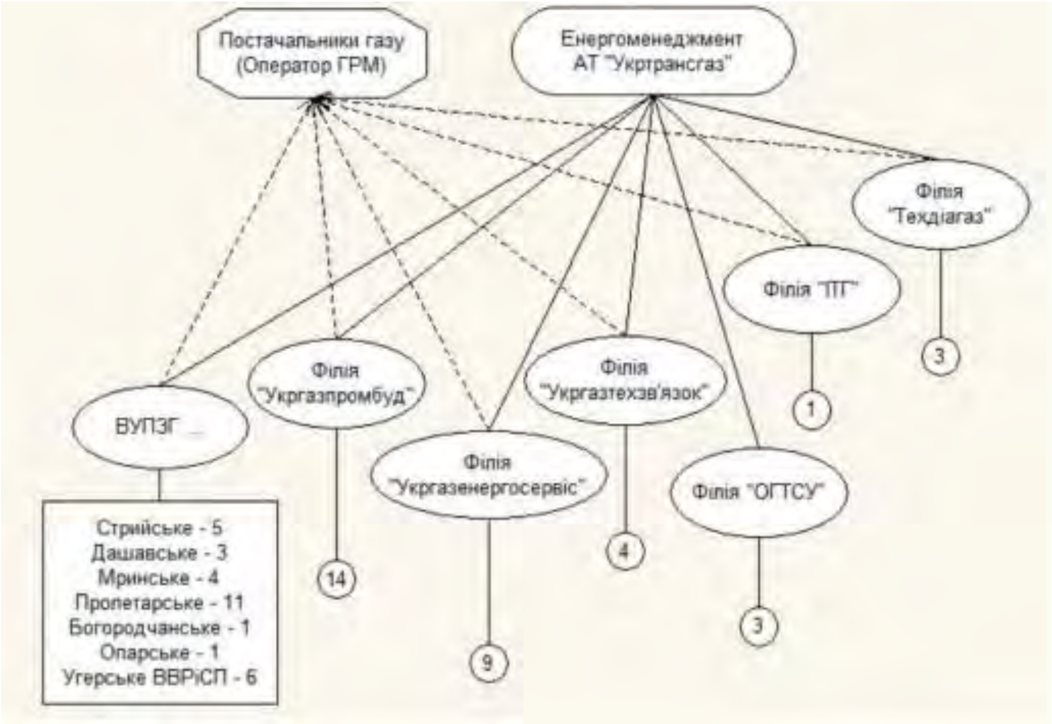


Рисунок 1: Організаційна схема збору даних

ПП ObliKot призначений функціонувати в автоматичному режимі із забезпеченням безперервного цілодобового режиму роботи, за винятком регламентних зупинок для проведення процедур технічного обслуговування.

Програмний продукт передбачає можливість підключення нових робочих місць без додаткових розробок. Крім того, передбачено можливість адаптації без істотного доопрацювання в разі змін в організаційній та штатній структурі підрозділів АТ "Укртрансгаз", регламентів роботи і схем звітності.

Програмний продукт має можливість подальшого розвитку інформаційної основи (модифікація існуючих і створення нових довідників, таблиць даних тощо). Структурну схему та компонентний склад ПП ObliKot наведено на рис. 2.

Таким чином, АТ «Укртрансгаз» ставить основною метою розроблення програмної платформи для побудови на базі єдиного композиційного підходу автоматизованих систем для вирішення комплексних завдань з автоматизації і моніторингу процесів енергообліку. Ця платформа може використовуватися для комерційного і технічного обліку таких ресурсів як електроенергія, природний газ, вода, пара, тепло тощо.

Як результат роботи в цьому напрямку прогнозується зниження трудомісткості і вартості робіт зі збирання, передавання, оброблення та документування інформації, підвищення швидкості обробки інформації, досягнення повноти, достовірності і оперативності інформації про різні параметри видачі і прийняття енергоресурсів, оптимізація режимів їх споживання на основі аналізу отриманих даних, зниження втрат енергоресурсів.

Література References

[1] Віктор Луценко, Сергій Бондарев, Юрій Пономарьов Автоматизація обчислень кількості природного газу в енергетичних одиницях // Информационные системы и технологи: материалы Международной науч.-техн. конф., Харьков, 18-20 листопада 2020 г.: тезисы докладов / [редкол.: А.Д. Тевяшев (отв. ред.) и др.].– Харків:, 2020. – С. 88-91.

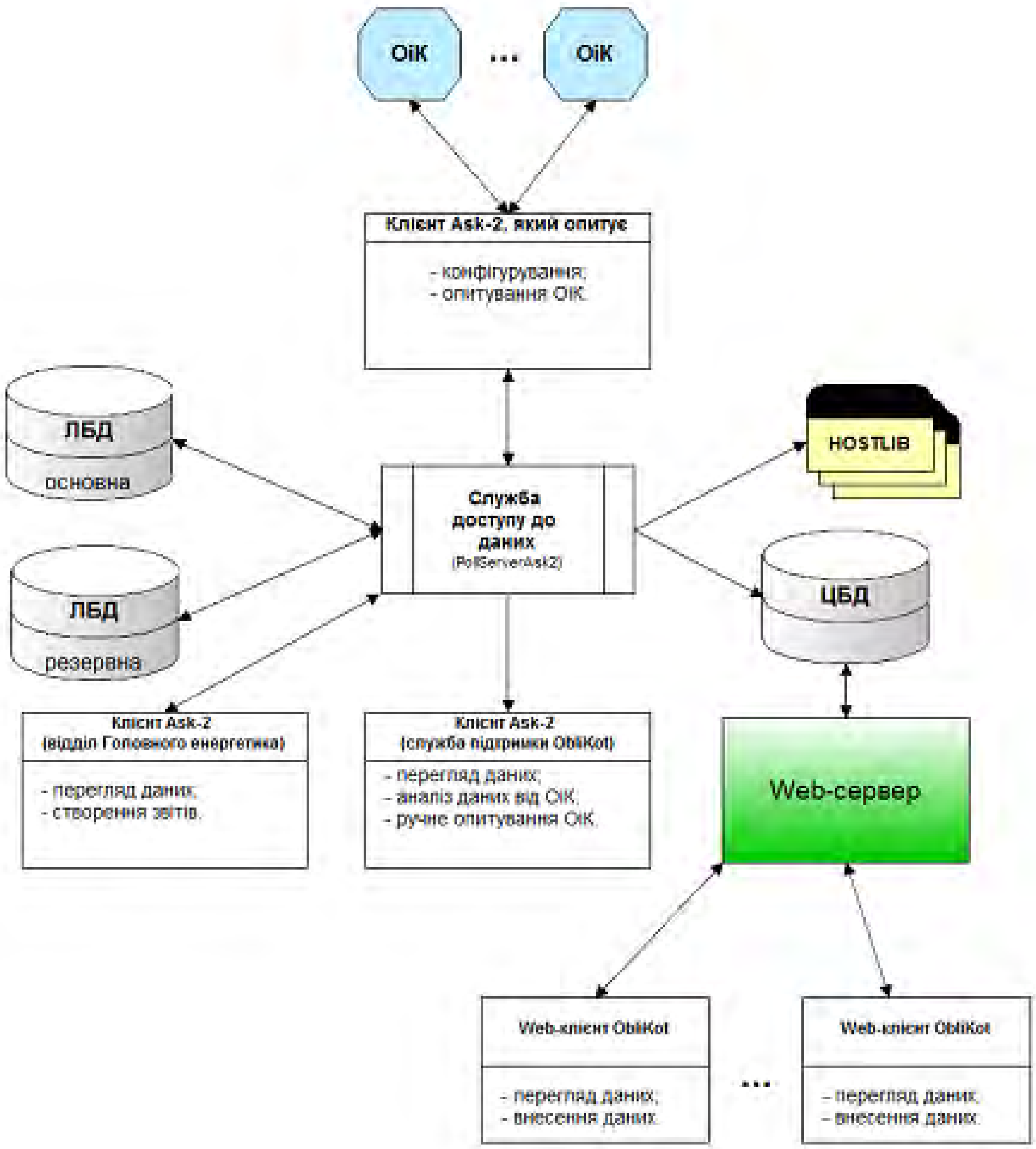


Рисунок 2: Структурна схема компонентів ПП ObliKot

Аналіз завдань маршрутизації транспорту інтернет-магазину з власним підрозділом доставки

Вадим Селезньов [1] , Олександра Дудка [1]

[1] *Харківський національний університет радіоелектроніки, пр. Науки, 14, м. Харків, 61166, Україна*

Анотація
Розглянуто актуальність задач маршрутизації транспорту. Наведено класифікацію типових варіацій даної задачі. Визначено вимоги і обмеження для маршрутизації транспорту інтернет-магазину та визначено клас даної задачі.

Ключові слова
Маршрутизація, логістика, доставка

Vehicle routing task of the online store with its delivery department

Vadym Seleznov [1], Oleksandra Dudka [1]

[1] *Kharkiv National University of Radio Electronics, Nauky Ave. 14, Kharkiv, 61166, Ukraine*

Abstract
Considered relevance of the vehicle routing task. The classification of typical variation of this task is given. Requirements and restriction for vehicle routing of online store are defined and the class of the given task is defined.

Keywords
Routing, logistics, delivery

1. Вступ

Сфера інтернет-послуг і торгівлі в сучасних реаліях є дуже популярною. Однак однією з головних проблем, що властиві цій сфері, є організація доставки товарів кінцевому споживачу. Ефективно організована з точки зору логістики доставка товарів зменшить витрати коштів і часу і може стати конкурентною перевагою бізнесу.

Наразі близько 10% вартості товару складають витрати на його транспортування [1]. При цьому доведено, що використання комп’ютерних програм з оптимізації маршрутів може надати близько 5% економії на транспортуванні товару. Відповідно, будь-які оптимізації, що дозволять скоротити витрати навіть менше ніж на 5% є значущими [2].

Оптимізацію процесу доставки ускладнює той факт, що задіяні особи (менеджер, кур'єри і клієнти) географічно рознесені. Для досягнення найкращих результатів необхідне програмне забезпечення, яке дозволить їм ефективно взаємодіяти, грамотно розподілить роботу виходячи з розрахункових показників витрат, побудує оптимальні маршрути та вбереже від помилок у комунікації.

Слід зазначити, що найбільше часу у процесі існування інтернет-замовлення припадає безпосередньо на перевезення товару. Саме тому доцільним є використання

EMAIL: seleznev.vadim28@gmail.com (A.1), oleksandra.dudka@nure.ua (A.2)
ORCID: 0000-0001-9279-7536 (A.1), 0000-0002-0949-3353 (A.2)

картографічних сервісів у поєднанні з відстежуванням місцезнаходження кур'єра чи товару. З одного боку це задовольнить інформаційні потреби клієнтів стосовно статусу та орієнтовної дати доставки, а з іншого дозволить автоматично будувати оптимальні маршрути доставки кур'єрам.

Тому основною задачею є розробка алгоритму оптимального розподілу замовлень для автоматизованої системи доставки з урахуванням географічних даних.

2. Визначення і класифікація задач маршрутизації транспорту

Задача маршрутизації транспорту – Vehicle Routing Problem (VRP) – типова задача комбінаторної оптимізації. Вона полягає у визначенні оптимальних маршрутів для парку транспортних засобів, що мають пройти через зазначені точки, задовольняючи певним обмеженням. Визначення оптимального рішення для задачі маршрутизації транспорту є NP-складним [3], тобто розмір задач, які можна вирішити оптимально за допомогою математичного програмування або комбінаторної оптимізації, може бути обмеженим.

Виходячи з різних потреб та обмежень при побудові маршрутів існує багато варіацій цієї задачі. До найбільш популярних з них можна віднести наступні класи:

- Задача маршрутизації транспорту з прибутком (VRP with Profits): задача максимізації, де не є необхідним відвідати усі точки, а ціль отримати найбільший прибуток відвідавши довільну кількість точок та задовольнивши обмеження на час використання транспорту [4].
- Задача маршрутизації транспорту з самовивозом та доставкою (VRP with Pickup and Delivery): певні товари необхідно доставити з одних точок у інші, за мету береться знаходження оптимальних маршрутів для парку транспортних засобів з урахуванням точок самовивозу та доставки [5].
- Задача маршрутизації транспорту з часовими проміжками (VRP with Time Windows): для кожної точки маршруту задано певний часовий інтервал, протягом якого транспортний засіб має відвідати цю точку [6].
- Задача маршрутизації транспорту з обмеженням вантажності (Capacitated VRP): на транспорт накладаються обмеження на кількість товарів, що можуть бути перевезені, виходячи з об'єму чи вантажопідйомності [7].
- Задача маршрутизації транспорту з багатьма рейсами (VRP with Multiple Trips): транспорт може робити більше ніж один рейс [8].

Варто зазначити, що підходи до розв'язання кожного класу задачі суттєво відрізняються.

3. Визначення рішень маршрутизації транспорту

Виходячи зі специфіки доставки інтернет-замовлень визначимо обмеження, які необхідно врахувати в алгоритмі:

1. Кожна одиниця транспорту має обмежену вантажність.
2. Кожен транспортний засіб може здійснювати декілька рейсів.
3. Кожен клієнт може обрати інтервал часу, протягом якого йому зручно отримати замовлення.

У якості критерія оптимальності доцільно встановити мінімізацію витрат на перевезення будь-якого товару від складу до клієнтів.

Таким чином, до задачі доставки інтернет-замовлень найкраще підходить комбінований алгоритм маршрутизації транспорту з часовими проміжками, обмеженням по вантажності та можливістю здійснення багатьох рейсів (Capacitated VRP with Multiple Trips and Time Windows).

З огляду на те, що переважним видом транспорту задіяним у процесі доставки є автомобілі, доцільно при розрахунках маршрутів враховувати поточну та прогнозовану завантаженість доріг. Відповідно, при реалізації алгоритму необхідно налаштувати інтеграцію з картографічними сервісами для оцінки часу переміщення з однієї точки маршруту в іншу.

4. Перспективи розвитку

Алгоритм маршрутизації транспорту за часовими проміжками поляже в основі нового хмарного додатку, що буде розповсюджено як SaaS (програмне забезпечення як послуга).

Така сучасна система доставки з урахуванням географічних даних допоможе не тільки оптимізувати доставку для підприємств, що не використовували таких рішень раніше, але й мати гнучку схему оплати такого рішення через можливість сплачувати через підписку до сервісу.

5. Висновки

В роботі розглянуто одне з завдань організації доставки інтернет-замовлень до кінцевих споживачів. У ході дослідження проведено огляд наявних рішень задачі маршрутизації транспорту та їх класифікацію. У якості найбільш оптимального рішення запропоновано використати алгоритм маршрутизації транспорту з часовими проміжками, обмеженням по вантажності та можливістю здійснення багатьох рейсів і інтегрувати його з картографічними сервісами для прогнозування часу переміщення між точками маршруту.

6. Література References

[1] C. Comtois, B. Slack, J.-P. Rodrigue, The geography of transport systems (3rd ed.). London: Routledge, Taylor & Francis Group, 2013.

[2] G. Hasle; K.-A. Lie; E. Quak, Geometric Modelling, Numerical Simulation, and Optimization: Applied Mathematics at SINTEF. Berlin: Springer Verlag, 2007.

[3] P. Toth, D. Vigo, The Vehicle Routing Problem. Monographs on Discrete Mathematics and Applications. Philadelphia: Society for Industrial and Applied Mathematics, 2002.

[4] C. Archetti, G. Sperenza, D. Vigo, "Vehicle routing problems with profits" Vehicle Routing: Problems, Methods, and Applications, 2nd. ed. (2014) 273–297.

[5] M. Drexl, On the one-to-one pickup-and-delivery problem with time windows and trailers. Cent Eur J Oper Res 29 (2021), 1115–1162.

[6] M. M. Solomon, Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints. Operations Research (1987), 254–265.

[7] T. Ralphs, L. Kopman, W. Pulleyblank, On the capacitated vehicle routing problem. Math. Program., Ser. B 94 (2003), 343–359.

[8] T. G. Crainic, Y. Gajpal, M. Gendreau, Multi-Zone Multi-Trip Vehicle Routing with Time Windows. CIRRIELT (2012), 34–36.

Інформаційні технології в процесі підготовки майбутніх інженерів

Волков Віктор[1], Лобода Юлія[2]

[1]Одеський національний університет ім. І.І. Мечникова, вул. Дворянська, 2, Одеса, 65082, Україна
[2]Національний університете «Одеська юридична академія», вул. Академічна, 2, Одеса, 65009, Україна

Анотація
Розглянуто проблеми використання інформаційних технологій в процесі підготовки майбутніх інженерів у вишах. Доведено, що впровадження в методичний та науково-методичний супровід процесу підготовки майбутніх інженерів (в усі його елементи і на всіх етапах) сучасних інформаційних (комп'ютерних та комп'ютерно-інтегрованих) технологій і досягнень педагогічної інформатики з одного боку забезпечує високу якість навчального процесу, а з іншого – само собою значно підвищує рівень володіння новітніми інформаційними технологіями випускниками технічних вишів.

Ключові слова
Інформаційні технології, науково-методичний супровід, підготовка майбутніх інженерів, дидактичні принципи, новітні технології

Information technologies in the process of training future engineers

Volkov Victor[1], Loboda Yuliia[2]

[1]Odessa I.I.Mechnikov National University, Dvoryanskaya str.,2, Odessa, 65082, Ukraine
[2]National University "Odesa Law Academy", Academichna str., 2, Odessa, 65009, Ukraine

Abstract
The problems of using information technologies in the process of training future engineers in universities are considered. It is proved that the introduction of modern information technologies (computer technologies and computer-integrated technologies) and achievements of pedagogical informatics in the methodical and scientific-methodical support of the process of training future engineers (in all its elements and at all stages) on the one hand provides high quality of educational process, and on the other - by itself significantly increases the level of mastery of the latest information technology by graduates of technical universities..

Keywords
Information technologies, scientific and methodical support, training of future engineers, didactic principles, latest technologies

EMAIL: viktor@te.net.ua (A. 1); jul.loboda@gmail.com (A. 2)
ORCID: 0000-0002-3990-8126(A.1); 0000-0001-7083-552(A. 2)

1. Вступ

Сучасні темпи науково-технічного прогресу настільки високі, що за 4-5-річний період навчання майбутнього інженера зазнає значних змін технічне устаткування, швидко розвиваються нові технологічні напрями, суттєво оновлюються як програмне забезпечення, так і апаратні засоби комп'ютерів, розширюється спектр високих технологій. При цьому випускники технічних вишів мають володіти найновітнішими інформаційними технологіями, а не зупинятися на рівні підготовки з відповідної дисципліни, що викладається на першому курсі. В цьому є певне протиріччя, розв'язанню якого присвячено цю роботу.

2. Науково-методичний супровід підготовки майбутніх інженерів

На думку О.В. Співаковського "швидкість зміни освітніх технологій повинна бути досить жорстко пов'язана зі швидкістю зміни комп'ютерних технологій" [1]. Тому процес підготовки майбутніх інженерів треба організовувати таким чином, щоб результатом була своєчасна адаптація студентів до інновацій виробництва, їх готовність до використання новітніх технологій з урахуванням особового та професійного зростання. Визначну роль в цьому процесі відіграють саме сучасні комп'ютерні технології з різноманітністю їх можливостей [2].

Особливість сучасної освітньої ситуації, що задає стратегію професійної підготовки майбутніх фахівців полягає в тому, що теперішній час вимагає іншого визначення норм щодо тієї чи іншої сфери розвитку людини (раніше такі норми були майже незмінними, статичними, виробленими у достатньо тривалому практичному досвіді).

За таких умов значно зростає значення науково-методичного супроводу протягом всього періоду навчання у процесі підготовки майбутнього інженера.

Термін "супровід" згідно з тлумачним словником має декілька значень: "це дія за значенням: супроводжувати, супроводити разом із ким-, чим-небудь", або "те, що супроводить яку-небудь дію, явище".

Ідеї супроводу людини ґрунтуються на гуманістичній теорії, що розглядає унікальність, неповторність кожної людини як умову розкриття його особового потенціалу та творчих можливостей [3].

Використання інформаційних технологій в процесі навчання сприяє розвитку творчих можливостей людини, бо застосування будь-якої інформаційної технології – на відміну від технології виробничої, коли вимагається максимально точне дотримання норм технологічного регламенту, – потребує від людини певного рівня творчості.

Науково-методичний супровід визначається:

- узгодженням зовнішніх чинників (освітні стандарти) і внутрішніх чинників (потреби, інтереси та можливості самих суб'єктів освітнього процесу), наявність яких зумовлює реалізацію професійно-освітньої програми, що неможливо без використання комп'ютерних технологій;
- вибором індивідуальної траєкторії навчання студента, що забезпечує досягнення необхідного рівня готовності до використання комп'ютерних та комп'ютерно-інтегрованих технологій, засобів управління і організаційних форм навчання майбутніх інженерів, і передбачає широке використання сучасних Internet-технологій;
- розробкою навчально-методичного забезпечення, що має спиратися на інтеграцію науки та освіти, а також на повноцінне використання сучасних інформаційних технологій.

В результаті роботи викладачів різних спеціальностей над науково-дослідними проектами у співпраці із студентами, магістрантами та аспірантами їх рівень володіння сучасними інформаційними технологіями значно підвищується.

Для того, щоб науково-методичний супровід був адекватним своїй сутності, необхідно дотримуватися дидактичних принципів скерованості, науковості навчання, систематичності та послідовності, які становлять основу організації навчального процесу взагалі.

3. Висновки

Спираючись на вищезазначене, можна стверджувати, що науково-методичний супровід підготовки майбутніх інженерів до використання новітніх технологій забезпечує:

1. Створення нової наукової інформації;
2. Збагачення змісту освіти новими науковими ідеями та відкриттями;
3. Формування у розробників та користувачів програмного забезпечення нових наукових знань, дослідної культури, а отже і культури професійної діяльності;
4. Розвиток професійної майстерності викладача вишу;
5. Активізацію пізнавальної діяльності майбутніх інженерів завдяки об’єднанню дослідною діяльності викладачів і студентів;
6. Індивідуалізацію вищої освіти, що враховую неповторну людську сутність, талант, обдарованість, можливості та здібності індивіда;
7. Цілісність вищої освіти завдяки єдиному процесу отримання, засвоєння, обробки та використання професійно необхідної наукової інформації;
8. Підсилення навчальної інформації новими ідеями, концепціями, педагогічним досвідом та підвищення якості такої інформації.

Розв’язання всіх цих проблем є неможливим без впровадження в науково-методичний супровід процесу підготовки майбутніх інженерів – в усі його елементи і на всіх етапах – сучасних інформаційних (комп’ютерних та комп'ютерно-інтегрованих) технологій і досягнень педагогічної інформатики. В той же час таке впровадження само собою значно підвищує рівень володіння випускниками технічних вишів новітніми інформаційними технологіями. При цьому, звичайно, також підвищується рівень володіння сучасними інформаційними технологіями і тими викладачами, хто не є фахівцями в галузі інформаційних технологій та комп’ютерних наук.

Таким чином суттєво підвищується конкурентоспроможність випускників технічних вишів на ринку праці. Найбільш технологічно розвинені виробництва і торгівельні компанії забезпечуються, з одного боку, кваліфікованим штатом інженерно-технічного персоналу, а з іншого боку – фундаментальними науковими та науково-технічними (точніше, науково-практичними) розробками, що проводять у різноманітних науково-дослідних лабораторіях при провідних університетах, академіях, науково-дослідних інститутах, науково-технічних центрах, приватних фірмах різних напрямків діяльності тощо.

4. Література

[1] О.В. Співаковський, Впровадження концептуальних питань інтеграційних технологій у молодшу ланку освіти, Початкова школа №3, 2002, с. 36–44.

[2] В.П. Беспалько, Образование и обучение с участием компьютеров (педагогика третьего тысячелетия), Москва: Издательство Московского Психолого-социального института; Воронеж: Издательство НПО "МОДЭК", 2002.

[3] О. Третяк, Психологічна клініка у вищому навчальному закладі, Соціальна психологія, Вип. №4 (30), 2008, с. 144-151.

Адаптивний генератор акустичного зашумлення виділеного приміщення

Юрій Касьянов[1], Іван Демідов[2] , Золотченко Вікторія[3]

[1, 2, 3] *Національний університет кораблебудування імені адмірала Макарова, просп. Героїв України, 9, Миколаїв, 54000, Україна*

Анотація
Сформульовано принцип адаптивного акустичного зашумлення виділеного приміщення. Описано створений генератор шуму, побудований за даним принципом.

Ключові слова
Технічний захист мовної інформації, генератор шуму, адаптивне зашумлення, виділене приміщення.

Adaptive generator of acoustic noise for the confidential room

Yurii Kasianov[1], Ivan Demidov[2] and Viktoriia Zolotchenko[3]

[1, 2, 3] *National University of Shipbuilding named by admiral Makarov, Heroes of Ukraine ave., Mykolaiv,54000, Ukraine*

Abstract
The principle of adaptive acoustic noise of the confidential room is formulated. The created noise generator, built on this principle, is described.

Keywords
Technical protection of speech information, noise generator, adaptive noise, confidential room.

1. Вступ

Мовна інформація є одним з головних джерел отримання інформації з обмеженим доступом про різні сфери життєдіяльності людини – особисте життя, фінансову, адміністративну або виробничу діяльність організації тощо. Дуже часто ця інформація має оперативний характер. Тому природньо, що мовна інформація з обмеженим доступом є предметом особливої уваги воєнних і комерційних розвідок та інших зловмисників.

Захист мовної інформації від витоку прямими (повітряними) акустичними каналами базується на маскувальній здатності акустичних завад і реалізується пасивними та активними засобами. Найбільш ефективними є активні засоби, оскільки дають можливість регулювати рівень завад.

Наразі існує цілий ряд генераторів шумових завад, як сертифікованих ДССЗЗІ України ("МАРС-ТЗО-4-2", "РІАС-2ГС", "БАЗАЛЬТ-4ГА") [1], так і несертифікованих в Україні ("Спрут", "Шепот", VNK-012GL, "Барон", тощо). Створені та розробляються нові високоефективні генератори мовоподібних завад, впровадження яких в практику гальмується через ряд проблем, пов'язаних з існуючим критерієм нормування захищеності мовної інформації [2-5].

Але при використанні генераторів акустичного шуму, відбувається зашумлення всього об'єму виділенго приміщення, в тому числі безпосереднього місця переговорів. В результаті знижується комфортність ведення переговорів, підвищується нервова напруга

EMAIL: yukas.nik.ua@gmail.com (A. 1); pikasso.ivan@gmail.com (A. 2); zolotchenko98@gmail.com (A. 3) ORCID: 0000-0003-2417-0111 (A. 1)

учасників та з'являється ймовірність помилок в розумінні тих чи інших моментів.

Рівень акустичної завади, як правило, встановлюється, виходячи з характерного рівня мовленнєвого сигналу у виділеному приміщенні. При високих встановлених рівнях виникає природне бажання говорити гучніше, що може привести до порушення нормативних значень відношення сигнал/шум в контрольних точках та витоку інформації. Окрім того, з часом умови в потенціальних акустичних каналах витоку мовної інформації можуть змінюватись, що при відсутності оперативного контролю може привести до витоку інформації навіть при використанні активного зашумлення.

Це обумовлює доцільність створення адаптивних систем зашумлення.

2. Принцип адаптивного зашумлення виділеного приміщення

Принцип адаптивного зашумлення полягає в тому, щоб в реальному часі контролювати допустиме відношення сигнал/шум в контрольних точках на межі контрольованої зони (точках можливого знімання інормації) і регулювати рівень шумової завади таким чином, щоб його мінімізувати, але не порушити допустимі норми, забезпечуючи тим самим оперативний контроль захищеності мовної інформації й одночасно максимально можливий комфорт ведення переговорів.

Вимірювання відношення сигнал/шум проводяться в 7-ми октавних смугах мовного діапазону з середньогеометричними частотами 125, 250, 500, 1000, 2000, 4000 та 8000 Гц.

Регулювання рівня шумової завади виконується за різницею між максимальним поточним значенням відношення сигнал/шум у контрольній точці і його нормативним значенням.

3. Експериментальний зразок адаптивного генератора акустичного зашумлення

За вказаним принципом було створено експериментальний зразок адаптивного генератора акустичного зашумлення. Структурна схема генератора приведена на рис. 1.

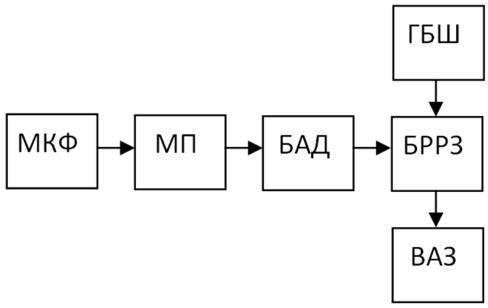


Рисунок1: Структура адаптивного генератора акустичного зашумлення: МКФ – мікрофон; МП – мікрофонний підсилювач; БАД – блок аналізу даних; ГБШ – генератор білого шуму; БРРЗ - блок регулювання рівня завади; ВАЗ – випромінювач акустичної завади.

Генератор складється з двох основних частин: аналізатора відношення сигнал/шум в контрольній точці, що включає МКФ, МП та БАД, і регульованого генератора акустичного білого шуму (ГБШ, БРРЗ та ВАЗ).

Аналізатор відношення сигнал/шум виконує наступні функції:

- визначення рівнів фонового шуму, сумарного сигналу, інформативного сигналу та відношення сигнал/шум в контрольній точці в 7-ми октавних смугах мовного діапазону;
- визначення максимального відношення сигнал/шум та порівняння його з нормативним значенням;
- формування керуючих імпульсів для регулювання рівня акустичної завади.

Для вимірювання сигналів використано модуль електретного мікрофона та підсилювач MAX9814.

Всі операції перетворення і обробки сигналів та математичних обчислень виконуються блоком аналізу даних (БАД), побудованому на мікроконтролері Arduino Nano R3.

Враховуючи, що за замовчуванням частота дискретизації АЦП даного мікроконтролера є 10 кГц, що, відповідно до теореми Котельникова, є недостатнім для перетворення мовного сигналу, було здійснено попереднє налаштування АЦП, яке дозволило отримати частоту дискретизації 38,4 кГц.

Для підвищення швидкості перетворень мовного сигналу і аналізу спектру в реальному часі було використано швидке перетворення

Хартлі. В результаті формується спектр сигналу у вигляді масиву зі 128 10-бітних значень. Далі обчислюються рівні сигналу в октавних смугах.

Одночасно з виконанням основного аналізу, контролюється рівень фонового шуму в моменти паузи між словами.

Цифровий генератор білого шуму виконаний також на мікроконтролері Arduino Nano R3. Максимальна частота створюваного шуму становить приблизно 18,5 кГц.

4. Висновки

Сформульовано принцип адаптивного акустичного зашумлення виділених приміщень, який реалізовано в експериментальному зразку адаптивного генератора акустичного шуму.

Побудова систем акустичного зашумлення за цим принципом дозволить забезпечити оперативний контроль захищеності мовної інформації й одночасно максимально можливий комфорт ведення переговорів.

5. Література References

[1] Засоби ТЗІ, які мають експертний висновок про відповідність до вимог технічного захисту інформації (станом на 12.08.2021) // ДССЗЗІ України. – Режим доступу: https://cip.gov.ua/ua/news/zasobi-tzi-yaki-mayut-ekspertnii-visnovok-pro-vidpovidnist-do-vimog-tekhnichnogo-zakhistu-informaciyi.

[2] Kasianov Yu.I., Nuzhnyi S.M. Otsiniuvannia efektyvnosti heneratora realnoi movopodibnoi zavady za kryteriem rozbirlyvist movy // Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Seriya: Avtomatyka, vymiriuvannia ta keruvannia. 2016. Issue 852. P. 105–110. URL: http://ena.lp.edu.ua:8080/handle/ntb/36473

[3] Касьянов Ю.І., Кисельова Ю.П. Методологія оцінювання генераторів реальних мовоподібних завад в захисті мовної інформації // Інформаційні системи та технології ІСТ-2018: Матеріали 7-ї міжнародної науково-технічної конференції. – Харків: ХНУРЕ, 2018. – С. 333–336.

[4] Blintsov, V., Nuzhniy, S., Parkhuts, L., Kasianov, Y. (2018). The objectified procedure and a technology for assessing the state of complex noise speech information protection. Eastern-European Journal Of Enterprise Technologies, 5 (9 (95)), 26–34. doi: http://dx.doi.org/10.15587/1729-4061.2018.144146

[5] Blintsov, V., Nuzhniy, S., Kasianov, Y., Korytskyi, V. (2020). Mathematical model of the system of active protection against eavesdropping of speech information on the scrambler generator. «EUREKA: Physics and Engineering», Computer Sciences, 3, 11-22. doi: https://doi.org/10.21303/2461-4262.2020.001241

Modeling of gas-dynamic processes in the inner cavity of gas gathering pipelines

Volovetskyi V.B. [1], Doroshenko Ya.V. [2], Doroshenko J.I. [2], Shchyrba O.M. [1]

[1] *Ukrainian Scientific Research Institute of Natural Gases Public Joint Stock Company "Ukrgasvydobuvannya", Kharkiv, 20 Himnaziina Naberezhna str., 61010, Ukraine*
[2] *Ivano-Frankivsk National Technical University of Oil and Gas, Ivano-Frankivsk, 15 Karpatska str., 76000, Ukraine*

Abstract
CFD modeling showed that the accumulation of liquid contaminants in the low sections of gas pipelines affects the gas-dynamic processes and leads to pressure losses above the values provided by the technological regime. With the increase in liquid contaminants volume the pressure losses occur. Moreover, with a small amount of contamination (up to 0.006 m^3), liquid contaminants do not have a significant effect on pressure loss. If the contaminants volume in the lowered section of the pipeline is greater than the specified value, the pressure loss increases by parabolic dependence. The increase in mass flow leads to an increase in the value of pressure loss at the site of liquid contamination.
Keywords:
Gas, gas gathering pipeline, hydraulic efficiency, liquid contaminations.

1. Introduction

During the development of the Yuliyivskyi oil, gas and condensate production facility (OGCPF), in particular Yuliivskyi, Skvortsivskyi, Narizhnyanskyi, Nedilnyi and others, there are many complications that negatively affect the volume of production, including during the transportation of hydrocarbons by pipelines.

One of the important tasks – is to ensure reliable transmission of hydrocarbons to gas collection and preparation systems. Given this, there is a need to maintain the maximum possible values of hydraulic efficiency of flowlines and gas gathering pipelines.

2. Literature review

In world practice, many different methods of cleaning gas pipelines are known, consider the main ones, namely:

− creating a high-speed gas flow;
− use of various chemical reagents (surfactant solution);
− use of various cleaning devices;
− use of various devices for liquid drainage.

Of the above methods, consider the second, which is often used in practice to improve the efficiency of operation of both wells and pipelines of the gas collection and preparation system.

In [1] the technology of cleaning the inner cavity of the flowlines at gas condensate wells Yuliivskyi OGCPF with foam was investigated. According to the results of experimental studies, as a result of cleaning the flowlines of wells 85 and 60 of Yuliivskyi oil and gas condensate field (OGCF) from fluid accumulations, the coefficients of their hydraulic efficiency increased by 12 % and 7 %, respectively. The measures taken to clean the inner cavity of the flowlines from the liquid justified their effectiveness and were recommended for other wells flowlines at different fields [2].

In [3] comprehensive measures are proposed to increase the efficiency of gas condensate wells operation. They are monitoring of operational parameters of wells by pressure and temperature gauges installed at the wellhead and at the inlet gas pipelines of the gas treatment unit; calculation of the volume of accumulated fluid in the wellbore and flowline; installation of a complex of automated feeding a surfactant solution of both in the annulus of the wells and in the flowline. For this purpose, two options for the complex

EMAIL: vvb11@ukr.net (A. 1); ya.doroshenko@nung.edu.ua (A. 2); doroshenk77@gmail.com (A. 3); omschyrba@ukr.net (A. 4)
ORCID: 0000-0001-8575-5143 (A. 1); 0000-0002-5853-3286 (A. 2); 0000-0002-7196-9383 (A. 3); 0000-0002-4584-7446 (A. 4)

arrangement are proposed. The proposed options involve the use of various equipment and have a different principle of operation.

In [4] it was stated that in order to increase the efficiency of the hydraulic characteristics of the wells' flowlines at the Yuliivskyi OGCPF, the injection of surfactant solution was used. This measure made it possible to improve the removal of liquid by the gas flow. Pumping the surfactant solution allowed to reduce gas losses on blowing the flowlines, in some cases to eliminate them completely, as well as to increase the productivity of wells.

Therefore, in order to ensure the stable production of hydrocarbons at Yuliivskyi OGCPF, comprehensive measures were taken to maintain the maximum value of hydraulic efficiency of gas gathering pipelines.

Given the above, it is important to find new approaches to increase the hydraulic efficiency of gas gathering pipelines. One of such approaches is CFD simulation of gas-dynamic processes by means of modern software complexes.

Apparently, the simulation of gas pipelines is becoming more widespread today, due to the increasing capacity of computer technology. Increasing the power of computer equipment allows to significantly reduce simulation time, solve more complex problems. CFD simulation is the closest to reality and makes it possible to consider the studied processes in three-dimensional positioning, as well as understanding the dynamics of the pipeline flows. It allows to see in detail a complex three-dimensional single-phase or multiphase flow inside gas pipelines and to study the distribution of pressure, flow rate. In [5, 6] by using CFD simulation the patterns of flow velocity distribution, pressure in the inner cavity of the shaped elements of gas pipelines were identified. The regularities of the influence of all factors on the pressure losses of single- and two-phase flows in the inner cavity of the studied shaped elements were determined. Also, CFD simulation made it possible to forecast the erosive wear of shaped elements of gas pipelines [7], which is an extremely difficult and currently less studied task due to a wide range of parameters that affect its location and intensity. Erosion, corrosion wear, uneven pressure distribution in the internal cavity, geometric shape are factors that affect the stress-strain state of gas pipelines and its separate elements. Therefore, the assessment of gas pipelines and their individual elements strength requires a synergy of hydro-gas-dynamic processes in their inner cavity, processes of erosion, corrosion wear and stress-strain state in a three-dimensional setting, i.e. it is necessary to perform multidisciplinary simulation. Such synergy for shaped elements of gas pipelines was performed by simulation in [8]. Trenchless technologies, in particular, the technology of pulling a new polyethylene pipe into a defective steel pipeline, are effective to restore the efficiency of sections of pipelines that have lost strength [9].

3. Formulation of the problem

At present, there is a system of gas pipelines between production facilities at Yuliivskyi OGCPF, which provides gas supply from various fields at CGTU-2 of Yuliivskyi OGCF, where liquid hydrocarbons (hydrocarbon condensate and propane-butane fraction) are extracted, and gas is prepared and transmiddes to consumers. Therefore, it is necessary to monitor the hydraulic state of gas pipelines to ensure stable transmission of natural gas and, accordingly, the planned volume of production.

The hydraulic efficiency of six gas gathering pipelines of Yuliivskyi OGCPF was calculated, the results of which are given in Table 1.

Table 1

Gas gathering pipelines at Yuliivskyi OGCPF and calculated hydraulic efficiency coefficients

Pipeline name	Hydraulic efficiency coefficient E,%
1	88
2	77
3	74
4	72
5	80
6	86

The results analysis shows that the lowest value of hydraulic efficiency coefficient was obtained on gas gathering pipelines (4, 3, 2). Therefore, the optimal mode of operation should be chosen to prevent the contaminants accumulation.

4. Results of the research

It should be noted that hydraulic efficiency was studied on gas gathering pipelines (1-6) at different times of the year. The purpose of these studies was to determine the volume of fluid that accumulates in the internal cavity, both by calculation and experimentally. Thus, according to the results of experimental studies, the amount of accumulated pollution in gas gathering pipelines was higher than the calculated one – 6 - 15%.

During experimental studies on the measuring lines CGTU-2 fluid samples from the segregator measuring RZ-1, RZ-2 were taken to analyze the content of contaminants. The result of contaminants analysis indicates that they are a complex multicomponent composition that contains: stratal and condensation water, hydrocarbon condensate, mechanical impurities, methanol in different ratios.

It should be noted that during the measurement of liquid volume transmitted together with the gas on the measuring line CGTU-2, its composition changed in percentage. A number of studies have shown that the main component is stratal and condensation water, which takes a significant volume and varies in the range from 50% to 65%. Given that the hydrocarbon condensate has a density within the range from 0.736 g/cm^3 to 0.748 g/cm^3, less than water (the density of which is more than 1.010 g/cm^3), it can be assumed that the actual mode of operation of the gas flow from the internal cavity of the gas gathering pipelines carries more gas condensate than water. Therefore, there is a need to take appropriate measures to clean the inner cavity of gas gathering pipelines from liquid contaminants.

In-depth knowledge of the relationship between flow geometry, pressure field, and flow kinematics is required to investigate the effect of liquid contaminants accumulated in lowered sections of the pipeline route on pressure losses. This information will help to understand the mechanisms of pressure loss. Flows in such places are complex and three-dimensional, and therefore they must be studied experimentally or hydraulic analysis of CFD simulation must be performed. For the in-situ conditions of gas pipelines such experiments cannot be performed because:

– it is impossible to determine the precise value of speed, pressure at any 3D point of the pipeline;

– gas pipelines are under high pressure and are explosive.

CFD simulation gives an understanding of the dynamics of gas flows, allows to observe the flow in the inner cavity of the pipeline and study the pressure loss, the laws of change of its speed and so on. Therefore, CFD simulation of gas-dynamic processes by the finite volume method was used to determine the magnitude of the influence of liquid contaminants accumulated in the lowered sections of the pipeline route on the magnitude of pressure loss and flow rate. CFD simulation was performed by means of ANSYS Fluent 2020 R2 Academic software package. Modern software is one of the best options for complex and accurate calculations and allows to save a lot of time and minimize the number of experiments.

Three-dimensional CFD simulation of gas flows in low sections of the pipeline route, where liquid contaminants were accumulated, was performed by numerical solving the Navier-Stokes equation (1), which expresses the law of momentum conservation and flow continuity (2), which is the law mass conservation:

$$\frac{\partial}{\partial t}\left(\rho\mathbf{u_i}\right)+\frac{\partial}{\partial x_j}\left(\rho\mathbf{u_i}\mathbf{u_j}\right)+\frac{\partial}{\partial x_j}\left(\rho\mathbf{u_i'}\mathbf{u_j'}\right)= \\ =-\frac{\partial p}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\left(\frac{\partial\mathbf{u_i}}{\partial x_j}+\frac{\partial\mathbf{u_j}}{\partial x_i}\right)\right)+f, \quad (1)$$

$$\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x_j}\left(\rho u_j\right)=0. \quad (2)$$

where xi, xj, – coordinates; t – time; ui, uj – velocity components; ρ – density of gas; μ – molecular dynamic viscosity of the gas; fi – a term that takes into account the effect of mass forces; p – pressure [10].

To describe turbulence in CFD simulation, a standard and the most common k – ε (k - turbulent kinetic energy, ε - dissipation rate of turbulent kinetic energy) turbulence model was used, which involves solving the following equations:

– equation of transfer of turbulent energy k

$$\frac{\partial(\rho k)}{\partial t}+\nabla(\rho u k)=\nabla\left(\left(\mu+\frac{\mu_t}{\sigma_k}\right)\nabla k\right)+\mu_t G-\rho\varepsilon\ ; \quad (3)$$

– equation of transfer turbulent dissipation ε

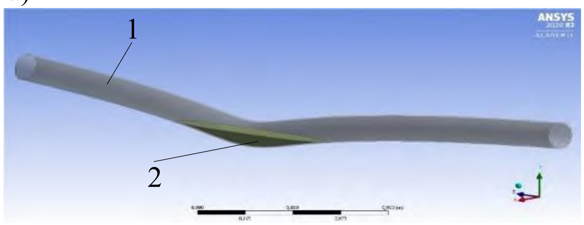
$$\frac{\partial(\rho\varepsilon)}{\partial t}+\nabla(\rho u\varepsilon)=\nabla\left(\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\nabla\varepsilon\right)+C_1\frac{\varepsilon}{k}\mu_t G-C_2\rho\frac{\varepsilon^2}{k}, \quad (4)$$

where u – flow rate of gas; μ_t– turbulent dynamic viscosity of gas; σ_k – coefficient equal to one; G – design variable; σ_ε – coefficient equal to σ_ε=1.3; C_1 – coefficient equal to C_1=1.44; C_2 – coefficient equal to C_2=1.92 [10].

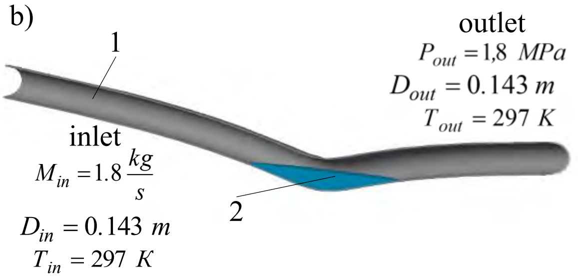
To study the gas-dynamic processes in the lowered sections of gas pipelines, where liquid contaminants accumulate, it is necessary to take into account the influence of the amount of contaminants on the formation of the flow. For this purpose, three-dimensional models of lowered sections of the gas pipeline with an internal diameter of 143 mm with liquid contaminants were drawn (Figure 1, a). The outer diameter of the investigated section of the pipeline was 159 mm, and the nominal wall thickness was 8 mm. These characteristics correspond to the gas gathering pipeline № 6 (table 1).

To study the influence of the contaminants amount on the flow parameters, geometric models of the internal cavity of the gas pipeline sections with different amounts of contamination were drawn. Four different volumes of liquid contaminants were selected (0.0035 m^3, 0.0051 m^3, 0.0084 m^3 and 0.0118 m^3).

a)



b)



a) – geometrical model; b) – calculation chart;
1 – gas pipeline; 2 – liquid contamination

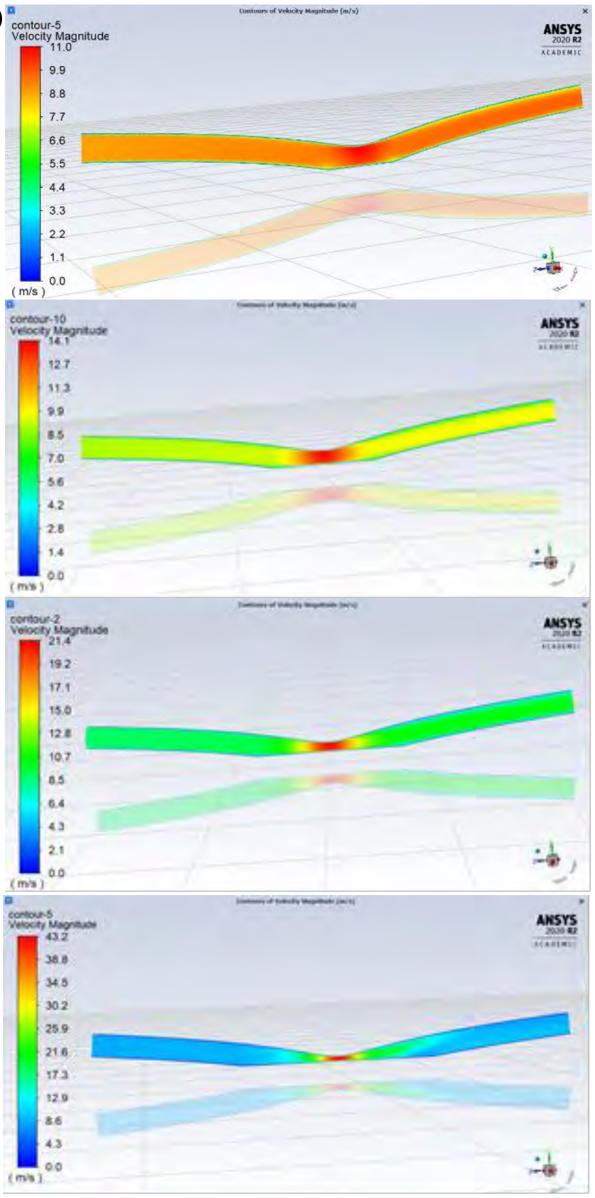
Figure 1: Lowered section of gas pipelines with liquid contamination

Limit conditions were set in the ANSYS Fluent preprocessor. Natural gas was set as the working medium, the movement of which was simulated by the gas pipeline. The equivalent roughness coefficient of the pipe wall was set equal to 0.03 mm. Mass flow was set at the inlet to the studied section of the gas pipeline, and pressure was set at the outlet. The calculation chart is shown in Figure 1, b. The mass flow rate at the inlet to the studied section of the gas pipeline varied and was assumed to be equal to 1.8 kg/s, 2.5 kg/s and 3.2 kg/s. The pressure was assumed to be 1.8 MPa. A separate simulation was performed for each value of mass flow. It should be noted that the initial data were set from the actual mode of operation of the gas gathering pipeline.

Also, at the inlet to the studied section of the gas pipeline the temperature of natural gas was set, which was assumed to be 297 K. At the inlet and outlet of the pipeline the turbulence intensity as for 5% (for this value the flow is considered to be completely turbulent) and hydraulic diameter were set. The hydraulic diameter was assumed to be equal to the inner diameter of the pipeline.

The velocity of the gas flow through the studied section of the gas pipeline and the pressure loss in it are two indicators that characterize the gas-dynamic processes. Therefore, in the postprocessor of the ANSYS Fluent software package, the simulation results were visualized by constructing a velocity field and a pressure field in the longitudinal crossections of the studied section. The visualized simulation results made it possible to see the structure of the flow in the investigated section of the gas pipeline and to collect comprehensive data about it. For example, the simulation results are considered when the mass flow rate of the gas flow at the inlet of the studied section of the pipeline was 1.8 kg/s. The velocity fields in the longitudinal crossections of the studied section for such mass flow are shown in Figure 2, and the pressure fields in Figure 3. It was determined that at the entrance to the study area the gas flow velocity along the axis is 8.9 m/s. From the flow axis in the direction of the wall there is a slight decrease in the flow rate, and near the wall the gas flow rate decreases sharply. The presence of contaminants in the lower (middle) part of the investigated section of the gas pipeline causes an acceleration of gas flow in this place, which is due to a decrease in the cross-section of the gas pipeline (Figure 2). A detailed analysis of the gas flow velocity fields shows a significant dependence of the gas flow velocity value in the average time of the studied section of the gas pipeline on the contamination volume. For the considered mode of transmission if the contamination volume in the inner cavity of the lowered section of the gas pipeline is 0.0035 m^3, the maximum value of the gas flow rate over the contaminants is 11 m/s (Figure 2, a). For the volume of contaminants of 0.0051 m^3, the maximum velocity is 14.1 m/s (Figure 2, b), for the volume of 0.0084 m^3 - 21.4 m/s (Figure 2, c), and for the volume of 0.0118 m^3 - 43.2 m/s (Figure 2, d).

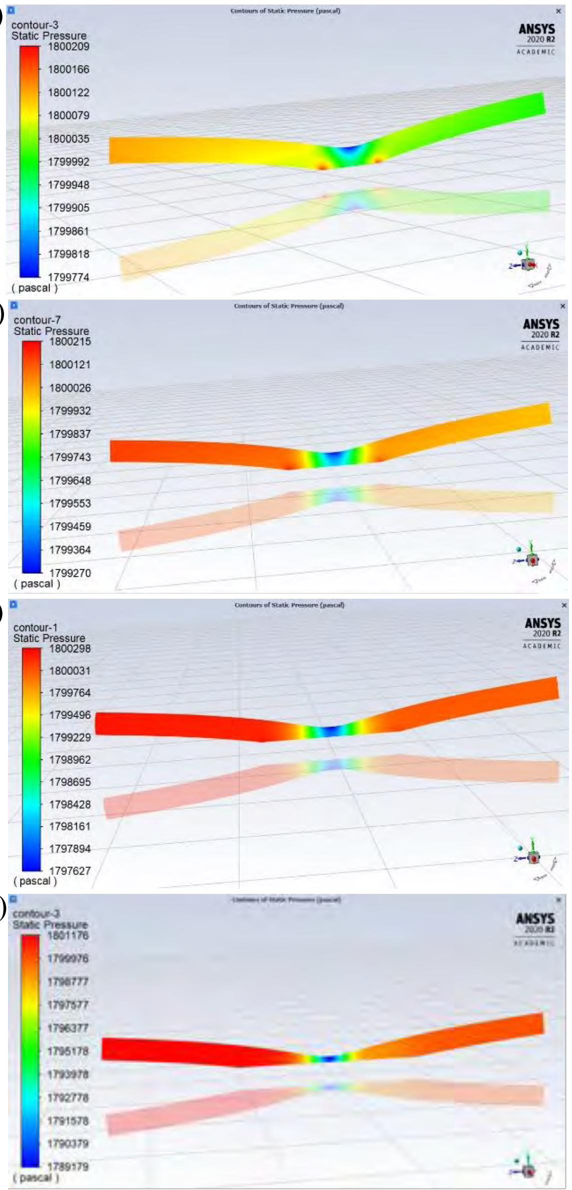
During the transmission of the gas flow through the lowered section of the pipeline, where liquid contaminants are accumulated, a complex pressure field occurs, and pressure loss is formed (Figure 3). Over liquid contaminants in the lowered (middle) part of the studied section of the pipeline there is a decrease in pressure, and the greater the volume of contamination, the greater the magnitude of the pressure drop.

a)
b)
c)
d)


a) – V_{liq}=0,0035 m^3; b) – V_{liq}=0,0051 m^3; c) – V_{liq}=0,0084 m^3; d) – V_{liq}=0,0118 m^3

Figure 2: Velocity fields in longitudinal crossections of the studied lowered section of the gas pipeline for different volumes of liquid contaminants in its inner cavity

The volume of contamination also affects the value of pressure loss in the studied section. The increase in the volume of contamination leads to an increase in the value of pressure loss. The value of pressure drop at the site of liquid contaminants can be determined by subtracting from the value of the pressure in the pipeline prior to the moment of liquid contaminants from the value of the pressure after its occurrence. Thus, for the considered mode of transmission, if the volume of contamination in the inner cavity of the lowered section of the gas pipeline is 0.0035 m^3, the pressure loss is 38 Pa (Figure 3, a), if the volume of contamination is 0.0051 m^3 - 99 Pa (Figure 3 , b), if 0.0084 m^3 - 873 Pa (Figure 3, c) and if 0.0118 m^3 - 1857 Pa (Figure 3, d).

a)
b)
c)
d)


a) – V_{liq}=0,0035 m^3; b) – V_{liq}=0,0051 m^3; c) – V_{liq}=0,0084 m^3; d) – V_{liq}=0,0118 m^3

Figure 3: Pressure fields in longitudinal crossections of the studied lowered section of the gas pipeline for different volumes of liquid contaminants in its inner cavity

The obtained simulation results showed that the increase in the volume of liquid contamination in the inner cavity of gas gathering pipelines leads to an increase in pressure losses above the value provided by the technological regime, and, accordingly, to irrational use of reservoir energy. This is confirmed by the obtained low coefficients of hydraulic efficiency of the gas gathering pipelines of the Yuliivskyi OGCPF due to the presence of accumulated liquid in the internal cavity. Therefore, it is necessary to monitor the hydraulic condition of gas pipelines and clean them in a timely manner.

5. Conclusions

The results of CFD simulation showed that the accumulation of liquid contaminants in the lowered sections of gas pipelines affects the gas-dynamic processes and leads to pressure losses above the value provided by the technological regime. As the volume of liquid contaminants increases, so does the pressure loss. Moreover, with a small volume of contamination (up to 0.006 m^3), there is not significant effect on pressure loss. If the volume of contamination in the lowered section of the pipeline is greater than the specified value, the pressure loss increases by parabolic dependence. The increase in mass flow leads to an increase in the value of pressure loss at the site of liquid contamination. Moreover, the greater the mass flow, the greater the impact of its changes on the pressure loss.

The performed CFD simulation made it possible not only to determine the patterns of pressure loss in places of liquid contaminants accumulation in the inner cavity of gas pipelines, but also to understand gas-dynamic processes in such places, which is an unconditional advantage of this method over experimental.

6. References

[1] V.B. Volovetskyi, A.V. Uhrynovskyi, Ya.V. Doroshenko, O.M. Shchyrba, Yu.S. Stakhmych, Developing a set of measures to provide maximum hydraulic efficiency of gas gathering pipelines, Journal of Achievements in Materials and Manufacturing Engineering 101/1 (2020) 27-41.

[2] V.B Volovetskyi, Ya.V. Doroshenko, O.S. Tarayevs'kyy, O.M. Shchyrba, J.I. Doroshenko, Yu.S. Stakhmych, Experimental effectiveness studies of the technology for cleaning the inner cavity of gas gathering pipelines, Journal of Achievements in Materials and Manufacturing Engineering 105/2 (2021) 61-77.

[3] V.B. Volovetskyi, Ya.V. Doroshenko, G.M. Kogut, I.V. Rybitskyi, J.I. Doroshenko, O.M. Shchyrba, Developing a complex of measures for liquid removal from gas condensate wells and flowlines using surfactants, Archives of Materials Science and Engineering 108/1 (2021) 24-41.

[4] V.B. Volovetskyi, O.Yu. Vytiaz, V.I. Kotsaba, O.M. Shchyrba, Methods of cleaning the flowlines of gas and gas condensate wells, Oil and Gas Power Engineering 2 (24) (2015) 32-43 (in Ukrainian).

[5] Doroshenko Y. Modeling computational fluid dynamics of multiphase flows in elbow and T-junction of the main gas pipeline / Y. Doroshenko, J. Doroshenko, V. Zapukhliak, L. Poberezhny, P. Maruschak // Transport. – 2019. – Vol. 34 (1). – P. 19–29.

[6] Doroshenko Ya. Investigation of the influence of the gas pipeline tee geometry on hydraulic energy loss of gas pipeline systems / Ya. Doroshenko, I. Rybitskyi // Eastern-European Journal of Enterprise Technologies. – 2020. – Vol. 1, No 8 (103). – P. 28–34.

[7] Doroshenko Ya. Numerical simulation of the stress state of an erosion-worn tee of the main gas pipeline / Ya. Doroshenko, V. Zapukhliak, Ya. Grudz, L. Poberezhny, A. Hrytsanchuk, P. Popovych, O. Shevchuk // Archives of Materials Science and Engineering. – 2020. – Vol. 101, No 2. – P. 63–78.

[8] Doroshenko Ya. V. Modeling of stress-strain state of piping systems with erosion and corrosion wear / Ya. V. Doroshenko, A. P. Oliynyk, O. M. Karpash // Physics and chemistry of solid state. – 2020. Vol. 21, No 1. – P. 151–156.

[9] Doroshenko Ya. Development of trenchless technology of reconstruction of “Pulling pig P” pipeline communications / Ya. Doroshenko, V. Zapukhliak, K. Poliarush, R. Stasiuk, S. Bagriy // Eastern-European Journal of Enterprise Technologies. – 2019. – Vol. 2, No. 1(98). – P. 28–38.

[10] Squires K. Particle response and turbulence modification in isotropic turbulence / K. Squires, J. Eaton // Phys. Fluid. – 1990. – Vol. 2, № 7. – P. 1191–1203.

Зміст

Наукове видання

«Інформаційні системи та технології» ICT-2021

П Р А Ц І
10-ї Міжнародної науково-технічної конференції

13-19 вересня 2021

Харків - Одеса, Україна

«INFORMATION SYSTEMS AND TECHNOLOGIES» IST-2021

Proceedings
of the 10-th International Scientific and Technical Conference

September 13-19, 2021

Kharkiv - Odesa, Ukraine

Наукові редактори: А.Д. Тевяшев, Л.Б. Петришин, В.В. Безкоровайний, В.Г. Кобзєв
Коректор – Ю.В. Міщеряков
Комп'ютерна верстка – К.Г. Онищенко, О.О. Олійник

Харківський національний університет радіоелектроніки
61166, Харків, пр. Науки, 14,
ХНУРЕ

Підписано до друку 19.09.2021.
Папір офсетний. Гранітура Times New Roman/
Умов.-друк. арк. – _____. Обл.-вид. арк. – ____.
Наклад ____. Замовлення _______

Віддруковано з готового орігінал-макета

Видавництво