Akademia
Techniczno-Humanistyczna
w Bielsku-Białej

Projekt interdyscyplinarny projektem XXI wieku

Tom 2

Ministerstwo Nauki
i Szkolnictwa Wyższego

Polska Akademia Nauk
Komitet Budowy Maszyn

Polska Akademia Nauk
Komitet Inżynierii Produkcji

IFToMM

International Federation
for the Promotion of Mechanism
and Machine Science

Bielsko – Biała 2017

Redaktor Naczelny Wydawnictwa: prof. dr hab. n.t. Iwona ADAMIEC-WÓJCIK
Redaktor Działu: prof. dr hab. inż. Jacek STADNICKI
Redakcja: dr inż. Jacek RYSIŃSKI
Sekretarz Redakcji: mgr Grzegorz ZAMOROWSKI

Adres Redakcji – Editorial Office – Adresse de redaction –
Schriftleitungadresse:

WYDAWNICTWO NAUKOWE
AKADEMII TECHNICZNO - HUMANISTYCZNEJ
W BIELSKU-BIAŁEJ

PL 43-309 Bielsko-Biała, ul. Willowa 2

ISBN 978-83-65182-70-8
ISBN 978-83-65182-81-4 (Tom 2)

XVIII Beskidzki Festiwal Nauki i Sztuki – zadanie finansowane w ramach umowy nr 866/P/DUN/2017 ze środków Ministra Nauki i Szkolnictwa Wyższego na działalność upowszechniającą naukę.

Bielsko – Biała 2017

KOMITET NAUKOWY - SCIENTIFIC COMMITTEE

CZECH Piotr	Politechnika Śląska
ČUBOŇOVÁ Nadežda	Žilinská Univerzita v Žiline, Słowacja
DANIELCZYK Piotr	Akademia Techniczno-Humanistyczna w Bielsku-Białej
DREWNIAK Józef	Akademia Techniczno-Humanistyczna w Bielsku-Białej
DROBINA Robert	Akademia Techniczno-Humanistyczna w Bielsku-Białej
FOMIN Aleksey	École Polytechnique Fédérale De Lausanne, Switzerland
GREGOR Milan	Žilinská Univerzita v Žiline, Słowacja
GRYŚ Sławomir	Politechnika Częstochowska
HOLUB Sephii	Czerkaski Narodowy Uniwersytet im. Bohdana Chmielnickiego, Ukraina
HOMIŠIN Jaroslav	Technical University of Košice, Słowacja
JAROSŁAW Janusz	Akademia Techniczno-Humanistyczna w Bielsku-Białej
JĘDRZEJCZYK Dariusz	Akademia Techniczno-Humanistyczna w Bielsku-Białej
KARPIŃSKI Mikołaj	Akademia Techniczno-Humanistyczna w Bielsku-Białej
KAZAKOVA Nadiia	Odeska Państwowa Akademia Regulacji Technicznej i Jakości, Ukraina
KŁOSIŃSKI Jacek	Akademia Techniczno-Humanistyczna w Bielsku-Białej
LUZHETSKYI Volodymyr	Winnicki Narodowy Uniwersytet Techniczny, Ukraina
MADEJ Jerzy	Akademia Techniczno-Humanistyczna w Bielsku-Białej
MARTSENYUK Vasyl	Akademia Techniczno-Humanistyczna w Bielsku-Białej
MATUSZEK Józef	Akademia Techniczno-Humanistyczna w Bielsku-Białej
MIČIETA Branislav	Žilinská Univerzita v Žiline, Słowacja
NOWAKOWSKI Jacek	Akademia Techniczno-Humanistyczna w Bielsku-Białej
PARKHUTS Lyubomyr	Narodowy Uniwersytet - Politechnika Lwowska, Ukraina
PLINTA Dariusz	Akademia Techniczno-Humanistyczna w Bielsku-Białej
RAJZER Izabella	Akademia Techniczno-Humanistyczna w Bielsku-Białej
RYSIŃSKI Jacek	Akademia Techniczno-Humanistyczna w Bielsku-Białej
SKOŁUD Bożena	Politechnika Śląska
STADNICKI Jacek	Akademia Techniczno-Humanistyczna w Bielsku-Białej
VASILIU Yerhen	Odeska Narodowa Akademia Łączności im. O.S. Popowa, Ukraina
VLASYUK Anatoliy	Międzynarodowy Uniwersytet Ekonomiczno-Humanistyczny im. akad. Stepana Demianczuka, Ukraina
WIĘCEK Dariusz	Akademia Techniczno-Humanistyczna w Bielsku-Białej
WIĘCEK Dorota	Akademia Techniczno-Humanistyczna w Bielsku-Białej
WOJNAR Grzegorz	Politechnika Śląska
ZAWIŚLAK Stanisław	Akademia Techniczno-Humanistyczna w Bielsku-Białej

Początki **Wydziału Budowy Maszyn i Informatyki** sięgają 1969 roku, kiedy utworzono oddział Wydziału Mechanicznego Politechniki Łódzkiej. W dniu 1 października 1976 roku stał się samodzielnym wydziałem zamiejscowym Politechniki Łódzkiej. Aktualnie jest jednym z pięciu wydziałów tworzących Akademię Techniczno-Humanistyczną w Bielsku-Białej, która powstała w 2001 roku. Wydział ma pełne prawa akademickie wynikające z uprawnień do nadawania stopni naukowych doktora i doktora habilitowanego w dyscyplinie budowa i eksploatacja maszyn oraz doktora w dyscyplinie inżynieria produkcji. Tworzy go osiem jednostek wydziałowych, w tym sześć katedr i dwa zakłady.

Na wydziale prowadzone są studia na kierunkach:

- mechanika i budowa maszyn,
- zarządzanie i inżynieria produkcji,
- automatyka i robotyka,
- informatyka,

na trzech poziomach studiowania: inżynierskim, magisterskim i doktoranckim. W swojej 48-letniej historii na wydziale wypromowano ponad. 8,1 tys. inżynierów i magistrów inżynierów, którzy zasilili kadrę techniczną wielu firm przede wszystkim Bielska-Białej i regionu, przyczyniając się istotnie do ich rozwoju. Kadrę Wydziału stanowi: 10 profesorów tytularnych, 26 doktorów habilitowanych, 52 doktorów, 18 magistrów i 29 pracowników administracyjnych i inżynieryjno-technicznych. Podstawowe obszary badań uprawianych na Wydziale związane są z prowadzonymi kierunkami kształcenia i obejmują zagadnienia z zakresu: projektowania, analizy i badań doświadczalnych konstrukcji mechanicznych; projektowania procesów technologicznych; zarządzania i organizacji tych procesów; metrologii, ergonomii i logistyki; automatyzacji i sterowania maszynami i urządzeniami; projektowania, analizy i badań doświadczalnych pojazdów; układów napędowych, silników, a także systemów przetwarzania danych, administrowania sieciami komputerowymi i bezpieczeństwa informacji.

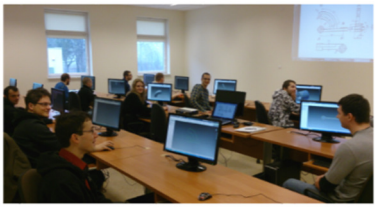
KOŁO NAUKOWE "INŻYNIER XXI WIEKU"

Inżynier XXI wieku

Jeżeli fascynują Cię nowe rozwiązania techniczne, masz własne pomysły na małe projekty badawcze i chcesz podjąć wyzwanie w ich realizacji z kolegami z innych dziedzin nauki - **dołącz do naszego koła!** W ramach działalności koła naukowego zespoły projektowe złożone ze studentów różnych kierunków studiów realizują zadania badawcze na styku mechaniki, automatyki i informatyki.

Nauka może być zabawna! Przekonaj się o tym osobiście biorąc udział w konkursach, np. w programowaniu i budowie minirobotów. Na najlepszych czekają naprawdę bardzo atrakcyjne nagrody.

Studenci w ramach koła naukowego mają dostęp do specjalistycznego laboratorium, wyposażonego między innymi w: skaner 3D, drukarkę 3D, laser pomiarowy z oprzyrządowaniem, miniroboty, mikrofabrykę. Mogą korzystać ze specjalistycznego oprogramowania z dziedziny projektowania, obliczeń wytrzymałościowych, oprogramowania sterowników przemysłowych oraz robotów.



Więcej informacji na stronie: **www.EngineerXXI.ath.eu**

facebook

Dołącz do nas na Facebook'u !
Strona: Koło naukowe "Inżynier XXI wieku"

ASTOR
gdzie technologia spotyka człowieka

SHILOH®

EATON
Powering Business Worldwide

STRUMET

PM POLMOTORS

BLUMENBECKER

ELOKON
Bezpieczeństwo. Postęp.

REKORD
SYSTEMY INFORMATYCZNE

FCA
FIAT CHRYSLER AUTOMOBILES

ALFA ROMEO

FIAT

FIAT PROFESSIONAL

LANCIA

CENTRUM SPRZEDAŻY
FCA POLAND

Bielsko-Biała, ul. Katowicka 24
tel. 33 813 44 42
www.centrumsprzedazy.fiat.pl

Fuks
GROUP

Przetwarzanie, transmisja i bezpieczeństwo informacji
Processing, transmission and security of information

Przetwarzanie, transmisja i bezpieczeństwo informacji

Processing, transmission and security of information

Anastasiia ABAKUMOVA[1], Mariia ROSHCHUK[2]

Opiekun naukowy: Roman ODARCHENKO[3]

STUDY THE PROBLEM OF SERVICE PROVISION QUALITY ASSESSMENT IN CELLULAR NETWORKS

Summary: The paper discusses key performance indicators to understand which part of the data is key to supporting business goals. The concept of key quality indicators and the selecting methodology of KPI/KQI functioning of telecommunication resources and services are presented. A scalar measure of the telecommunication network effectiveness with an arbitrary structure is proposed.

Keywords: efficiency, indicator, quality, telecommunication network.

ANALIZA PROBLEMU OCENY JAKOŚCI USŁUG W SIECIACH KOMÓRKOWYCH

Streszczenie: W niniejszym artykule, analizowane są wskaźniki działania sieci mających istotne znaczenie przy ocenie ich działania oraz efektywności. Zaprezentowano koncept tzw. kluczowych wskaźników jakości oraz metodologię wyboru KPI/KQI związaną z funkcjonowaniem sieci telekomunikacyjnych - oceny ich zasobów oraz usług. Zaproponowano skalarną miarę efektywności sieci telekomunikacyjnych o dowolnej strukturze.

Słowa kluczowe: efektywność, wskaźnik, jakość, sieć telekomunikacyjna

1. Introduction

The problem of telecommunications networks effectiveness assessment in communication theory is given quite a lot of attention. Finding an adequate solution to this problem will largely determine the possibility of improving the efficiency of the telecommunication networks functioning. The relevance of this problem increases with the development of telecommunication networks, the structures of which are

[1] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD student, nastia.abakumova@gmail.com
[2] National Aviation University, Department of Constitutional and Administrative Law Education and Research Institute of Low, PhD student, roshchukmv@gmail.com
[3] PhD, associated professor, National Aviation University, Institute of Air Navigation, odarchenko.r.s@ukr.net

rather complex or unknown, or have the properties of dynamically changing configurations. In addition to the difficulties associated with multi-pole networks effectiveness assessment, there are certain problems in justifying the indicators and performance criteria, because of their abundance and variety arising in the stages of creation and various telecommunication networks operation.

2. Setting objectives

The goal of key performance indicators (KPI) detailed description is to establish a close relationship between generally accepted directives for network management indicators and the current aggressive "business focus" in the telecommunications industry. The environment is aware that the influence of processes determines at the organizational level business goals, such as income (growth and protection), cost reduction, etc.

The important component of the telecommunication sector objectives implementation is the identification and acceptance of verifiable performance indicators. These indicators will help assess the level of decision-makers achievement and company executives responsible for implementing these decisions. Therefore, the choice of indicators for research depends on the specific focus of the problem. In the telecommunications sector, the problems to be studied may include [1]:

- Large unmet demand for services and lack of next-generation telecommunications services required for business and commerce;
- Poor quality of service (QoS);
- Weak financial support and lack of financial resources;
- Lack of skilled manpower.

One of the challenges associated with networks management and their services is the ability to understand what part of the data is key to supporting business goals. For example, in terms of revenue growth, the availability of services or networks can provide revenue in the future. However, network availability may be a necessary condition, which is established by state regulation to guarantee the service.

Key performance characteristics are those measures that directly support key business objectives and can be quantified in real time. Key performance characteristics are identified as:

- Availability;
- Reducing the impact of network failures;
- Data correctness;
- Protection;
- Downtime;
- Notification;
- Reliability;
- Frequency;
- Effectiveness;
- Productivity;
- Security.

Performance indicators should be used to monitor progress in the main direction and then address various problems affecting the performance of work.

3. The concept of key quality indicators

Business considers the information technology (IT) as a means of increasing its productivity and improving competitiveness. The effectiveness of business processes implementation depends on the quality of IT services. The increasing number of IT services required to automate business technologies, the complexity of applications and the increasing number of IT infrastructure components lead to a decrease in the efficiency of IT departments and increase the costs of maintaining the regular mode of the IT infrastructure functioning. The IT services provision is governed by the package of service level agreements (SLAs) concluded between business units and the IT department. The SLA defines the values of key performance indicators (KPIs) and quality (KQI) – a limited set of objectively measurable parameters, allowing to assess the quality of IT services [2]. To maintain the KPI and KQI values at the level fixed in the SLA, administrators ensure the continuous functioning of the IT infrastructure, perform maintenance and repair using automatic, automated and manual control methods.

Key Performance Indicators (KPI) represents the results of tests and measurements or statistics obtained directly from network technical resources or applications (Fig. 1). Nowadays, there is a large list of KPIs for each type of network technology and services/applications. For a more detailed account of the network features, this list can be extended by additional KPIs defined directly by the representatives of the IT service.

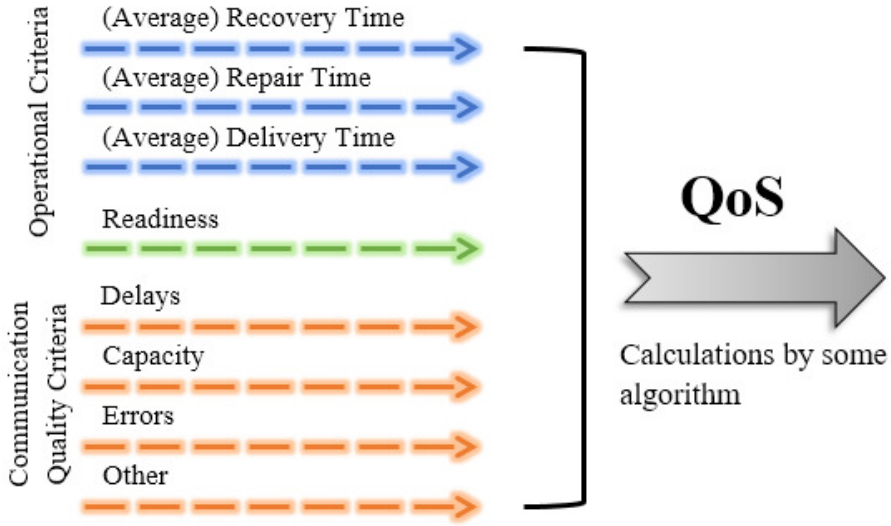


Figure 1. KPI communication with QoS

In the process of further analytical processing, based on the service model, the methods and technologies of their provision, KPIs are aggregated into KQI – key quality indicators of the service or its component parts. The relationship between KQI and its defining set of KPIs, as well as their threshold values, is established both from theoretical calculations and experimentally (Fig. 2).

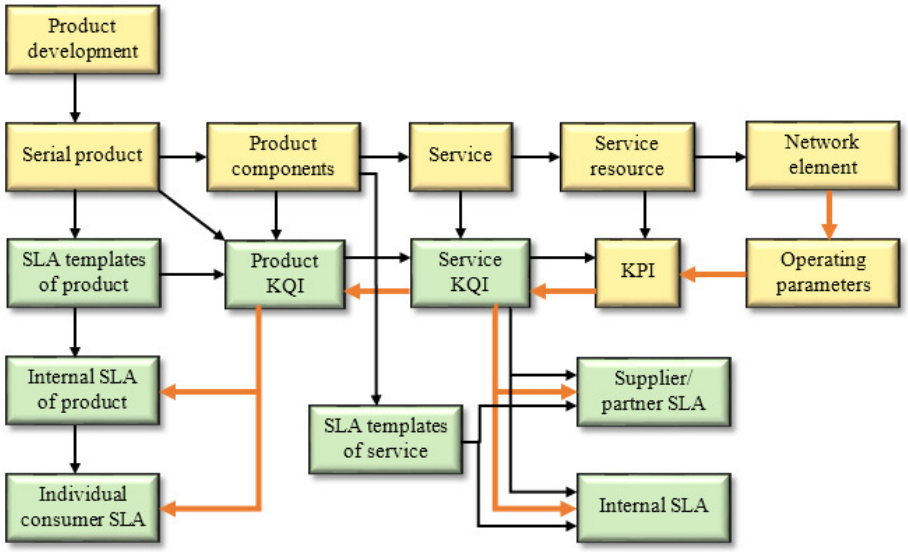


Figure 2. The sequence of actions in determining KQI

In turn, the set of relevant KQIs defines the Product Key Quality Indicator (PKQI), which is the main metric in the SLA definition (Fig. 3).

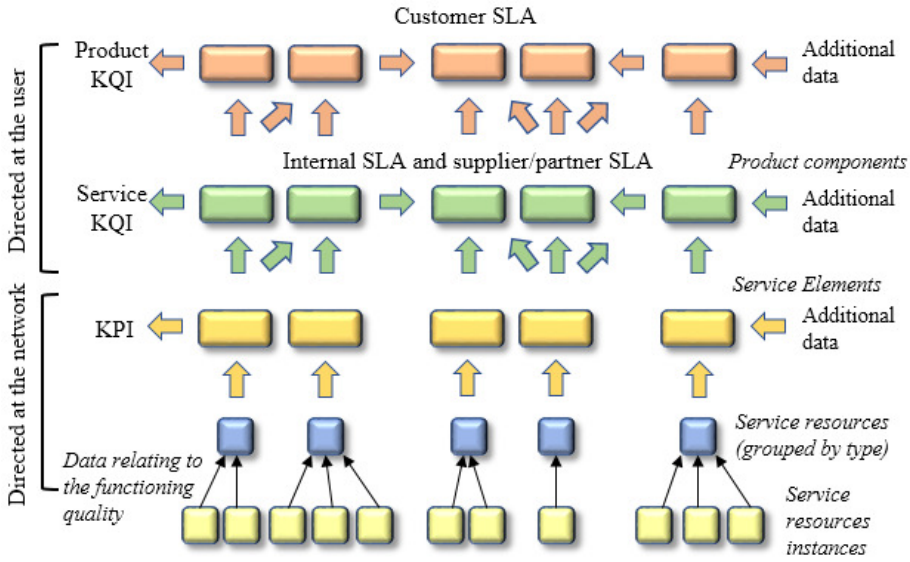


Figure 3. Hierarchy KPI/KQI/PKQI

When choosing the indicators necessary for an adequate assessment of the QoS, it is necessary to minimize their number and take into account the possible "cross" effect of a separate KPI on several different KQIs.

Choice methodology of KPI/KQI functioning of telecommunication resources and services (Fig. 4) consists of the following steps:

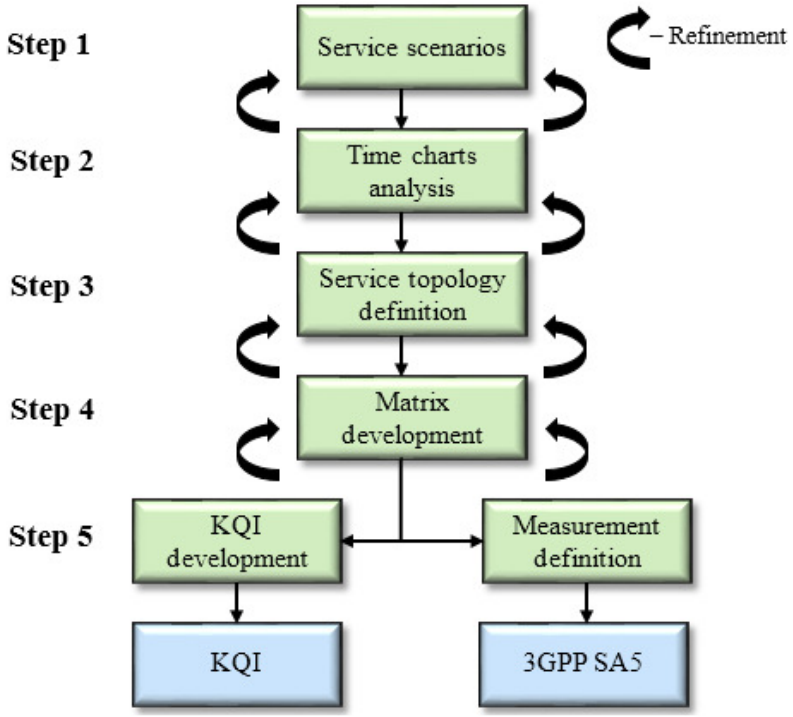


Figure 4. Choice methodology of KPI/KQI functioning of telecommunication resources and services

Step 1. Description of the supplier services spectrum and service components.

Step 2. Description of all supplier interactions with the consumer at all stages of the service delivery.

Step 3. Description of all service resources that make up its complete topology.

Step 4. Detailing all supplier interactions with the service consumer to the network resource level.

Step 5. Determination of the indicators needed to evaluate KQI and methods for their measuring.

4. Evaluation of the telecommunication network efficiency

Among the QoS network performance indicators, the most commonly used data rate, packet delay, the deviation from the average packet delay, the level of packet loss and distortion. It is quite obvious that in the presence of several quality indicators it is

desirable to have a generalized efficiency indicator suitable for telecommunication networks of any kind of their functioning organization.
The problem of generalization of particular indicators is usually solved by moving to the indicator in the form of a weighted mean [3]:

$$F = \sum_{i=1}^{k} h_i F_i \,, \tag{1}$$

where F_i – private quality indicator; h_i – weight coefficient of the private indicator F_i, determined on the basis of subjective considerations and normalized by k private quality indicators, wherein

$$\sum_{i=1}^{k} h_i = 1 \,. \tag{2}$$

The generalized efficiency index F obtained in this way is a scalar quantity that has a significant drawback, namely, the subjectivity of the assessment, which is a consequence of the subjective determination of private quality indicators weight.
The paper presents an alternative approach to obtaining a scalar indicator of the telecommunication network efficiency that reflects the interests of users and provides quantitative estimates of the network effectiveness directly in the process of its operation.
Let's imagine that the network structure and the principles of its functioning are arbitrary and unknown to us (Fig. 5).

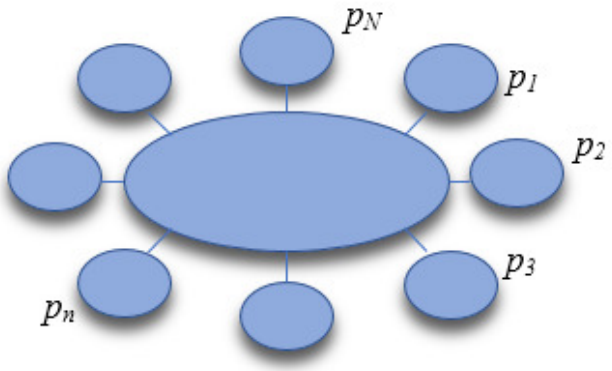


Figure 5. Telecommunication network model

In assessing the number of users in the network N is known. Each user p_i, $i = \overline{1,\ N}$ has the potential to arbitrarily communicate with any other network user p_j, $j = \overline{1,\ N}$, $j \neq i$. We assume that the communication session p_i user with p_j user is considered successful when the user p_i sends and the user p_j receives a confirmation of the successfully received information message.
Assume that the users' interests in the telecommunication network efficiency are set of quality indicators that are presented to each individual information message in the form of an admissible accuracy index of the k-th message α_k received by the p_j user, wherein

$$\alpha_k \geq A_k\,, \tag{3}$$

where A_k – minimum admissible accuracy reproduction of the user's p_j message of this type.
The second quality indicator of the communication session is the requirement for the admissible transmission time of the *k*-th message t_k, presented by the sender as an inequality:

$$t_k \leq T_k\,, \tag{4}$$

where T_k – maximum allowed time for message delivery of this type.

In addition, previously formed success conditions of the communication session between users put forward certain requirements for success sending confirmation of the message. It is obvious that the confirmation, as well as the information message, must correspond to the requirement of reproduction accuracy, which can be written in the form of inequality

$$\beta_k \geq B_k\,, \tag{5}$$

where B_k – minimum acceptable accuracy of confirmation playback, and the request for an acceptable timeout for confirmation

$$\tau_k \leq T_l^D\,, \tag{6}$$

where T_k^D – confirmation directive timeout.

With a detailed consideration of inequality (6), it becomes evident that the acceptable timeout for confirmation includes the message transmission time t_k, the preparation time and the confirmation directive timeout. Proceeding from this, it is fair to assume that the requirements of inequality (6) are more stringent than the requirements of inequality (4) and take into account the time constraints imposed both on the message transmission and on the confirmation transmission, the excess of which is classified as an unsuccessful communication session.
The foregoing assumptions of the telecommunications network functioning make it possible to formulate scenarios for a separate information message transmission. Suppose that at an arbitrary moment of time t_k, the network user p_i sent a *k*-th message to the recipient p_j, $j \neq i$. This moment of time is fixed by the user p_j as the moment of sending the message t_k^i.
Further, as a result of message sending to the recipient p_j, the following group of events can be formed:

- The message is accepted, the verification of the received message for compliance with the condition (3) was successful. The recipient of the message p_j fixes the receipt time of the message t_k^i, which is transmitted in confirmation to the sender p_i;
- The message is accepted, but the verification of the received message for compliance with the condition (3) was unsuccessful. This fact by the recipient of the message p_j can be transmitted to the sender p_i in the confirmation, which allows the sender to fix in the memory the network failure. Or confirmation

of message receiving by the receiver p_j may not be transmitted at all, which will also be qualified by the sender p_i as a network failure after the time indicated by the condition (6);
- The message is not received. In this case, the sender fixes the system failure due to the condition (6) failure.

The following group of events is formed on the sender's side p_i and is related to the fact of receiving confirmation from the recipient of the message p_j:
- Confirmation of the successful transmission of the *k*-th information message was obtained in accordance with conditions (5) and (6). The sender fixes the fact of the *k*-th information message transmission;
- Confirmation is accepted in accordance with condition (6), but non-compliance with condition (5) leads to impossibility of its reproduction, as a result of which the sender fixes the fact of the *k*-th information message loss. The sender registers a network failure and resends the *k*-th message;
- Confirmation is not accepted due to exceeding its waiting time, regulated by the condition (6). The sender p_i registers a network failure and resends the *k*-th message.

Thus, as a result of the review, the concept of denial in the network is completely specified in terms of a separate information message transmission, where the fact of the network failure is recorded by the sender.

Further, suppose that for some time interval each p_i-th network user transmitted informational messages K_i with a certain number of successful s_i to arbitrary network subscribers. The ratio of these values is a statistical probability estimate of successful transmission of P_i messages for the p_i-th user:

$$P_i^* = \frac{s_i}{K_i}. \tag{7}$$

It can be assumed that this estimate is an exhaustive characteristic of the telecommunication network efficiency from the message sender point of view, since along with the itself network characteristics, random states of arbitrary message recipients are also taken into account. A generalized estimate, obtained taking into account the private statistical probabilities P_i^* of each network user *N*, could fully characterize the telecommunication network effectiveness from each user point of view.

It is obvious that the contribution of each user to the network information flow will be different, which is generally determined by the quantitative indicators of the messages transmitted in the network. Proceeding from this, it is fair to assume that the weight of the private statistical probabilities in the generalized efficiency index can be determined by the proportion of sent messages in the general network information flow. To form an integrated indicator of the telecommunication network efficiency, we will use the quantitative measure of each user weight in the general network information flow [4]:

$$w_i^* = \frac{K_i}{\sum_{i=1}^{N} K_i}. \tag{8}$$

In the presence of a sufficiently large information flow in the network, that is for $K_i \to \infty$, $i = \overline{1,\ N}$, the measure of the weight w_i^* converges to the posteriori probability w_i of the interaction of the i-th user with the telecommunication network in the messaging mode, and always have a fair relationship

$$\sum\nolimits_{i=1}^{N} w_i \ . \tag{9}$$

Thus, the integral indicator of the telecommunication network efficiency can be represented as:

$$P = \sum\nolimits_{i=1}^{N} w_i P_i \ . \tag{10}$$

The presented indicator is devoid of the subjectively determined estimate weight (1), which in this case is determined on the basis of the network load objective characteristics. In addition, the statistical probability of the success message delivery, in contrast to (1), reflects the evaluation of the network functioning from the user point of view.

Thus, at the physical level, indicator (10) represents the estimate probability that any message sent by an arbitrary network user at a random time will be successfully delivered to the recipient.

5. Conclusion

Thus, the goal of the indicators is to help the management to monitor the performance of work and to perform corrective action as needed. Timely introduction of these indicators is particularly important when there are problems. The system of performance indicators should have inherent mechanisms that can immediately signal serious problems.

The obtained index completely satisfies the conditions of the problem posed at the beginning of the article. It is scalar, provides an exhaustive objective and physically understandable assessment of the telecommunications network efficiency. The maximum measure indicator takes into account the requirements of network users and is based on quite natural physical premises.

REFERENCES

1. ITU-T E.419 (2006): Business oriented key performance indicators for management of networks and services. Telecommunication standardization sector of ITU. Series E: Overall Network Operation, Telephone Service, Service Operation And Human Factors. Network management – International network management
2. BROOKS P..: Metrics for IT Service Management, Van Heren Publishing 2008.
3. ЩЕРБИНА Л.П. Основы теории военной связи, Л.П. Щербина. – Л.: ВАС, 1984. – 170 с.

4. KORNIENKO S., KORNIENKO I.: the evaluation of space-allocated Telecommunication networks, І.В. Корнієнко, С.П. Корнієнко, Чернігівський науковий часопис. Серія 2. Техніка і природа: електронний збірник наукових праць. – Чернігів: ЧДІЕУ, 1(2011)1, 96–101.

Zhibek ALIBIYEVA[1], Anar TASHIMOVA[2]

Scientific Supervisor: Ihor TEREIKOVSKYI[3]

REDUKCJA SZUMU SYGNAŁU GŁOSOWEGO W BIOMETRYCZNYCH SYSTEMACH UWIERZYTELNIANIA

Streszczenie: Artykuł poświęcony jest opracowaniu metody redukcji szumów sygnału głosowego przeznaczonego do wykorzystania w biometrycznym uwierzytelnianiu użytkowników ogólnych systemów informacyjnych. Wskazano jakie procedury analizy widma sygnał są zalecane do zastosowania. Określono kolejne etapy zaproponowanej metody, ponadto generowane jest niezbędne oprogramowanie. Badania doświadczalne pokazują dalsze perspektywy metody, jak również główne kierunki jej dalszego rozwoju.

Słowa kluczowe: uwierzytelnianie biometryczne, sygnał głosowy, redukcja szumu

VOICE SIGNAL'S NOISE REDUCTION IN THE BIOMETRIC AUTHENTICATION SYSTEMS

Summary: The article is devoted to the development of a method for noise reduction in a voice signal intended for use in biometric authentication, general purpose information systems users. The usage of the spectral subtraction of the signal is substantiated in the method of the procedure. The stages of the method were determined, the necessary software was generated. Experimental studies show the prospects of the method, as well as the main directions of a development are identified.

Keywords: biometric authentication, voice signal, noise reduction

1. Introduction

Qualitative noise reduction of the voice signal is one of the main factor ensuring the effective functioning of systems for the analysis of voice signals of any type [3, 6]. Especially, the quality of noise reduction is important in biometric authentication

[1] Kazakh National Research Technical University named after K.I.Satpayev, Institute of Information and Telecommunication Technologies: Computers and software, Doctoral Student, alibieva_j@mail.ru

[2] Aktyubinsk regional state university named after K. Zhubanov, Institute of Information and Telecommunication Technologies: Computers and software, Doctoral Student, tashimova_a@mail.ru

[3] Doctor of Technical Sciences, Assoc., National Technical University of Ukraine, Igor Sikorsky Kyiv Polytechnic Institute, Proffesor of the Department of System Programming and Specialized Computer Systems, terejkowski@ukr.net

systems, the recognition accuracy of which in many cases is critical [3].Although the analysis [1-6] points out significant scientific and technical developments in this direction, it also notes the absence of a detailed description of noise reduction method, which is optimal from the point of view of the conditions of use for biometric authentication for general purpose information systems users. This determined the purpose of the study - the development of a method of voice signal's noise reduction, optimal from the point of view of its use in general purpose information systems for biometric user authentication.

2. Analysis of common methods for noise reduction

In accordance with [4], the following methods have been widely used in modern systems for processing voice signals:

- Spectral subtraction;
- Active noise cancellation;
- Frequency limits.

Using the data [1,2,4,5], was carried out an analysis of the main characteristics, merits and demerits of these methods .

Method of spectral subtraction. The essence of the method is that the spectrum of pure noise is subtracted from the amplitude-frequency spectrum of the useful signal. The specified spectrum is either determined in advance before the signal is translated, or is determined at the time of translation. In this case, the number of frequency bands to which the signal is divided, depending on the implementation of the method, can range from several tens to several thousand units. Thus, the bandwidth of the signal processing can be adapted to the operating conditions of the voice signal processing system. Due to this, it is possible to achieve an optimal compromise between the efficiency of noise reduction and the computational costs of its implementation.

Active noise cancellation method. The essence of this method is to get a synchronous audio signal synchronously with the signal being cleared, which contains only the noise component (without the voice signal of interest).Subsequently, the received audio signal is inverted into the so-called antiphase and superimposed on the signal being cleared. As a result, noise and anti-noise negate each other, and you can get a clear voice signal at the output. It should be noted that under the expected conditions for using biometric authentication for general purpose information systems users, it is extremely difficult to obtain a signal corresponding only to the noise component. This shortcoming points to the need for significant improvements of this method in case of biometric authentication.

Method of frequency limitation. The essence of the method consists in the complete removal of frequencies from the sound signal that are not characteristic of a human's voice. The disadvantage of this method is that the noise remains at frequencies coinciding with the frequency range of a human's voice. This greatly complicates the application of this method in biometric authentication systems.

As a result was determined the expediency of choosing the method of spectral subtraction as the basic method of noise purification.

3. Characteristic of the spectral subtraction method

The method is based on the procedure of spectral subtraction of signals, the theoretical basis of which is the assertion that the sum of spectrum of the voice signal and noise is equal to the sum of the spectra of this signal and noise. Therefore, in order to clear an existing noisy signal, it is necessary to realize the spectral subtraction of the noise spectrum's amplitudes from each spectrum's amplitude of this signal at each fixed time. Obviously, in order to obtain the above amplitude spectra, it is necessary to have a spectrogram of both the analyzed signal and the noise spectrogram. The spectrogram is a graph of the signal amplitude versus the sampling frequency. The spectrogram can be calculated by using a fast Fourier transform. In the base case, in order to obtain the initial data in the calculation of the noise spectrogram, it can be assumed that the noise is a manually set and averaged sound signal at a certain time interval.
After the considered process of subtraction from the purified spectrogram, the synthesis of the purified signal is realized, which is represented a purified voice signal. Note that after subtraction, the phase spectrum of the voice signal remains unchanged. This disadvantage is due to the probabilistic nature of the noise, which does not allow us to determine its phase spectrum. However, the results of [2, 3] indicate a minor effect of this deficiency on the effectiveness of voice signal analysis in the biometric authentication of users.
The considered procedure of subtraction has a certain similarity with the multiband gate procedure based on the suppression of signals whose level is below a certain threshold and skipping unchanged signals whose level is above this threshold.
From this position, the spectrogram can be treated as a set of horizontal bands, each of which is a separate sub-band signal. And over each such band, a subtraction operation is performed from its amplitude of a certain constant-the noise level in this band. When the noise level in the band is large, subtraction of a small constant practically does not change the signal in this band. When the signal level is close to the noise level, reducing the signal level by the amount of the noise level is actually equivalent to using a gate with a "soft threshold". Therefore, the whole process of spectral subtraction can be considered as a multiband gate with a large number of "soft thresholds" [4].
At the present time, there are many different realizations of the spectral subtraction method, differing in different characteristics (the type of transfer characteristic for individual gates, the number of frequency bands, methods of combating various artifacts) and, accordingly, the areas of use. This indicates its flexibility and confirms the prospects of its adaptation to biometric authentication of general-purpose information systems users.
In the basic version, the implementation of this method consists of performing 6 steps.

1. To the input of method, a raw (noisy) $y[n]$ signal is supplied in digitized form, which consists of a voice signal $x[n]$ and an additive noise:

$$y[n] = x[n] + d[n],\ n = 1,...N\ , \tag{1}$$

Where: n – number of signal sampling points, N - number of sampling points.

As an illustration of method's input data, Figure 1 shows a plot of the amplitude of an arbitrary initial signal as a function of time, and Figure 2 shows a graph of the voice signal's amplitude.

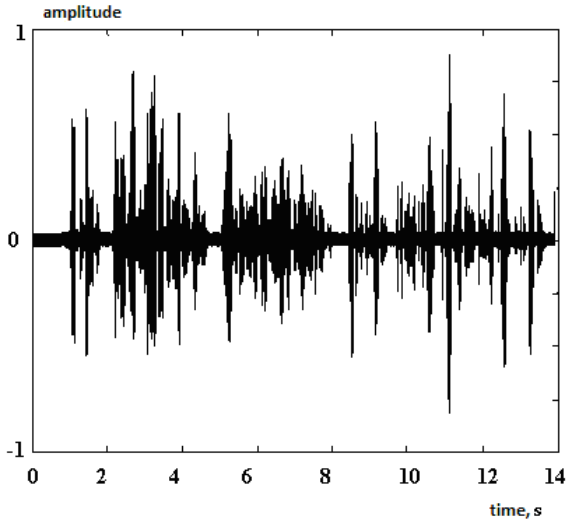


Figure 1. Noisy signal

2. In order to reduce bursts of energy, which entail a distortion of signal, it is divided into quasi-stationary sections (frames). The length of the frame is 10-20 ms. Typically, the frames have a half overlap. The separation process is illustrated in Figure 3.

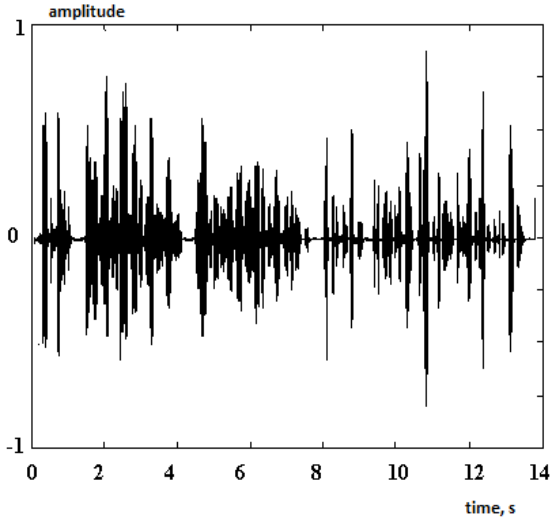


Figure 2. Voice signal

3. For each of the frames, a discrete Fourier transform is used whose mathematical apparatus is determined by an expression of the form:

$$Y(f)=\sum_{n=0}^{N-1} y[n]\bullet \exp(-2j\pi kn/N), k=0,1,...N-1, \quad (2)$$

where $Y(f)$ - value of signal amplitude at frequency *f*.

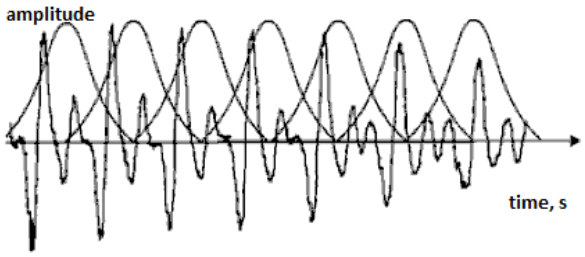


Figure 3. Overlapped signal

An example of a spectrogram for a certain frame of a noisy signal, obtained as a result of applying a Fourier transform, is shown in Figure 4.

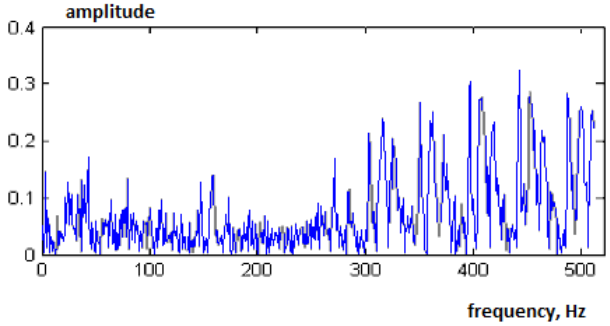


Figure 4. The spectrum of the noisy signal (1 frame)

4. For each of the frequencies, calculated the noise envelope:

$$Noise = p\bullet|d(i-1)|+(1-p)\bullet|d(i)| \quad (3)$$

5. The subtraction of the noise component from the signal

$$S=Y[f]-\alpha\bullet Noise[f], \quad (4)$$

where α - coefficient of noise suppression (in the first approximation).

6. Restore signal is implemented using an inverse discrete Fourier transform of the form:

$$S[n]=\sum_{m=0}^{N-1} S[m]\bullet\sum_{k=0}^{N-1}\exp\left(j\bullet\frac{2\pi}{N}\bullet(n-m)\bullet k\right). \quad (5)$$

Elements of an array correspond to the values of the cleared voice signal.

A generalized scheme of the subtraction method is shown in Figure 5.

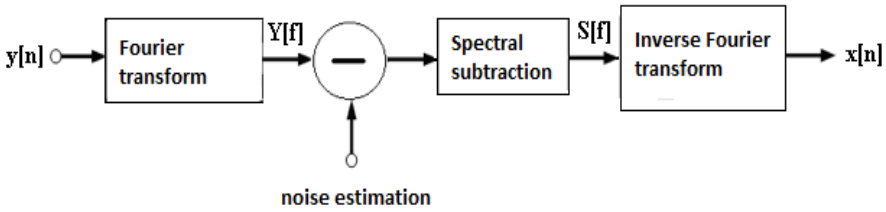


Figure 5. The scheme of spectral subtraction method

The described method was used in the development of a software complex for noise reduction in the voice signal. The program code is written in the Python programming language using additional Qt and PyQt libraries.
The display of the program, which has just started, is shown in Figure 6.

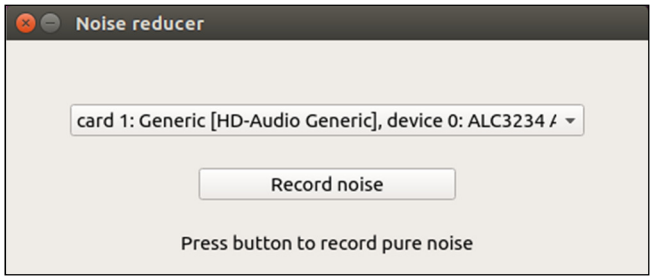


Figure 6. The window of the running program

This window displays a list of available audio recorders (microphones), a button for initial recording of background noise, and a text field for displaying the current state. The function code, which is called when the "Record noise" button is clicked, is shown in the listing:

```
def button_noise_callback():
    nonlocal self
    def on_process_end():
    nonlocal self
    def on_process_end_inner():
        self.btn.setText('Record voice')
        self.btn.setEnabled(True)
        self.btn.clicked.disconnect()
    self.btn.clicked.connect(button_voice_callback)
    self.label.setText('Press button to record voice')
    self.inner_process = QProcess()
    self.inner_process.finished.connect(on_process_end_inner)
    self.card_num = int(re.match(r'^card (?P<card_num>\d*):
    .*$',
    self.combobox.currentText()).groupdict()['card_num'])
    self.combobox.setEnabled(False)
    self.btn.setEnabled(False)
```

```
self.label.setText('Please wait 5 sec for noise to be
recorded')
self.process = QProcess()
self.process.finished.connect(on_process_end)
self.process.start('ffmpeg -f alsa -i hw:{0} -t 5 -y
noise.wav'.format(str(self.card_num)))
```

Note that this function is intended to:
- change the state of the dynamic list of devices, button, text box;
- run the ffmpeg library to record sound with noise;
- save sound to a text file;
- start the procedure for spectral conversion of the noise signal recorded in the file.

Therefore, by using this developed program were conducted experiments aimed at verifying the basic version of this method. Experiments have shown that the results of noise purification are satisfactory in case of stable uniform noise. Furthermore, in case of dynamic nature of noise, the quality of noise reduction is dramatically reduced. This indicates the need to adapt the method to this kind of noise.

4. Conclusion

It is shown that an important step in the development of biometric authentication systems is the development of a method for removing noise in the voice signal, which can be performed on the basis of the spectral subtraction procedure. Therefore, method steps and mathematical apparatus are determined which are based on the use of discrete Fourier transform. Moreover, the corresponding software has been developed, which makes it possible to carry out noise reduction from the voice signal. Thereby, conducted experiments confirmed the promisingness of the proposed method and allowed to determine that an important direction of its improvement is the adaptation to the dynamic nature of noise.

BIBLIOGRAPHY

1. LUKIN A., TODD J.: Adaptive Time-Frequency Resolution for Analysis and Processing of Audio: Databook IEEE Transactions on speech and audio processing, 2006.
2. MUSTIERE F., BOUCHARD M.: Efficient SNR-based subband post-processing for residual noise reduction in speech enhancement algorithms: Databook 18-th European Signal Processing Conference, 2010.
3. KORCHENKO A., TEREYKOVSKIY I., KARPINSKIY N., TYNYMBAYEV S.: Neural network models, methods and security options assessment tools Internet-oriented information systems. – Nash Format, Kiev 2016. (in Russian).
4. Analyzer of speech and sound signals. Methods, algorithms and practice, Under the editorship of. A.A. Petrovsky. Mn.: Theprint, is white 2009. 397-430.

5. Noise Estimation for Speech Enhancement in Non-Stationary Environments – A New Method. : Databook World Academy of Science Engineering and Technology 70, 2010.
6. TEREYKOVSKA L.O.: Neyromerezhevi modeli ta metodi rozpiznavannya fonem v golosovomu signali v sistemi distantsiynogo navchannya: dís. kand. tekhn. nauk: 05.13.06, Kiev 2016. (in Ukrainian).

Karyna ALIEKSIEIEVA[1]

Scientific Supervisor: Serhii TOLIUPA[2]

THE INFORMATION SYSTEM OF DECISION SUPPORT AS THE CORE ELEMENT OF INCIDENT MANAGEMENT

Summary: Analysis of existing systems of management of modern information and telecommunication networks have shown that their level does not fully meet the modern requirements to the management of next generation networks, does not allow to obtain information of the right quality for making decisions on property management, exchange of information between the management system and also makes it impossible to quickly manage situations on networks in an automated mode.

Keywords: incident, decision support, incident management, information system, security.

INFORMATYCZNY SYSTEM WSPOMAGANIA DECYZJI JAKO PODSTAWOWY ELEMENT ZARZĄDZANIA INCYDENTAMI

Streszczenie: Analiza istniejących systemów sterowania nowoczesnych technologii informacyjnych i telekomunikacyjnych, sieci wykazała, że ich poziom nie w pełni odpowiada wymogom zarządzania sieci nowej generacji. Na przykład, nie pozwalają one na otrzymywanie informacji o pożądanej jakości do podejmowania decyzji w zakresie zarządzania obiektami, wymiany informacji pomiędzy systemami sterowania, a także sprawia, że niemożliwe jest, aby szybko poradzić sobie z sytuacjami w sieciach w trybie automatycznym.

Słowa kluczowe: incydent, wspomagania decyzji, zarządzanie incydentami systemu informacyjnego, bezpieczeństwa.

[1] Taras Shevchenko National University of Kyiv, Faculty of Information Technology, Department of Cybersecurity and Information Protection, specialty: Information Security Management, Student, karinaalekseeva@ukr.net

[2] Ph.D., Professor of Cybersecurity and Information Protection, Taras Shevchenko National University of Kyiv, faculty of Information Technology, tolupa@i.ua.

1. Formulation of the problem

The development of information technology, the Internet and the growth of the number and capacity of computing systems directly affect the modern business environment. The new types of threats to information introduce new approaches to building protection systems, the protection against threats in real-time. The arising number of security issues in the information environment makes you look for new methods and approaches to the management of information security incidents.

2. Problem definition

Incidents of information security (IS) is a separate subclass of crisis and emergency situations that can happen in the info-socio-technical infrastructure of the country, and, as a special case, in organizational-technical systems (OTS) and communication networks (CN), affecting the status of state information resources and national security.
The organization of information security incidents management is a complicated and time-consuming task without the use of automation. However, it is still possible to identify general principles that can be used by the information system of decision support (ISDS) developer.

3. Statement of the main material

The first step in making decisions related to the management of incidents should be the systematization and inclusion the incident classes definitions to the Service Level Agreements (SLA). It is proposed to constantly modernize the laws, methods and institutions in the field of information technology management and information security.
Specific issues of the management of information security incidents are described in the following documents: ISO/IEC 27001:2005 Information security management system. Requirements; ISO/IEC TR 18044 Information security incident management [1]; CMU/SEI-2004-TR-015 Defining incident management processes for CISRT.
It is well-known that the intelligent systems have recently become a common commercial product that finds wide demand from users-specialists in various fields of engineering-technical and scientific spheres of activity [2].
Thus, the process of managing events and incidents can be carried out with solutions of Security information and event management (SIEM) class and well-designed organizational process.
Before selecting a SIEM system, you must define the criteria by which to make comparison. First and foremost, is the ability to collect events. The most significant in system of events and incidents management are the opportunities for intelligent analysis of events and their correlations.
Providing tools for investigation and analysis of incidents is one of the priority tasks implemented with SIEM solutions.

The final stage of any SIEM system is an appropriate response to identified incidents. The main criterion for an appropriate response is the number of alert ways (console, email, SNMP), as well as the quantity and quality of preset actions.
By this way, the information system of decision support is proposed as informational system, which can operate in the following algorithms:

1. Operational data of security sensors should be analyzed and evaluated for the presence of signatures of known incidents using knowledge base of information system of decision support (ISDS).
2. ISDS, on the basis of the analysis should provide the instruction on elimination of the causes and consequences of the incident, if the signature is in the knowledge base.
3. If there was an incident of the signature which is not in the knowledge base, ISDS gives several hypotheses regarding further actions the administrator of IS.
4. It is advisable to perform actions to prevent a recurrence of the incident. In order to have the procedure performed correctly and efficiently, all these steps must be constantly and consistently repeated, being fit into the cycle of the PDCA model, process approach defined in the relevant international standards ISO/IEC [3].

It is important to note that the procedure for incident management is closely linked to all other procedures of safety management in the company. Since the incident first of all is impermissible event, it should be forbidden by someone, therefore, you must have documents that clearly describe all the actions that you can perform in the system and perform illegal [4].

4. Conclusions

In conclusion, it is worth mentioning that the main objective of incident management is to restore normal services and minimize the negative impact of the incident on the organization to maintain the quality and availability of services at the highest of the possible level [5]. It's considered normal, that is not beyond the scope of the agreement about level of service. During this task we identified the main factors, criteria and indicators of automated procedures of decision-support for managing incidents.

REFERENCES

1. Standard ISO/IEC TR 18044:2004: Information technology - Security techniques - Information security incident management.
2. GLADYSH S. V.: Support decision-making for managing information security incidents to the organizational and technical systems. Expert systems and decision support. 116-124.
3. TOLIUPA S. V.: Design of the decision support system in the recovery process and ensure comprehensive protection in information systems. Scientific-technical journal “Modern information security. 4(2012). 69-74.

4. HOLOV A.: Responding to information security incidents, Intelligent Enterprise. 22(2005) E-resource.
5. MIELIEKNIN I.: Incidents management, Jet Info 7(2006) E-resource.

Zhuldyz ALIMSEITOVA [1], Nazym ZHUMANGALIYEVA[2]

Scientific Supervisor: Anna KORCHENKO[3]

SYSTEM DO IDENTYFIKACJI ANOMALNYCH STANÓW W SYSTEMACH INFORMATYCZNYCH

Streszczenie: Systemy komputerowe coraz częściej narażone są na działanie czynników stanowiących różnorodne zagrożenia. Nowe typy tychże systemów dają nowe możliwości przeprowadzania cyber-ataków na zasoby sieciowe. W celu zwiększenia poziomu bezpieczeństwa potrzebne są odpowiednie specjalne przeciwdziałania, które mogą być skuteczne dla nowych rodzajów zagrożeń i to w rozmytych warunkach w celu identyfikacji cyber-ataków na różne zasoby systemu informatycznego. Istnieje wiele modeli, technik i metod stosowanych do rozwiązywania zadań ochrony w warunkach rozmytych. Dla ich skutecznej realizacji wymagany jest odpowiedni sprzęt realizujący w celu wykrycia nieprawidłowego stanu. Opracowano systemu zorientowany na zadania wykrywania ataków komputerowych w systemach informatycznych. Zaproponowany system oparto na matematycznych modelach i metodach logiki rozmytej. Zatem, implementowano następujące podsystemy: tworzenia rozmytych wzorców, formowania reguł decyzyjnych, wstępnego procesowania, a także stworzono moduły: arytmetyki rozmytej, logiki wykrywania, wizualizacji i sterowania.

Słowa kluczowe: cyberatak, anomalie, system wykrywania włamań, system wykrywania anomalii, sieć komputerowa

A SYSTEM FOR IDENTIFYING ANOMALY STATE IN INFORMATIONAL SYSTEMS

Summary: Computer systems are increasingly exposed to impacts of threats, new types of them can caused new cyber attacks on their resources. For increasing the security level it needs appropriate special counteraction that can be effective in the emergence of new types of threats and allow in fuzzy terms to identify cyber attacks targeted at a variety of resources of informational systems. There are a number of models, methods and approaches used for solving protection tasks in fuzzy conditions. For their effective implementation it requires an appropriate system implements technology to identify the abnormal condition. In order to get this aim a system focused on the tasks detect cyber attacks in the informational systems, which is based on mathematical models and methods of fuzzy logics and is implemented through

[1] Kazakh National Research Technical University named after K.I. Satpayev, Department of Information security, lecturer, zhuldyz_al@mail.ru

[2] Kazakh National Research Technical University named after K.I. Satpayev, doctoral student, nazym_k.81@mail.ru

[3] National Aviation University, Academic Department of IT-Security, Associate professor, annakor@ukr.net

subsystems: the formation of fuzzy standards, the formation of decision rules, primary processing, as well as modules: fuzzy arithmetic, logical deduction, visualization and control.

Keywords: cyber attacks, anomalies, intrusion detection system, anomaly detection systems, intrusion detection system, anomaly detection system, computer networks

1. Introduction

The use of intrusion detection systems is directly related to the rapid development of cyberspace, where new types of threats resources of informational systems (RIS), such as 0-day attacks and non-signature types of cyber attacks. Expanding impact of cyber attacks aimed at different RIS initiates the creation of countermeasures, which are able to be effective in the emergence of new types of threats from unknown or ill-defined properties and actually operate in weakly formalized, fuzzy environment [1]. Using the methods, models and systems based on fuzzy sets [1-21] for the construction of means of detection of anomalies generated by the implementation of cyber threats which will improve the existing intrusion detection systems in computer systems and networks. In this context, the development of appropriate technical solutions, operating in fuzzy conditions and enable to identify new and modified types of cyber attacks is an actual scientific task.

2. The aim and tasks of the research

There are some, quite effective development used to solve these problems, the identification of cyber attacks, such as: fuzzy approaches to intrusion detection [2 -[3] and the detection of anomalies [4]; corresponding fuzzy models [1, 5], and methods [6,7]; intrusion detection systems [8-15]sets of fuzzy rules [2, 3], [7,9]; methods for constructing linguistic variables[16,[18] and fuzzy standards[19] as well as other developments that are used to address the protection problems in fuzzy terms[19]. These studies demonstrated the efficacy of the respective application of the mathematical equipment of fuzzy sets, and its use for the formalization of the approach to the identification of cyber attacks, will improve the process of creating the corresponding intrusion detection systems.However, in these studies there is no structured approach and a generalized solution for the construction of appropriate intrusion detection tools.

Based on an analysis of existing studies and the relevance of the task, as well as for the effective application of well-known models, methods and systems of the aim of this work is to develop a system which implements technology to identify abnormal states for intrusion detection systems, which can be used to improve network security oriented to control activity in surrounding environment.

With the help of such a system (in solving problems identifying cyber attacks) can effectively identify certain types of cyber attacks on the specific media environment in a given time period, as well as extend the functionality of modern systems of intrusion detection through efficient identification of new (0-day) and nonsignature types of cyberattacks.

3. The review of the previous researches

For solving this problem, a system implementing the technology for identifying the abnormal condition in the informational systems and networks (JIAS), which is based on a number of modules and subsystems. The first subsystem is aimed at measuring the current values of the quantities of network traffic, and the second by the formation of decision rules (designed to test the truth of the relationship of reference and current values for the evaluation of network activity), identifying the abnormal condition. The system includes:

1. Subsystem of forming fuzzy standards (SFFS)of network variables, oriented to obtain all necessary terms for each fuzzy variable [6] for the purpose of measuring the current values of network variables, which includes:
1) Register intrusions and values (RGIV) for receiving and storing the current values of identifiers I_i ($i = \overline{1,n}$) intrusions and values V_i($i = \overline{1,m}$);
2) switching unit values (SUB), carrying out the formation of the flux corresponding type of invasion;
3) forming unit intrusion pairs and value (BPIV) for binding pair identifier intrusion and the corresponding quantities;
4) forming unit aggregate terms T_{ij}^{g} (FUAT) used to generate a given set;
5) forming unit standards (FUS) performs the calculation for each respective reference fuzzy number (FN);
6) Register of standards (RGS), which serves for receiving and temporarily storing the calculated reference FN;
7) standards of imaging processor (SIP)is intended for displaying in graphical form obtained by reference FN;

2.Subsystem of formation of decisive rules (SFDR) is used to create a set of rules for monitoring network activity [6], which includes:
1) Register of standard (RGS);
2) switching unit (SU), which serves for the formation of streams $\underset{\sim}{t}_i$ [1] on the unit for forming conjugate pairs (UFCP);
3) UFCP designed to transform the logical reference T_{ij}^{g}[19];
4) Ranging unit (RU) carrying out the formation of the importance of factors (IF);
5) initialization block rules (BIR), which forms a matrix FI(i, r_r) and MP (i, r_r) [3];
6) rule base (RB), which serves to store the relevant data sectors (SDi, $i = \overline{1,d}$)), sets the rules $SR_{i_{r_i}}$ ($i = \overline{1,n}$) [3];
7) register of current values (RGCV) and fuzzy identifiers (RGFI) designed respectively for storage in the process of computing the values $\underset{\sim}{t}$ and FI_i($i = \overline{1,d}$);
8) Register Regulations (RGR) for receiving and storing subsets of rules SR_i ;

3. The primary processing subsystem (PPS) for forming sets intrusion, values and their fuzzification[6];

4. Modules fuzzy arithmetic (MFA), inference (MLI) and visualization (MLV) designed to generate results in a fuzzy and graphical representation [6];

5. The control unit (CU), which serves to control switching (CC), as well as the transferring system into standards of mode adjustment (SMA) and the adjustment of the rules (AR). The system operates as follows (see. fig. 1).

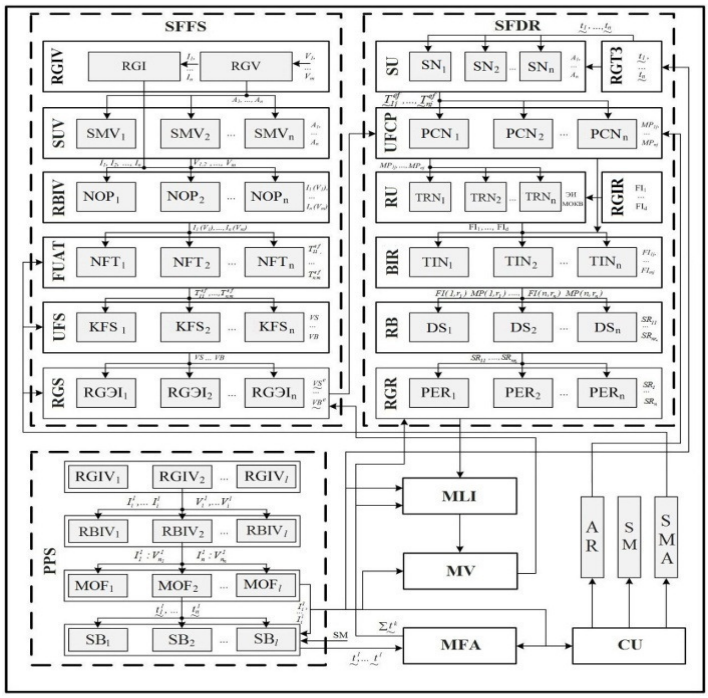


Figure 1. Detection of anomalous state system

The input register intrusion (RGI) and register values (RGV), referred RGIV PFNE pre-recorded and stored during the entire computing process, respectively the current values I_i ($i = \overline{1,n}$) and input values V_j($j = \overline{1,m}$).

On the basis of I_i($i = \overline{1,n}$)identifiers (received from RGI RGIV) and sets the values V_j ($j = \overline{1,m}$), (received with RGV RGIV through the corresponding values switching unit (UKV_i, $i = \overline{1,n}$)BKV for signal I_i ($i = \overline{1,n}$))in BPIV into the steam units (UP_i,) carried out forming arrays (vectors) I_i ($V(A_i)$) ($i = \overline{1,n}$)(wherein each fixed value i A_i signal is switched to a different set of variables $V(A_i)$, for example, when i = 1 $UP_1$with RGI enters I_1, and from RGV by signal A_1 through UKV_1 on UP_1 group values V (A_1) = $V_{1,2}$ = {V_1, V_2} = {NVC, VCA} similarly for i = 2, V (A_2) = $V_{3,4,5}$ = {V_3, V_4, V_5} = {NCC, SPR, DVR}, and when i = 3, V (A_3) = $V_{3,6}$ = {V_3, V_6} = {NCC, NPSA}, that is I_1 (V (A_1)) = I_1 ($V_{1,2}$) = SCANNING (NVC, VCA), I_2 (V (A_2)) = I_2 ($V_{3,4,5}$) = DOS (NCC, SPR, DVR), and I_3 (V (A_3)) = I_3 ($V_{3,6}$) = SPOOFING (NCC, NPSA), where

1. «Scanning of ports (SCANNING)» - «Port scanning"

2. «Denial of service (DOS)» - «Denial of service"
3. «Spoofing (SPOOFING)» - «Spoofing"
4. «Numbers of Virtual channels (NVC)» - «The number of virtual channels"
5. «Virtual Channel Age (VCA)» - «virtual channel Age"
1) «Number of concurrent connections to the server (NCC)» - «The number of simultaneous connections to the server"
2) «Speed of processing requests from the clients (SPR)» - «The speed of processing requests from clients"
3) «The delay between requests from the single user (DVR)" - "The delay between requests from a single user"
4) «Number of packages with the same sender and receiver address (NPSA)" - "The number of packets with the same address of the sender and the recipient."
The names of these arrays correspond to identifiers such as incursions and elements – are the values used for the detection of anomalies generated by the corresponding intrusion.
Further, in terms of forming the respective nodes ($TFRN_i$,) FUST generated values T_{ij}^{f} $(f=\overline{1,r};\ i=\overline{1,n};\ j=\overline{1,m})$ for all $V_i(i=\overline{1,m})$..
The number of such terms and fuzzy interpretation is determined by the expert information (EI), obtained on the basis of opinions of experts on the subject field [19]. For each TRF_i FUST regarding EI determined by their values and according to which the required form for all sets V_i. For example, the array (vector) is generated when m = 2 and n = 5 when i = 3 = 3 = 2 and the output UFT_1:

$$UFT_1(\{T_{I1}^{e1}, T_{I1}^{e2}, T_{I1}^{e3}, T_{I1}^{e4}, T_{I1}^{e5}\}, \{T_{I2}^{e1}, T_{I2}^{e2}, T_{I2}^{e3}\})=$$
$$UFT_1(\{T_{NVC}^{e1}, T_{NVC}^{e2}, T_{NVC}^{e3}, T_{NVC}^{e4}, T_{NVC}^{e5}\}, \{T_{VCA}^{e1}, T_{VCA}^{e2}, T_{VCA}^{e3}\}).$$

Upon receipt of the required set of terms for each Ii in UFE defines specific values for each T_{ij}^{f}FN. When implementing this procedure is necessary to set limit values for all $V_i(i=\overline{1,m})$ i.e. and (for example, V_1 and V_2 border $min\,v_1 = min(NVC)$, $max(v_1) = max(NVC)$ and $min\,v_2 = min(VCA)$, $max(v_2) = max(VCA)$, and choose a method of forming membership functions (IFPI) (see stage 1 in [6] according to established criteria[19].
Thus, the inlet is formed TRF_i array (vector), for example, values of n, and similar to a FUST are as follows:
$TRF_1(\{\underset{\sim}{T}_{NVC}^{e1}, \underset{\sim}{T}_{NVC}^{e2}, \underset{\sim}{T}_{NVC}^{e3}, \underset{\sim}{T}_{NVC}^{e4}, \underset{\sim}{T}_{NVC}^{e5}\}, \{\underset{\sim}{T}_{VCA}^{e1}, \underset{\sim}{T}_{VCA}^{e2}, \underset{\sim}{T}_{VCA}^{e3}\})=TRF_1(\{\underset{\sim}{VS}^{e}, \underset{\sim}{S}^{e}, \underset{\sim}{A}^{e}, \underset{\sim}{B}^{e}, \underset{\sim}{VB}^{e}\}, \{\underset{\sim}{Y}^{e}, \underset{\sim}{A}^{e}, \underset{\sim}{O}^{e}\})$ (see stage 4 in [16])

Further formed FN sets $T_{ij}^{f}$for all overwritten in RGS, while the reference value corresponds to the value of i-th invasion entered in the appropriate register of standards i-th intrusion ($RGEI_i$) and kept there during the entire computing process.
For every means of imaging coprocessor ($IC_i, i=\overline{1,n}$) PVS formed graphical representation standards values for each Ii. In other words IC_1 visualizes reference to I_1, IC_2 - to I_2, and IC_n- for I_n, for example, when n = 3 IC_1 visualizes standards for

SCANNING (ICSCANNING) IC_2 - for DOS (ICDOS) and IC_3 - SPOOFING (ICSPOOFING).

Also, in each i-th RGS intrusion ($RGSI_i, i = \overline{1,n}$)) SFDR recorded and stored during the whole computing process values of group standards $\underset{\sim}{T}_{ij}^{ef}$ $(i = \overline{1,n})$ () corresponding to the values characteristic of i- th invasion, as well as the current values $\underset{\sim}{t}_j$ $(i = \overline{1,n})$ comes in RGCM.

The nodes forming pairs (NFP_i,) UFCP based reference values $\underset{\sim}{T}_{ij}^{ef}$ $(i = \overline{1,n})$ coming from $RGSI_i(i = \overline{1,n})$ and a subset of the current values $\underset{\sim}{T}_{ij}^{ef}$ $(i = \overline{1,n})$ received a RGCM through switching nodes (SU_i,) BC by the control A_i signal (s) (eg, for values i = 1, i = 2 and i = n in UFP_1, UFP_2 and UFP_n with RGCM through UK_1, UK_2 and UK_n arrive respectively the values $t_{1,2} = \{\underset{\sim}{t}_1, \underset{\sim}{t}_2\} = \{\underset{\sim}{t}_{NVC}, \underset{\sim}{t}_{VCA}\}$, $t_{3,4,5} = \{\underset{\sim}{t}_3, \underset{\sim}{t}_4, \underset{\sim}{t}_5\} = \{\underset{\sim}{t}_{NCC}, \underset{\sim}{t}_{SPR}, \underset{\sim}{t}_{DBR}\}$ and $t_{3,n} = \{\underset{\sim}{t}_3, \underset{\sim}{t}_n\} = \{\underset{\sim}{t}_{NCC}, \underset{\sim}{t}_{NPSA}\}$), respectively, will form and go to the output NFP_i conjugate pairs MP_{ij}, for example. $MP_{21} = (\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^{e})$.

Note that all values $FI_i(i = \overline{1,d})$ are recorded in RGNI and kept there throughout the process of formation of the rules.

Each node ranking ($UR_i, i = \overline{1,n}$) BR for each $MP_{ij}(i = \overline{1,n})$ as a possible outcome in turn are associated with all the fuzzy identifiers $FI_i(i = \overline{1,d})$, received a RGNI. Further, based on the method of determining the IF (MDIF) and SI from the thus formed set of alternative rules defined set of FI_{ij}, you need to initialize the DR.

Further, the nodes of initialization (NI_i,) BIK based UR_i and UFP_i pairs formed elements of the MR data matrices (1, r_1) and FI (1, r_1), on the basis of which the initialization is required rulesets.

Generated in NI_i pairwise matrix are recorded in the data sector (SD_i, $i = \overline{1,n}$) iBP, thus forming a set of rules $SR_{ij}(i = \overline{1,n}, j = \overline{1,r_i})$ designed to detect abnormal state generated by the i-th invasion. Further, these rules $SR_i(i = \overline{1,n})$ are overwritten registers SR_i ($PER_i, i = \overline{1,n}$)) and kept there throughout the process of functioning of the system.

Before computing process in SFFS on the basis of the values of network traffic (VCT) according to the corresponding BIE [2] is formed by the sets I_i intrusions$(i = \overline{1,n}$ and value $V_i(i = \overline{1,m})$, by means of which, using a selected (in accordance with established criteria) IFPI (see stage 2in [6], the generated samples of6to certain LV for each term T_{ij}^{ef}.

4. Statement of the main material

According to the received measurement standards in SFDR created templates $SR_i(i = \overline{1,n})$sets of decisive rules, [5] that are used to monitor network activity for possible manifestations of attack on a computer network. These templates and reference values are not changed during the entire operation process CISN, but if necessary, can be modified by transferring them to SCE or RCP.

Further, considering that SFDR oriented to control abnormal, state to l nodes (workstations, servers, etc.) A computer network, in parallel l registers intrusions and values ($RGIV_k\ k = \overline{1,l}$) PPSrecorded intrusion identifiers $I_i^k\ (i = \overline{1,n},\ k = \overline{1,l})$. and (at specified intervals) the present values $V_i^k\ (i = \overline{1,m},\ k = \overline{1,l})$).

For example, if n = 3 and m = 6 is performed forming I_i and V_i [6] for the k-th host, allowing to identify the abnormal condition, generated by three types of intrusions I_1^k, $I_2^k$and $I_3^k(SCANNING^k, SCANNING\ DOS^k$ and $SPOOFING^k)$,and (and) based on six variables V_1^k, V_2^k, V_3^k,V_4^k, $V_5^k$and $V_6^k\ NVC^k, VCA^k, NCC^k, SPR^k, DBR^k$ and $NPSA^k$. It should be noted that if the network nodes are heterogeneous in their characteristics, that is for certain types of anomalies generated by relevant attacking actions the reference value will be different.

For the formation of a pair of concrete intrusion necessary for its identifying the values of [6] is used in the PPSl intrusion pairs of blocks and size of BPIVk (), which are specially organized memory. For example, for the same n = 3 and m = 6 to k-th node identifiers Invasion:

$$(I_1^k),\ (I_2^k) \text{ and } (I_3^k)$$

respectively, are formed with a pair of values:

$$V_{n_1}^{\ k} = (V_1^k,\ V_2^k),\ V_{n_2}^{\ k} = (V_3^k, V_4^k,\ V_5^k) \text{ and } V_{n_3}^{\ k} = (V_3^k,\ V_6^k), \text{ that is}$$

$$SCANNING^k : \{NVC^k,\ VCA^k\},\ DOS^k : \{NCC^k, SPR^k,\ DBR^k\} \text{ and}$$

$$SPOOFING^k : \{NCC^k, NPSA^k\},\ (k = \overline{1,l}).$$

In this example, concerning the organization BPIV, it can be noted that identifiers and addresses are specially organized memory device, {,}, {,}, and { }, respectively, the contents of these addresses.

Upon completion in $BPIV_k\ (k = \overline{1,l})$) procedure of formation of pairs $I_i^k : V_\eta^k$ with modules fuzzification of $MF_k\ (k = \overline{1,l})$) carried out the conversion (using IFPI) $(i = \overline{1,n})$,[6] of the set current values of the values (observed over a certain period of time) by means of a fuzzy number (FN) $\underset{\sim}{t}_i^k\ (i = \overline{1,n})$ [6], and thus the output MF_k obtain n bass () associated with the corresponding I_i. For example, when the value n = 6 $\underset{\sim}{t}_1^k = \underset{\sim}{t}_{NVC}^k$, $\underset{\sim}{t}_2^k = \underset{\sim}{t}_{VCA}^k$, $\underset{\sim}{t}_3^k = \underset{\sim}{t}_{NCC}^k$, $\underset{\sim}{t}_4^k = \underset{\sim}{t}_{SPR}^k$, $\underset{\sim}{t}_5^k = \underset{\sim}{t}_{DBR}^k$ and $\underset{\sim}{t}_6^k = \underset{\sim}{t}_{NPSA}^k$.

Further, in turn derived $\underset{\sim}{t}_i^k\ (i = \overline{1,n},\ k = \overline{1,l})$through k-th switching units $SU_k(k = \overline{1,l})$ on the switching control signal (CS) according to the type of intrusion ($I_i^k\ (i = \overline{1,n},\ k = \overline{1,l})$Current value $(i = \overline{1,n})$from all SUk$(k = \overline{1,l})$come from the PPS in fuzzy arithmetic unit (MFA) $\sum \underset{\sim}{t}^k$ for aggregate indicators characterizing the

activity in all the nodes of the network. The most appropriate method that can be used to implement the fuzzy arithmetic operations (fourteen fixed) is selected according to specific criteria and implemented in MFA [56].
If the abnormal condition detection process according VCT is only one node on the computer network, the MFA is transparent, i.e. no totals of variables in it is formed.

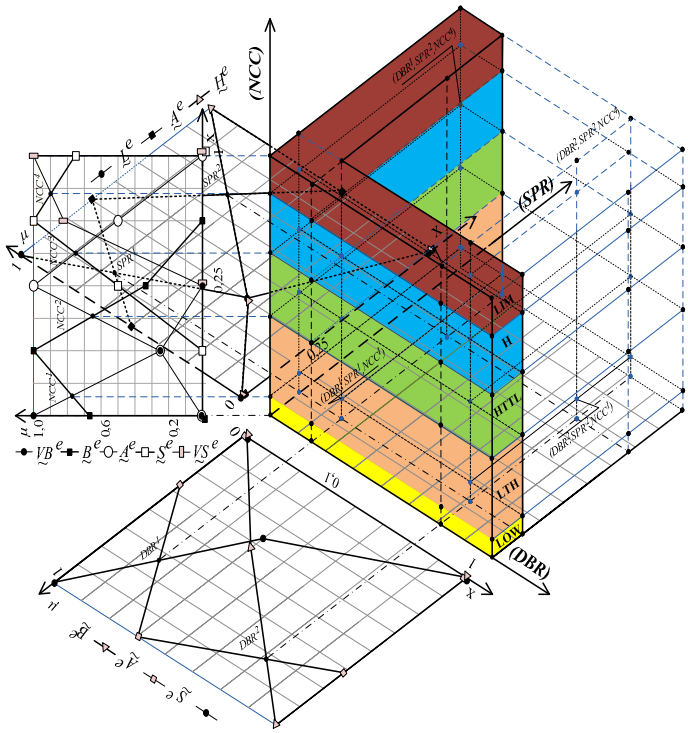


Figure 2. Graphical interpretation of abnormal state generated by I_2

Based on the obtained MFA totals $\underset{\sim}{t}^{k}$ and using initiated in SFDR plurality SR_i rules$(i = \overline{1,n})$ corresponding to certain I_i in MLI, according to known technology [6] by $FI_i(i = \overline{1,d})$, a determination of the current level of the abnormal state in VCT, which may be generated as a specific type of cyber attacks.
This level may be provided in the fuzzy formand be identified by CF in graphical form as a corresponding FA (fig.1), displayed on a background formed PFRP reference values of linguistic variables.

5. Conclusions

Thus, on the basis of CISN proposals based on the subsystem of formation of fuzzy standards, decisive rules and primary processing, as well as modules of fuzzy arithmetic, logical inference, visualization and control module can be developed algorithmic, software and firmware used for the detection of an abnormal state

generated action non-signature cyberattacks. This software can be used independently or as extender functionality of modern systems of intrusion detection in computer networks.

REFERENCES

1. AKHMETOV B.S., KORCHENKO A.A., ZHUMANGALIEVA N.K.: Model of decision rules to detect anomalies in information systems. NEWS Of The National Academy Of Sciences Of The Republic Of Kazakhstan Physico-Mathematical Series, (2016)307, 91-100.
2. YAO J.T., ZHAO S.L., SAXTON L.V.: A study on fuzzy intrusion detection. Proc. of SPIE Data Mining, Intrusion Detection, Information Assurance, And Data Networks Security, Orlando, Florida, USA (2005)5812, 23-30.
3. FRIES P.: A Fuzzy-Genetic Approach to Network Intrusion Detection Terrence. Genetic and Evolutionary Computation Conference, GECCO (Companion) July 12-16, 2008, 2141-2146.
4. LIM M.J., NEGNEVITSKY M., HARTNETT J.: A Fuzzy Approach For Detecting Anomalous Behaviour in E-mail Traffic [Electronic resource] - About Research Online @ ECU. – Electronic data. – Perth Western Australia] : Edith Cowan University, 2006. http://ro.ecu.edu.au/adf/29/, (viewed on May 26, 2015).
5. AKHMETOV B.S., ABDRAKHMANOV R.B., KORCHENKO A.A., ZHUMANGALIEVA N.K.: Base models reference values for intrusion detection system. - Bulletin of International Kazakh-Turkish University named H.A. Yasawi, (2015)5-6 (97-98), 15-26. (in Russian)
6. AKHMETOV B.S., KORCHENKO A.A., ZHUMANGALIYEVA N.K.: Technology of abnormal states for intrusion detection systems. - Al-Farabi Kazakh National University Kaznu Bulletin Mathematics, Mechanics, Computer Science Series, (2016)1 (88), 106-113. (in Russian)
7. WIJAYASEKARA D., LINDA O., MANIC M., RIEGER C.G.: Mining Building Energy Management System Data Using Fuzzy Anomaly Detection and Linguistic Descriptions. IEEE Trans. Industrial Informatics, (2014)3, 1829-1840.
8. EINIPOUR A.: Intelligent Intrusion Detection In Computer Networks Using Fuzzy Systems. Global Journal of Computer Science and Technology Neural & Artificial Intelligence (GJCST), (2012)12, 19-29.
9. SHANMUGAVADIVU R., NAGARAJAN N.: Network Intrusion Detection System Using Fuzzy Logic. Indian Journal of Computer Science and Engineering (IJCSE), (2011)1, 101-111.
10. LINDA O., VOLLMER T., WRIGHT J., MANIC M.: Fuzzy Logic Based Anomaly Detection for Embedded Network Security Cyber Sensor. Proc. IEEE Symposium Series on Computational Intelligence, Paris, France, April 2011, 202-209.
11. LINDA O., MANIC M., MCJUNKIN T.R.: Anomaly Detection for Resilient Control Systems Using Fuzzy-Neural Data Fusion Engine. Proc. IEEE Symposium on Resilience Control Systems, ISRCS 2011, Boise, Idaho, Aug. 9-11, 2011.

12. BRIDGES S.M., VAUGHN R.B.: Fuzzy data mining and genetic algorithms applied to intrusion detection. Proceedings of the 23rd National Information Systems Security Conference. October 2000, 13-31.
13. SHAMSHIRBAND S., ANUAR N.B., LAIHA M., KIAH M., MISRA S.: Anomaly Detection using Fuzzy Q-learning Algorithm. Acta Polytechnica Hungarica. (2014)8, 5-28.
14. DICKERSON J.E., JUSLIN J., KOUKOUSOULA O., DICKERSON J.A.: Fuzzy Intrusion Detection. IFSA World Congress and 20th NAFIPS International Conference, (2001)3, 1506-1510.
15. TSANG C.-H., KWONG S., WANG H.: Genetic-Fuzzy Rule Mining Approach and Evaluation of Feature Selection Techniques for Anomaly Intrusion Detection. Pattern Recognition, (2007)9, 2373-2391.
16. ZADEH L.A.: Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, (1973)1, 28-44.
17. GÓMEZ J., GONZÁLEZ F., DASGUPTA D.: An Immuno-Fuzzy Approach to Anomaly Detection. The 12th IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 25-28 May 2003, 1219-1224.
18. ZADEH L.A.: The concept of a linguistic variable and its application to approximate reasoning. I - Information Sciences, July (1975)3, 199-249.
19. KORCHENKO A.G. The development of information protection systems based on the fuzzy sets. The theory and practical solutions, Kiev, 2006. (in Russian)
20. SHAIKHANOVA A., ZOLOTOV A., DUBCHAK L., KARPINSKI M., KARPINSKYI V.: Access Distribution Scheme to the Computer System Based on Fuzzy Logic. Graph-Based Modelling in Engineering (Eds. S. Zawiślak, J. Rysiński). – Springer, 2017, 39-50.
21. SHAIKHANOVA A.K, ZOLOTOV A.D., STEPANOVA O.A., KARPINSKI M.P., DUBCHAK L.O.: Fuzzy System of Access Distribution within a Computer Network. Journal of Theoretical and Applied Information Technology, Vol. 80(2015)1, 105-113.

Suliko ASABASHVILI[1], Daria KONOTOP[2], Stepan SHUPROVYCH[3]

Supervisor: Oleksii FRAZE-FRAZENKO[4]

POPRAWA POZIOMU OCHRONY POJAZDÓW WYKORZYSTUJĄC TECHNOLOGIĘ NFC ORAZ SZYFROWANIE Z KLUCZEM PUBLICZNYM

Streszczenie: Artykuł dotyczy bezpieczeństwa algorytmów szyfrujących stosowanych w alarmach samochodowych, ataków na nie oraz metod ich ochrony. Udowodniono doświadczalnie teoretyczne wyniki uzyskane podczas implementacji ataku na statyczny kod w rzeczywistym alarmie samochodowym i wykonano symulacje kodu dynamicznego Keeloq w systemie bezpieczeństwa ISIS. Przeprowadzano analizę algorytmów szyfrowania i opracowywano odpowiednie rekomendacje. Uwzględniono również nową metodę szyfrowania opartą na smartfonie z NFC oraz szyfrowaniem z kluczem publicznym RSA.

Słowa kluczowe: Keeloq, RSA, NFC, alarm samochodowy, protokół, szyfrowanie

CAR ALARM SECURITY LEVEL INCREASE ON NFC BASED TECHNOLOGY AND ASYMMETRIC ENCIPHERING

Summary: Safety issues of encryption algorithms which are used in car alarm systems, the attacks to them and methods of their protection are considered in the article. The received theoretical results at implementation of the attack to a static code on real car alarm systems and modeling of the dynamic Keeloq code in ISIS are experimentally confirmed. The analysis of encryption algorithms is carried out and recommendations are developed. The new cryptography technique based on the smartphone with NFC and RSA asynchronous enciphering is also considered.

Keywords: Keeloq, RSA, NFC, car alarm, protocol, enciphering.

---

[1] Doctoral student, Odesa State Academy of Technical Regulation and Quality, the department of information and measurement technologies, as.sulico@gmail.com

[2] Student, Odesa State Academy of Technical Regulation and Quality, the Faculty of Information and Measurement Technologies, dariakotop96@gmail.com

[3] Student, Odesa State Academy of Technical Regulation and Quality, student of the Faculty of Information and Measurement Technologies, amazingstepa@gmail.com

[4] Engineering Science Ph.D., Odesa State Academy of Technical Regulation and Quality, an associate professor at the department of computer, information and measurement technologies, frazenko@gmail.com

1. Introduction

Problem of automobile anti-theft systems is prevention of illegal physical access to cars, however, recently, opposition of car alarms vendors with malefactors amplified. The purposes of potential criminals are rather various. Thieves seek for the car and expensive automobile component parts stealing, or things left in salon. Vandals try to damage property. Professional criminals − to get access to the car for use in criminal purposes. The hackers aim, in addition to listed above, is often entertainment. The successful bypass of protection systems gives criminals not only physical access to the car, but also to onboard systems. Thieves, as a rule, carry out technologically easy attacks (to reject the lock, to execute search of the keys left in the car, to damage signal conductors and to carry out direct start of the engine). At the same time the number of the high-tech attacks to wireless systems, especially on high-class cars is growing. Thereafter, modern automobile signalings can not guarantee complete safety of the vehicle. At the same time, there are number of technical capabilities of the described problem solution which are given in this work.

2. Review of the attacks to signalings

At present there is a certain types classification of the attacks to cars anti-theft systems by the criteria of characteristics and properties given below. The analysis of a number of sources on subject [1,2,3,4] allowed to select the most widespread attacks and to make their detailed analysis. Short systematization of the most widespread types of the attacks is given in table 1 [1].

Table 1. Attacks on antitheft systems

Attack	Tool	Vulnerability	Action	Target	Unauthorized Rezult
Keeloq attack	Info. exchange	Design (short key; short block size; and similar key schedule)	Read, authenticate	Data (encryption key)	Disclosure of information (encryption key)
DST attack	Info. exchange	Design (short key; cipher function; and structure)	Read, authenticate	Data (cipher function; encryption key)	Disclosure of information (encryption key and cipher function)
Relay attack	Toolkit	Design	Spoof		Resource theft
Bypass kit	Toolkit	Design	Bypass		Resource theft
Jamming	Toolkit	Design	Flood	Network	Denial of service
RollJam	Toolkit	Design	Scan, copy	Data	Resource theft

It is possible to study the detailed description of the specified attacks in [1, p. 141-142].

There are car alarms with feedback coupling − with a key fob which informs on a car status, and rather simple car alarms which use the unidirectional communication link that results in low safety of system in general. In such devices, as a rule, code combination does not change in general or the quantity of such changes options is

limited. Systems with backward communication channel have much higher level of protection but, owing to the high cost, found broad commercial application quite recently.
Based on this is possible to formulate two basic rules which will allow to remote control system to be called safe with the unidirectional communication link:

- number of possible code combinations has to be big;
- coder should not create the same code twice.

In new generations of car alarms systems the authentication technology, widely known in cryptography, for one-time passwords through insecure channel with use of the so-called unidirectional functions - cryptographically resistant hash functions or symmetric block encryption algorithms is used.
For cracking of car alarms (substitution of code radio parcel) criminals use the code grabbers.
Three types of such devices are distinguished: for static codes, on the principle of a code substitution and algorithmic. For earlier systems of car alarms which use a static code the device which intercepts this code and remembers it suffices. The work algorithm which demands repeated clicking key by the owner of fob button is characteristic of the code grabber using the principle of a code substitution, using at the same time radio suppression and interception of key fob sending. An algorithmic code grabber- the device which will recognize signaling type (the device vendor) on key fob digital sending and, using a so-called "manufactory code", becomes an owners key fob clone. This principle is applied to car alarms which use cryptoresistant algorithms (Keeloq and other).
The vast majority of car alarms systems work at the Microchip Technology chips realizing Keeloq algorithm. In Keeloq technology the system of reverse identification is used. In [5] check on this attack algorithm known as Avalanche Effect is described. Despite detection of Keeloq vulnerabilities because of bulk copy of encrypting keys to speak about a total failure from use of radio key fobs at the moment early, even taking into account development of alternative owner identification systems (by means of SMS, with use of mobile applications or GSM systems).
Due to the above, along with development of new systems in the field of cars protection vendors began to pass to other algorithms, the dialogue code became basic of them.

3. Analysis of encryption algorithms and their vulnerabilities

In the work, for researching of a static code vulnerability based on real car alarm system the Arduino UNO platform and 2 transmitter-receiver modules RF 315 (433) MHz were used.
By means of the RF module and Arduino Uno scanning of frequency range, with use of the sketch based on Rcswitch library [6] is carried out.
As a experiment result with scanning of air all commands of key fob buttons were received, namely: unblocking (1268130) to block (1268129), to turn off sound indication (1268136), to include sound indication (128132).

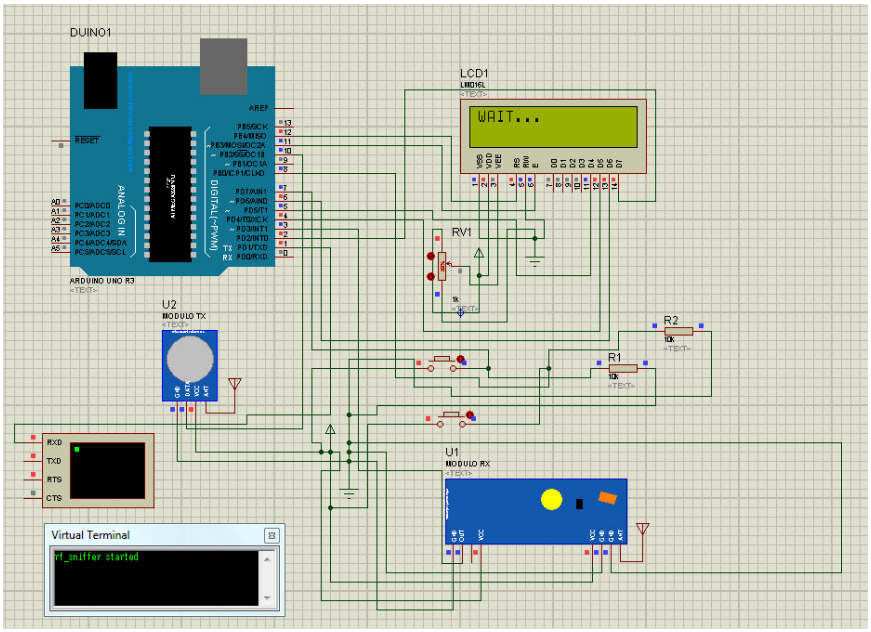


Figure 1. Simulation stand

For check of received data correctness, by means of Arduino and the RF module the created packets were sent for blocking and unblocking of car alarm systems control block.

```
#include <Rcswitch.h>
Rcswitch myswitch = Rcswitch();
void setup() {
  Serial.begin(9600);
  myswitch.enabletransmit(10);
   myswitch.send(1268130, 24);
}
void loop() {
  delay(10000);
  delay(10000);
  myswitch.send(1268129, 24);
  delay(10000);
}
```

For modeling of the dynamic Keeloq code software environment of ISIS was used. The sketch for Arduino (TX - the Transmitter):

```
#include <Keeloq.h>
#include <EEPROM.h>
#include <Softeasytransfer.h>
#include <Easytransfervirtualwire.h>
Easytransfervirtualwire ET;
#define LED 13
Keeloq k(0x01320334,0x05063708);///keys
unsigned int count = 65535;
struct SEND_DATA {
```

```
unsigned long enc;//counter
byte id = 1;//id
byte cmd = 1;//team
};
SEND_DATA data;
void setup () {
delay(200);
Serial.begin(9600);
pinmode (LED, OUTPUT);
digitalwrite (LED, HIGH);
ET.begin(details(data));
  data.id = 1;
  vw_set_ptt_inverted(true);
  vw_set_tx_pin(12);
  vw_setup(2000);
EEPROM.get (0, count);//we receive from EEPROM int
count-;//we take away 1
data.enc = k.encrypt(count);//we code
ET.senddata ();//We send
EEPROM.put (0, count);//we save int in EEPROM
digitalwrite (LED, LOW);
}
void loop () {
//Lowpower.powerdown (SLEEP_FOREVER, ADC_OFF,
BOD_OFF);//SLEEP_FOREVER
}
```

The sketch for Arduino (RX - the Receiver):

```
//Control block
#include <Virtualwire.h>
#include <Easytransfervirtualwire.h>
#include <Keeloq.h>
#include <EEPROM.h>
Easytransfervirtualwire ET;
Keeloq k(0x01320334,0x05063708);///keys
#define LED 13
unsigned int oldcount = 65535;
unsigned int count;
struct RECEIVE_DATA {
unsigned long enc;//counter
byte id = 1;//id
byte cmd = 1;//team
};
RECEIVE_DATA data;
void setup () {
ET.begin(details(data));
Serial.begin(9600);
 vw_set_ptt_inverted(true);
  vw_setup(2000);//
  vw_set_rx_pin(12);
  vw_rx_start ();
pinmode (LED, OUTPUT);
}
void loop () {
if (ET.receivedata) {//if was received a packet
```

```
if (data.id = 1) {//also matched id
EEPROM.get (0, oldcount);//we receive value of the counter
from EEPROM
count = k.decrypt(data.enc);//we decode
if (<= oldcount) {//if value of the counter is more than count
or equal to value of saved
count-;//we take away 1
EEPROM.put (0, count);//we write in EEPROM
digitalwrite (LED! digitalread(LED));
Serial.println (OK);//
}
else Serial.println ("ALARM!!!");//or we try to send the saved
packet
}
//Variables for debugging
Serial.println (" data:");
Serial.print ("enc:"); Serial.println(data.enc);
Serial.print("id:"); Serial.println (data.id);
Serial.print ("count:"); Serial.println(count);
Serial.print ("oldcount:"); Serial.println(oldcount);
Serial.println ();
}
}
```

Let's note that alternative libraries are necessary for work of a code: Keeloq.h, Softeasytransfer and Easytransfervirtualwire.

For simulation of grabber work, count value, for an output for borders of a condition of +16 combinations (fig. 2) is changed.

```
EEPROM.get(0, count);// дістаємо із EEPROM int
count++;//
data.enc = k.encrypt(count);// кодуємо
ET.sendData();//Відправляємо
EEPROM.put(0, count);// зберігаємо int в EEPROM
digitalWrite(LED, LOW);
}
```

Figure 2. Modification of the count variable - on count ++

Warning as the imitated confirmation code does not meet a condition (fig. 3) is as a result received.

Modeling showed that the dynamic Keeloq code has the average vulnerability level which is exposed to the attacks with social engineering elements and does not provide the sufficient security level.

The dialogue code [7] as a car alarms protection method, uses for identification of a key fob authentication technology through insecure channel, widely known in cryptography. Enciphering of this type is conducted in the dialog mode between a key fob and car alarm control block located in the car. When clicking the button by the owner, from a key fob the request for command execution is sent. The control block sends a random number to a key fob. This number is processed on a certain algorithm and sent back in control block. At the same time the control block processes that number and compares the result to the result received from a charm. At equality of values, the control block executes the command. The calculations algorithm on a key fob and control block is individual for each car alarm and is established in course

of production. To intercept and decrypt a data packet of a dialogue car alarm is impracticable. For coding of a signal a hash functions are used.

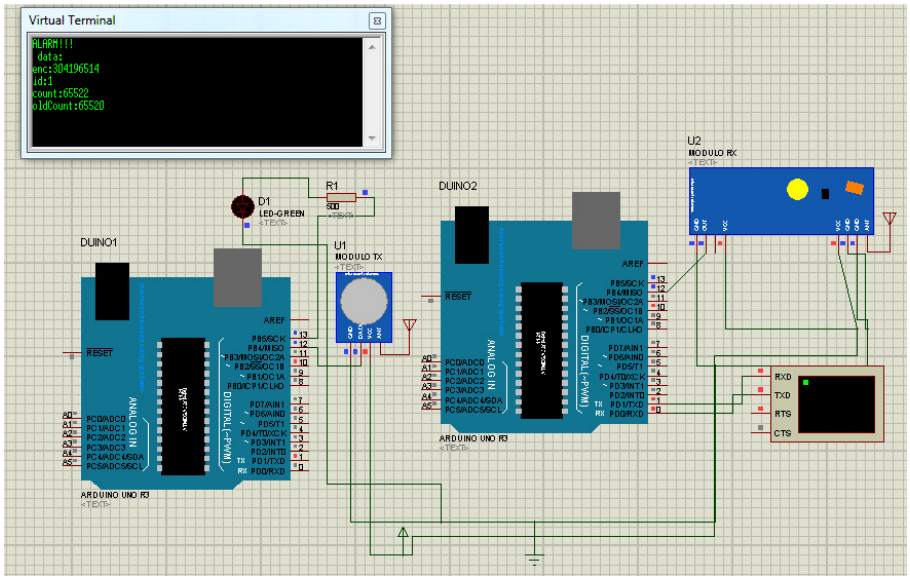


Figure 3. Simulation of grabber work

The analysis result of encryption algorithms is given in fig. 4.

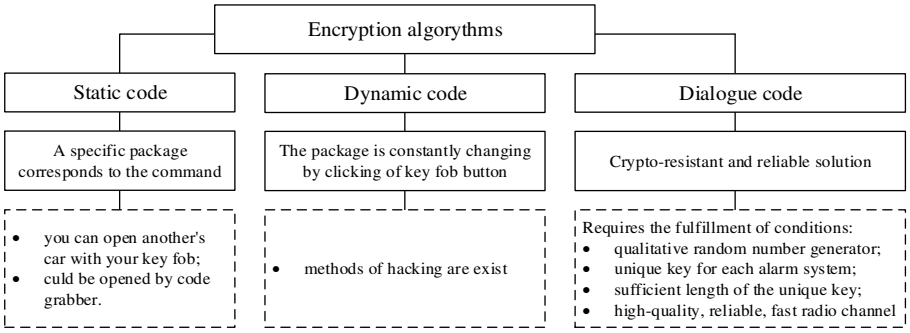


Figure 4. Analysis of car alarms encryption algorithms

4. Development of the security algorithm

In an analysis result of benefits and shortcomings of encryption algorithms the option of implementation of protection of cars, using asymmetric enciphering of RSA and NFC technology is offered.

The structural scheme of the offered security algorithm is shown in the figure 5.

As "remote controller " is a key fob or the smartphone which are not suitable for difficult calculations, the majority of calculations are carried out on the managing

module, namely decoding and key generation. The password is ciphered by public key before it is sent to the management system (MS). The management system carries out two tasks at the same time: (1) decrypts the message to receive the initial password, and checks the blocking status; (2) generates new couple of public key and private key, and sends public key to the remote controller for the following access. In consequence of which the public key constantly changes.

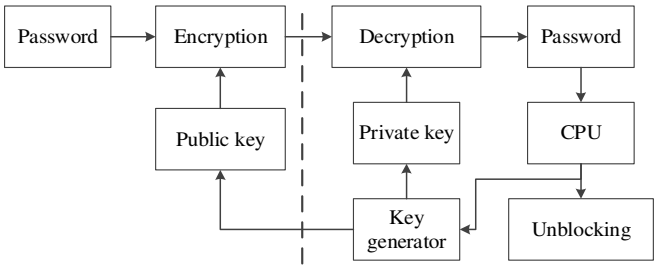


Figure 5. Structural scheme of access control system algorithm

The interaction mechanism provides the sufficient security level that allows to lower requirements to strength of enciphering algorithm. The optimum encryption algorithm for this task is RSA, because the RSA is often used for transfer of the ciphered general keys which, in turn, can carry out enciphering decoding transactions at great speed. MS generates numbers n and e - public key, then sends it to the remote controller, for further enciphering of the password. At the same time, MS saves a private key d which is used for decoding of the ciphered text of the password received from the panel compares the received results and builds the conclusions, concerning unblocking.
Besides, each system has to have identification number (identifier). The identifier of system can be a part of public key, allowing working only with certain remote controllers. The operation scheme of remote controller is provided on the figure 6.

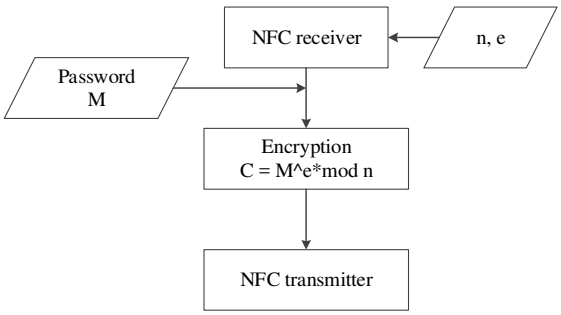


Figure 6. Operation scheme of remote controller

The remote controller receives numbers n and e (public key) from a management system, based on these data ciphers the password and sends the received value C back. The management system generates keys which generation process consists of the following steps:
1. Creation of two random prime numbers p and q.
2. Calculation: $n = p \cdot q$ и $\phi(n) = (p-1) \cdot (q-1)$.

3. Generation of number e (an enciphering indicator) which is mutually simple page $\phi(n) i\, e$. At the same time $1 < e < \phi(n)$.

4. Search of values d_p, d_Q и q_{Inv} (decoding indicator): $e \cdot d_p \equiv 1(\text{mod}\ (p-1))$, $e \cdot d_Q \equiv 1(\text{mod}\ (p-1))$, $q \cdot q_{Inv} \equiv 1(\text{mod}\ p)$.

After receiving C from the remote controller, the control block begins the decoding procedure (fig. 7):

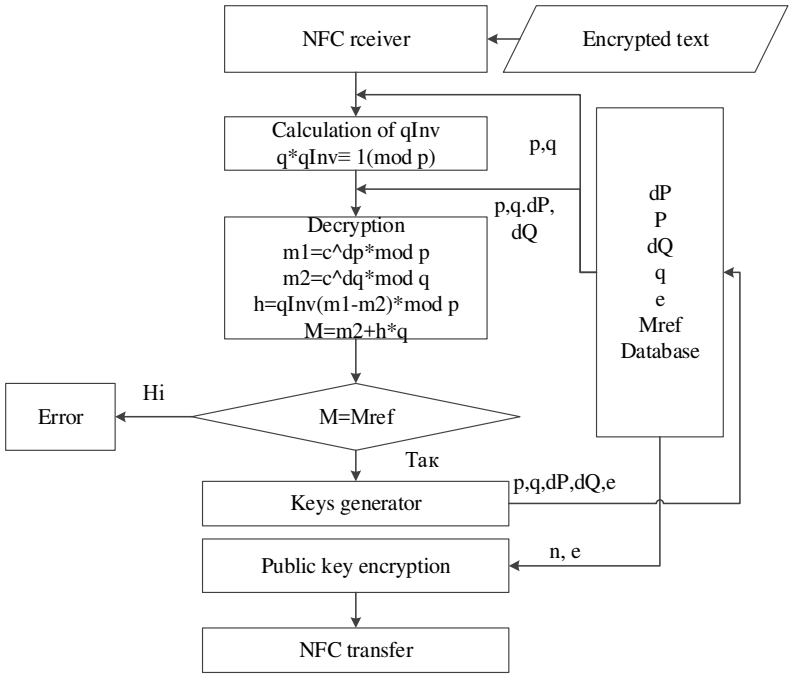


Figure 7. Procedure of key decoding and management

Variable value M is compared to value of the password. If they are equal, there is an unblocking.

The result of RSA enciphering modeling is presented in the figure 8.

```
#include <RSA.h>
char msg[PLAINTEXT_SIZE] = "TEST RSA by IVT";
char plain[PLAINTEXT_SIZE];
int publickey[2] = {14351, 11};
int privatekey[2] = {14351, 1283};
void setup()
{
  Serial.begin(9600);
  char cipher_msg[CIPHERTEXT_SIZE];
  rsa.encrypt(msg, cipher_msg, publickey);
Serial.println(cipher_msg);
  rsa.decrypt(plain, cipher_msg, privatekey);
```

```
  Serial.println("-------------------------");
  Serial.println(plain);
}
void loop()
{
}
```

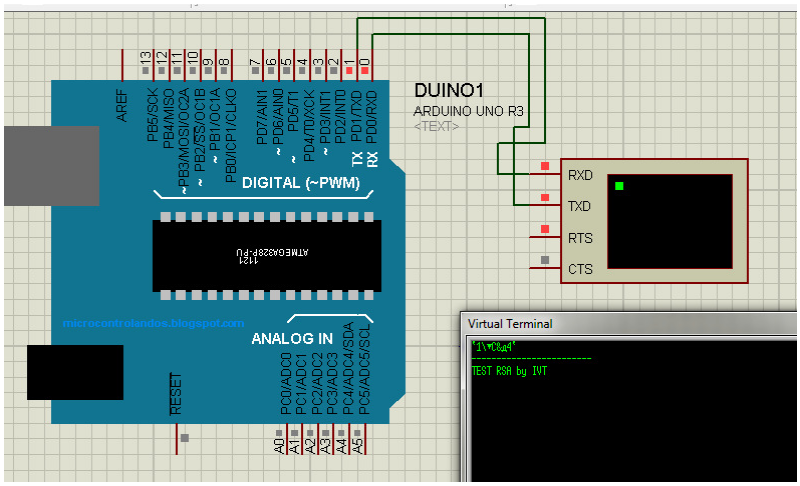


Figure 8. Modeling of RSA asymmetric enciphering on ARDUINO

For the algorithm effectively working in the built-in systems, it is necessary to optimize it [8].

Generation of p and q consists of the following stages:

Creation of an array with prime numbers: $A = [a_1, a_2, a_3 ... a_n]$.

Choice in a random way integer number i $(0 < i < n)$.

The number p is calculated by a formula: $p = \prod_{k=1}^{i} a_k + 1$ is also a prime number as it is not divided into any integer number less then $p/2$.

The number q is similarly calculated by a formula: $q = \prod_{k=1}^{j} a_k + 1$, at the same time the integer number j is selected in a random way and satisfies conditions $(0 < j < n,\ j \neq i)$.

It is the main part of public key generation process and a private key for RSA enciphering. The size of two prime numbers p and q depends on the size of an array $\boldsymbol{A}$ and a method of the indexes choice.

Using generated p and q, it is possible to draw easily conclusion that $\phi(n)$ is divided into all integer numbers from a_1 to a_l, where $l = \max(i, j)$. Thus, coprime numbers from function $\phi(n)$ can be found on a formula: $e = \prod_{i=(l+1)k}^{k \le n} a_i$.

It is reasonable to check a prototype of the car alarm protection system on the Arduino Mega 2560 platform which contains the Atmega2560 microcontroller, and on NFC RFID PN532 modules.

The smartphone which has to have the built-in NFC module can act as the remote controller.

The maximum number which can process the microcontroller is 6,8×1038 (double) therefore it is necessary to use Bignumber.h library.

5. Conclusion

The analysis of the encryption algorithms used in automobile car alarm systems showed that the static code is the most vulnerable. The dynamic Keeloq code is the subject to the attacks with elements of social engineering owing to what also is vulnerable. The dialogue code has the highest security level and is recommended for use. Enciphering of this type is conducted in the dialog mode between a key fob and control block located in the car. Concerning systems of keyless access (a long hand) reduction of detecting zone to 1 meter and, the most important, response time to 6 nanoseconds, at flowing 4-400mks, depending on the vendor will be effective protection. Time for relaying occupies 300 nanoseconds that allows to carry out the attack without problems. The offered implementation method of protection of cars is based on asymmetric enciphering of RSA and NFC technology, and gives the chance to create the reliable anti-theft systems, steady against the existing types of threats. It is also necessary to note that reliability of anti-theft systems depends on a complex of the actions including as regular blocking devices (for example, doors and a cowl) and auxiliary security measures that will allow to raise the overall level of security.

REFERENCES

1. DENG J., YU L., FU Y.: Oluwakemi Hambolu and Richard R. Brooks, Chapter 6 - Security and Data Privacy of Modern Automobiles, In Data Analytics for Intelligent Transportation Systems, Elsevier, 2017, 131-163.
2. INDESTEEGE S., KELLER N., DUNKELMAN O., BIHAM E., PRENEEL B.: A practical attack on Keeloq, in: N. Smart (Ed.), Eurocrypt'08,volume 4965 of LNCS, Springer-Verlag, 2008, 1-18.
3. BONO S.C., GREEN M., STUBBLEFIELD A.: Security analysis of a cryptographically-enabled RFID device, Proceeding of 14th conference on Usenix Security Symposium, 14(2005), 1-15.
4. FRANCILLON A., DANEV B., CAPKUN S.: Relay Attacks on Passive Keyless Entry and Start Systems in Modern Carsin: A. Perrig (Ed.), NDSS, 2011.

5. KASPER M., KASPER T., MORADI A., PAAR C.: Breaking KeeLoq in a flash: on extracting keys at lightning speed. Progress in Cryptology–AFRICACRYPT 2009, 403–420, 2009.
6. DIVER_SANT: Interception of the coding Princeton by means of Arduino [Electronic resource]: *http://www.phreakerclub.com/1547,* 01.10.2017
7. MIGAL V.D. Engineering cybernetics of transport/EL Migal, V. P. Volkov. - Kharkiv: HNADU, 2007.
8. TRAN B.A., TRAN VH.: A Novel Encryption Mechanism for Door Lock to Resist Jam-and-Relay Attack. In: Dang T., Wagner R., Küng J., Thoai N., Takizawa M., Neuhold E. (eds) Future Data and Security Engineering. FDSE 2016. Lecture Notes in Computer Science, vol 10018. Springer, Cham.

Vladyslava DMYTRUK[1]

Scientific Supervisor: Oleksandr OKSIIUK[2]

ANALYSIS OF GENERAL PROVISIONS OF ESTABLISHING A SYSTEM OF INFORMATION SECURITY IN ENTERPRISES

Summary: Due to the constant expansion of computer networks, the informatization of most processes in the enterprise and the penetration of information technology in all spheres of society, there is an urgent need to create a reliable system of protection of confidential information belonging to organizations from internal and external threats. In order for the system of information protection to be effective, it is necessary to approach its creation in a complex way, so it is possible to use all components that can affect the state of enterprise security.

Keywords: information security, management, enterprise, protection of information

ANALIZA OGÓLNYCH PRZEPISÓW DOTYCZĄCYCH BUDOWY SYSTEMU BEZPIECZEŃSTWA INFORMACJI W PRZEDSIĘBIORSTWACH

Streszczenie: Ze względu na stały rozwój sieci komputerowych, informatyzację większości procesów w przedsiębiorstwie i powszechne stosowanie technologii informacyjnych we wszystkich sferach społeczeństwa, istnieje pilna potrzeba stworzenia niezawodnego systemu ochrony poufnych informacji należących do organizacji ze względu na wewnętrzne oraz zewnętrzne zagrożenia. Aby system ochrony informacji był skuteczny, konieczne jest podejście do jego tworzenia w sposób złożony, dzięki czemu możliwe jest wykorzystanie wszystkich elementów, które mogą wpływać na stan bezpieczeństwa przedsiębiorstwa.

Słowa kluczowe: bezpieczeństwo informacji, zarządzanie, przedsiębiorstwo, ochrona informacji

[1] Taras Shevchenko National University of Kyiv, Faculty of Information Technology, Department of Cybersecurity and Information Protection, specialty: Information Security Management, Student, vladislava.dmitruk@gmail.com

[2] Ph.D., Professor of Cybersecurity and Information Protection, Taras Shevchenko National University of Kyiv, faculty of Information Technology, oksiuk@ukr.net.

1. Formulation of the problem

Information security (IS) is one of the fundamental components of nation-wide security in each country. At the same time, the state of the IS is influenced by a large number of objective and subjective factors, for example, the political situation, the state of technical support of specific structures, economic development of the state. In addition, the information sphere is constantly evolving, which leads to an increase in its social significance.

2. Problem definition

Each large company can not work without IT, the creation of electronic archives and data arrays, on the one hand, is a prerequisite for improving the efficiency of each enterprise, and, on the other, a source of new dangers. The creating of advanced information security systems at the enterprises is the basis for implementing an information security strategy for the whole country [1]. That is why the development of new methods to ensure the information security is an extremely important and topical issue.

3. Statement of the main material

In order to coordinate the actions of various agencies to provide the IS, it is necessary to create a sufficient regulatory apparatus, and to develop a legislative framework. Measures that are being taken at this level should operate in two directions:

- the establishment of a tendency towards the development of the society in the field of information technology (at this stage also should be the training of profile specialists);
- the establishment of a stable negative attitude to violations of legislation in the field of information security [2].

The next level is an administrative, which purpose is to develop a definite program of actions aimed at providing the IT with current realities and future changes. At the same time criteria for determining the expediency of carrying out certain actions and stages of control should be formulated. The basis for working on this level is a security policy that is built upon an analysis of the risk of an IS for a specific organization. It also includes a detailed description of the resources that will be involved in the implementation of the chosen protection strategy.

The procedural and software-technical levels are the next stages in the development of a system for providing IS. These include work with users and the construction of an advanced technical device, respectively. In the process of developing secure system at these levels, it is necessary to apply such methods, which will help to reduce the probability of intentional violations and random errors from the local users by technical means.

When building a system for information security at an enterprise, the specifics of the organization itself should be taken into account in order to ensure that the set of measures developed can withstand internal and external threats. But in the general

case, it is possible to distinguish several functions that must be in one degree or another in each system of providing IS.
Firstly, it is necessary to develop elements of the system for protection of information, and to ensure their correct and smooth functioning. To do this, it is necessary to prepare a regulatory legal apparatus, provide a systematic approach (attention should be paid to the real possibilities of the enterprise with resources, personnel components, analytical support, material and technical basis, etc.). It also involves making managerial decisions regarding the implementation and improvement of the built-in system [3].
After a complex of protection measures is developed, it is necessary to continuously manage its work. To do this, the company must constantly analyze threats and hazards, predict the emergence of new ones, work with personnel to identify insider trading, and monitor the external environment. At this stage the well-known model PDCA (Plan-Do-Check-Act) is used in order to react promptly to the change of the state of enterprise IS. The management stage also assesses the costs, the feasibility of the IS and resource allocation efficiency.
The department of Management of Information Security at the enterprise is also engaged in planning and operational activities. For each unit, the main interests of the IS are determined and given priority in order to ensure these interests. Persons who access information to varying degrees in accordance with their job responsibilities must provide written confirmation that they are familiar with the status of the information and are responsible for not disclosing confidential data. All staff changes should also be monitored in order to prevent excessive use of information resources.
While maintaining the work of the information security system, we can not forget about the constant development of material and technical measures. All software must be updated in a timely manner in order to prevent the use of vulnerabilities by offenders, and technical measures must be certified and undergo diagnostics in a time-bound document. These actions are aimed at preventing the leakage of confidential information over the perimeter of protection, as well as to prevent the violation of the integrity of information in the enterprise and its distortion [4].
And the last function of the system of providing the IS - is the implementation of control and oversight of the work of those units in the enterprise, whose work is related to information with restricted access. The same category also applies to the supervision of compliance with the legislation in the sphere of IS, control over the implementation of the plans defined by the company strategy of information security.

4. Conclusions

So, in order to solve the problem of protecting the company's information resources from external and internal threats, an integrated approach to the creation of a security system is required. It should contain a technical component, a software solution, organizational arrangements and work with personnel - an integral part of the information security process. The constant control measures is also inportant. Only in this case, the security system can be effective and perform the functions assigned to it.

REFERENCES

1. MOROZOV O.: Information security in the conditions of the present state and prospects of statehood development, 23-25, (in Ukrainian).
2. KUSTOVSKAYA O.V.: Methodology of system approach and scientific Research: Course of lectures, (in Ukrainian).
3. NIZHNIK N.R.: National Security of Ukraine (methodological aspects, state and development trends), (in Ukrainian).
4. NASHINETS-NAUMOVA A.Y.: Institute of false information in information law, (in Ukrainian).

Lesia DUBCHAK[1], Myroslav KOMAR[2]

Opiekun naukowy: Anatoliy SACHENKO[3], Volodymyr KOCHAN[4]

SPEEDY PROCESING METHOD OF FUZZY DATA FOR INTELLIGENT SYSTEMS OF INTRUSION DETECTION

Summary: A method of processing fuzzy data, based on the classic mechanism of Mamdani's fuzzy inference was proposed. Implementation of the method differs in dividing into stages of learning and exploitation, which allows to reduce the number of operations on the exploitation stage of the fuzzy information means processing and improves its performance. It can be used to build intrusion detection systems.

Keywords: fuzzy data, Mamdani's fuzzy inference, intrusion detection system

METODA PRZYŚPIESZONEGO PRZETWARZANIA ROZMYTYCH DANYCH DLA INTELIGENTNYCH SYSTEMÓW WYKRYWANIA WŁAMAŃ

Streszczenie:. Zaproponowano metodę przetwarzania rozmytych danych, opartą na klasycznym mechanizmie rozmytego wnioskowania. Wdrożenie proponowanej metody różni się podziałem na etapy uczenia i wykorzystania, co umożliwia zmniejszyć liczbę operacji na etapie wykorzystania środków do przetwarzania rozmytej informacji i zwiększenia prędkości działania. Może być stosowana do budowy systemów wykrycia włamań.

Słowa kluczowe: rozmyte dane, rozmyte wnioski Mamdaniego, system wykrycia włamań

1. Introduction

The mathematical theory of fuzzy sets and fuzzy logic are generalizations of classical set theory and classical formal logic. These concepts were first proposed by the

[1]Research Institute for Intelligent Computer Systems, Ternopil National Economic University, Ukraine, dlo@tneu.edu.ua
[2]Research Institute for Intelligent Computer Systems, Ternopil National Economic University, Ukraine, mko@tneu.edu.ua
[3]Doctor of Technical Sciences, Professor, Research Institute for Intelligent Computer Systems, Ternopil National Economic University, Ukraine, as@tneu.edu.ua
[4]PhD Associated Professor, Research Institute for Intelligent Computer Systems, Ternopil National Economic University, Ukraine, vk@tneu.edu.ua

American scientist Lotfi Zade in 1965 [1]. The main reason for the emergence of a new theory was the presence of fuzzy and approximate reasoning in describing processes, systems and objects.

The main advantages of fuzzy systems in comparison with others are [1, 2]:

- the ability to operate with input data that is set vaguely, for example, values that constantly change over time (dynamic tasks),
- the possibility of fuzzy formalization of evaluation and comparison criterias,
- the ability to qualitative assessment of not only the actual data values, but also their degree of reliability and its distribution,
- the possibility to perform rapid simulation of complex dynamic systems and their comparative analysis with a given degree of accuracy.

As a rule, the Mamdani fuzzy inference mechanism is used in engineering tasks [1]. It uses the mini-max composition of fuzzy sets. This mechanism includes the following actions [3-5]:

- fuzzification procedure: determine the degree of truth, that is the values of the membership functions for the left parts of each rule. For a rule base of m rules, the truth degrees are denoted by $A_{ik}(x_k)$, $i=1 \quad m$, $k=1 \quad n$,
- fuzzy inference. First, the cut-off levels for the left side of each rule are determined: $alfa_i = \min_i\left(A_{ik}\left(x_k\right)\right)$. Next, there are "truncated" output membership functions: $B_i^*(y) = \min_i\left(alfa_i, B_i(y)\right)$,
- combining the obtained truncated functions, for which the maximum composition of fuzzy sets is used: $MF(y) = \max_i\left(B_i^*(y)\right)$, where MF(y) - membership function of the final fuzzy set,
- defuzzification or reduction to clarity. There are several methods of defuzzification (for example, the method of the middle center or the centroid method. The geometric meaning of this value is the center of gravity for the curve MF(y)).

The main drawback of the fuzzy conclusion built on Mamdani's classical mechanism is that for any input data it is necessary to process the entire rule base, that mean to carry out three steps (determining the values of membership functions for input variables, mini-max composition and defuzzification). This way of processing fuzzy data reduces the speed of the system and requires memory, so the developing a method for processing fuzzy data, based on the classical Mamdani method, which would satisfy the requirements for speed, is an actual task.

2. Method of fuzzy data processing

The essence of the proposed method is based on dividing of the incoming fuzzy information processing into the training and operation stages.

During the study of the fuzzy data processing device, the function areas of the output, which belonged to each rule, are defined.

During the operation process, the following steps are taken:

- comparison of the input data with the values of the exit membership functions in the memory areas defined by the rules base,

- cut-off values of output membership functions, which exceed the original data,
- choosing of the minimum exit membership functions values obtained after clipping, and figure construction from them,
- searching for the center of gravity of the obtained figure [6].

The analysis of the proposed method shows that all operations of the proposed method are close to the operations of the classical Mamdani mechanism and do not exceed them in complexity. However, the number of operations in the proposed method is smaller, which increases its productivity. The reduction in the number of operations is due to the fact that in the training phase, the areas of exit membership functions for each of the rules are defined. Such preliminary preparation actually avoids the operation of finding the lowest value of the input membership functions provided in the Mamdani method.

The fuzzy data processing tool should perform the following functions:

- receiving from the server the corresponding output membership functions for each of the fuzzy inference rules,
- receiving from the server the specified current values of the input variables,
- calculating of the gravity center of the final output figure,
- receiving from the server a signal for starting the calculation of the gravity center and feeding to the server the signal of the calculation end.

Calculations of the gravity center can be realized by assuming that the membership functions are a flat figure of the same thickness. Then the center of gravity is determined by only two coordinates, and its radius-vector can be calculated by the formula [7]:

$$r_{gc} = \frac{\sum r_i m_i}{\sum m_i} \tag{1}$$

where r_{gc} – coordinate of the gravity center,

r_i – the gravity center coordinate of the *i*-th rectangle, from which the output figure consists,

m_i – mass *i*-th rectangle.

According to the formulated requirements, the structural scheme of the means for processing fuzzy data has been synthesized (Figure 1). It determines the gravity center of the figure constructed according to Mamdani rules based on the input membership functions.

The computing system that implements the assigned application task named as implementation block (IB). Before operation, the BI writes to the multi-channel memory unit (MMU) the values of the previously processed data of the input membership functions. The control unit (CU) generates the corresponding addresses. The values of the input membership functions are fed to the memory register MR8.

Suppose that the fuzzy system has three inputs and one output. Then the set of membership functions values of the first input is written into the memory device MD1, the set of values of the second input membership functions is stored in the memory MD2, and the set of values of the third input membership functions is stored in the memory device MD3, respectively. It is advisable to use flash-memory, which is stable to power supply failures, as MD1-MD3.

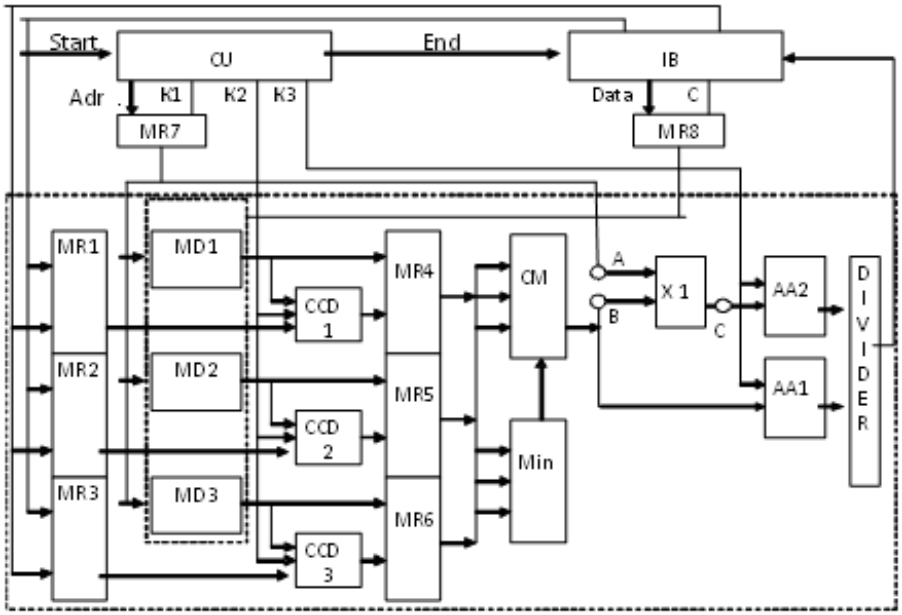


Figure 1. Block diagram of the fuzzy data processing mean

When an applied task must be solved, the IB sets its input values. These parameters enter the registers of memory MR1 - MR3 and set the signal "start", which is also formed by the BI. The "start" signal starts the control unit of the CU of this fuzzy data processing means.

According to this algorithm, the IB first writes the initial address input membership functions values set to the memory register MR7 by the signal K1. According to this address, the codes of the current value of the output accessory functions (coming from MD1 - MD3) and their permissible values (coming from MR1 - MR3) are received by the code comparison device CCD1-CCD3.

At the signal K2, the code comparison devices CCD1-CCD3 are activated. If some of the values coming from MR1 - MR3, are greater than the corresponding values coming from MD1 - MD3, then the last ones are written to the registers MR4 - MR6. If any of the values arriving from MR1 - MR3 are less than the corresponding values coming from MD1-MD3, then writing to the registers of MR4-MR6 is not carried out. Further, according to the Mamdani fuzzy inference mechanism, the minimum value selector Min selects the minimum value of the input membership function from the values recorded in the MR1 - MR3, and, correspondingly addressing the CM switch, applies this value to the inputs of the multiplier X1 and the accumulator adder AA1.

The minimum value of the input membership function from the values recorded in MR1-MR3 corresponds to the mass of the current rectangle mi in formula (1). The product of the address and the minimum value from the registers MR1-MR3, obtained at the output of the multiplier X1, corresponds to the product rimi in the numerator of this formula. Summation over the numerator and denominator of formula (1) takes place in the accumulators AA2 and AA1, respectively. The value of the gravity center coordinate obtained as a result of the Mamdani fuzzy inference mechanism implementation, according to formula (1), is obtained by dividing the

accomplices by a block. This value is fed to the BI unit, where it is compared with the values of the output membership functions that characterize the results of the task solution.

It should be noted that in the structural diagram presented in Fig. 1, the calculation of the length of vectors is realized using a simplified method - only the abscissa of the current point is taken into account, and the ordinate is ignored. The geometrical interpretation of the method of calculating the length of vectors according to the structural scheme of Fig. 1 is shown in Fig. 2. Such simplification substantially reduces the time of determination of the gravity center coordinate. However, in some cases, ignoring the ordinate of the vector length can lead increasing of the methodical error by several times.

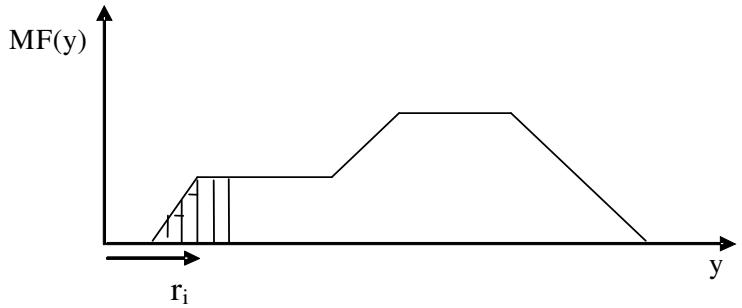


Figure 2. The figure gravity center without taking into account the coordinates of the vectors

To eliminate this error in the vector length calculating, it is proposed to introduce into the structural diagram of Fig. 1, instead of the multiplier X1, a node, which structural diagram is shown in Fig.3. It consists of a divisor by 2, multipliers X1 and X2, an adder A3 and a square root extractor SRE. The node is connected to the circuit of Fig. 1 by buses A, B, C, according to the notations in Fig. 1 and 3. This node determines the vector length by the Pythagorean Theorem, according to Fig.4.

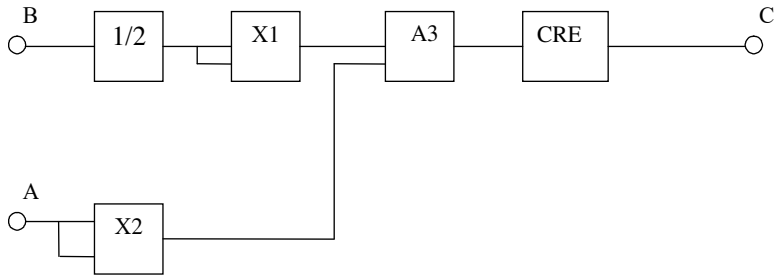


Figure 3. Node for increasing the accuracy of the gravity center calculation

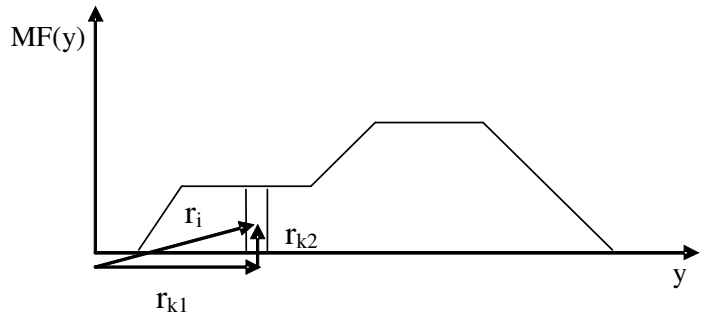


Figure 4. Searching of the gravity center, taking into account its coordinates

3. Intelligent system of intrusion detection

The intelligent system of intrusion detection is the set of "intelligent" immune detectors and rules that describe their behavior. The system consists of modules that perform the control of immune detectors. The immune detectors are going through the different stages during the lifetime. There is creation, training, selection, detection etc. stages. Each stage can be represented as a module of the system of intrusion detection [8-14].

Selecting the structure of the immune detector is very important since it influences directly to the detection ability. In our opinion the neural network structure of immune detectors is preferable and it enables to construct more powerful detectors. We designed immune detectors which are based on the feed forward counter propagation neural network [15]. The counter propagation network is guaranteed that it finds the correct weights during the learning process (Fig.5).

The neural network includes the three layers of nodes. Input layer's nodes connect to each node in the hidden layer and receive data from outside (it can be different files, system processes, network traffic and etc.). The number of inputs nodes *n* defines the size of the window for data inputs to neural network.

The hidden layer consists of *m* Kohonen neurons and it represents a vector quantization layer [8, 16], which gives the cluster label of the input pattern. The competitive learning rule (winner-takes-all) is used for training the hidden layer. The number of neurons

$$m = p + r \tag{2}$$

where *p* is the number of the first neurons which corresponding to legitimate files;
r is the number of last neurons, their activity characterizes the class of malicious files.

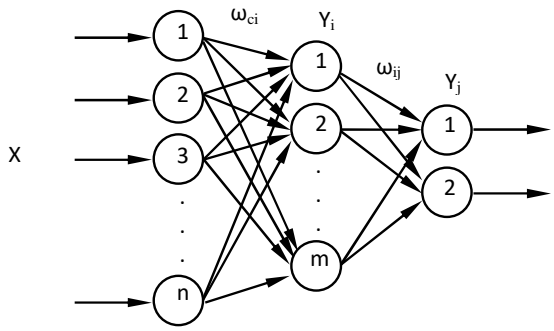


Figure 5. Feed forward Counter propagation Neural Network as the basis of Immune Detector

The ratio of p to r should be multiple of 4 to 1 (for example $p = 8$, $r = 2$).
This ratio is related to the algorithm of forming the learning sample which was received experimentally and showed best results.
The output layer consists of linear units and carries out mapping of clusters into classes. It includes the two nodes: the activity of the first node indicates the legitimate object; the activity of the second node represents a computer attack. Weights between layers and relationships and classes are defined within the learning phase.
During the lifetime the immune detector goes through a several stages: creation, learning, selection, cloning and mutation, notably the immune detector evolves during its "life" [9, 10].
The detectors' life starts from its creation. At this stage the neural networks with random weights are generated.
At the training stage the created neural network immune detectors are subjected to the training process. As a result the ensemble of various detectors is created, and each detector can detect different computer attacks.
Let's consider the N is the data which belongs to the certain type of computer attacks, and the M is the data which belongs to the class of legitimate objects. Then the ensemble of input images can be formed casually for the i detector training:

$$X_i = \begin{bmatrix} X_i^1 \\ X_i^2 \\ \dots \\ X_i^L \end{bmatrix} = \begin{bmatrix} X_{i1}^1 & X_{i2}^1 & \dots & X_{in}^1 \\ X_{i1}^2 & X_{i2}^2 & \dots & X_{in}^2 \\ \dots & & & \\ X_{i1}^L & X_{i2}^L & \dots & X_{in}^L \end{bmatrix} \tag{3}$$

Accordingly, we get a plural of reference images

$$e_i = \begin{bmatrix} e_i^1 \\ \mathrm{e}_i^2 \\ \dots \\ e_i^L \end{bmatrix} = \begin{bmatrix} e_{i1}^1 & e_{i2}^1 \\ e_{i1}^2 & e_{i2}^2 \\ \dots & \\ e_{i1}^L & e_{i2}^L \end{bmatrix} \tag{4}$$

where L is a dimension of the training sample.

Reference output values for the *i* detector are formed as the following:

$$e_{i1}^k = \begin{cases} 1, & if\ X_i^k \in N \\ 0, & otherwise \end{cases}$$
$$e_{i2}^k = \begin{cases} 1, & if\ X_i^k \in M \\ 0, & otherwise \end{cases} \quad (5)$$

First of all a training of neuron network is running to the moment of the total quadratic error minimization:

$$E_i = \frac{1}{2}\sum_{k=1}^{L}\sum_{j=1}^{2}(Y_j^k - e_j^k)^2, \quad (6)$$

where Y_j^k is the *j* output value of detector for the *k* input image;

e_j^k is the *j* standard value for the *k* reference image.

Then the training is running unless the amount of trained immune detectors becomes equal to the set value of *F*.

After training the set of detectors scan an environment and classify objects. The proposed structure of immune detectors enables to use the small dataset for training and classify correctly real-world patterns after training.

During the lifetime neural network immune detectors evolve continually. The evolution of immune detectors is an important part the intelligent system running, because it enables to expose new regularities and features of continually appearing novel attacks as well as adapt to it. As a result the system is evolving and improving its defense's abilities. Let's examine the process of adaption.

If the *i*-th immune detector detects the attack then it activates the alarm, and the cloning and mutation of the given detector are performed. As a result the set of clones is generated and each clone is trained using the attack code (the process of mutation). The algorithm of evolution consists of following steps:

- creating a set D of copies (clones) of the detector that found the computer attack,
- creating the learning sample L from the data of the detected attack,
- training the clones,
- calculating the fitness F of clones. If the fitness increases the clone then it is "good". In the opposite case the clone is eliminated.

As a result, the set of detectors-clones D_i are produced, which are aimed to detect the given attack.

The goal of the mutation is to explore a novel computer attack, and find samples of novel attack's techniques, and elaborate robust detectors. Therefore we apply the relearning process of clones. During this process clones adapt to the novel attack and provide an effective defense from the active cyber attack.

The fitness function is used to determine the detection quality. The mean-squared error between input and output vectors (5) for the immune detector can be used as the fitness.

The immune memory is creating in the final step of immune detectors evolution. The immune memory consists of the "best" neural immune detectors which confirmed the

highest fitness during the detecting the certain cyber attacks:

$$M_k = D_i, if\ E_i < E_j. \tag{7}$$

Detector-clones with the minimal value of mean-squared error are transformed into the memory detector with the "unlimited" lifetime.
To make decisions on the use of a particular set of detectors, it is proposed to use the method described above.

4. Conclusions

The proposed method for fuzzy data processing on the basis of the Mamdani mechanism makes it possible to improve the speed of systems solving applied engineering problems by distributing the process of its implementation to the stages of training and operation.
A method for accelerated fuzzy data processing has been developed, which can be implemented using modern FPGAs, which will allow it to be used for decision-making in intelligent intrusion detection systems.

REFERENCES

1. ROSS T.J.: Fuzzy Logic with Engineering Applications. McGraw-Hill Inc.(USA), 1995.
2. SHTOVBA S.D.: Introduction to the theory of fuzzy sets and fuzzy logic: *http://matlab.exponenta.ru/fuzzylogic/book1* (in Russian).
3. SHTOVBA S.D.: Securing accuracy and transparency of fuzzy Mamdani's models in learning of experemental Data. Management and Informatics, 4 (2007), 102–114 (in Ukrainian).
4. SHTOVBA S.D.: Ensuring Accuracy and Transparency of Mamdani Fuzzy Model in Learning by Experimental Data. Journal of Automation and Information Sciences, 39 (2007), 39-52.
5. DUBCHAK L., VASYLKIV N., KOCHAN V., LYAPANDRA A.: Fuzzy Data Processing Method. Proc. Inter. Conf. Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS-2013), Berlin 2013, 373-375.
6. DUBCHAK L.: Method for processing fuzzy information. Visnyk Skhidnoukrayins′koho natsional′noho universytetu im. V.Dalya, 8 (2012), 306-309 (in Ukrainian).
7. KARASEV A.I. Probability Theory and Mathematical Statistics. Statistica, 1979.
8. GOLOVKO V., BEZOBRAZOV S., KACHURKA P., VAITSEKHOVICH L.: Neural Network and Artificial Immune Systems for Malware and Network Intrusion Detection. Studies in computational intelligence, Springer Berlin/Heidelberg, 263 (2010): Advances in machine learning II, 485–513.
9. KOMAR M., GOLOVKO V., SACHENKO A., BEZOBRAZOV S.: Development of neural network immune detectors for computer attacks recognition and classification. Proc. Inter. Conf. Intelligent Data Acquisition and

Advanced Computing Systems: Technology and Applications (IDAACS-2013), Berlin 2013, 665-668.
10. Pat. Number 109640 Ukraine, IPC (2012) H04W 12/08, G06F 21/00, G06F 12/14. Method of detection of computer attacks by the neural network artificial immune system, Komar M., Sachenko A., Golovko V., Bezobrazov S., patent holder Ternopil National Economic University. – № a201205350; appl. 28.04.12, publ. 25.09.15, № 18 (In Ukrainian).
11. KOMAR M., SACHENKO A., BEZOBRAZOV S., GOLOVKO V. Intelligent Cyber Defense System. Proc. Inter. Conf. on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer (ICTERI 2016), Kyiv 2016, 534-549.
12. KOMAR M., GOLOVKO V., SACHENKO A., BEZOBRAZOV S. Intelligent system for detection of networking intrusion. Proc. Inter. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS-2011), Prague 2011, 374-377.
13. BEZOBRAZOV S, SACHENKO A., KOMAR M., RUBANAU V. The Methods of Artificial Intelligence for Malicious Applications Detection in Android OS. International Journal of «Computing». – 2016. – Vol. 15, no. 3. – P. 184 – 190.
14. KOTENKO I. Multi-agent Simulation of Attacks and Defense Mechanisms in Computer Networks. International Journal of «Computing». – 2008. – Vol. 7, no. 2. – P. 35 – 43.
15. HAYKIN S. Neural Networks: A Comprehensive Foundation. Prentice Hall, 1999.
16. KOHONEN T. The self organizing map. Proc. of the Institute of Electrical and E1ectronics Engineers, 78 (1990), 1464 – 1480.

Liliia GALATA[1]

Scientific advisor: Bogdan KORNIYENKO [2]

MODELLING OF INFORMATION SECURITY SYSTEM IN COMPUTER NETWORK

Summary: This article presents simulation modelling process as the way to study the behaviour of the Information Security system. Graphical Network Simulator is used for modelling such system and Kali Linux is used for penetration testing and security audit. The main approaches to simulation of computer networks are considered. The functional capabilities of the GNS3 package are explored. When building an imitation model, the main components of information protection were used. The Kali Linux package implements a number of attacks. Using simulation in the design of computer systems done the following: estimated bandwidth network and its components identified bottlenecks in the structure of computer systems; compared different options for computer systems; made a promising forecast for the development of computer systems; provides future requirements for bandwidth.

Keywords: mathematical model, simulation model, security, threats, computer network

MODELOWANIE SYSTEMU BEZPIECZEŃSTWA INFORMACJI W SIECIACH KOMPUTEROWYCH

Streszczenie: W niniejszej pracy przedstawiono symulację procesu jako sposób na badanie zachowania Systemów Bezpieczeństwa Informacji. Graficzny Symulator Sieci jest używany do modelowania, natomiast Kali Linux jest stosowany do przeprowadzania testów penetracji oraz audytu bezpieczeństwa. Funkcjonalne możliwości pakietu GNS3 zostały przebadane. W trakcie budowania modeli, używano głównych komponentów do zabezpieczania informacji. Za pomocą Kali Linux przeprowadzono serię symulowanych ataków. Na podstawie powyższych symulacji określono m.in. sieci szerokopasmowe oraz zidentyfikowano ich ograniczenia (ang. bottleneck) w strukturze sieci, porównano różne opcje budowy sieci komputerowych, przeprowadzono przewidywanie/prognozę rozwoju sieci komputerowych oraz określono przyszłe wymagania/specyfikacje dla sieci szerokopasmowych.

Keywords: model matematyczny, symulacja, bezpieczeństwo, zagrożenie, sieć komputerowa

[1] National Aviation University, Kyiv, Ukraine, e-mail: galataliliya@gmail.com

[2] Dr.habil., associate professor, National Aviation University, Kyiv, Ukraine, email: bogdanko@i.ua

1. Introduction

Efficient construction and usage of corporate information systems have become an extremely important task, especially in insufficient funding of information technology in enterprises. Evaluation criteria for efficiency are the cost reducing of the information system implementation, current and nearest future requirements compliance, the opportunity and the cost of further development and transition to new technologies [1-3]. The information system core is a computing system that includes the following components: cable network and active network equipment, computer and peripheral equipment, data storages (libraries), system software (operating systems, database management systems), special software (monitoring and network management) and in some cases the applied software [4-9].
The purpose of research is to construct and study mathematical model for Information Security System in Computer Network by using modern software.

2. Modelling

Now the most common approach in information systems design is to use expert estimates. According to this approach, experts in the field of computing tools, active network equipment, cable networks, design computing system to solve the specific task or class of tasks, based on their experience and expert estimates. This approach minimizes the cost of the design stage, quickly estimate the cost of implementing the information system. However, decisions obtained by using expert estimates are subjective, hardware and software requirements as the assessment of guarantees for efficiency of proposed system project are subjective too.
As an alternative may be used approach, which involves the development of models and modelling (simulation work - simulation) of computing system behaviour. The modelling is a fundamental method for studying the behaviour of complex systems.
The modelling is one of the main methods of knowledge, and a form of reflection of reality. The modelling is to clarify or reproduction of certain properties of real objects, things and events through other objects, processes, events, or through abstract descriptions such as image, plan, map, set of equations, algorithms and applications.
The model is defined as "a system that is provided or material implemented, that is replaced the real object (system) in the process of cognition or analysing, while retaining some of the most important features for its research, and its study gives us new information about the object.
Here are the main types of models used in practice to describe the different processes and systems are as follows:

- the conceptual model – the model describes the system using special characters, symbols, operations or using natural or artificial languages;
- the physical model - the system reproduces based on the ratio of similarity, that is resulting from the similarity of physical phenomena;
- the structural and functional model - as a model uses scheme (block diagram), tables, graphs, diagrams and drawings with special rules of their union and transformation;
- the math model is a math representation of reality, the description of some phenomenon or system using math concepts and symbols;

- the simulation model - economic and math model uses in the experimental study of system or phenomena by using personal computers.

These types of models can be used both individually and a few at a time, also, when you use simulation modeling, it involves all of these types or their separate techniques. The simulation model allows us to visualize the final or intermediate result dynamically, that is an important aspect for a successful understanding the received results by persons who did not participate in its development.

3. Simulation model

Common definitions of term “simulation modelling”:

- it is the method allows to build models that describe the processes as they would take place in reality. Such model can be "plaied" for a single test or set of tests. The results will be determined by the random behavior of the process;
- it is the research method in which studied system is replaced by a model that accuracy describes the real system, with which experiments are conducted to obtain information about this system;
- it is the special case of math modeling. There is a class of objects, for which have not developed analytical models or solution methods of resulting model for various reasons. In this case, the analytical model is by the simulator or simulation model;
- it is logical and mathematical description of an object that can be used to experiment on your computer in order to design, analysis and assessment of the object.

Simulation modelling is used to study the behaviour of the system by using math tools and computing equipment. The calculation of the required results may be automated by using computing technology, with only initial data, for example, been obtained statistically. It is especially important, when complex system that consist of many components is being used, because for calculation of required results you need to use cumbersome formulas usually. Simulation modelling is applied to study the behaviour of various systems, including the information one [1] [2].

Mainly in information systems modelling there is an aim to achieve information about request processing time or resource load level. As for computing networks, their simulation models reproduce the processes of message generation by applications, of splitting messages into specific protocols packets and frames, of delays in processing messages, packets and frames within the operating system, of computer access to the shared network environment, of router incoming packets processing and etc. No need to buy expensive equipment by using simulation modelling network - its work simulates by programs that accurately reproduce the equipment main features and options.

4. Simulation environment

It was offered to use simulation modelling for determination the actual security threats [3]. GNS3 Cisco Systems software have been selected as the simulation environment (Fig 1).

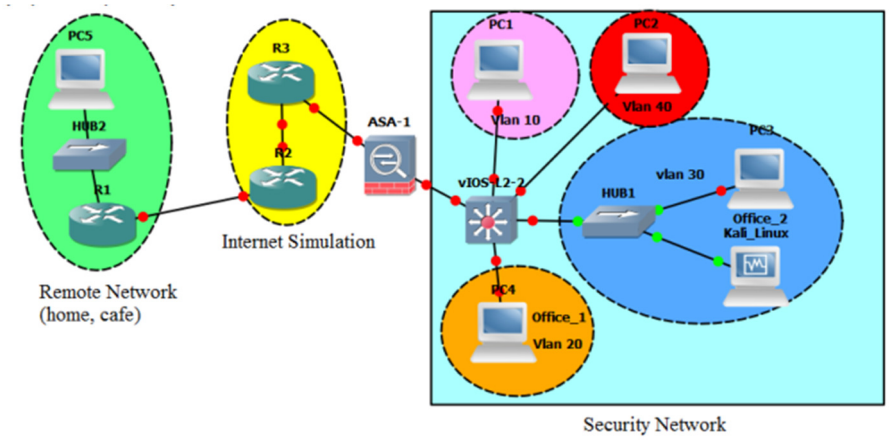


Figure 1. The simulation model of information security system in computer network

The choice of this software was done due to the following factors:

- a process of creating models is facilitated by using a graphical development environment;
- previously created modules and libraries can be used to create new models;
- object-oriented approach to model building;
- a large number of built-in libraries for creating simulation models;
- models can be run at any software and hardware platform;
- a simulation model can be run without development tools.

Graphical Network Simulator GNS3 is a cross-platform program with open source. It is based on popular Dynamips (CISCO IOS emulator), Dynagen (Dynamips text interface) and Pemu (Cisco PIX emulator).

GNS3 provides easy to use GUI, and a range of other features. You can model the new configuration, various images of IOS, or perhaps, make fully reconstruction of some complex network parts. That is much easier with this program than its be in a real network. The product processes the installation and configuration of essential utilities automatically. The GNS3 installation package includes all emulators. In the case of GNS3 installation on the Microsoft Windows operating system, you must also install WireShark, it is necessary to intercept monitor network packets and libraries.

Sometimes, it is not enough network devices in the company and there is no router, which deals with internal networks routing, but there is only L2-switch and security device Cisco ASA with IOS 8.4.2 version. So, it is necessary to set additional

functionality on Cisco ASA, such as routing. Similarly, we have only one interface to connect with the L2-switch. Also we need to configure remote users' connections by VPN.

There is the next task: the networking between internal networks should be organize to meet the requirements of security. The guest network access should be organize only to "INTERNET" with limited speed of 1024 Kb. The remote users connect should be organize through Remote Access VPN, therefore remote users should connect to the internet via Cisco ASA device and have access to the internal resources of the company. Internal website should be available on "INTERNET". All of this tasks should be done through CLI. We have a central office with installed Cisco ASA device and L2-switch. Four networks (VLANs) have been created at switch, which submitted to security device through Trunk.

There are the characteristics of each VLAN-s:

- Vlan_Office_1 - network 192.168.2.0/24.
- Security Level is 100. It is uses for first part of employees. There are the Internet access from this subnet, and Vlan_Office_2, Vlan_DMZ and Vlan_Guests access;
- Vlan_Office_2 - network 192.168.3.0/24.
- Security Level is 100. It is uses for second part of employees. There are the Internet access from this subnet, and Vlan_Office_1, Vlan_DMZ and Vlan_Guests access;
- Vlan_DMZ - network 192.168.1.0/24.
- Security Level is 50. There is a Web-Server (WWW-SRV) with company's website in this network. Accordingly, it is available from Internet at port 80 (TCP) and there are access from Vlan_Office_1, Vlan_Office_2 subnets to this network and from Vlan_Guests subnet at 80th port;
- Vlan_Guests - Guest subnet 192.168.4.0/24.
- It is uses for "guests" who came to our office. Security Level is 10. There are the Internet access from this subnet with limit speed of 512 Kb and access to the internal website (SRV-WWW) only at 80th port.

There is a network, that simulates "INTERNET", which uses two routers (Router_1 and Router_2). There is loopback-interface (IP-address 1.1.1.1) at Router_1, which will be use to check for the "INTERNET". Dynamic routing protocol OSPF is used for routes exchange. Also there is a remote user, which is placed in the subnet 192.168.5.0/24 (say it is internet-cafe) behind Remote Router. This remote user has access to the Internet but he is considered dangerous without VPN connection to the central office.

A simulation model consists of protected and unprotected networks. The main element of information security system is the firewall ASA 8.4, a platform for atak-set Kali Linux. Kali Linux is modern Linux-distribution for penetration testing and security audit. Kali is a complete reassembly BackTrack Linux, fully according to Debian development standards.

All new infrastructure has been revised, all the tools were analysed and packaged, and we switched to Git for our VCS.

- There are more than 300 tools for penetration testing: After considering each tool that was included in BackTrack, we have removed a large number of tools that either do not work or duplicate other tools with similar functionality.

- Kali Linux is completely free and always will be free. You will never have to pay for Kali Linux.
- Git tree with open source code: our tree is open to all, and all of sources available for set up or rebuild packages.
- FHS compliant: Kali was designed to observe the Filesystem Hierarchy Standard, which allows all Linux users easily find executable files, support files, libraries, etc.
- Wide support for wireless devices: Kali Linux was built to support many wireless devices, allowing it to work correctly with a wide range of hardware devices and making it compatible with many USB and other wireless devices.
- Special core patches from injection: developers often need to audit wireless networks, so our core includes the latest patches for them.
- Secure Development Environment: Kali Linux development team consists of a small group of trusted persons who can add packages or interact with storage only by using several secure protocols.
- GPG signed packages and repositories: All packages are signed by each individual Kali developer when they are created and recorded, and then the repository signed packages also.
- Multilingual: Kali has a true multi-language support, allowing most users to work in their native language and to find the tools needed for the job.
- Customizable: You can as easy as possible customize Kali Linux to your taste, down to the core.
- Support ARMEL and ARMHF: Kali supports ARM-systems and has installations for ARMEL and ARMHF systems. Kali Linux ARM repository is integrated with the main distribution.

LOIC was used to implement attacks. The program performs a distributed attack such as "denial of service" by TCP-, UDP-packets or HTTP-requests regular transfer to the certain site or host with a goal to destroy the target node. There is also an edition of the program LOIC Hive Mind, that automatically receive the task to attack via IRC, RSS or Twitter, which allows centralized DDoS-attacks.

Attacks occur from a remote location to internal subnet (Office_1, Office_2, DMZ, Guests) with different security levels. There are customized security levels: offices - security level is 100, the traffic between offices is not filtered. DMZ security level – 50, Guests -10. Offices traffic is unrestricted with other subnets, there is access from DMZ to Guest subnet, there is access from the guest subnet only to the Internet. Internet working is emulated by two routers with loopback-interface.

5. Conclusions

Using modelling in the design of computing systems, you can:

- estimate the bandwidth of the network and its components;
- identify vulnerability in the structure of computing system;
- compare different organizations of a computing system;
- make a perspective development forecast for computering system;
- predict future requirements for network bandwidth;
- estimate the performance and the required number of servers in the network;
- compare various options for computing system upgrading;
- estimate the impact of software upgrades, workstations or servers power, network protocols changes on the computing system.

Research computing system parameters with different characteristics of the individual components allows us to select the network and computing equipment, taking into account its performance, quality of service, reliability and cost. As the cost of a single port in active network equipment can vary depends on the manufacturer's equipment, technology used, reliability, manageability.

The modeling can minimize the cost of equipment for the computing system. The modeling becomes effective when the number of workstations are 50-100, and when it more than 300, the total savings could reach 30-40% of project cost.

REFERENCES

1. KORNIYENKO B.Y. Model of Open Systems Interconnection terms of information security, B. Korniyenko, Science intensive technology. – 2012, № 3 (15), P. 83 – 89., doi.org/10.18372/2310-5461.15.5120 (ukr).
2. KORNIYENKO B.Y. Open systems interconnection model investigation from the viewpoint of information security /B. Korniyenko, O. Yudin, E. Novizkij, The Advanced Science Journal. – 2013. - issue 8. - P. 53 – 56.
3. KORNIYENKO B.Y. Implementation of information security a model of open systems interconnection, B. Korniyenko, O. Yudin, Abstracts of the VI International Scientific Conference "Computer systems and network technologies» (CSNT-2013), 11-13 June 2013, - P.73.
4. KORNIYENKO B.Y. Information security and computer network technologies: monograph, B. Korniyenko, ISBN 978-3-330-02028-3, LAMBERT Academic Publishing, Saarbrucken, Deutschland. − 2016. – 102 p.
5. KORNIYENKO B.Y. Modeling of security and risk assessment in information and communication system /B. Korniyenko, L. Galata, O. Kozuberda/ Sciences of Europe. – 2016. – V. 2. – No 2 (2). – P. 61 -63.
6. KORNIYENKO B.Y. The classification of information technologies and control systems, B. Korniyenko, International scientific journal. – 2016. –№ 2. – P. 78 - 81.
7. KORNIYENKO B.Y. Risk estimation of information system, B. Korniyenko, A. Yudin, L. Galata, Wschodnioeuropejskie Czasopismo Naukowe. – 2016. –№ 5. – P. 35 - 40.

8. KORNIYENKO B.Y. Simulation of information security of computer networks, B. Korniyenko, L. Galata, B. Udowenko, Intellectual decision making systems and computing intelligence problems (ISDMCI'2016): Collection of scientific papers of the international scientific conference May 24-28, 2016 Kherson, Ukraine, pp. 77 - 79.
9. KORNIYENKO B.Y. Cyber security - operating systems and protocols, B. Korniyenko, ISBN 978-3-330-08397-4, LAMBERT Academic Publishing, Saarbrucken, Deutschland. − 2017. – 122 p.

Viktor GNATYUK[1], Nadiia DYKA[2] Vitalii KOTELIANETS[3],
Serhii DAKOV[4]

Opiekun naukowy: Roman ODARCHENKO[5]

ARCHITEKTURA SYSTEMU IoT DLA SYSTEMU MONITORINGU ZANIECZYSZCZENIA POWIETRZA

Streszczenie: Świat technologii rozwija się obecnie gwałtowni. Od pewnego czasu rozwijana jest koncepcja Internetu Rzeczy. W ogólności technologię Internetu Rzeczy można podzielić na dwa aspekty: elektroniczny oraz telekomunikacyjny. W odniesieniu do konceptu Internetu Rzeczy rozważa się popularne protokołu komunikacyjne – takie jak: CoAP, MQTT, HTTP, AMQP, XMPP. Są one używane do przesyłania informacji z wszelkiego rodzaju sensorów do chmury serwera. Wszystkie rozważane protokoły różnią się znacząco między sobą. Najłatwiejszym sposobem, aby dokonać ich klasyfikacji, jest wyróżnienie ich kluczowych parametrów np.: jakości usług, adresowania oraz zakresu zastosowań. W pracy przeprowadzono szczegółową analizę technologii używanych w zakresie Internetu Rzeczy (IoT). Porównano ich główne wady oraz określono zakresy użytkowania. Został omówiony system typu Internet Rzeczy do monitorowania jakości powietrza. Dodatkowo, zaproponowano praktyczne wdrożenie sieci opartej o oprogramowanie Arduino oraz platformę sprzętową.

Słowa kluczowe: Internet Rzeczy, komunikacja, protokół, internet, technologie bezprzewodowe, ZigBee, Bluetooth, Wi-Fi, sieci rozległe

IoT ARCHITECTURE FOR AIR POLLUTION MONITORING SYSTEM

Summary: The world of technology is rapidly developing. For a long time already such a concept has appeared as the Internet of Things (IOT). Conditionally IOT can be divided into 2 parts: electronics and telecommunications. Therefore, the concept of the Internet of things and the most popular communication protocols of this concept such as: CoAP, MQTT, HTTP, AMQP, XMPP, which are used for sending information from any sensor to the cloud server were considered in this paper. All protocols which have been considered considerably differ

[1] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD, associate professor, victorgnatyuk@ukr.net
[2] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD student, Nadin_dyka@ukr.net
[3] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD student, v.kotelianets@ukr.net
[4] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD student, Dakov@ukr.net
[5] National Aviation University, Institute of Air Navigation, Department of Telecommunication Systems, PhD, associate professor, odarchenko.r.s@ukr.net

among themselves. The easiest way to classify them according to several key parameters: quality of service, addressing and application. A detailed analysis of the technologies used in the IoT was carried out, their comparisons were made, their main drawbacks and application features were determined. The architecture of IoT for monitoring of quality of air was developed. In addition, the practical implementation of the IoT network architecture based on the Arduino software and hardware platform was proposed..

Keywords: IoT, communication, protocol, internet, wireless technologies, ZigBee, Bluetooth, Wi-Fi, LoRaWAN.

1. Introduction

The world of technology is rapidly evolving that soon anything that cannot be contacted over the network will be morally obsolete. Therefore, such a concept as Internet of Things (IOT) was appeared [1,2]. The Internet of things covers practically all aspects of human life and existence now. Real objects, which provided by processors will transmit data about their condition and environment to the network, forming their virtual image. Things will be able to receive information from the world around, interact, share data. Computers will receive own means of collection of information, will begin to see, hear and smell.
The artificial intelligence, embedded in the program, will allow you to estimate what is happening, take into account the information and experience accumulated previously for providing decision-making both with the participation of the person, and in the automatic mode.
IoT will be used also in technological areas, including telemetry, telematics, M2M-communications (Machine to machine communication), intellectual networks, the intellectual systems of transportation and laptops. In fact, it is an extremely complex system that has assembled all the latest high-tech technologies of our time.

In the general case, the Internet of things is understood as set of different instruments, sensors, devices, united in a network using any available communication channels that use different interaction protocols and a single protocol for access to the global network [3]. As the global network of Internet of things, the Internet is now being used. In other words, the Internet of things can be regarded as a network of networks, in which small, non-connected networks form larger ones.

2. Analysis of research and publications

Both domestic and foreign scientists worked on a research of methods of ensuring work and means for automation of processes in IoT. Works of such Ukrainian scientists devote to this problem: T. N. Sooner Pavel Mikhaylov, Alexander Yakunin, Dmitry Kuleshov, M.V. Dzyuba, Anastasia Semakina. In the list of scientists It should be noted Erica Brindzholfsona, Andrew Mac Afa, Daniel Abadovski, Vessels Dzhami, Peter Lucas, Peter Vakher, Don Norman.

3. Goals and research problem

The state of atmospheric air - one of the main factors that affect the health of the population. With the development of crowded industrial cities the problem of pollution has become more important.
Using automobiles and other machines made pollution steadily worse. Annually around 17 million tons of harmful substances are released into the atmosphere. About half of the Earth's population breathes air, which is officially recognized as harmful to health. According to data of World Health Organization, air pollution is the major ecological factor of increase in incidence and mortality in the world.
The enterprises of ferrous metallurgy, power, the coal industry, the chemical and petrochemical industry are the main air pollutants in Ukraine. However, emissions from CHP and cars, which are increasing every year, have a significant impact on this. For measurement of a status of air pollution, it will be expedient to use sensors of the concept of the Internet of things.
Therefore the purpose of this article is the analysis of protocols and IoT technologies and as well as the development of the architecture and a concept of system which will be necessary for measurement the pollution of the environment.

3. Statement of the main material of a research

The concept of IoT plays a decisive role in the further development of the info communication industry. This is confirmed by both the position of the International Telecommunication Union (ITU) and the European Union in the matter, and inclusion of the Internet of things in the list of breakthrough technologies in the USA, China and other countries [3].
Conditionally IOT can be divided into 2 parts: electronics and telecommunications. IOT will create tendencies to combining of different telecommunication technologies that will open opportunities for provision of services of the new type.
IoT will create tendencies to combining of different telecommunication technologies that will open opportunities for provision of services of the new type. Integration the global integration of digital mobile communication GSM with Near Field Communication (NFC), personal area networks on the basis of Bluetooth, wireless local area networks, ZigBee standard wireless sensor networks, in combination with the system of global positioning and technology of subscriber identification (SIM card) is supposed [4].
The Application layer of IoT in Recommendation Y.2060 is not considered in detail. Service support and Application support layer include the general opportunities for various objects of IoT for processing and data storage as well as the capabilities required for some IoT applications or groups of such applications.
The network layer includes networking capability (function of resource management of an access network and transport network, control of mobility, functions of authorization, authentication and calculations) and transport opportunities (ensuring network connectivity for the transfer of information of applications and services IoT). Finally, device layer includes device capabilities and gateway capabilities [7]. The device features assume direct exchange with a communication network, exchange

via the gateway, exchange through wireless dynamic ad-hoc network, as well as temporary shutdown and restoration of the device for energy-saving purposes.

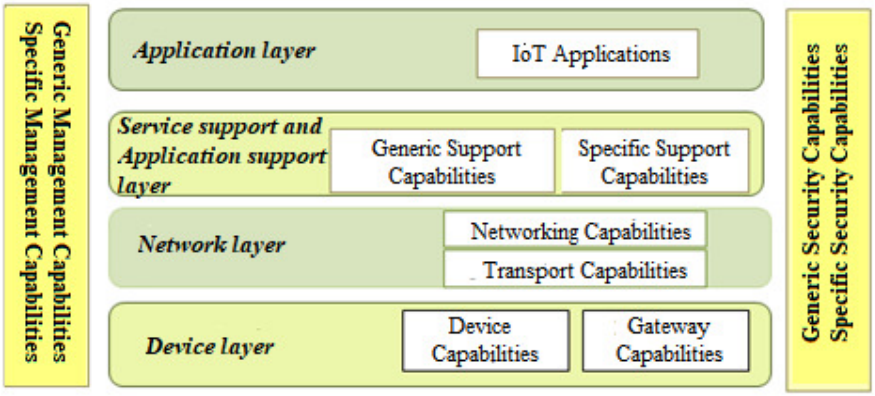


Figure 1. Layers model for IoT

Gateway capabilities include support for many interfaces for devices (CAN, ZigBee, Bluetooth, WiFi, etc.) and for access / transport networks (2G / 3G, LTE, DSL, etc.). Another possibility of the gateway is support of conversion of protocols, provided that the protocols for the device and network interfaces differ from each other.

There are also two vertical layers - layer of management capabilities and security capabilities layer covering all four horizontal levels.

The capabilities of the vertical level of operational management include managing the consequences of failures, network capabilities, configuration, security, and billing data. The main objects of control are devices, local area networks and their topology, a traffic and overloads congestion. The capabilities of the vertical level of security depend on the horizontal level.

At the device layer - the capabilities of authorization, authentication, access control and privacy of data.

IoT protocol stack. For the interaction of a huge number of different devices in the IoT, standardized interfaces, data formats and communications protocols are required. Figure 2 shows the protocol stack IOT [18,19].

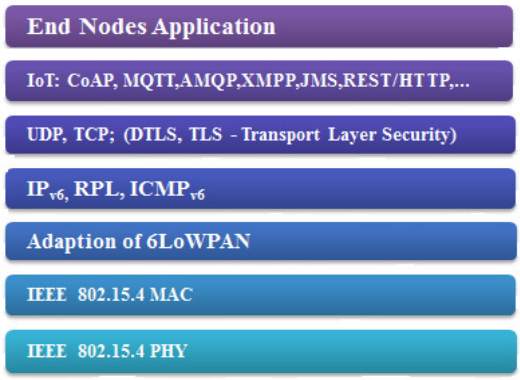


Figure 2. IoT protocol stack

Further, we will in more detail consider protocols of each level.
In principle for data transfer in the modern concept of IoT practically all possible and widespread protocols of the data link layer can be used (Ethernet, Wi-Fi, PLC, etc.), however wireless technologies in most cases attract the greatest interest. As can be seen from Fig. 3, there are numerous such technologies that can be used for the needs of IOT. Therefore, we will stop with the analysis of the most widespread and perspective.

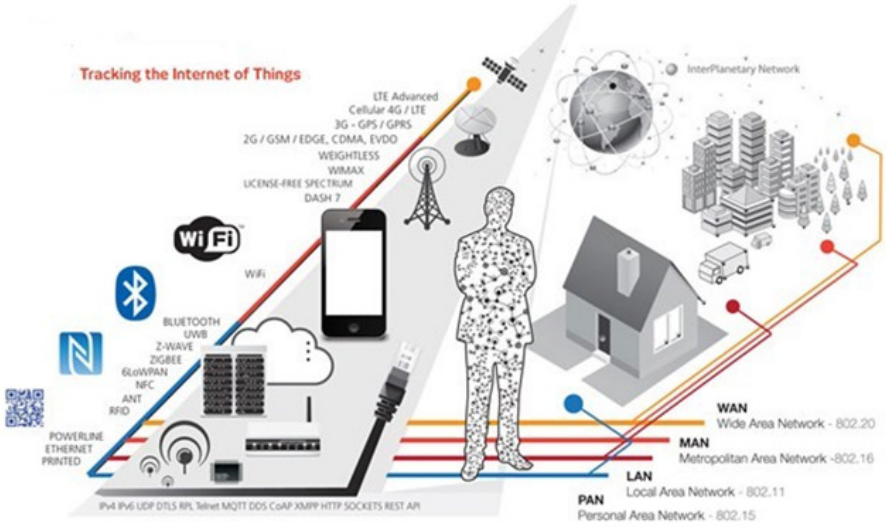


Figure 3. Wireless network protocols for the IOT concept

As a rule, the choice of network technology in many respects depends on a network coverage.
When data is to be transmitted over short distances (for example, within a room), devices may use Personal Area Network (PAN) provided by such technologies of wireless data transfer as BLE (Bluetooth Low Energy), ZigBee, 6LoWPAN and leading USB interface. If it is about data transfer on rather long distance (for example, at the office), it is possible to involve a local area network (Local Area Network, LAN). The leading local area networks are in most cases built on the basis of Ethernet technology and optical fiber, and wireless - on the basis of Wi-fi technology. For the organization of the Wide area network (Wide Area Network, WAN) uses technologies WiMax, LTE, etc. In the last two years technologies of communication for connection of devices with low energy consumption to a wide area network - LPWAN appeared. So, in detail, consider each of the technologies, analyzing their main advantages, disadvantages and features of the application.
Bluetooth LE technology. Bluetooth is the wireless technology, providing data transfer at small distances and allows devices which use this technology, to exchange data [3,8]. Bluetooth allows to communicate with such devices when they are in radius up to 10 meters from each other.
The great advantage of Bluetooth LE is low energy consumption and energy consumption in sleep mode, as well as super-low peak energy consumption.

The devices using Bluetooth LE consume less energy and are capable to transfer data 50 times faster, in comparison with Bluetooth devices of the previous generations. For comfortable use in the fields of house entertainments, health care and security systems this Bluetooth version gives support of a wide range of applications and allows to reduce the size of the end device.

BLE devices work in the range of 2,4 GHz. Control safety devices, control of electric devices and displays of indications, sensors on batteries, home medical devices, sports simulators are the main scopes of the BLE.

Wi-fi technology. Wireless LAN of the WiFi standard 802.11 was developed for the provision of broadband wireless access to data communication networks at high speeds. The stack of protocols of the IEEE 802.11 standard consists of the physical layer of PHY and the data link layer with the media access control sublayers of MAC and logical data transfer of LLC [9].

By means of Wi-Fi it is possible to develop a network without laying of a cable, to have access to network of mobile devices. Within zone Wi-Fi several users from computers, laptops, tablets, phones, etc. can connect to the Internet.

Besides, for IoT also the new Wi-fi HaLow standard (IEEE 802.11ah specification) [3] with low energy consumption is created. This type of communication will function in the range less than 1 GHz, the IEEE 802.11ah specifications. For connection of Wi-fi HaLow the unlicensed frequency of 900 MHz will be used. It noticeably will increase penetration of a signal in urban development, and the radius of its action will be much bigger, than at the modern wireless standard - to 1 kilometer. At the same time, the fee for for "long-range capability" is the low power of a signal. Throughput of Wi-fi HaLow will be much lower, than Wi-Fi maximum 802.11ac (7 Gbit / c), estimated speed: 50 kbps - 18 Mbps [10].

ZigBee Technology (6LoWPAN). ZigBee is a technology based on the IEEE 802.15.4 standard and is designed for the creation of wireless personal networks (WPAN) with use of low-power radio transmitters of the small sizes.

The ZigBee technology is targeted at applications that require longer battery life and greater security in case of data transfer at small speeds [11, 12].

The ZigBee specification is oriented on programs demanding the guaranteed safe data transfer at rather low speeds and the ability to last a long time network devices from standalone power sources (batteries). It provides low consuming of energy and data transfer with a speed up to 250 Kbps on distance to 75 meters in the conditions of direct visibility [13].

The main feature of ZigBee technology is in what it in case of small energy consumption supports not only simple network topologies ("point-to-point", "tree" and the "star") but also which is self-organized and the self-healing mesh topology with relaying and routing of messages. Besides, the ZigBee specification provides simplicity of deployment, service and upgrade.

Z-Wave Technology. Z-Wave is the first open wireless standard for home automation (smart home system), based on a mesh network [10]. This standard uses a 908 MHz frequency and is currently being marketed as a cheaper alternative to ZigBee technology.

Unlike Wi-Fi and other standards of data transfer of IEEE 802.11 which are intended primarily for large streams of information, the Z-Wave standard works in the range of frequencies up to 1 GHz and is optimized for transmission of ordinary managing directors of commands (for example, to include / switch off, change the loudness,

brightness, etc.). The choice of the low radio-frequency range for Z-Wave is caused by a small quantity of potential sources of electromagnetic fields (unlike the loaded range of 2,4 GHz in which it is necessary to resort to measures that reduce the possible interference from working various wireless devices - Wi-Fi, ZigBee, Bluetooth).Other advantages of the standard include low power consumption, low cost of production and the integration of Z-Wave modules into various household appliances.
The data transfer rate is 9.6 kbps or 40 kbps with full compatibility. The radius of action about 30 meters in the conditions of direct visibility, indoors decreases depending on the form and material of the walls and also from a type of the antenna.
The apparent disadvantage of Z-Wave is the lack of scalability required when increasing the number of supported functions in home and industrial networks.
LoRaWAN technology. LoRa (Long Range) is becoming an increasingly popular solution for the IoT concept. LoRaWAN is a "conventionally global" network, which is designed for long-range networks, with the goal of transmitting telemetry data from various accounting devices (water, gas, etc.) to remote locations. An example of LoRa network implementation is shown in Fig. 4 [14].

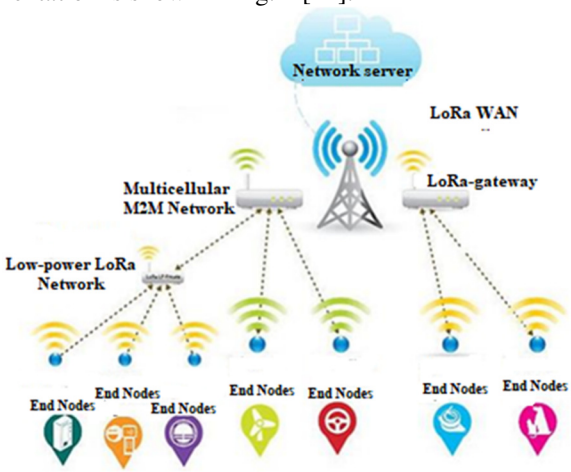


Figure 4. Implementation of the LoRaWAN network

It consists of end nodes (nodes) that send data to a hub (gateway). Each hub has access to the Internet and sends the received data to the server, which in turn sends them to signed clients.
The LoRaWAN technology represents a network of the sizes of the district or even the whole city. The LoRaWAN network has a simple architecture like "star" without retranslators and mesh-communications and is intended for data collection from a large number of sensors which are uniformly located on the significant area. To nodes of a network are characterized by low power consumption (up to 10 years of operation from ordinary AA batteries), low data rate, high communication distance (15 km in rural areas and 5 km in dense urban buildings) and low cost of finite equipment.
Protocol 6LoWPAN. 6LoWPAN (IPv6 Low-Power Wireless Personal Area Network) - the standard, which provides interaction of small wireless networks with the IP networks, according to the IPv6 protocol with small energy consumption [15]. Data transfer in the 6LoWPAN standard implies use of the sub-gigahertz range and provides a transmission rate of 50 to 200 kbps at a distance of up to 800 meters.

Consequently, 6LoWPAN is a network technology or level of adaptation that allows the efficient transmission of IPv6 packets in small channel-level frames defined in the wireless IEEE 802.15.4 standard.
First of all, networks 6LoWPAN is subnets of IPv6-networks, that is they can interact with other networks and nodes of an IP network, but are not transit for its network traffic.
6LoWPAN is that it originally was developed to support the low-power wireless networks of 2,4 GHz, built on the basis of IEEE 802.15.4, but now, this standard is adapted and is used in a set of other environments of network transmission, including wireless networks in the ranges which is lower than 1 GHz, Smart Bluetooth, data transmission over power lines (PLC) and low-power Wi-Fi networks [16].
Thus, we can see some advantages and disadvantages of wireless technologies that can be used in the IoT concept, as well as their field of potential application (tab. 1).

Table 1. Comparison of the core technologies of the concept of IoT

Specifications	**Wi-Fi**	**Wi-Fi HaLow**	**ZigBee**	**LoRaWAN**	**Z-Wave**	**Bluetooth LE**
Standard	IEEE 802.11	IEEE 802.11 ah	IEEE 802.15.4	LoRaWAN	Z-Wave	Bluetooth 4.0
Frequency	2,4 GHz, 5 GHz	900 MHz	915 MHz / 2,4 GHz	863-870 MHz	900 MHz	2,4 GHz
Range of action	Up to 100 m	Up to 1 km	100 m / Mesh	2-5 km in the city; up to 15 km outside the city	30 m / Mesh	80 m
Transfer rate	7 Gbps	50 kbps – 18 Mbps	250 kbps	290 bit/s - 50 kbps	10-100 kbps	< 1 Mbps
Energy consumption	High	Lowered	Low	Low	Low	Low
Scalability	Yes	Yes	Yes	Yes	Limited	Yes
ISM range	Yes	Yes	Yes	Yes	Yes	Yes
Authentication	Yes	Yes	Yes	Yes	Yes	problematicaly
E2E encryption	Yes	Yes	Yes	Yes	Yes	Так
Equipment cost	High	High	Low	Low	High	Low
Is the sensor location known?	Yes	Yes	-	Yes	-	No
Complete two-directional	Yes	Yes	Yes	Yes, depending from mode	Yes	Yes
Support for sensors moving between hubs	Yes	Yes	Yes, mesh	Yes	Yes, mesh	Yes

The detail analysis of protocols of the network layer (IP, routing protocols), the transport layer (TCP, UDP, SCTP) of the OSI model, that are provided in fig. 2 and are involved in a data interchange in the concept of IoT was carried out, but only the HTTP and MQTT protocols are best suited for solving a task from all application-layer protocols. Therefore, we will concentrate only on these protocols.

HTTP protocol. At the application layer on the Internet, the protocol - HTTP, which is a symbol-oriented client-server protocol is widely used, running over TCP [18,21]. HTTP uses XML - a text language with a large amount of service information. Therefore, it is not optimal to use HTTP on many 6LoWPAN systems. However, HTTP can still be very useful for communicating between 6LoWPAN and the Internet.

MQTT protocol. MQTT (Message Queuing Telemetry Transport) is simple and easy messaging protocol that implements the publish / subscribe model and is intended for communication between computerized devices connected to a local or global network with each other and various public or private web services [20].

The MQTT protocol was originally created for sensors that track the state of the pipes, but later the sphere of its activities was expanded and it has been used in a variety of built-in solutions, including smartphones. So the social network Facebook uses this messaging protocol (Facebook Messenger).

In the network on the basis of the protocol MQTT distinguish three objects (Figure 5) [17]:

1) Publisher - MQTT-client, which, in the event of certain events, transmits information about it to the broker;

2) Broker (broker) - MQTT server, which receives information from publishers and passes it to the relevant subscribers, in complex systems can perform various operations related to the analysis and processing of received data;

3) Subscriber - the MQTT-client, who after subscribing to the appropriate broker most of his time "listens" to him and is always ready to receive and process an incoming message from the broker.

dobrej jakości. Opis rysunku – styl „Rysunek” – tekst centrowany.

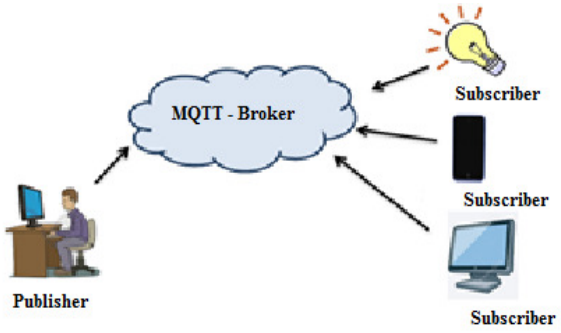


Figure 5. The main structure of the MQTT protocol

The difference to HTTP is that a client doesn’t have to pull the information it needs, but the broker pushes the information to the client, in the case there is something new. Therefore, each MQTT client has a permanently open TCP connection to the broker.

If this connection is interrupted by any circumstances, the MQTT broker can buffer all messages and send them to the client when it is back online.

4. Statement of the main material of a research

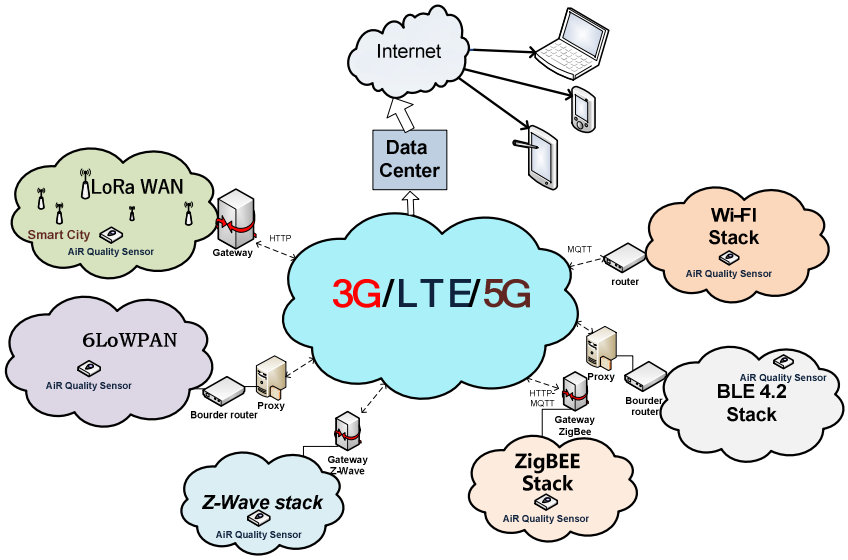


Figure 6. Architecture of IoT for monitoring of air pollution

This figure examines a specific application scenario in which IoT devices and networks are used for applications such as structural monitoring of air pollution. Fig. 6 provides a generic view of an IoT network architecture using different wireless technologies, such as: LoRAWAN, 6LoWPAN,Z-Wave,ZigBee, Wi-Fi, BLE4.2 in which diverse IoT components are being connected to the 3GPP network components for proper operation and data transfer.

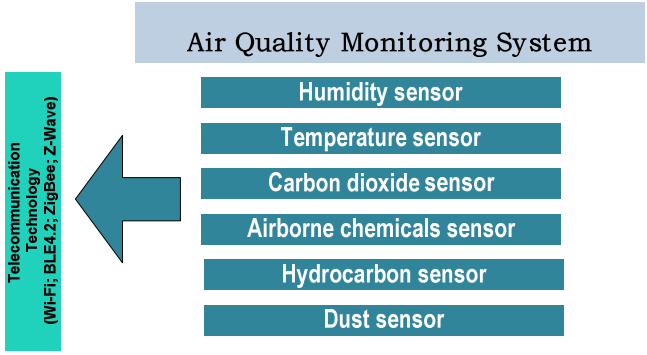


Figure 7. Air Quality MonitoringSystem

The figure shows the formation of IoT networks, also called M2M area networks. The architecture presented in Figure 6 represents a kind of a capillary network architecture,

in which all devices are transferring their collected data to the IoT server through an intermediate entity that is an IoT gateway.
The system comprises of a wireless sensors and a control panel unit to adjust the ambient air quality in a room depending on user parameters. The sensor monitor consists of CO2 sensor (300 – 5000 ppm), temperature sensor (0 – 50 deg C) and humidity sensors (0% to 100% rH). The sensor unit has an inbuilt transmitter module that constantly sends the sensor information to the receiver control unit.
will be necessary for measurement the pollution of the environment.

5. Practical implementation

For the practical implementation of the proposed network architecture of IoT can be used software and hardware platform Arduino, intended for operation with various physical objects and it is a simple board with a microcontroller, it is also a special development environment for writing software of the microcontroller. Arduino can be used to develop interactive systems controlled by different sensors and switches. Such systems, can manage the work of various indicators, engines and other devices. Arduino projects can be either independent or interact with software running on a personal computer (for example, Flash applications, Processing, MaxMSP). Any Arduino card can be agregate manually or you can buy a finished device; the development environment for programming such a board has an open source code and it is completely free. The programming language of Arduino is the implementation of the similar hardware platform "Wiring", which based on the medium of programming of multimedia called "Processing".
The proposed system consists of a temperature and humidity sensor, for example, we use the DHT11 sensor module. The program code for its connection can look like this:

```
#include <dht.h>
DHT sensor = DHT();
void setup()
{
    Serial.begin(9600);
    sensor.attach(A0);
    delay(1000);
}
void loop()
{
    sensor.update();
    switch (sensor.getLastError())
    {
        case DHT_ERROR_OK:
            char msg[128];
            sprintf(msg, "Temperature = %dC, Humidity = %d%%",
                    sensor.getTemperatureInt(),
sensor.getHumidityInt());
            Serial.println(msg);
            break;
        case DHT_ERROR_START_FAILED_1:
            Serial.println("Error: start failed (stage 1)");
            break;
        case DHT_ERROR_START_FAILED_2:
```

```
            Serial.println("Error: start failed (stage 2)");
            break;
        case DHT_ERROR_READ_TIMEOUT:
            Serial.println("Error: read timeout");
            break;
        case DHT_ERROR_CHECKSUM_FAILURE:
            Serial.println("Error: checksum error");
            break;
    }
    delay(2000);
}
```

The MQ135 air quality sensor module is designed to determine the content and the amount of harmful and dangerous gases in the air such as: NH3, NOx, alcohol vapor, gas, smoke, CO2, etc. The program code for connecting it can look like this:

```
const int analogSignal = A0;
const int digitalSignal = 8;
boolean noGas;
int gasValue = 0;
void setup() {
    pinMode(digitalSignal, INPUT);
    Serial.begin(9600);
}
void loop() {
    noGas = digitalRead(digitalSignal);
    gasValue = analogRead(analogSignal);
    Serial.print("There is ");
    if (noGas) Serial.print("no gas");
    else Serial.print("gas");
    Serial.print(", the gas value is ");
    Serial.println(gasValue);
    delay(1000);
}
```

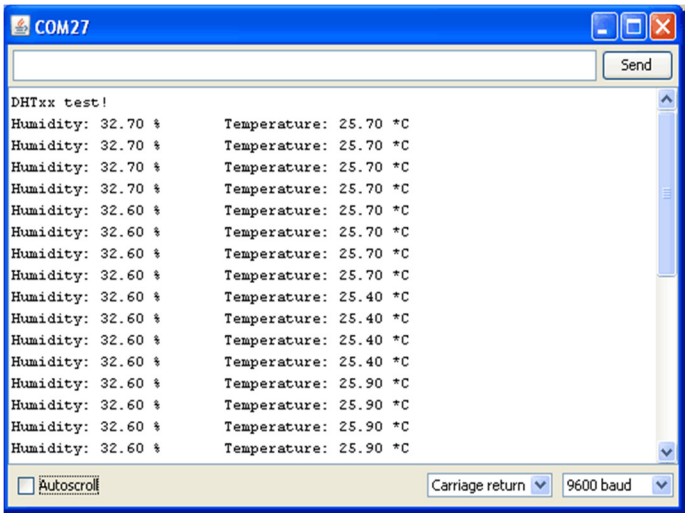


Figure 8. Results of measurements

Indicators from sensors can be transmitted to the main control panel using LoRAWAN wireless technology, 6LoWPAN, Z-Wave, ZigBee, Wi-Fi, BLE4.2, etc. Results of measurements of temperature, humidity and other indicators can be displayed on the screen (Fig 8).

6. Conclusion

The Internet of Things (IoT) is set to occupy a substantial component of the future Internet. The IoT connects sensors and devices that record physical observations to applications and services of the Internet. As the IoT represents the future state of the Internet, an intelligent and scalable architecture is required to provide connectivity between these silos, enabling discovery of physical sensors and interpretation of messages between the things. Therefore, an IoT architecture was developed to detect air pollution.

For this task, comparison of the main wireless technologies for the concept of the Internet of Things has been carried out. The main objectives of technologies - different, different architecture and opportunities. Therefore, for the application, it is necessary to approach the choice of optimum technology thoroughly, and objectively, to weigh all advantages and disadvantages of each of them.

In addition, the HTTP and MQTT protocols which are applied in this architecture were analyzed.

Finally, the practical implementation of the IoT network architecture based on the Arduino software and hardware platform was proposed. Program codes for different modules which measure quality of air have been written.

REFERENCES

1. ALEKSANDER M.B., KORCHENKO A.G., KARPINSKY N., ODARCHENKO R.: Research of vulnerabilities of touch subnets of architecture of the Internet of things to various types of the attacks, Information security: 22(2016)1, 12-19.
2. KORCHENKO A.G., ALEKSANDR M.B., ODARCHENKO R., NAJI A.A.A., PETRENKO O.Y.: The analysis of threats and mechanisms of ensuring information security in touch networks, Information security: 18(2016)1, 48-56.
3. ROSLYAKOV A.V., VANYASHIN S.V., YU A.: Internet of things: manual. Grebeshkov. Samara: PGUTI, 2015. – 200 pages.
4. What network "will catch" things? - [An electronic resource] - electronic text data the access mode: http://mikrotik.kpi.ua/index.php/courses-list/iot/104-what-network-catch-things.
5. Understanding IoT Protocols – Matching your Requirements to the Right Option. – [An electronic resource] - electronic text data the access mode: https://solace.com/blog/use-cases/understanding-iot-protocols-matching-requirements-right-option?utm_source=adroll&utm_medium=cpc&utm_campaign=iot&utm_content=25.

6. Recommendation Y.2060 – [An electronic resource] - electronic text data the access mode: https://www.itu.int/rec/T-REC-Y.2060-201206-I
7. ODARCHENKO R.; ABAKUMOVA A.; TKALICH O.; USTINOV O.: LTE and wireless sensor networks integration in the concept of ”Smart Home”, 2016 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC) 35-38.
8. Bluetooth Protocols Technology – [An electronic resource] - electronic text data the access mode: http://wiki.kspu.kr.ua/index.php/Технологія_Bluetooth
9. Modern telecommunications: networks, technology, security, economy, regulation. - Edition 2. - By general ed. Long SO - K .: "Azimuth-Ukraine". - 2013. – 608 p.
10. How to choose the standard of communication for IoT network - [An electronic resource] - electronic text data the access mode: http://savepearlharbor.com/?m=201602&paged=59.
11. FAHIER N., FANG W.C.: An Advanced Plug-and-Play Network Architecture for Wireless Body Area Network Using HBC, ZigBee and NFC, IEEE International Conference on Consumer Electronics (ICCE). 2014. 165–166.
12. KIRICHEK R. V., PARAMONOV A. I., PROKOPYEV A. V., KUCHERIAVI A. E.: Evolution of researches in the field of wireless touch networks, Information technologies and telecommunications. 4(2014)8. 29-41. URL: http://www.sut.ru/doci/nauka/review/4-14.pdf.
13. ODARCHENKO R.S.: The concept of a sensor network of collection of meteorological data for the system of regulation of emitting power of the radio-transmitting devices of cellular networks//Problems of creation, test, application and maintenance of difficult information systems//Issues 8. - 2013. 53-61.
14. Architecture of LoRaWAN of networks. – [An electronic resource] - electronic text data the access mode: http://lorawan.lace.io/lorawan-networks/
15. We reveal secrets 6LoWPAN, electronics news – №. 11 – 2015. – Page 30-36.
16. What is LoRa? – [An electronic resource] - electronic text data the access mode: http://lorawan.lace.io/faqs/lora/.
17. GEPKO I.A., OLEYNIK V.F., CHAIKA YU.D., BONDARENKO A.V.: Modern wireless networks: state and prospects of a development, navchalniya pos_bnik. – Kiev, 2009 – 672 pages.
18. Understanding The Protocols Behind The Internet Of Things – [An electronic resource] - electronic text data the access mode: http://www.electronicdesign.com/iot/understanding-protocols-behind-internet-things.
19. MQTT, CoAP, IoT Protocols – [An electronic resource] - electronic text data the access mode: https://eclipse.org/community/eclipse_newsletter/2014/february/article2.php
20. IoT - modern telecommunication technologies– [An electronic resource] - electronic text data the access mode: http://www.lessons-tva.info/articles/net/013.html.

Yuliana GRUZDIEVA[1]

Scientific Supervisor: Ivan TYSHYK[2]

ZASTOSOWANIE SYGNAŁÓW NIESTACJONARNYCH W SYSTEMACH OCHRONY SYGNALIZACJI

Streszczenie: W celu zwiększenia odporności na zakłócenia w systemach alarmowych fal radiowych, proponuje się użycie zarówno próbkowania niestacjonarnych sygnałów ciągłych z dalszym przetwarzaniem danych próbkowania oraz analizę częstotliwościowo-czasową odbitego sygnału (wavelet). Przedstawiono i przeanalizowano wyniki symulacji przetwarzania wspomnianych sygnałów poprzez transformatę falkową.

Słowa kluczowe: Efekt Dopplera, sygnał niestacjonarny, systemy alarmowe, lokalizacje częstotliwościowe, transformacja falkowa

APPLICATION OF NON-STATIONARY SIGNALS IN PROTECTIVE SYSTEMS OF SIGNALIZATION

Summary: In order to increase the noise immunity of radio wave alarm systems, it is proposed to use both probing non-stationary continuous-signal signals with the further processing of the probing and reflected signals in the time-frequency (wavelet) region. The results of simulation of the processing of mentioned signals by wavelet transform are presented and analyzed.

Keywords: Doppler Effect, Non-stationary signal, security alarm systems, Frequency-time localization, wavelet transform

1. Formulation of the problem

The use of Doppler radio-wave devices in alarm systems has certain features related to the fact that, when the object is approaching in a controlled area of space in the direction of the receiver, the width of the spectrum of the reflected signals due to the presence of the Doppler effect increases with respect to the width of the spectrum of the radiation signal. In this case, the operation of such devices occurs, as a rule, under the influence of external and internal noise, whose energetic spectrum is mainly linear,

[1]National University «L'viv Polytechnic», Department of Cyber security, specialty: Information security management, night2505@gmail.com
[2]National University «L'viv Polytechnic», Department of Cyber security, PhD, tyshyk_iy@polynet.lviv.ua

narrowband and located in the frequency domain of useful signals. This leads to a decrease in the stability of radio wave alarm systems to false alarms. The increase to some extent of their resistance to false alarms by known methods occurs, usually, by reducing the reliability of the detection of moving objects in a controlled area of space. In such radio-wave devices of security systems, when selecting the appropriate probing signal, are looking for a compromise between these parameters. However, the qualitative characteristics of the mentioned systems, obtained as a result of the chosen compromise, often do not satisfy the consumer.

2. Analysis of recent research and publications

The advantage of Doppler radio waves narrowband systems is to provide them with high selectivity of reflected signals. At the same time, a small number of signs of a useful signal, by which can be carried out selection of the reflected signal against the background of noise, leads to a deterioration of the probabilistic characteristics of the detection of a moving object, and the presence in the selection of noise, in addition, significantly reduces the stability of such systems to the false alarms. To improve the noise immunity of the mentioned systems to some extent is possible by the way of increasing the lower boundary of Doppler frequency selection by their filter systems, however, this may lead to exceeding the permissible limit of registration of the minimum speed of objects. Finding a compromise between the specified parameters, when choosing the appropriate type of probing signal, is an important problem for such narrowband radio-wave devices of security systems. At present, when using by radio-waves devices harmonic probing signals of continuous type, their range of registration of object velocities is in the range of $0{,}1 \div 1$ m/s, which is unacceptable for many applications [1].

Modern radio-waves devices of the security systems for detecting the movement of objects widely use probing broadband signals, since such signals allow obtaining the required resolution for distances, provide higher noise immunity and accuracy of detection for narrowband location signals. Such security broadband systems mainly use the following types of broadband probing signals: short-range radio and video pulses; non-stationary signals of a continuous type of long duration [2, 4].

Broadband short-range pulse signals of location can significantly increase the resolution and accuracy of measuring the distance to the object of observation, reduce the "dead zone" of the system, increase its resistance to the effects of all types of passive interferences and simplify the observation of moving objects against the background of powerful reflections from stationary objects. However, their use as a probing in the devices of security systems has its disadvantages, which are associated with the low energy of such signals, difficulties in their generation, radiation and processing [2- 5].

The advantage of probing non-stationary continuous-type signals with respect to short-term pulsed signals is the relative simplicity of their generation and radiation, as well as the possibility of obtaining the required energy. For the selection of the informative parameter of such reflected signals, the correlation-filter methods are used for their processing and processing on the basis of the time-frequency filter, whose work is based on the discrete window of the Fourier transform However, the effectiveness of these methods is lost when the high requirements for linear deviation

of the frequency of mentioned emitted signals are not fulfilled and the impossibility of performing on the basis of the provided methods the qualitative and accuracy estimation of reflected signals of such signals within their wide frequency range, which leads to deterioration of the reliability of detection of moving objects by the appropriate radio-waves systems [5, 6].

3. Problem definition

The purpose of the work is to explore the possibility of using a wavelet transform to process probed non-stationary signals of radio-wave security alarm systems, which should lead to an improvement in the stability of such systems to false alarms, while ensuring their high reliability of detection of moving objects.

4. Statement of the main material

It is known that wavelet transformation is particularly advantageous to use in cases where the result of an analysis of a some signal should include not only a simple list of its characteristic frequencies, but also information about the local coordinates in which these frequencies are manifested. The wavelet transformation is a very convenient tool for the adequate representation of signals with localized frequencies, since the elements of its basis are well localized and have a moving time-frequency window. Due to the constant change in the size of the window, the wavelet transform can provide a proportional resolution in each frequency band, which allows creation of windows with constant fractal resolution of bandwidth, which makes it possible to analyze and compare broadband signals. Thus, the analysis and processing of broadband non-stationary signals in time is the main field of application of wavelet transformation. In the conditions of non-stationary signals and the presence in them the background of a forced noise of wavelet functions are the most suitable basis for solving the problem of efficient filtration of such signals, in conjunction with good time localization of their features. In view of the above, in the work are proposed to process probed broadband non-stationary signals of radio-wave alarm systems by wavelet transformation, which should increase the noise immunity of such systems and the reliability of their detection of moving objects against the background of interferences [7, 8].

In this case, the radiated (standard) signal $s(t)$ is presented as a complex signal, the carrier frequency of which consists of two sine waves of different frequencies, and the reconfiguration from one carrier frequency to another is jump-like. The mathematical model of such a signal has the form:

$$s(t) = \begin{cases} S_0 \cos(\omega t + \varphi) \ \ by \ \ t \le t_e \\ S_0 \cos(k\omega t + \varphi) \ \text{by} \ \ t > t_e \end{cases}, \tag{1}$$

where S_0 – the value of the amplitude of the emitted signal; ω and φ – the frequency of its carrier and the initial phase, respectively; t_e – region of a jump-like change in the carrier signal frequency; k – scale factor of frequency.

Reflected from a moving object, the signal is modeled as a delayed and noisy version of the emitted signal $s'(t)$. The case when the object reflecting the signal approaches the observer frontally is considered, therefore, the region of the jump-like change in the frequency of the carrier of each next reflected signal will shift in time relative to the region of the jump-like change in the carrier frequency of its previous version by a value proportional to the distance passed by the object of location for a certain period of time (Fig. 1*a*, *b*).

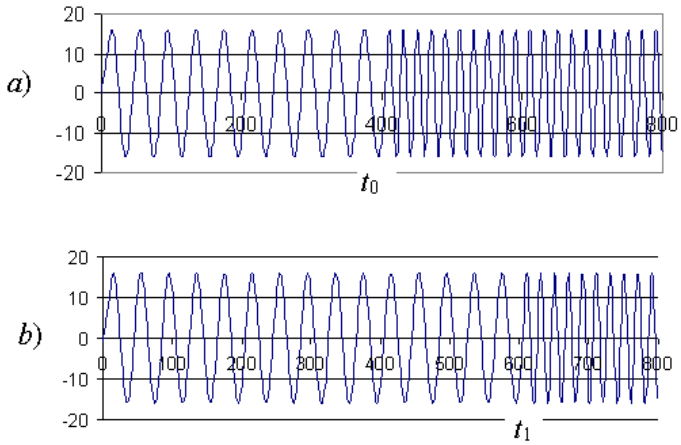


Figure 1. Time-domain representation of reflected broadband signals:
a) – at a certain moment of the spatial position of the object of the location;
b) – at the next moment of the spatial position of the object of the location.

The realization of the wavelet transform method is based on the operation of the decomposition of the reflected signal on sub-bands using one of the known algorithms, which provides sub-coding of the discrete sequences of this signal. According to the theory of wavelet transformation, scale and wavelet-functions are considered as functions of filters, which are derived from the conditions of multiple-scale analysis. The decomposition on wavelet-component sequences of discrete values of received signal $s'[k]$ is due to the operation of convolution of its values with filter functions [8, 10]:

$$d_{j,n} = \sum_k s'[k] h_j \left[k - 2^j n \right] \tag{2}$$

$$c_{J,n} = \sum_k s'[k] g_J \left[k - 2^J n \right] \tag{3}$$

where $h_j \left[k - 2^j n \right]$ and $g_J \left[k - 2^J n \right]$ – analyzing the discrete wavelet and scale functions respectively; k – sample number; $d_{j,n}$, $c_{J,n}$ – sequences of detailed and approximated wavelet coefficients of decomposition of received signal, obtained at different levels of transformation j, (j = 1,2,3,..., J; n = 1,2,3,..., 2^j).
Thus, according to (2) and (3), the received signal at the j-level of conversion will be represented by the corresponding set of wavelet coefficients.

The concept of filtration based on the wavelet transform method consists in thresholds of the noise quantities of the detailed wavelet coefficients, which are mainly localized on high-frequency sub-bands of decomposition. By setting a certain threshold for a particular level of decomposition and rejecting detailed coefficients on it, it is possible to reduce the noise level in the received signal in one way or another.

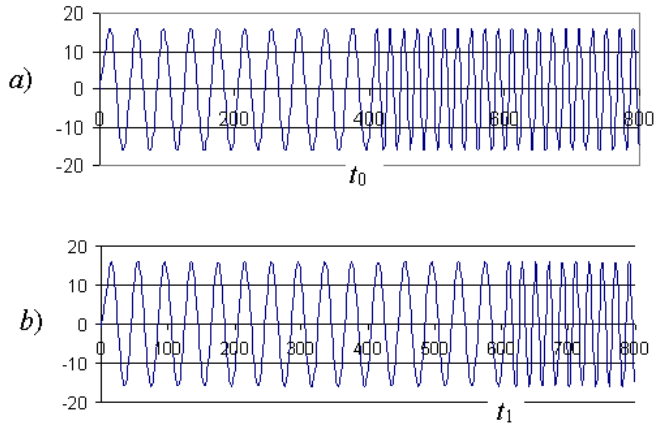


Figure 2. Time-domain representation of reflected broadband signals:
a) – at a certain moment of the spatial position of the object of the location;
b) – at the next moment of the spatial position of the object of the location.

To implement the filtering $s(t)$ an adaptive threshold was used. At each decomposition level, the value of the threshold λ_j was set, the value of which was calculated on the basis of a "universal" criterion with ignoring noise estimation as [8]:

$$\lambda_j = \text{sqrt}\left(2 * \log\left(length\ \left(N_j\right)\right)\right), \tag{4}$$

where N_j – the length of the wavelet coefficients of the input sequence on the j-th level of the decomposition, which is calculated from the total length N:

$$N_j = \frac{N}{2^j} \tag{5}$$

The values of those wavelet coefficients will be informative $d_{j,n}^{S}$, which will satisfy the following condition:

$$d_{j,n}^{s} = \begin{cases} d_{j,n}^{s} & by \quad d_{j,n}^{s'} > \lambda_j \\ 0 & by \quad d_{j,n}^{s'} \le \lambda_j \end{cases} \tag{6}$$

The resulting value of the wavelet components of the reflected signal is given as:

$$d_n^s = \sum_{j=1}^{J} d_{j,n}^{s} \tag{7}$$

The main feature of such a representation is that it enables to effectively filter the noise of the received signal and to carry out the time localization of a set of its weighty wavelet components, which significantly improves the signal-to-noise ratio at the output of such a filter system and allows to evaluate the distance to the object of observation directly in time-frequency domain.

The simulation of the processing of received location signals was carried out using the MATLAB application package. Accepted (simulated) signals were broadband non-stationary signals with a jump-like change in carrier frequency (Fig. 1a, b). In fig. 1a such a signal reflected from the object at some point in time is shown that was located in a certain place of the controlled area. The region of the jump-like change in the carrier frequency of this signal corresponds to a certain time point t0. In fig. 1*b* the next signal reflected from the object of location at another time is shown, during which there was a change in the position of the object. The region of the jump-like change in the carrier frequency of this signal corresponds to the time point *t*1. On the reflected signals, in addition, Gaussian noise was superimposed, the value of which was 30% of the signal level. According to the theory of radiolocation, the magnitude of the displacement of the region of the jump-like change in the frequency of the reflected signal relative to a certain support will depend proportionally on the magnitude of the movement of the object of location in the controlled zone.

From the theory of wavelet transformation, [7-9] is known that the result of wavelet analysis of some signal contains not only a simple list of its characteristic frequencies (scales), but also information about certain local coordinates, at which these frequencies themselves manifested. Thus, the wavelet coefficients of the decomposition of the above-mentioned location signals can provide information about the location in time in certain peculiarities of these signals. Such features, in this case, are the region of a jump-like change in the frequency of these signals.

In fig. 2*a*, *b* the result of the processing of mentioned signals by wavelet transformation is shown. The set of wavelet coefficients of the decomposition of the signal given in Fig. 1a is shown in Fig. 2*a*. According to the above theory, the total value of the weighted wavelet coefficients, in this case, grouped at the point n0, which corresponds to the time domain t0 jump-like change in the frequency of the reflected signal. Similarly, the set of wavelet coefficients of the decomposition of the next reflected signal, given in Fig. 1*b*, is shown in Fig. 2*b*. Total value, in this case, grouped at the point *n*1, which corresponds to the time domain *t*1 jump-like change in the frequency of the reflected signal.

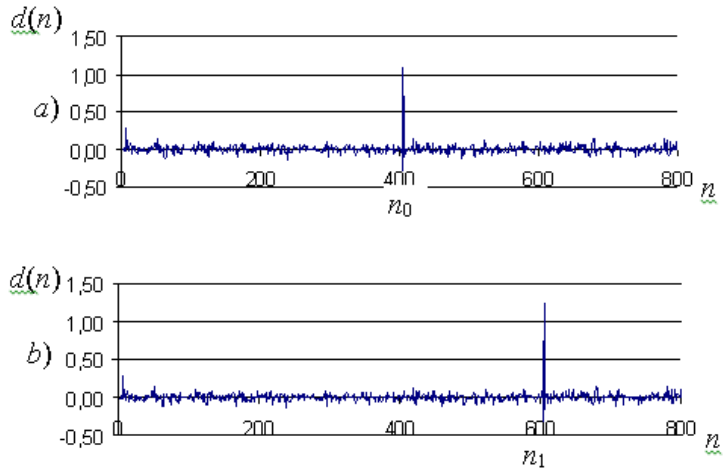


Figure 3. The values of the resulting wavelet-components of decomposition of reflected location signals:
a) – wavelet-components of decompositions of the non-stationary signal, is given in Fig. 1a;
b) – wavelet-components of decompositions of the non-stationary signal, is given in Fig. 1b

From Fig. 2*a*, *b* it is obvious that the signal-to-noise ratio (the ratio of the informative components of the decomposed signals to the residual noise components) is quite significant, which makes it possible to clearly identify the change in the position of the object of the location in the controlled area of space by the difference in time values of the weighty wavelet coefficients of the reflected signals, without the implementation of backward wavelet transformation.

The processing of received signals in the wavelet domain was carried out using the base wavelet functions of Daubechies 4th order and the pyramidal algorithm of Mallat. The number of frequency sub-bands of the decomposition j of the input signal (for N = 1024) was 10. Processing of incoming sequences s'[k] by wavelet transformation took place in two stages: at the first stage, the decomposition was made for the frequency sub-bands of the input signal components, which represented the values of wavelet coefficients (2, 3). According to (4), the magnitudes λj were calculated for these sub-bands; on the second one - the values of wavelet coefficients were estimated on the corresponding decomposition according to (6) throughout the processing interval and the final wavelet coefficients according to (7) were summed up, taking into account their difference of time offsets for each sub-band of the decomposition.

The results of the conducted simulation show that the processing of the proposed non-stationary broadband signals of location by wavelet transforms allows to combine their effective filtration from noise and the time localization of the informative components that represent the region of change in the frequencies of these signals and serve the informative parameters upon which the conclusion is drawn regarding the movement of objects of location.

The advantage of this representation is that it is possible to estimate the distance to a moving object of observation at each subsequent time which allows with high

reliability to detect its movement against the background of stationary objects and to ensure the high stability of the corresponding location system to false alarms.

5. Conclusions

1. It is shown that for modeling of probed non-stationary signals of the mentioned type radio-wave devices of security alarm systems it is expedient to use models based on wavelet transformation. Such representations make it possible to solve the problems of efficient filtration of these signals, in conjunction with good time localization of their features, which represent the region of change of their frequencies and serve the informative parameters, which leads to the conclusion about the movement of objects of location directly in the time-frequency region.
2. The results of the simulation show the expediency of using wavelet transform technology to process the mentioned location signals, which allows improving the informative of the appropriate radio-wave alarm systems and increase their resistance to false alarms.

LITERATURA

1. LEVCHUK S.A., MAKAROV S.B., PETROV A.YU.: Doppler radio-wave detectors for security alarm systems. Problems of Information Security, 1(2000).
2. VOLKOV A.: Ultrasonic security alarm sensor.Radio 5(1996), 54-56.
3. IMMOREEV I.YA. Ultra-wideband radars: new opportunities, unusual problems, system features, Bulletin of the MSTU, №4, 1998, - P. 25-56.
4. SUDAKOV A.A.: Signals used in UWB radio systems, High technology, April 2005.
5. SIMON M.K.: Spread spectrum communication handbook. New York: McGraw n Hill 1994, 7-9.
6. ALY O.A.M., OMAR A.S.: Detection and localization of RF-radar pulses in noise environments using wavelet packet transform and higher order statistics. Progress In Electromagnetics Research, PIER 58,301 –317, 2006.
7. NAKONECHNY A.Y.: Theory of short-wave transformation and its application. Lviv. “Phoenix”, 2001. P.93.
8. DYAKONOV V.P.: Wavelets. From theory to practice. SOLON-R. Moscow 2002.
9. VOROBOV V.I., GRIBUNIN V.G.: Theory and practice of wavelet transform. VUS, 1999.

Mariya GRYGORAK[1], Tamara OLESHKO[2], Tetiana KUZNETSOVA[3]

MODELOWANIE 3D W TECHNOLOGII INFORMACYJNEJ DLA PRZEDSIĘBIORSTWA PRZEWOZÓW LOTNICZYCH

Streszczenie: W przemyśle lotniczym ważnym problem z zakresu podejmowania decyzji przez dyrekcję jest dążenia do całkowitego wyeliminowania katastrof samolotowych, a tym samym ochrona życia ludzkiego. Autorzy proponują nowe koncepcje w tym zakresie, a także podkreślają znaczenie modelowania 3D w technologiach informacyjnych dla przedsiębiorstwa przewozów lotniczych. Ponadto, Autorzy przeprowadzili modelowanie 3D podejmowania decyzji z zakresu problemów związanych z obszarem przestrzeni powietrznej. Proponowany system może być rozpatrywany jako symulator w małej skali tzw. Flexible Time Scale (FTS) system. W rozważaniach określono środki modelowania matematycznego podejmowania decyzji zarządczych w problemach przemysłu lotniczego. Mianowicie, określono (opracowano) moduły oprogramowanie, które spełniają wymagania projektantów systemów lotniczych. W systemie zastosowano dokładny opis/rejestr przylotów oraz wylotów (i odpowiednich tras) za pomocą specjalnych segmentów. Tę samą metodę zaproponowano do opisu stref postoju. Dzięki temu systemowi jest możliwe nie tylko przechowywanie i integracja danych, ale także odzwierciedlenie operacji na obiektach poprzez modele 3D (trójwymiarowe).

Słowa kluczowe: modelowanie; model 3D (trójwymiarowy); technologie informacyjne; przedsiębiorstwo

3D-MODELING IN INFORMATION TECHNOLOGY OF AIR ENTERPRISES

Summary: In the aviation industry the problem of managerial decision-making completely eliminating aircraft crashes and save human lives is particularly acute issue. The authors introduced new concepts and determined the role of 3D-modelling in information technology of air enterprises from their point of view. And also the authors made 3D-modeling of decision-making in airspace issues, which can be regarded as a small-scale version of the simulation in Flexible Time Scale (FTS) using information technology. The study defined means of mathematical modeling of managerial decision-making in air industry problems, consisting of software modules that are used to meet the needs of airspace designers. Modeling tools typically do not use curved segments. The exact description of the arrival and departure routes defined by curved segments could be determined using approximated linear model segments. The same method is proposed to use for describing standby areas. Through their use one can not only store and integrate data, but also reflect the process of objects' operation on 3D-models.

Keywords: modeling; 3D-model; information; technology; enterprise

[1] National Aviation University, 1 Komarov Ave, Kyiv, 03680, Ukraine
[2] National Aviation University, 1 Komarov Ave, Kyiv, 03680, Ukraine, ti_oleshko@ukr.net
[3] National Aviation University, 1 Komarov Ave, Kyiv, 03680, Ukraine

1. INTRODUCTION

Successful engineers are trying to "program" as many as possible product decisions to improve the effectiveness. In aviation industry the issue of making product decisions completely eliminating air crashes and saving lives is extremely acute. This is why the authors have considered new engineering concepts and from their point the role of 3D-modeling in information technology of airlines has been determined..

2. Overview of the known solutions

The lack of theoretical positions and practical recommendations for 3D-modeling in information technologies of air enterprises determines the particular urgency of the problem.

3. The purpose of the article

The aim of this research is the consideration of 3D-modeling in information technology of air enterprises.

4. Results

Complication of production tasks needs informative support as at strategic so at the operative planning. Especially it shows up in the conditions of instabilities of supplying with acquisition, absence of necessary financial resources and other.

In the process of operative management operations a master, technologist, controller, et cetera, that, it will be a person which makes decision (PMD) at workshop level to settle the semi structured problems. And here the programmatic modules of information technology (IT) of support of processes of management in computerintegrated automated system (CIAS) of the production setting have an important value as: Computer-Aided Design (CAD)/Computer-Aided Manufacturing (CAM)/Computer-Aided Engineering (CAE)/Computer-Aided Process Planning (CAPP)/Product Data Management (PDM)/Enterprise Resource Planning System (ERP)/Manufacturing Enterprise Solutions (MES)-systems and other, what will allow to prompt the variants of operating on the process of production for achievement of the put aims and the end-point are possible. The modules of the programs of IT of support of management processes in CIAS must be in a position to adapt oneself to the change of calculable models, «socialize» with an user on specific for the guided area a «language», present results in such form which would be instrumental in more deep understanding of results [6]. That, the function of IT of support of management processes consists not in that, to replace a leader, but in that, to promote his efficiency. IT of support of management processes must support intuition, able to recognize ambiguity and incompleteness of information, and have facilities for their overcoming.

Some success of the air company in all areas, is a function of "information and choice" like a Marxist production function of "labor and capital". In addition, "work" and "choice" refers to the action, and "capital" and "information" – to cash assets (tangible and intangible). Moreover, undesirability to make a choice is a choice too as a category is always present (in space, time and intellectual-intuitive activity), i.e. completely independent [4].

For efficient activity of air enterprises functional dependence

$$EIT = f(V_1, V_2, C) \tag{1}$$

for 3D-modeling has been defined by the authors, where:

EIT – effectiveness of information technologies,
V_1 – volume of information,
V_2 – velocity, time, period,
C – choice,

In this context information is seen as the knowledge base (conscious) and spiritual base (super conscious), namely ideas, intuition, "tips from the top", gift, charisma, etc.

The authors have developed three-factor (factor 1 – volume of information, factor 2 – velocity, factor 3 – choice) forecasting model of impact on the effectiveness of information technologies in air enterprises:

$$EIT = \lim_{V_2 \to 0} ((1-\alpha) \cdot V_1 + C) \tag{2}$$

Herein α is a coefficient of information obsolescence risk [2].

It is important to take into account that V_1 and *C* are not numbers (rational and/or irrational), but integrated software (individual specific) systems (matrixes), within which the signals and symbols of the internal and external environment are transformed into knowledge, thoughts, ideas, intuition, etc.

Naturally, wrong, i.e. negative, choice decreases (weakens) *EIT* and right (positive) one increases (enhances) it.

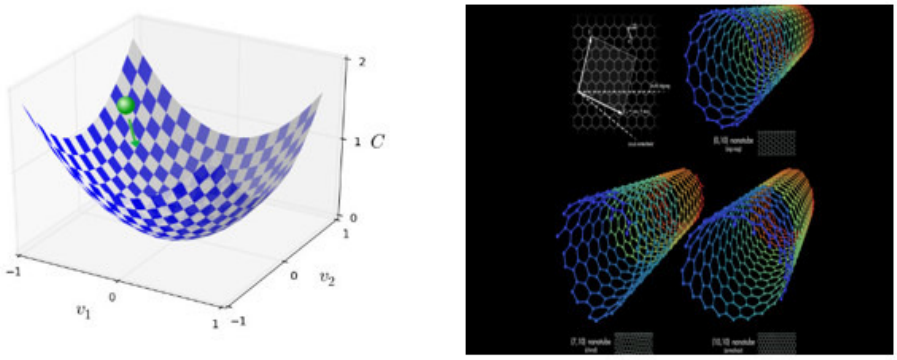


Figure 1. The 3D-model of functional dependency EIT = f (V_1, V_2, C) for efficient activity of air enterprises

Note: obtained by the authors

With the implementation of 3D-modeling in information technologies have obtained a synergistic effect in the activity of air enterprises.

Figure 2. The trajectories of the information in the 3D-model of functional dependency EIT = f (V_1, V_2, C)

Note: obtained by the authors

Nowadays, within the current technological and informational economics, consistent patterns of technology of airlines creation are closely associated with forecasting and 3D-modeling.

By the authors, which can be regarded as a small-scale version of the modeling in Flexible Time Scale (FTS). The main objective of this mathematical modeling was to create appropriate routes and structural elements (sectors), as well as analysis of their interaction with different models of air traffic. Mathematical modeling tools should generate 2D-trajectory (location + time) according to flight plans describing the patterns of air movement in the relevant airspace organization. These trajectories and airspace structure elements are used for calculating statistical data such as: load on sectors, load on the route segments, conflicts and so on. Airspace mathematical modeling by information technologies of air enterprises allows obtaining accurate data concerning load on the sector and its capacity [8].

Mathematical modeling tools of information technologies in airspace problems have been identified by the authors, which consist of software modules used to meet the needs of airspace designers:

- graphical tools used to determine airspace organization and its visualization in 2D and 3D;
- tools for manipulating trajectories used to determine the air traffic model (distribution of air traffic, time for control transmission, 4D-generation of trajectories);
- data analysis and processing tools (distribution of air traffic, load on the sectors, checking for conflicts).

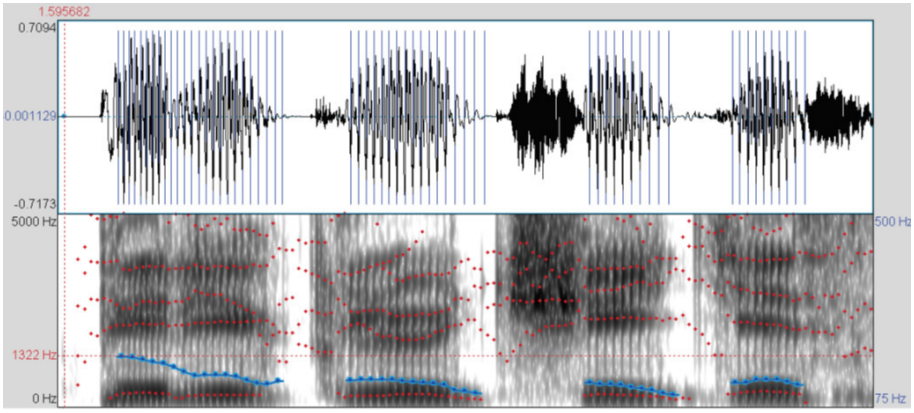


Figure 3. 4D-generation of trajectories of information technologies

Note: obtained by the authors

The first step of decision-making mathematical modeling in airspace problems is to convert design project, processed by a airspace design working group, in a simplified for presentation version based on computer technologies. In most cases, the routes are described as 2D-network of segments. These segments may have some peculiarities related to air traffic: direction and type of movement.

Modeling tools typically do not use curved segments. The exact description of the arrival and departure routes defined by curved segments can be determined using approximated linear model segments. The same method can be used to describe the standby areas. Sectors represent the block of airspace, defined by horizontal and vertical contours.

The horizontal configuration of sectors is described as a closed polygon. If the horizontal configuration of sectors is defined by the curved segments, it can be described by approximated linear segments. If the sector has a complex vertical configuration, it should be divided into component parts, i.e. basic geometrical blocks to be connected together for the purpose of analysis [10].

After completing the simulation by technology of air enterprises, the designer should verify configuration of sectors on the image correctness and absence of holes between sectors in horizontal and vertical plans.

Approval or rejection of each design project can not be based only on the results of quantitative modeling data in flexible time scale without considering prospects of development of air enterprises.

Both aircrew of air force and civil aviation are its users. The purpose of the development is to construct not only the flight path of the aircraft (aircraft groups), followed by visualization in three-dimensional computer space, but to have opportunity to take the most effective solution in the case of unforeseen situation (including catastrophic), using initial conditions via the information base, appropriate choice of the necessary data from this database and possible time (Fig. 4).

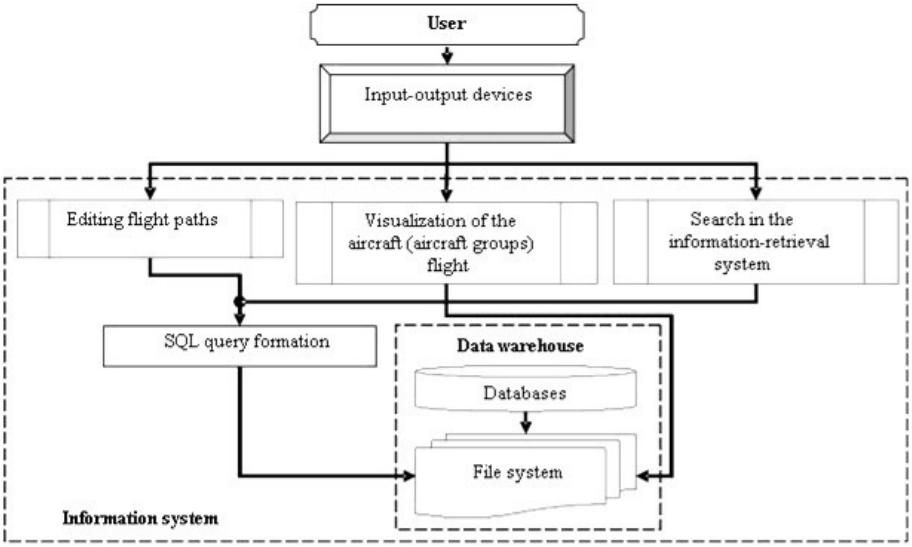


Figure 4: Conceptual model of information technologies of air enterprises

Source: created by the authors

Subsystem "Data warehouse" consists of two elements: "Databases" and "File system".

Subsystem "Editing flight paths" consists of blocks "Territory" – adding and change a file of 3D-underlying surface model; "Aircraft" – adding, deleting of file 3D-aircraft model; "Battle array" – creation of new or editing existing fighting system a of aircraft group; "Flight path" – creation of new or editing existing flight path of the aircraft (aircraft groups).

Figure 5. 3D-underlying surface model "Aircraft"

Note: obtained by the authors

Visualization of the aircraft (aircraft groups) flight subsystem is responsible not only for the formation of three-dimensional images on the screen, but also for the flight visualization. The division into subsystems ("Editing flight paths" and "Visualization of the aircraft (aircraft groups) flight") is caused by the fact that the time to run the program and visualize the flight should be minimal (a few seconds).

The user builds once the flight path, and then visualizes the flight of the aircraft (aircraft groups) on this trajectory at any time [5].

Subsystem of search in information retrieval system provides the user with full access to information systems resources: text, graphic, audio and video information, virtual review of three-dimensional air models. User creates a query as a row of descriptors and information retrieval system displays all the information – available, in databases, and one that responds to a user's query.

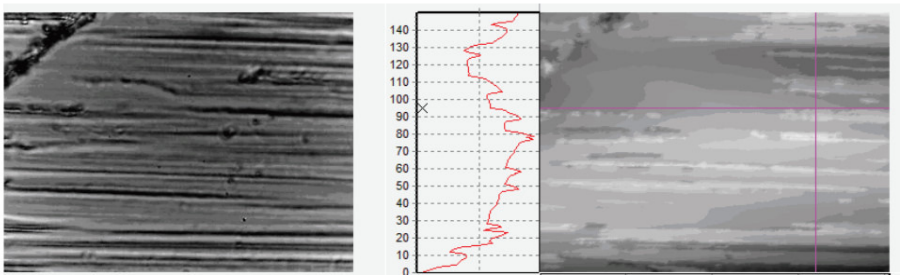


Figure 6. Visualization of the aircraft from 3D-model of rapid transmission of information

Note: obtained by the authors

Thus, information technologies of air enterprises allows modeling the flight path with execution of all shapes of aerobatics. Information system is universal as it provides the ability to add different resources. For example, a group of fighter aircraft or group of sports aircraft can participate in aerobatics. The main thing is to add appropriate three-dimensional models to the system and specify their characteristics.

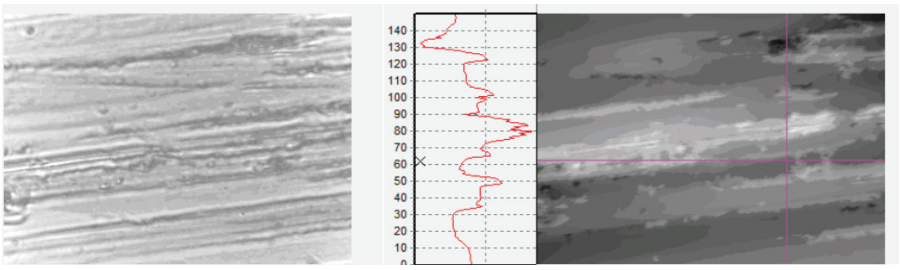


Figure 7. Visualization of the aircraft from 3D-model of slow transmission of information

Note: obtained by the authors

The advantage of the system is that it allows creating flight paths and then carrying out their review without fuel costs and costs of other resources. This system demonstrates how a group of aircraft could perform aerobatics shape under various conditions without pilots in aircraft piloting. Using the information system can avoid errors in aircraft group piloting and save pilots' lives. An analogue of such a system has not yet been identified by the authors.

The technology of three-dimensional image mapping C3-Technologies practically excludes manual work. To collect information airplanes equipped with high-level

digital SLR cameras are used. Four cameras located in the direction of sides of the world, film images of the earth surface at a certain angle. Other cameras, the quantity of which is not named, are located under precisely measured angles; they make shots of the surface in such quantity, which is enough to create three-dimensional models. The last operation is carried out via C3-software solution established by the experts that compares images, determining the depth of objects like the stereoscopic view mechanism of the human brain, and automatically creates highly detailed three-dimensional objects (10-20 cm resolution) [4].

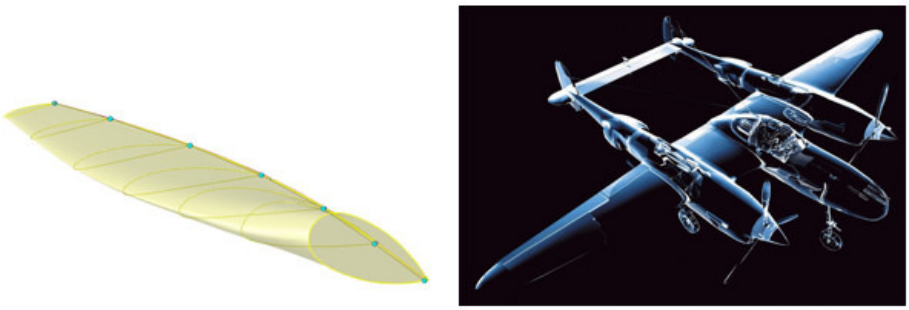


Figure 8. The C3-Technologies of three-dimensional image mapping

Note: created by authors

Except for the known requirements to the informative systems (powerful SMDB which provides effective access to information, their integrity and defense; developed analytical and calculable procedures which provide treatment and analysis of data; transportability, reliability, flexibility, possibility of including of new technological procedures), IT of support of management processes must have such specific lines, as:

- possibility of forming of variants of decisions in the special, unexpected for PMD situations;
- models, applied in the system, must be in a position to adapt oneself to concrete, specific reality as a result of dialog with an user;
- a subsystem must interactively generate models in the process of their exploitation [7].

Thus, developed of IT of support of management processes in management information is computer-integrated, that arose up as natural development and continuation of the administrative informative systems and control system by bases information, must decide the unstructured and semi structured multicriterion tasks.

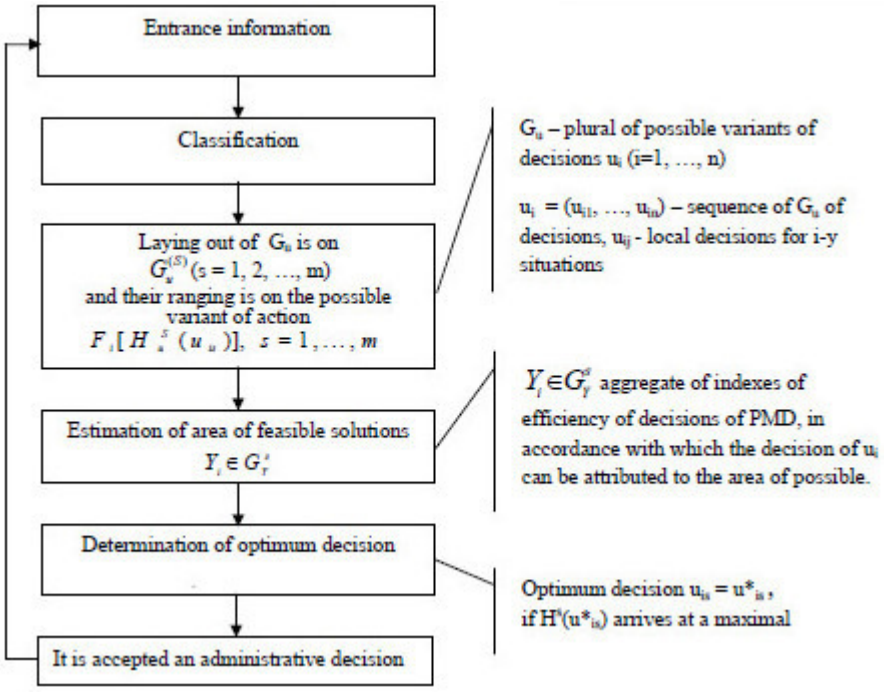


Figure 9. Formalized project procedure of information technologies for effective 3D-modeling

Note: created by authors

At the same time, one of major resources of air enterprise, the more so in the conditions of large nomenclature of wares of enterprise, knowledge is. The competitiveness of enterprise straight depends on organization of management these knowledge's. The wide use of information technologies is provided by the transfer of source of knowledge's from paper transmitters in electronic databases technological setting of industrial enterprises [3].

The process of getting of knowledge's focuses on application of algorithms of search of templates, that allows to attain higher results in comparing to traditional the methods of treatment of information [9].

Knowledge about technological processes is one of the most meaningful areas of knowledge's for a modern production. It has the manned analysis and interpretation the traditional method of transformation of information.

With introduction of IT of support of management processes in IAS on the enterprise of knowledge about the processes of management these bases accumulate in bases given can become the basic source of knowledge's. A receipt of information is about the process of management, based on the models of production process can be the effective mean of automation of processes of management of operations. A process of forming of knowledge's on the basis of getting (extraction) of information from a database IT of support of management processes in CIAS is specific is the process of application of specific algorithms for the receipt of information from bases given (DB).

All stages of this process (for example, preparation, selection, and revision of information, and also interpretation of results) are key for the receipt of complete knowledge's, got from DB.

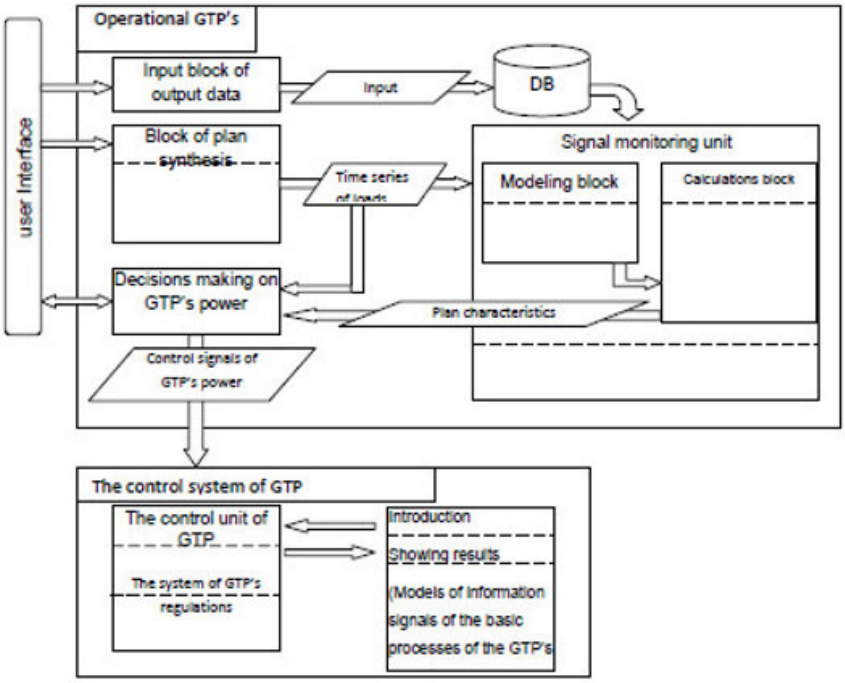


Figure 10. The architecture of information technologies for effective 3D-modeling

Note: created by authors

Realization of IT of support of management processes is in CIAS, in basis of which there are the guided models of the object-oriented platform, creation of upgradable universal and adaptive CIAS guarantees. Such CIAS can dynamically change the structure of presentation of this DB and source of data. In-use in such system a general informative model changes the object-oriented approach as design method. This approach is a fundamental mean for formalization of area of knowledge's and description of elements of informative model in style of human thought.

In traditional control system the base of knowledge's mainly contains information for making decision. At offered IT of support of management processes in CIAS a database contains the system informative model of the specialized additions – base of these administrative decisions also (fig. 11).

The modern generation of bases given is created mainly for support of business-additions. Success of language, applied in all modern SMDB, is based on the use of two-bit of simple elements, sufficient for description majority of production-additions.

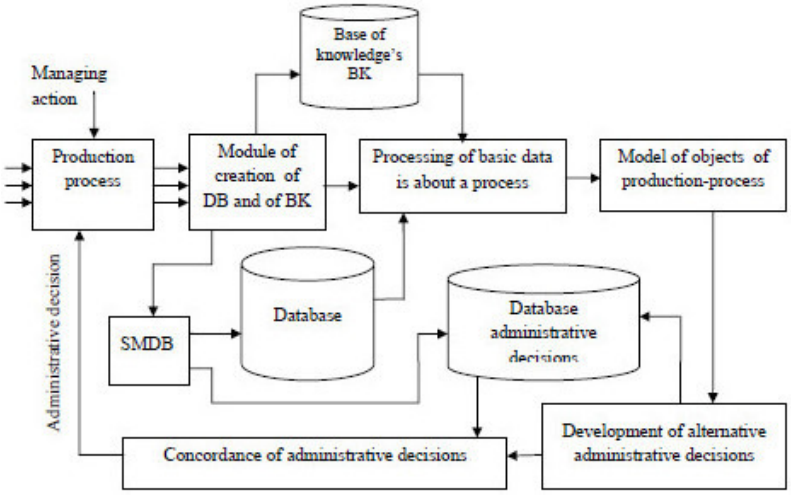


Figure 11. Structural functional chart of informative support of processes in the 3D-modeling. Note: created by authors

Unfortunately, to the set of these elements not enough for description of class of the systems, that appears, which (systems) work with knowledge's. The process of extraction of knowledge's from bases given must inherit basic principles which modern SMDB is based on, must be more concrete, what process of creation of queries. It links with that objects bases of knowledge's are more complex in comparing to the records in DB.

Thus, there is a necessity of creation of language, alike from SQL and intended for description of objects knowledge's. Such language must be semantically alike with the language of SQL and to be in a position to translate the elements of language in the elements of SQL for the language of knowledge's in DB.

5. Conclusions and suggestions

Currently, 3D-technologies evolve towards increasing the number of coordinates. These solutions are particularly in demand in the modeling design works in aviation industry. 4D-model is formed through consolidation of work by calendar network schedule of design works with corresponding elements of the design three-dimensional model, and thus includes 4 parameters: three spatial coordinates and time. 4D-model can be used both for virtual air building modeling and tracking actual progress of construction and installation air works.

REFERENCES

1. ГРИГОРАК М.Ю.: Формування професійних компетенцій менеджерів з логістики у віртуальних лабораторіях з використанням хмарних технологій Збірник наукових праць Державного економіко-технологічного університету: Серія „Економіка і управління". Вип.29. К.: ДЕТУТ, 2014. 123- 127.

2. ГРИГОРАК М.Ю. Теоретичні положення інтелектуально зорієнтованої логістики, Бізнес-інформ. 2(2015)445. 20-29. (Index Copernicus)
3. GRYGORAK M.Y., VARENKO Y. V.: Intelligent logistics systems, Proceeding: the seventh world congress „Aviation in the XXI-st century". Safety in Aviation and space technologies: 23-24 sept. 3(2014), 3.118-3.122.
4. Економічна кібернетика як наука. Проблеми та перспективи розвитку економічної кібернетики: Монографія, Уклад. Т.І. Олешко, Г.Ф. Іванченко, В.В. Гавриленко та ін. К.: ВД ТОВ „Agrar Media Group", 2013. 232
5. КУЗНЄЦОВА Т.В.: 3D-modelling in angstrommanagement technology of air enterprises, Т.В. Кузнєцова, Економічний часопис-XXI. 2014. Вип. 5-6. 234-245. (Index Scopus)
6. КУЗНЄЦОВА Т.В.: Angstrommanagement technology and 3d-modelling in of airlines, Т.В. Кузнєцова, VI Всесвітній конгрес „Авіація у XXI столітті" „Безпека в авіації та космічні технології" НАУ, 23-25 вересня 2014 року К., 2014. 27-30.
7. КУЗНЄЦОВА Т.В.: Нейромережева модель прогнозування впливу ризиків рейдерства на авіапідприємства, Т.В. Кузнєцова, Т.Г. Остапенко, АВІА-2015: XII міжнародна науково-технічна конференція, 28-29 квітня 2015 р.: тези доп. К., 2015. 137-143.
8. ОЛЕШКО Т.І.: Формування та аналіз семантичних параметрів в інформаційній технології, Т.І. Олешко, Моделювання та інформаційні технології: зб. наук. праць. К.: НАНУ, 2009. Вип. 50. 93-101.
9. ОЛЕШКО Т.І.: Організація програмних засобів управління системою зв'язку, Т.І. Олешко, Моделювання та інформаційні технології: зб. наук. праць. К.: НАНУ, 2010. Вип. 57. 105-113.
10. ОЛЕШКО Т.І.: Еволюційний спосіб функціонування системи, Т.І. Олешко, Моделювання та інформаційні технології: зб. наук. праць. К.: НАНУ, 2011. Вип. 58. 179-185.
11. ОЛЕШКО Т.І.: Сучасні методи інтелектуального аналізу даних, Т.І. Олешко, Інформаційні технології, системний аналіз і моделювання соціо-екологоекономічних систем: III міжнародна науково-практична конференція, 19-21 жовтня 2011р.: тези доп. К., 2011. 7-8.
12. ОЛЕШКО Т.І.: Formation of the information portrait of the airport, Т.І. Олешко, Вісник НАУ. зб. наук. пр. К.: НАУ, 2013. 34-39
13. ОЛЕШКО Т.І.: Оцінка найбільш ймовірного розподілу польотів ентропійним інструментарієм, Т.І. Олешко, О.М. Горбачова, О.Л. Лещинський, Інформаційні технології, системний аналіз і моделювання соціоекологоеконо-мічних систем: V міжнародна науково-практична конференція, 19-20 березня 2014 р.: тези доп. К., 2014. 34-37.
14. ОЛЕШКО Т.І.: Побудова графової моделі живучості нечіткої мережі аеропортів, Т.І. Олешко, О.Л. Лещинський, О.М. Горбачова, Журнал „Проблеми економіки": Харківський національний економічний університет МОН України, НДЦ індустріальних проблем розвитку НАН України ІНЖЕК, 1(2015).

Łukasz HAMERA[1], Anna GAŁUSZKA[2]

Opiekun naukowy: Szymon WĄSOWICZ[3]

KONWOLUCYJNE SIECI NEURONOWE NA PRZYKŁADZIE ROZPOZNAWANIA CYFR

Streszczenie: Mózg jest wciąż niedoścignionym organem, który analizuje i rozpoznaje biliardy obiektów w ciągu naszego życia. W dniu dzisiejszym wszystkie okazy sztucznej inteligencji nie dorównują tak idealnemu organowi, jednakże konwolucyjne sieci neuronowe [1] oferują w pewnych czynnościach lepszą skuteczność. W pracy opisany zostanie algorytm rozpoznawania cyfry na podstawie mapy bitowej.

Słowa kluczowe: sieci neuronowe, konwolucja, wykrywanie wzorców

DIGITS RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORKS

Summary: The brain is still a transcendent organ that analyses and recognizes billiards objects throughout our lives. Today, all articial intelligence is not as good as the ideal organ, but convolutional neural networks offer better performance in certain operations. The algorithm which recognises digits from their bitmaps will be presented.

Keywords: neural networks, convolution, pattern recognition

1. Wstęp

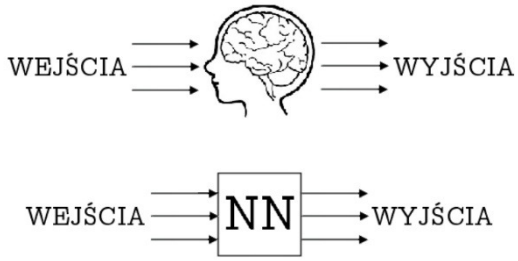
W ostatnich latach, w związku z potrzebą wprowadzania danych do wielu aplikacji, pismo odręczne stało się bardzo ważnym obiektem badań. Obecnie istnieje wiele technologii zajmujących się tym problemem, kilka z nich odniosło duży sukces na (np. OCR). Termin rozpoznawanie cyfr odnosi się do tłumaczenia map bitowych przedstawiających cyfry na format zrozumiały przez komputer. Podchodząc do tego zagadnienia możemy wyróżnić dwie możliwości wczytywania danych do programu.

[1] Student, Akademia Techniczn-Humanistyczna w Bielsku-Białej, Wydział Budowy Maszyn i Informatyki, kierunek: Informatyka
[2] Student, Akademia Techniczn-Humanistyczna w Bielsku-Białej, Wydział Budowy Maszyn i Informatyki, kierunek: Informatyka
[3] Dr hab. prof. ATH, Akademia Techniczn-Humanistyczna w Bielsku-Białej, Wydział Budowy Maszyn i Informatyki, Katedra Matematyki, swasowicz@ath.bielsko.pl

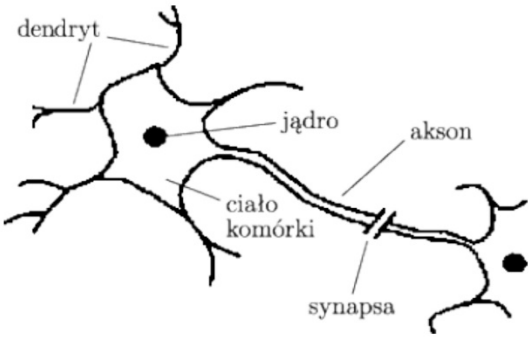
Odczytać można istniejący skan z pamięci lub ręcznie wprowadzić dane za pomocą ekranu dotykowego. Celem artykułu jest przestawienie metody rozpoznawania cyfr za pomocą konwolucyjnych sieci nerownowych. Przedstawiono budowę pojedynczego sztucznego neuronu wraz z opisem poszczególnych warstw oraz metody nauki.

2. Sztuczna sieć neuronowa

Ludzki mózg można opisać jako biologiczną sieć neuronów – zbiór komórek nerwowych transmitujących sygnał elektryczny.
Dendryty otrzymują sygnał, który w oparciu o wejścia jest przetwarzany. Następnie sygnał wyjściowy trafia do aksonu, który jest połączony z następną komórką nerwową. Wszystkich neuronów jest ok 10^{11}, a połączeń między nimi jest 10^{15}.



Rysunek 1. Idea działania mózgu oraz sztucznej sieci neuronowej



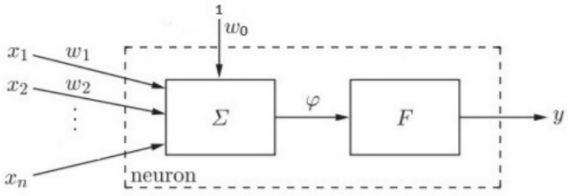
Rysunek 2. Biologiczny neuron

Taka skomplikowana struktura jest niemożliwa do odwzorowania nawet na najlepszych superkomputerach. Biliardy równoległych operacji uniemożliwiają przełożenie działania ludzkiego organu na algorytm. Dlatego naukowcy zaproponowali uproszczony matematyczny model, który nazwali *sztuczną siecią neuronową*.
Budulcem sieci neuronowej jest odpowiednik biologicznej komórki nerwowej – sztuczny neuron. Cechuje się on jednym lub wieloma wejściami $x_0, x_1, ..., x_n$ oraz

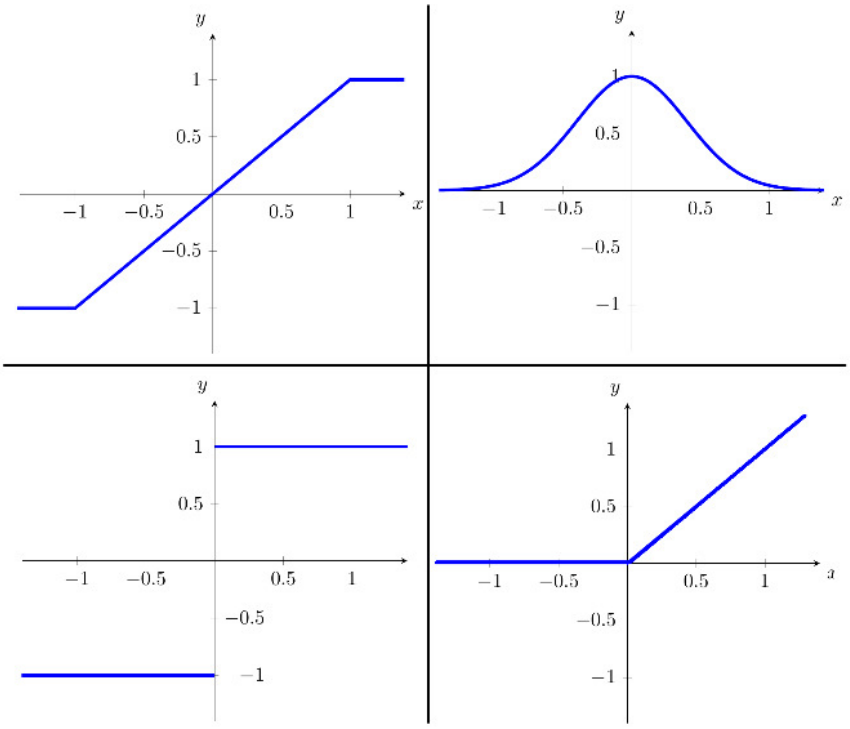
jednym wyjściem y. Dodatkowo każdemu sygnałowi wejściowemu x_i jest przyporządkowana waga w_i. Komórka oblicza wartość:

$$\rho = \sum_{i=0}^{n} x_i w_i \tag{1}$$

która jest argumentem *funkcji aktywacji*. Waga w_0 jest tzw. *wagą progową*, przyporządkowany jest jej stan wejściowy wysoki. Funkcja aktywacji przybiera różne wzory. Wpływa bezpośrednio na szybkość nauki, odwzorowanie rzeczywistego sygnału wyjściowego oraz możliwość wyszkolenia sieci. Źle dobrana może całkowicie zakłócić odpowiednie wykrywanie wzorców. Połączone ze sobą sztuczne neurony tworzą sieć.



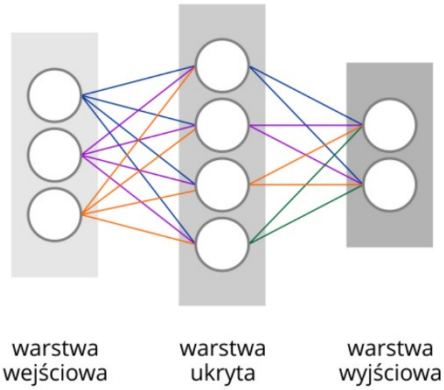
Rysunek 3. Sztuczny neuron



Rysunek 4. Funkcje aktywacji

Zazwyczaj zbudowana jest z bardzo dużej ilości neuronów, które można pogrupować w następujące warstwy:

1. warstwa wejściowa – zbiór neuronów odpowiedzialnych za poprawne przetworzenie danych wejściowych
2. warstwy ukryte – wszystkie warstwy pośrednie między wejściowymi, a wyjściowymi. Ich zadaniem jest stworzenie ogromnej ilości połączeń między warstwami oraz poprawę skuteczności sieci. Zbyt duża ilość może skutkować odwrotnym procesem - skuteczność sieci może drastycznie spaść.
3. warstwa wyjściowa – warstwa mająca na celu zwrócenie odpowiedniego wyniku.

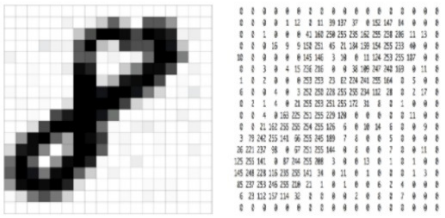


Rysunek 5. Sztuczny neuron

3. Konwolucyjna sieć neuronowa

Sztuczna sieć neuronowa jest świetnym narzędziem w klasyfikacji danych, jednak jej podstawowa odmiana potrzebuje prostych danych wejściowych. Jej skuteczność drastycznie spada jeżeli dane wejściowe są w formie map sygnałów.

Każdy obraz przechowywany na nośnikach danych może być reprezentowany przez macierze z wartościami kolorów. Liczba tych macierzy zależy od reprezentacji kolorów w obrazie rastrowym. W trybie kolorowym RGB każdy piksel dysponuje trzema wartościami (czerwony, zielony, niebieski), a w odcieniu szarości jedną wartością.

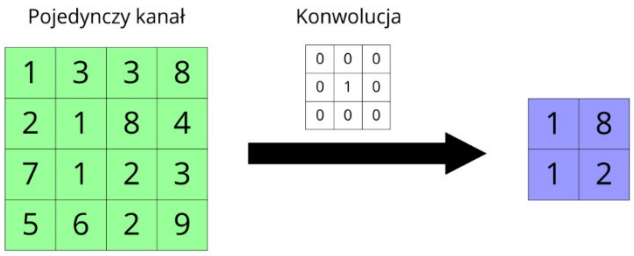


Rysunek 6. Reprezentacja obrazu w formie macierzy

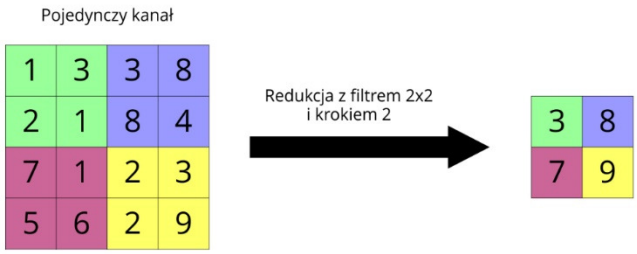
Przetworzenie obrazu w konwolucyjnej sieci neuronowej opiera się na filtracji, czyli takim przekształceniu, które przez odpowiednią modyfikację pozwala uwidocznić niektóre informacje (np. wzmocnienie krawędzi) lub pozbyć się niechcianych efektów (np. zniekształcenie, refleksy świetlne). Jedną z podstawowych metod filtracji jest wyznaczenie obrazu wynikowego, w którym każdy piksel h[c,m,n] został utworzony na podstawie jego sąsiedztwa . W literaturze metoda ta nazywana jest liniową filtracją kontekstową obrazu. Realizowana jest przez operator splotu, który prezentuje się następująco:

$$h[c,m,n] = (f \cdot g)[c,m,n] = \sum_{j} \sum_{k} f[c, j, k]\, g[c, m-j, n-k] \tag{2}$$

Funkcja f jest dwuwymiarową macierzą zawierającą wartości każdego piksela w danym kanale, natomiast funkcja g (zwana filtrem) jest obrazem o mniejszych rozmiarach. Wartości otrzymane w wyniku konwolucji (słowo to pochodzi od angielskiego terminu convolution czyli splot) są argumentami funkcji aktywacji, która najczęściej przybiera formę RELU [2].



Rysunek 7. Splot

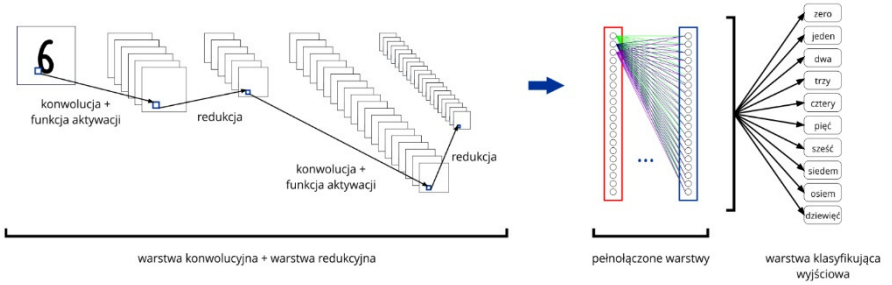


Rysunek 8. Redukcja

Wybierania jest z powodu znacznego przyspieszenia nauki. Może osiąga wartości większe niż 1, zaś obliczenie gradientu nie jest kłopotliwe. Sieć konwolucyjną cechują też warstwy redukujące mające na celu redukcję analizowanego obszaru. Neurony tej warstwy są podłączone do pewnej ilości neuronów z warstwy poprzedniej. Pozwala to tworzenie mniej skomplikowanych następnych warstw, ale wiąże ze sobą pogorszenie przekazywanej informacji.

Warstwy redukujące bezpośrednio nie biorą udziału w procesie uczenia, ponieważ nie posiadają wag. Podczas działania algorytmu wstecznej propagacji błędu przekazują błąd bez jakiejkolwiek modyfikacji. Na ich wyjściu mogą być różne formy redukcji takie jak: średnia arytmetyczna czy maksymalna wartość.

Konwolucyjna sieć neuronowa jest połączeniem sztucznej sieci neuronowej z wstępnymi warstwami konwolucyjnymi oraz redukcji. Zazwyczaj stosowany jest schemat przedstawiony na rysunku 9.



Rysunek 9. Konwolucyjna sieć neuronowa

Obraz jest wczytywany przez warstwę wejściową, następnie przechodzi na zmianę przez warstwę splotu oraz redukcji. Ilość takich powtórzeń jest zależna od rozległości i złożoności problemu. Następnie dane z części konwolucyjnej są wprowadzane do części struktury odpowiadającej klasyfikacji obiektów.

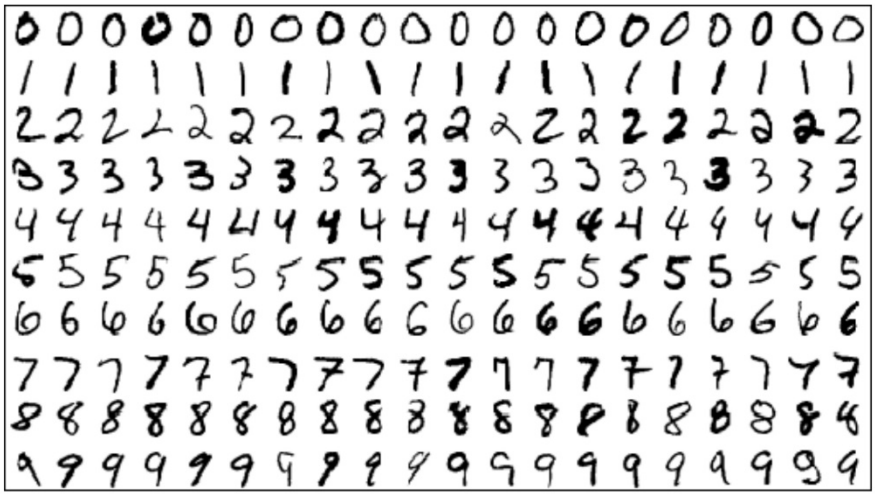
4. Sieć neuronowa rozpoznająca cyfry

Budowanie sztucznej sieci neuronowej służącej do rozwiązania konkretnego zagadnienia jest bardzo indywidualnym procesem. Nie ma jasno zdefiniowanych modeli oraz praw tworzenia tej struktury, więc każdy problem trzeba rozpatrywać osobno. W wyborze parametrów trzeba się liczyć z możliwością niedotrenowania, a nawet przetrenowania sieci.

Rozpoznawanie cyfr jest skomplikowaną procedurą opartą na przetworzeniu odręcznie narysowanego obrazu. Konwolucyjny typ sieci neuronowych idealnie poradzi sobie z tym zagadnieniem. Biorąc to pod uwagę, uczymy sieć rozpoznawania obrazów o jednym danym rozmiarze, zaś wszystkie obrazy mające inny wymiar są skalowane. Wynika to ze stałej ilości neuronów w warstwie wejściowej, której nie można zmienić po procesie nauki.

Baza MNIST jest zbiorem obrazów ręcznie pisanych cyfr w liczbie 70 000 rekordów, z których 60 000 wykorzystywanych jest przez proces uczenia, a reszta do testowania sprawności sieci.

Proces uczenia polega na dobraniu wag wszystkich wejść neuronów tak, aby uzyskiwane wartości wyjściowe w jak najmniejszym stopniu odbiegały od poprawnego wyniku. Do ustawienia wag użyty został algorytm wstecznej propagacji błędu, polegający na minimalizacji sumy kwadratów błędów z wykorzystaniem algorytmów optymalizacji, w naszym przypadku ADAM [3].



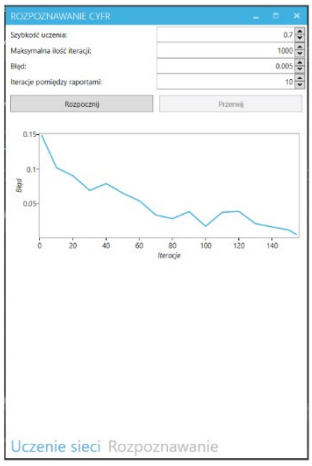
Rysunek 10. Wykaz 20 rekordów MNIST odpowiadających każdej liczbie

Zbudowana sieć składa się z warstw przedstawionych w poniższej tabeli.

Tabela 1: Budowa opisywanej sieci

Lp.	Typ	Rozmiar	Dodatkowe parametry
1	Wejściowa	1	Obraz 28×28
2	Konwolucyjna	6	Filtr 3×3, RELU
3	Redukcyjna	6	2×2, skok 2
4	Konwolucyjna	20	Filtr 3×3, RELU
5	Redukcyjna	2	2×2, skok 2
6	Pełno połączona	120	Sigmoidalna
7	Pełno połączona	100	Sigmoidalna
8	Wyjściowa	10	Klasyfikacyjna

Przy parametrach znajdujących się w Tabeli 1 sprawność sieci dla 10 prób wyniosła 97%, jednak zmiany rozmiaru oraz różne pochylenie cyfr drastycznie zaniżają tą wartość. Konwolucyjne sieci rozpoznają korelacje, a nie kształty. Powyżej opisana struktura została zaimplementowana wśrodowisku .Net przy pomocy narzędzia WPF. Aplikacja posiada możliwość rysowania cyfry i śledzenia zmian wyniku sieci w czasie rzeczywistym. Podzielona jest na dwie zakładki, pierwsza z nich oferuje możliwość zmiany parametrów nauki oraz wykonania tego procesu. Najważniejszym z nich jest szybkość nauki, definiująca jak wielki skok algorytm może wykonać podczas minimalizacji funkcji błędu (w wielu jest on pomijany i dobierany automatycznie). Dodatkowo podczas nauki uaktualniany jest wykres obrazujący wysokość błędu względem numeru iteracji.



Rysunek 11. Autorski program do rozpoznawania cyfr

Kolejny rysunek ilustruje główną cześć – panel do rysowania i wynik każdego wyjścia. Po wprowadzeniu danych aplikacja konwertuje powstały obraz do rozdzielczości 28×28 pikseli i uruchamia sieć.

Rysunek 12. Autorski program do rozpoznawania cyfr

LITERATURA

1. KRIZHEVSKY A., SUTSKEVER I., HINTON G.E.: ImageNet Classification with Deep Convolutional Neural Networks. Nevada, 2012.
2. NAIR V., HINTON G.E.: Rectified Linear Units Improve Restricted Boltzmann Machines. University of Toronto, Toronto, 2010.
3. KINGMA D.P., BA J.L.: A method for stochastic optimization. 3rd International Conference for Learning Representations, San Diego, 2015, https://arxiv.org/pdf/1412.6980.pdf

Andrii HORKUNENKO[1], Andrii SVERSTYUK[2], Serhii LUPENKO[3], Iaroslav LYTVYNENKO[4]

PAKIET OPROGRAMOWANIA DO SYMULACJI I PRZETWARZANIA SYNCHRONICZNIE ZAREJESTROWANYCH SYGNAŁÓW PRACY SERCA

Streszczenie: Opracowano pakiet oprogramowania umożliwiający spójne statystyczne przetwarzanie synchronicznie zarejestrowanych sygnałów pracy serca na podstawie modelu wektorowego cyklicznie rytmicznych procesów losowych.

Słowa kluczowe: synchronicznie zarejestrowane sygnały serca, metody przetwarzania statystycznego, oprogramowanie

SOFTWARE COMPLEX FOR MODELING AND PROCESSING OF SYNCHRONOUSLY REGISTERED CARDIOSIGNALS

Abstract: It has been developed the software complex, that allows to perform mutual statistical processing of synchronously registered cardiosignals on the basis of the vector model of the cyclic rhythmically connected stochastic processes.

Keywords: synchronously registered cardiosignals, statistic processing methods, software complex

Introduction

The effectiveness of modern cardio-diagnostic systems to a large extent depends on the hardware and software components on which they are based. Using of different processing methods, which are a part of the software, extends the functionality greatly and increases the reliability of diagnosis of human heart [1].

[1] Ph.D, I. Horbachevsky Ternopil State Medical University, Department of Medical Physics of Diagnostic and Therapeutic Equipment, associate professor, horkunenkoab@tdmu.edu.ua
[2] Ph.D, I. Horbachevsky Ternopil State Medical University, Medical Informatics Department, associate professor, sverstyuk@tdmu.edu.ua
[3] Prof. D.Sc., Ternopil Ivan Puluj National Technical University, Department of Computer Systems and Networks, professor, serhii.lupenko@gmail.com
[4] Ph.D, Ternopil Ivan Puluj National Technical University, Department of Computer Science, associate professor, d_e_l@i.ua

Whereas, methods are based on mathematical models which set the opportunity and specific of the processing. In particular, the simultaneous processing of cardiosignals of different physical nature is possible perform only if their mathematical models are correlated with each other in a certain way and have a similar structure.

Analysis of recent researches

We should mention Dreifus LS, Jorna PG, Kusick VA, Talbot S. A, Webb GN, Berkutov AM, Gurevich M. B., Dragan Y. P., Marchenko B. G., Fainsilberg L. S., Shulgina V. I., Shcherbak L. M. among the scholars who were involved in the development of software systems for the analysis of cardiosignals.

Objectives

Development and application of software for processing and modeling of synchronously registered cardiac signals as a component of cardiac diagnostic systems.

Method

According to paper [2], we give a definition of the vector of cyclic rhythmically connected stochastic processes.

Definition 1. If there is a function such as $T(t,n)$, which satisfies the conditions of the rhythm function that finitely measurable vectors

$$\{\xi_{i1}(\omega,t_1),\xi_{i2}(\omega,t_2),...,\xi_{ik}(\omega,t_k)\}$$

and

$$\{\xi_{i1}(\omega,t_1+T(t_1,n)),\xi_{i2}(\omega,t_2+T(t_2,n)),...,\xi_{ik}(\omega,t_k+T(t_k,n))\}.$$

$n\in \mathbf{Z}, i_1,...,i_k=\overline{1,N}$, where $\{t_1,...,t_k\}$- multiple separability of the vector $\mathbf{\Theta}_N(\omega,t)$, for all the integers $k\geq 1$ are stochastic equivalent in the broadest sense, we will call the vector $\mathbf{\Theta}_N(\omega,t)$ of cyclic stochastic processes $\{\xi_i(\omega,t), i=\overline{1,N}, \omega\in\mathbf{\Omega}, t\in \mathbf{W}\}$ as the vector of strictly rhythmically connected stochastic processes and the processes as strictly rhythmically connected.

Area of definition **W** vector of cyclic rhythmically connected stochastic processes can be as ordered discretely $\mathbf{W}=\mathbf{D}=\{t_{ml}\in R, m\in\mathbf{Z}, l=\overline{1,L}\}$ or continuous **W=R** set of real numbers. In the case of discrete domain definition **W=D** for its elements if $m_2>m_1$, or if $m_2=m_1$, a $l_2>l_1$, in other cases $t_{m_1l_1}>t_{m_2l_2}$; $m_1,m_1\in\mathbf{Z}$, $l_1,l_2\in\overline{1,L}$ there is a type of linear ordering: $t_{m_1l_1}<t_{m_2l_2}$. Moreover $0<t_{m,l+1}-t_{m,l}<\infty$.

The rhythm function $T(t,n)$ determines the law of changing the time intervals between the single-phase values of the vector of cyclic rhythmically connected

stochastic processes. The function of the rhythm satisfies such conditions according to the theorem which is proved in the paper:

a) $T(t,n)>0$, if $n>0$ ($T(t,1)<\infty$);

b) $T(t,n)=0$, if $n=0$;

c) $T(t,n)<0$, if $n<0$, $t\in \mathbf{W}$; (1)

for any $t_1\in \mathbf{W}$ and $t_2\in \mathbf{W}$, for which $t_1<t_2$, for function $T(t,n)$ should be performed a strict inequality:

$$T(t_1,n)+t_1<T(t_2,n)+t_2,\ \forall n\in \mathbf{Z}; \quad (2)$$

function $T(t,n)$ is the smallest modulo $\left(\left|T(t,n)\right|\le\left|T_\gamma(t,n)\right|\right)$ among all such functions $\{T_\gamma(t,n),\ \gamma\in \Gamma\}$, which satisfy (1) and (2).

In the partial case, if the rhythm function is $T(t,n)=n\cdot T$ $(T>0, n\in \mathbf{Z})$, we will call the vector $\mathbf{\Theta}_N(\omega,t)$ as the vector T- periodically connected stochastic processes.

Let us consider the properties of some probabilistic characteristics of the vector $\mathbf{\Theta}_N(\omega,t)$ cyclic rhythmically connected stochastic processes. So, for its compatible k - measurable distribution function takes place equation:

$$F_{k\xi_{i_1}\ldots\xi_{i_k}}(x_1,\ldots,x_k;t_1,\ldots,t_k)=$$

$$=F_{k\xi_{i_1}\ldots\xi_{i_k}}(x_1,\ldots,x_k;t_1+T(t_1,n),\ldots,t_k+T(t_k,n)),n\in \mathbf{Z}, i_1,\ldots,i_k=\overline{1,N},t_1,\ldots,t_k\in \mathbf{W}. \quad (3)$$

Combined central moments function of order $p=\sum_{j=1}^{k}R_j$:

$$r_{p\,\xi_{i_1}\ldots\xi_{i_k}}(t_1,\ldots,t_k)=\mathbf{M}\left\{\left(\xi_{i_1}(\omega,t_1)-m_{\xi_{i_1}}(t_1)\right)^{R_1}\cdot\ldots\cdot\left(\xi_{i_p}(\omega,t_k)-m_{\xi_{i_k}}(t_k)\right)^{R_k}\right\}= \quad (4)$$

$$=r_{p\,\xi_{i_1}\ldots\xi_{i_k}}(t_1+T(t_1,n),\ldots,t_k+T(t_k,n)),\ t_1,t_2,\ldots,t_k\in \mathbf{W},\ i_1,\ldots,i_k=\overline{1,N},\ n\in \mathbf{Z}.$$

Results and Discussion

It were made the series of experiments on processing of the cardiaosignals of the same and different physical origin which were investigated for the purpose of approving of the greater effectiveness of the simultaneous processing of synchronously registered cardiosignals (SRCS) based on the model of the vector of cyclic rhythmically connected stochastic processes in comparison with the well-known method of their processing.

As the example it is shown on the Figure 1 the realizations of SRCS electrocardiosignal (ECS) and phonocardiosignal (PCS) and on the figures 2-4 are represented the results of a comparative analysis.

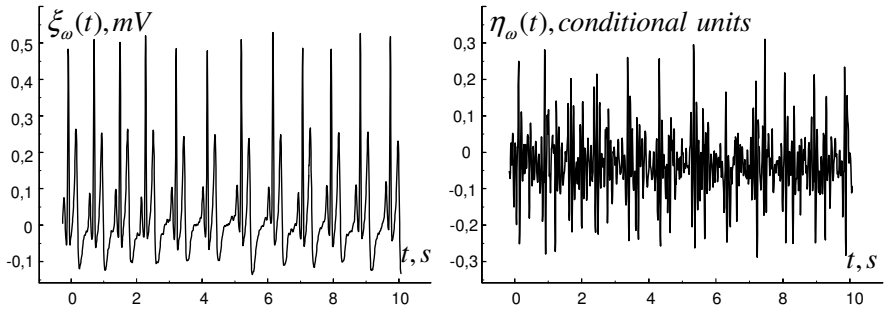


Figure 1. The realizations of SRCS: (a) electrocardiogram, (b) phonocardiogram

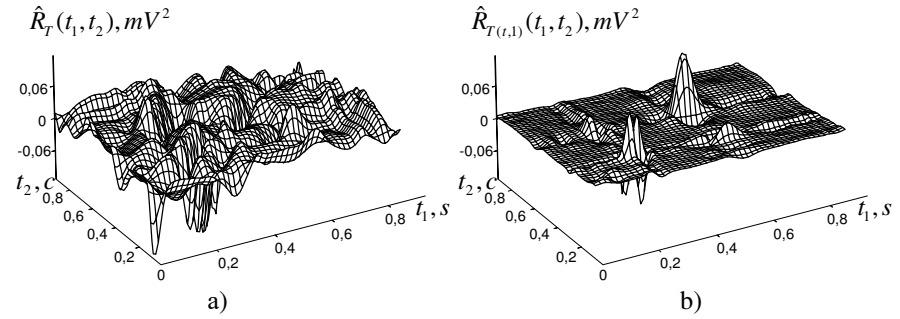


Figure 2. The realization graphs of statistical estimations of the autocorrelation function of the ECS while its processing on the basis of: (a) the vector of periodically connected stochastic processes; (b) the vector of cyclic rhythmically connected stochastic processes

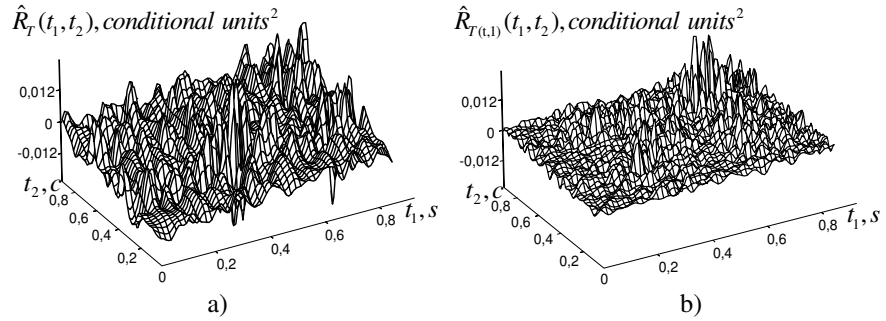


Figure 3. The realization graphs of statistical estimations of the autocorrelation function of PCS while its processing on the basis of: (a) the vector of periodically connected stochastic processes; (b) the vector of cyclic rhythmically connected stochastic processes

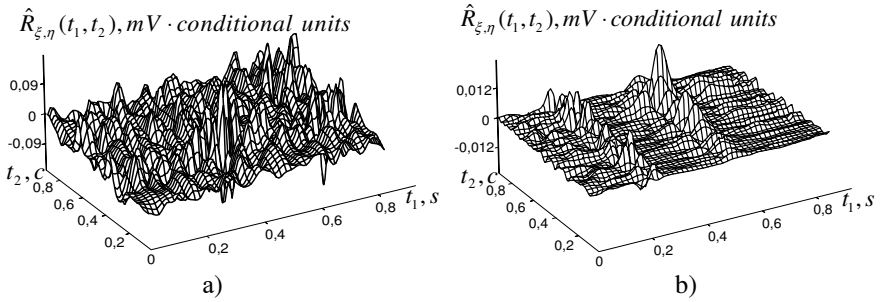


Figure 4. The realization graphs of statistical estimations of the mutual correlation function of the ECS and PCS while its processing on the basis of: (a) the vector of periodically connected stochastic processes; (b) the vector of cyclic rhythmically connected stochastic processes

It was set in the result of the comparative analysis of the statistical processing of sets of SRCS which was made that the method of statistical processing of the analyzed cardiosignals on the basis of the vector of cyclic rhythmically connected stochastic processes significantly reduces the negative effect of "blurring" of statistical estimations of mutual correlation functions as a part of set of synchronous cyclic signals of a heart that is strongly-pronounced in results of statistical processing of the cardiac signals which were investigated on the basis of the vector of periodically connected stochastic processes. It is proved by the fact that the new method of compatible statistical processing takes into account the variability of the SRCS rhythm, in contrast to the well-known methods.

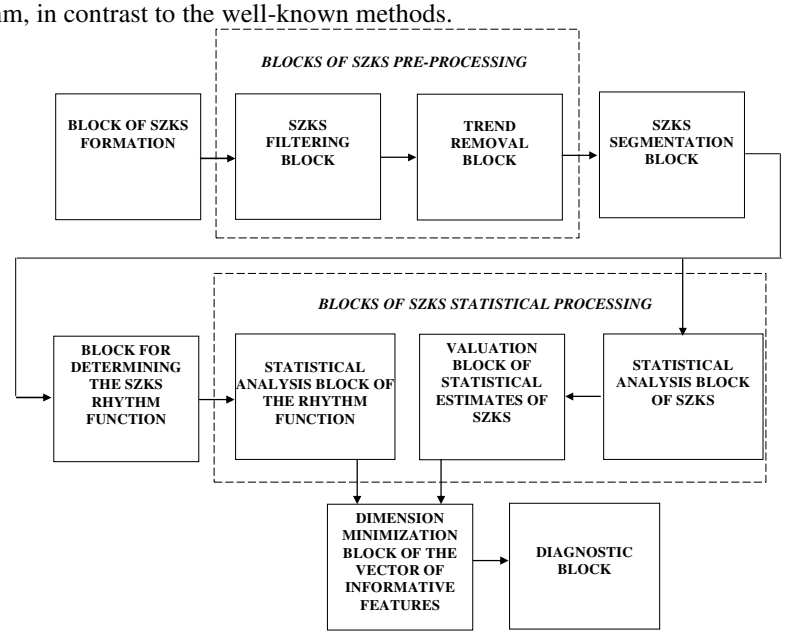


Figure 5. Software complex for statistical processing of synchronously registered cardiac signals

It was developed the software complex which allows to perform SRCS statistical processing and its structural scheme is shown in Figure 5.

In this complex was implemented: block of SZKS formation, SZKS filtering block, trend removal block, SZKS segmentation block, block for determining the SZKS rhythm function, statistical analysis block of the rhythm function, valuation block of statistical estimates of SZKS, statistical analysis block of SZKS, dimension minimization block of the vector of informative features, diagnostic block.

Conclusions

A complex of programs is developed that allows performing statistical processing of cardiosignals of different physical origin and obtain estimations of mathematical expectation, dispersion, autocorrelation and mutual correlation functions (taking into account the rhythm function and the period). This software complex can be used as a component of the human heart's diagnostic systems.

REFERENCES

1. BOULDAKOVA T.I., GRIDNEV V.I., KIRILLOV K.I., LANTCBERG A.V., SUYATINOV S.I.: Program-Analytical System for Model Processing of Biosignals. Biomedical Radioelectronics.- Moscow: “Radioteknika”, 1(2009), 71-77.
2. SVERSTYUK A. S.: Argumentation and verification of the mathematical model of synchronously registered cardiosignals using the vector of cyclic rhythmically connected stochastic processes. A. S. Sverstyuk - Measuring and Computing Devices in Technological Processes**. –** Khmelnitsky: Khmelnitsky National University, 1(2009), 143–147.
3. LUPENKO S. A.: The statistic common cardiosignals’ analysis on the basis of vector of cyclic rhythmically connected stochastic processes, S. A. Lupenko, I. V. Lytvynenko, A. S. Sverstyuk - Electronics and Control Systems. – Kiev: National Aviation University, 4(2008)18, 22–29.

Vladyslav HRIHA, Andrii GIZUN, Iryna SHCHUDLYK[1]

SYSTEM OPARTY NA INFORMACJACH DOTYCZĄCYCH WYKRYWANIA I IDENTYFIKACJI INFORMACYJNEGO I PSYCHOLOGICZNEGO ODDZIAŁYWANIA

Streszczenie: W artykule omówiono teorię informacji konfrontacji. Opracowano własne metody klasyfikacji informacji i oddziaływania psychologicznego. Opracowano sformalizowany i zintegrowany model informacji psychologicznych stosowanych podczas procedury wykrywania i identyfikacji. Po raz pierwszy zastosowano metodę i system oparty na informacjach dotyczących wykrywania i identyfikacji i psychologicznego oddziaływania oparte na logice rozmytej. Uzyskane rezultaty mogą być wykorzystywane do skutecznego wdrożenia środków zaradczych w celu zwalczania negatywnego wpływu informacji w działaniach wojennych.

Słowa kluczowe: model, informacja, psychologiczne metody wpływu informacji i oddziaływania psychologicznego, identyfikacja, logika rozmyta

INFORMATION PSYCHOLOGICAL IMPACT DETECTION AND IDENTIFICATION SYSTEM

Abstract: the article discusses the theory of information confrontation. Own classification of methods of information psychological impact were developed. The formalized and integrated models of information-psychological impact that are used during its detection and identification are developed. The method and system of detection and identification of information psychological impact based on fuzzy logic were developed for the first time. The results of work can be used for the effective implementation of countermeasures against the negative informational impacts in the conditions of information warfare.

Keywords: the model, information psychological impact, the methods of information psychological impact, identification, fuzzy logic.

1. Formulation of the problem

Information impact is becoming more and more important in the modern world. This process is facilitated by globalization and the transition to an information society. Exactly this factor affects on the wider use of the information tools for self-interest.

[1] National Aviation University, Kyiv, Ukraine, gsmgrey1@gmail.com

The most widespread areas of their implementation are military, political and economic. During the military operations, important aspects are the popular support of the military actions, deterioration and disorganization of the enemy morally-psychological state, in the political sphere – the incensement of the government trust percentage, the imposition of ideology, and in the economic sphere – gaining an advantage over a competing company or over the whole state. One of the methods to achieve this is the information psychological impact.

The issue of providing information psychological security in Ukraine gained the particular importance due to Russia's aggression against Ukraine, when the question of formation of support by the population of the territorial integrity of Ukraine, and maintenance a high level of moral combat spirit of ATO troops was raised. The information psychological impact implementation and providing information psychological security is impossible without a detailed examination of its theory and methods of realization [1].

2. Analysis of recent research and publications

The researchers in many countries of the world described the processes and created models of information psychological impact [2]. Taking into account the current situation with the information confrontation, the relevance of these studies will only increase. The summary of the examined models is presented in Table 1.

Table 1. Summary results of analytical studies

Name	Features
Shiyan Model [3]	Detection of information psychological impact by changing one of the characteristics or the entire information space
Technological aspects of information warfare [4]	Examination of the indicative classification of methods of information warfare. It has five characteristics: by the type of confrontation, by purpose, by the nature of the impact, by the source of distribution, by the target audience.
Information confrontation life cycle model[5]	Two-level cycle: "ATTACK" and "PROTECTION" with its own set of methods.
Scientific research council of the USA information psychological impact models [6]	Examination of four components: mathematical models of faith formation in response to message transmission, network persuasion, model of processing of social information and transactional memory.
Game representation of information warfare [7]	It is assumed that the information warfare is a game with two rational players, for which several scenarios are possible. The essence of the game is to choose a better scenario.
Johnson Informational Attack Model [8]	It is determined that attacks can occur both on the whole informational environment, and at its separate levels.

Project CEPA	Analysis of the information environment by keywords, detection and identifying information psychological effects by context.
Information warfare tactics [9]	Representation of the information environment as 5 subsystems: implementation, coordination, internal control, intelligence and processing, policy and strategy, attacks tactics construction and their protection.
Mathematical models of information warfare [10]	Formulas that describe the processes of information operations, the definition of critical assets and the detection of negative information impact have been developed.
Frame theory during the investigation of information warfare [11]	Consideration of information warfare as a set of frames, which will let to determine their source and implement countermeasures better.

3. Information psychological impact detection and identification method

Information psychological impact can be described via the tuple I = <Idp, Is, T, Q, R>. The above parameters of the tuple form the target model and are determined by the signs of information psychological impact.

Idp – methods of information psychological impact.

Is – the space on which the information-psychological impact is carried out.

It is important to limit the size of space, objects of information infrastructure and social groups that are exposed to IPI (aggression does not affect the whole information psychological space of the victim state, but only part of it).

T – time of impact. The duration of the use of information psychological impact methods on a particular information space.

Q – purpose of impact. The purpose is a local or partial goal, as a rule, aggression ceases after the aggressor has achieved the full achievement of the specific goals and rarely takes a protracted nature.

R – a set of countermeasures aimed at resisting information psychological impact.

The functional model of information-psychological impact can be represented as:

Iw = <Id, Pig, Pog, R>, where

Id – this is a set of methods of information psychological impact. During one information-psychological impact action several methods can be used simultaneously. During the modern warfare, the following classification can be made:

1) methods aimed at people who perceive the information critically:
   - change of opinion by persuasion;
   - psychological isolation of the object;
   - coercion;
   - propaganda.
2) methods aimed at people who perceive the information uncritically:
   - misinformation;
   - propaganda;
   - change of sights by suggestion;
   - infection;
   - manipulation;
   - reframing.

As we can see, there is a certain imbalance among the methods, which are directed at people who perceive information uncritically in most cases. This situation is due to the fact that it is much easier to achieve a result, to carry out an attack, if the attacker's actions are aimed at non-critical thinking, since they will bypass a certain "psychological shield" of a person.

Propaganda is attributed to both groups, because of the variety of means, it is evident that its use is equally effective for all people.

Concepts and classifications regarding information psychological impact analysis has shown that today there is no single classification that would cover all aspects and characteristics of its implementation during the information warfare.

During the research, the following basic characteristics of the IPI were identified:

- «information psychological impact implementation duration» - EL,
- «stage of manifestation» - DM,
- «economic losses level» - EL;
- «percentage of the population that watches foreign TV» - PP;
- «percentage of the population that reads the foreign press» - PN;
- «level of trust in government» - CG;
- «the level of population protest attitudes» - PM;
- «information infrastructure development degree» - II;
- «external factors impact degree» - IF.

So, the set of identifying parameters for the number of investigated situations when n = 8 can be presented as follows:

$$IPF = \left\{ \bigcup_{I=1}^{9} IPF_I \right\} = \{IPF_1, IPF_2, IPF_3, IPF_4, IPF_5, IPF_6, IPF_7, IPF_8\} =$$

$$= \{LT, DM, EL,, PP, PN, CG, PM, II, IF\}$$

Reference values are formed in accordance with [12, 13].

The parameter EL is characterized by the following linguistic assessments: {short (S), medium (M), long (L)}. Intervals for reference values determination = {[0-25], [26-50], [51-75]} days.

After making the transformation, obtain a set of parameter reference values EL = {short (S), medium (M), long (L)} and the terms of the linguistic variables for this parameter:

S = {0/0,25; 1/0,25, 0,3/0,5; 0,1/1; 0/1},
M = {0/0,25; 0,6/0,25; 1/0,5; 0,7/1; 0/1},
L= {0/0,25; 0,5/0,25; 0,7/0,5; 1/1; 0/1}.

The parameter DM is characterized by the following linguistic assessment: {primary (P), deployed (D), final (F)}. Intervals for reference values determination = {[0-10], [11-20], [21-30]} IPI events per year.

After making the transformation, obtain a set of parameter reference values DM=T_{log} = {primary (P), deployed (D), final (F)} and the terms of the linguistic variables for this parameter:

P = {0/0,33; 1/0,33, 0,6/0,67; 0,24/1; 0/1},
D = {0/0,33; 0,86/0,33; 1/0,67; 0,59/1; 0/1},
F= {0/0,33; 0,5/0,33; 0,93/0,67; 1/1; 0/1}.

The parameter EL is characterized by the following linguistic assessment: {small (S), middle (M), high (H)}. Intervals for reference values determination = {[0-20], [21-50], [51-80]} thousands of dollars per year..

After making the transformation, obtain a set of parameter reference values EL= {small (S), middle (M), high (H)} and the terms of the linguistic variables for this parameter:

S = {0/0,25; 1/0,25; 0,67/0,63, 0,13/1,0/1},
M = {0/0,25; 0,44/0,25; 1/0,63; 0,38/1; 0/1},
H= {0/0,25; 0,06/0,25; 0,42/0,63; 1/1; 0/1}.

The parameter PP is characterized by the following linguistic assessment: {small (S), medium (M), high (H)}. Intervals for reference values determination = {[0-33], [34-66], [67-100]} percent.

After making the transformation, obtain a set of parameter reference values PP= {small (S), medium (M), high (H)} and the terms of the linguistic variables for this parameter:

S = {0/0,33; 1/0,33; 0,92/1; 0,4/1; 0/1},
M = {0/0,33; 0,4/0,33; 1/0,66; 0,7/1; 0/1},
H= {0/0,33; 0,1/0,33; 0,44/0,66; 1/1; 0/1}.

The parameter PN is characterized by the following linguistic assessment: {small (S), medium (M), high (H)}. Intervals for reference values determination = {[0-33], [34-66], [67-100]} percent.

After making the transformation, obtain a set of parameter reference values PN=T_{log} = {small (S), medium (M), high (H)} and the terms of the linguistic variables for this parameter:

S = {0/0,33; 1/0,33; 0,92/0,67; 0,44/1; 0/1},
M = {0/0,33; 0,4/0,33; 1/0,67; 0/1},
H = {0/0,33; 0,58/0,67; 1/1; 0/1}.

The parameter CG characterized by the following linguistic assessment:{distrust (D), partial trust (P), full confidence (F)}. Intervals for reference values determination = {[0-33], [34-66], [67-100]} percent during year.

After making the transformation, obtain a set of parameter reference values CG= {distrust (D), partial trust (P), full confidence (F)} and the terms of the linguistic variables for this parameter:

D = {0/0,33; 1/0,33; 0,6/0,66; 0,27/1; 0/1},
P = {0/0,33; 0,69/0,33; 1/0,66; 0,82/1; 0/1},
F = {0/0,33; 0,2/0,66; 1/1, 0/1}.

The parameter PM characterized by the following linguistic assessment: {low (L), middle (M), high (H)}. Intervals for reference values determination = {[0-33], [34-66], [67-100]} percent during year.

After making the transformation, obtain a set of parameter reference values PM= {low (L), middle (M), high (H)} and the terms of the linguistic variables for this parameter:

L = {0/0,33; 1/0,33, 0,95/0,66; 0,78/1; 0/1},
M = {0/0,33; 0,42/0,33; 1/0,66; 0,39/1; 0/1},
H= {0/0,33; 0,13/0,33; 0,4/0,67; 1/1, 0/1}.

The parameter II characterized by the following linguistic assessment: {undeveloped (U), moderately developed (M), developed (D)}. Intervals for reference values determination = {[0-5], [6-10], [11-15]} score.

After making the transformation, obtain a set of parameter reference values II=T_{log} = {undeveloped (U), moderately developed (M), developed (D)} and the terms of the linguistic variables for this parameter:

U = {0/0,33; 1/0,33, 0,4/0,67; 0,12/1; 0/1},

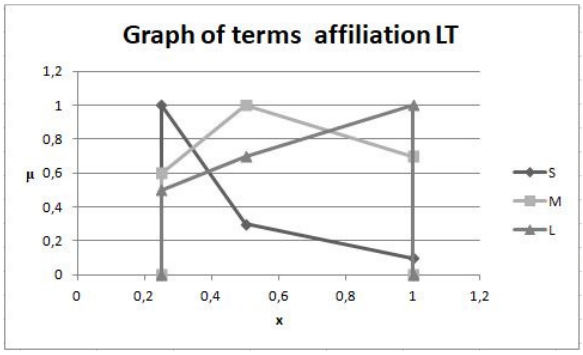
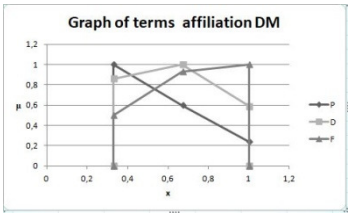
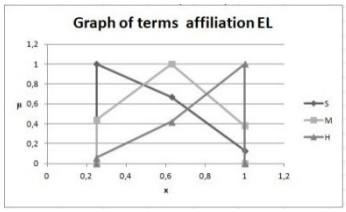
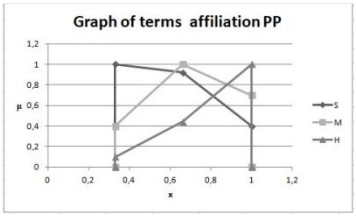
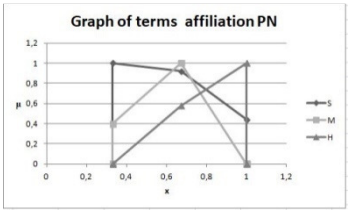
M = {0/0,33; 0,86/0,33; 1/0,67, 0,47/1; 0/1},
D= {0/0,33; 0,57/0,33; 0,9/0,67; 1/1; 0/1}.

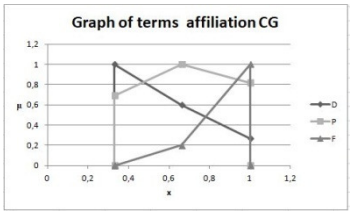
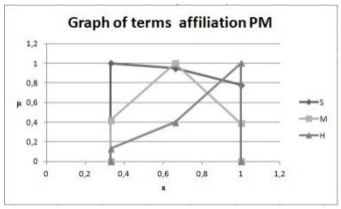
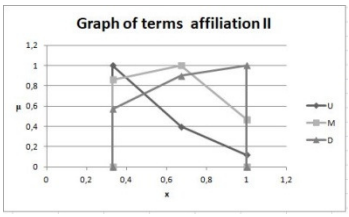
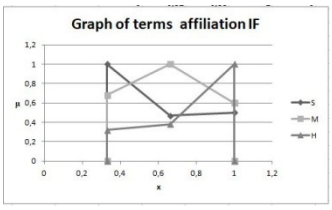
The parameter IF characterized by the following linguistic assessment: {low (L), middle (M), high (H)}. Intervals for reference values determination = {[0-33], [34-66], [67-100]} percent during the year.

After making the transformation, obtain a set of parameter reference values IF=T_{log} = {low (L), middle (M), high (H)} and the terms of the linguistic variables for this parameter:

L = {0/0,33; 1/0,33, 0,47/0,66; 0,5/1; 0/1},
M = {0/0,33; 0,68/0,33; 1/0,66; 0,6/1; 0/1},
H= {0/0,33; 0,32/0,33; 0,38/0,66; 1/1, 0/1}.

Let's represent the calculated reference values in the form of graphs.

<table>
<tr><td colspan="2">Graphic representation of fuzzy values</td></tr>
<tr><td colspan="2"> 
</td></tr>
<tr><td colspan="2">Duration of implementation of IPI</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>Stage of manifestation</td><td>Economic losses level</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td>percentage of the population that watches foreign TV</td><td>Percentage of the population that reads the foreign press</td></tr>
</table>

	
Level of trust in government	The level of protest attitudes of the population
	
Information infrastructure development degree	External factors impact degree

Therefore, in order to predict the possibility of realizing information-psychological impact or to define and identify it is necessary to develop a system that will monitor the basic characteristics and, basing on heuristic rules, to reveal information psychological impact.

Currently, there is a problem of identifying the methods of information psychological impact. For its solution it is expedient to use a mechanism of heuristic rules, which will help to conduct identification on the basic characteristics of the methods.

The construction of rules is usually carried out on the basis of an expert approach, and this is especially important in cases where one of the alternatives needs to be preferred. To select one solution from the set of alternatives, we use the methods of determining the coefficients of importance [13, 14]. We will use the method of rank transformations, because it allows to attract several experts, tabular forms are used as input, the output function is linear, and the complexity is low.

Apply the apparatus of the logical-linguistic connection "method-parameter" and the interdependence between the rates of parameters and the possibility of implementing an information attack, we will form heuristic rules for their detection and identification by analogy with [15]. Heuristic rules allow you to evaluate the possibility (degree) of the implementation of this or that attack, depending on the set of values of parameters at a certain point in time. During the research, the following basic characteristics of the IPI were identified: «stage of manifestation» - DM, «economic losses level» - EL; «percentage of the population that watches foreign TV» - PP; «percentage of the population that reads the foreign press» - PN; «the level of protest attitudes of the population» - PM; «information infrastructure development degree» - II; «level of trust in government» - CG; «the goal of IPI implementation» - RE; «external factors impact degree» - IF, by which the detection and identification process is carried out. Characteristics may acquire the values determined by the

linguistic variables – «Low» (L), «Medium» (M), «High» (H) for IF, EL, PM, «Primary» (P), «Deployed» (D), «Final» (F) for DM, «Undeveloped» (U), «Moderately developed» (M), «Developed» (D) for II; «Distrust» (D), «Partial trust» (P), «Full confidence» (F) for CG; «Small» (S), «Medium» (M), «High» (H) for PP and PN.

Let's formulate a set of heuristic rules for detection and identifying the IPI method "Persuasion" and present it in the form of a table (Table 2). It can be characterized by the following confidence indicators of its implementation: "Low" (L), "Average" (A), "High" (H), which is evaluated based on the value of the identifying parameters set for it (P_{DM}, P_{EL}, P_{II}, P_{PN}, P_{PP}). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 2.

p	P_{DM}	P_{EL}	P_{II}	P_{PN}	P_{PP}	Result
1	P	L	U	S	S	L
2	P	L	U	S	M	L
3	P	L	U	S	H	A
4	P	L	U	S	S	A
5	P	L	U	S	M	A
….						
142	D	M	M	M	H	A
143	D	M	M	M	S	L
144	D	M	M	M	M	L
145	D	M	M	M	H	L
....						
239	F	H	D	H	S	H
240	F	H	D	H	H	H
241	F	H	D	H	S	H
242	F	H	D	H	H	A
243	F	H	D	H	H	A

The IPI method "Psychological isolation" can be characterized by the following indicators of confidence in the fact of its implementation: "Low" (L), "Average" (A), "High" (H), estimated on the basis of the value of the identifying parameters specified for it (P_{EL}, P_{PP}, P_{II}, P_{PN}, P_{IF}).

Let`s form a set of heuristic rules for detection and identification of "Psychological isolation" and present it in the form of a table (Table 3). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 3.

P	P_{EL}	P_{PP}	P_{II}	P_{PN}	P_{IF}	Result
1	L	S	H	M	L	L
2	L	S	H	M	L	L
3	L	S	H	M	L	L

4	L	S	H	M	L	L
5	L	S	H	M	L	L
….						
142	M	M	M	M	M	A
143	M	M	M	M	M	A
144	M	M	M	M	M	A
145	M	M	M	M	M	A
....						
239	H	H	D	H	H	H
240	H	H	D	H	H	H
241	H	H	D	H	H	H
242	H	H	D	H	H	H
243	H	H	D	H	H	H

The IPI method “Coercion” can be characterized by the following indicators of assurance about the fact of its implementation: "Low" (L), "Average" (A), "Critical" (C), which is estimated on the basis of the value identifying parameters assigned to it (P_{DM}, P_{CG}, P_{PM}, P_{IF}).

Let`s form a set of heuristic rules for detection and identification of "Coercion" and present it in the form of a table (Table 4). Taking into account all possible combinations of controlled parameters states, it is possible to form 81 corresponding heuristic rules. Here are some of them.

Table 4.

P	P_{DM}	P_{CG}	P_{PM}	P_{IF}	Result
1	P	D	L	L	L
2	P	D	L	L	L
3	P	D	L	L	L
4	P	D	L	L	L
5	P	D	L	L	L
...					
35	D	P	M	M	A
36	D	P	M	M	A
37	D	P	M	M	A
38	D	P	M	M	A
39	D	P	M	M	A
...					
78	F	F	H	H	C
79	F	F	H	H	C
80	F	F	H	H	C
81	F	F	H	H	C

The IPI method “Misinformation” can be characterized by the following indicators of assurance about the fact of its implementation: "Low" (L), "Average" (A), "High" (H), "Critical" (C), which is estimated on the basis of the value identifying parameters assigned to it (P_{CG}, P_{EL}, P_{PP}, P_{PN}, P_{IF}).

Let`s form a set of heuristic rules for detection and identification of "Misinformation" and present it in the form of a table (Table 5). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 5.

p	P_{CG}	P_{EL}	P_{PP}	P_{PN}	P_{IF}	Result
1	D	L	S	S	L	L
2	D	L	S	S	L	L
3	D	L	S	S	L	L
4	D	L	S	S	L	L
5	D	L	S	S	L	L
….						
142	P	M	M	M	M	A
143	P	M	M	M	M	A
144	P	M	M	M	M	A
145	P	M	M	M	M	A
....						
239	F	H	H	H	H	C
240	F	H	H	H	H	C
241	F	H	H	H	H	C
242	F	H	H	H	H	C
243	F	H	H	H	H	C

The IPI method “Propaganda” can be characterized by the following indicators of assurance about the fact of its implementation: "Insignificant" (I), "Average" (A), "High" (H), which is estimated on the basis of the values given for this type of IPI characteristics (P_{DM}, P_{EL}, P_{PM}, P_{PN}, P_{CG}).

Let`s form a set of heuristic rules for detection and identification of "Propaganda" and present it in the form of a table (Table 6). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 6.

P	P_{DM}	P_{EL}	P_{PM}	P_{PN}	P_{CG}	Result
1	P	L	L	S	D	I
2	P	L	L	S	D	I
3	P	L	L	S	D	I
4	P	L	L	S	D	I
5	P	L	L	S	D	I
….						
142	D	M	M	M	F	A
143	D	M	M	M	F	A
144	D	M	M	M	F	A
145	D	M	M	M	F	A
....						

239	F	H	H	H	F	H
240	F	H	H	H	F	H
241	F	H	H	H	F	H
242	F	H	H	H	F	H
243	F	H	H	H	F	H

The IPI method “Suggestion” can be characterized by the following indicators of assurance about the fact of its implementation: "Low" (L), "Medium" (M), "High" (H), which is estimated on the basis of the value identifying parameters assigned to it (P_{EL}, P_{PP}, P_{II}, P_{PN}).

Let`s form a set of heuristic rules for detection and identification of "Suggestion" and present it in the form of a table (Table 7). Taking into account all possible combinations of controlled parameters states, it is possible to form 81 corresponding heuristic rules. Here are some of them.

Table 7.

P	P_{EL}	P_{PP}	P_{II}	P_{PN}	Result
1	L	S	U	S	L
2	L	S	U	S	L
3	L	S	U	S	L
4	L	S	U	S	L
5	L	S	U	S	L
…					
36	M	M	M	M	M
37	M	M	M	M	M
38	M	M	M	M	M
39	M	M	M	M	M
…					
77	H	H	D	H	H
78	H	H	D	H	H
79	H	H	D	H	H
80	H	H	D	H	H
81	H	H	D	H	H

The IPI method “Poisoning” can be characterized by the following indicators of assurance about the fact of its implementation: "Low" (L), "Medium" (M), "High" (H), which is estimated on the basis of the value identifying parameters assigned to it (P_{DM}, P_{EL}, P_{II}, P_{CG}, P_{PM}).

Let`s form a set of heuristic rules for detection and identification of "Poisoning" and present it in the form of a table (Table 8). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 8.

P	P_{DM}	P_{PM}	P_{II}	P_{EL}	P_{CG}	Result
1	P	L	U	L	D	L
2	P	L	U	L	D	L
3	P	L	U	L	D	L
4	P	L	U	L	D	L
5	P	L	U	L	D	L
….						
36	D	M	M	M	P	M
37	D	M	M	M	P	M
38	D	M	M	M	P	M
39	D	M	M	M	P	M
....						
77	F	H	D	H	F	H
78	F	H	D	H	F	H
79	F	H	D	H	F	H
80	F	H	D	H	F	H
81	F	H	D	H	F	H

The IPI method “Manipulation” can be characterized by the following indicators of assurance about the fact of its implementation: "Low" (L), "Raised" (R), "High" (H), which is estimated on the basis of the value identifying parameters assigned to it (P_{DM}, P_{PM}, P_{CG}, P_{IF}).

Let`s form a set of heuristic rules for detection and identification of "Manipulation" and present it in the form of a table (Table 9). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 9.

P	P_{DM}	P_{PM}	P_{CG}	P_{IF}	Result
1	P	L	D	L	L
2	P	L	D	L	L
3	P	L	D	L	L
4	P	L	D	L	L
5	P	L	D	L	L
…					
142	D	M	P	M	R
143	D	M	P	M	R
144	D	M	P	M	R
145	D	M	P	M	R
…					
239	F	H	F	H	H
240	F	H	F	H	H
241	F	H	F	H	H
242	F	H	F	H	H
243	F	H	F	H	H

The IPI method “Reframing” can be characterized by the following indicators of assurance about the fact of its implementation: «Uncritical "(U)," Medium "(M)," Critical “(C), which is estimated on the basis of the value identifying parameters assigned to it (P_{DM}, P_{EL}, P_{II}, P_{CG}, P_{IF}).
Let`s form a set of heuristic rules for detection and identification of "Reframing" and present it in the form of a table (Table 9). Taking into account all possible combinations of controlled parameters states, it is possible to form 243 corresponding heuristic rules. Here are some of them.

Table 10.

p	P_{DM}	P_{EL}	P_{II}	P_{CG}	P_{IF}	Result
1	P	L	U	D	L	U
2	P	L	U	D	L	U
3	P	L	U	D	L	U
4	P	L	U	D	L	U
5	P	L	U	D	L	U
….						
142	D	M	M	P	M	M
143	D	M	M	P	M	M
144	D	M	M	P	M	M
145	D	M	M	P	M	M
....						
239	F	H	D	F	H	C
240	F	H	D	F	H	C
241	F	H	D	F	H	C
242	F	H	D	F	H	C
243	F	H	D	F	H	C

It should be noted that there are possible cases when the same situation is agreed upon by several heuristic rules from different groups. In this case a collision occurs. Therefore, it is necessary to introduce a system of priorities. So the priority is a rule with a greater degree of criticality, and in the case of equality of degree - with a greater number of parameters [15].
Detection and identification of information psychological impact will help to select and implement means of countermeasures more effectively, which will help minimize losses from its negative effects. The method of detection and identification of information psychological effects was developed exactly for this purpose. It allows to detect and identify impact on a population of more than 1,000 people. The method uses such methods of fuzzy logic as a method of linguistic terms using statistical data (MELS) – for constructing reference values of parameters and evaluation benchmarks, linear approximation by local maximum (LALM), generalized Heming distance (HD) – to process fuzzy data and carry out fuzzy logic operations. In addition, expert methods of evaluation and ranking are used: the method of average grades (AG).

The method consists of the following steps:

Stage 1. Formation of sets of basic characteristics and information psychological impact methods.

Stage 2. Formation of bundles of basic characteristics with information psychological impact methods.
A set of bundles of the basic characteristics with information psychological impact methods is formed, on the basis of which the identification of a specific type of IPI and the construction of reference values is carried out [16,17].

Stage 3. Fuzzy parameters benchmarks formation.
This stage is aimed at obtaining reference values, which compares to the current values of controlled parameters [18].

Stage 4. Formation of a set of heuristic rules.
Creation of sets of heuristic rules that are used to detect information psychological impact on the basis of comparison of reference and current values of information space parameters using a set of identifiers of the current state, unique to each method of information psychological impact [15].

Stage 5. Pacification of parameters monitored for the purpose of revealing information psychological impact.
At this stage, the transformation of a current values plurality of parameters that are fixed at each t intervals over a certain period of time T in one fuzzy number occurs and thus we obtain fuzzy numbers characterizing the current values of the identifying parameters [13, 16, 17, 19].

Stage 6. Identifying parameters current values processing and result formation.

4. Experimental study of Information psychological impact detection and identification system

The identification of information impact is not currently enough, considering the importance of protecting social groups, coupled with the rapid development of modern means and methods of violating information security. There is a need for its identification, since the choice of countermeasures is more effective for a particular and previously known impact. Therefore, the primary purpose of this system is to predict or detect and identify the IPI. Nowadays, there are no analogous systems. An example is a system for detecting information-psychological impact, which is based on the theory of probability [7]. This approach does not allow to detect unknown methods of impact, to control a slightly formalized space. In addition, such systems require a long-term preparatory stage before they are put into operation. As part of such training, there is usually a sample of statistical data, system training, etc. Therefore, when developing this system, we will use approaches based on fuzzy logic and expert methods that devoid of such disadvantages [14, 20]. The purpose of DIIPIS is to detect and identify the IPI. The input data of the system are IPI identifiers, controlled parameters and their values. At the output of the system - information about IPI, its identifying data.

The architecture of DIIPIS is presented in Fig. 1. It includes such structural elements as: sensor system (SS); the module of initial processing of input parameters, containing the identifying parameters registers (IPR), IPI registers (IPIR), impact-parameter bundle formation block (IPBFB); module of secondary processing of

identifying parameters, consisting of a phasification parameters identification block (PPIB) and the phased parameters sets formation block (PPSFB); module for fuzzy arithmetic operations execution, which includes the current state identifier formation block (CSIFB) and decision block (DB); a module of benchmarks and heuristic rules formation, which includes the corresponding blocks of the same name BFB and HRFD; a result representation module containing a logical conclusion block (LCB) and visualisation block (VB); and mode control module (MCM) that puts the system in a benchmarks correction mode (BCM) or heuristic rules correction mode (HRCM). The SS is located in a cybernetic environment that is fuzzy and slightly formalized. The composition of SS depends on the goals set. The SS monitors the context in the cybernetic space.

In IPBFB are formed bundles of a specific IPI type with the parameters necessary for its detection. So, IPF_i subsets are created for individual IPI methods.

Module of secondary processing of identifying parameters is aimed at the phasification of identifying parameters and their subsequent grouping according to the IPI methods they determine. In PPIB the phasing procedure is carried out, which consists in converting measured current parameters for a certain period in fuzzy terms, resulting in the formation of a subset of $PIPF_i$. In the PPSFB, the already phased identifying parameters are grouped into subsets in accordance with the bonds formed in the IPBFB. In a module of benchmarks formation and heuristic rules, reference values that are needed to measure the current values of controlled parameters and heuristic rules for decision making are formed. BFB intended for experts to create a set of identifying parameters benchmarks. Standards are described using terms as linguistic variables.

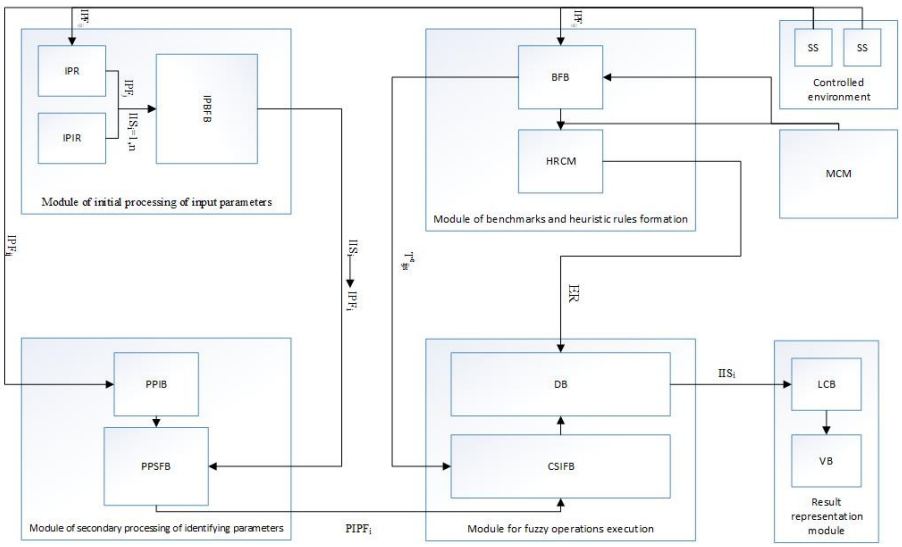


Figure 1. DIIPIS architecture

In the HRCM, in the process of matching the current status identifiers, determined by a combination of values of the current parameters, and the linguistic identifiers of the implementation of the IPI, a set of rules for all the given methods of impact is formed.

The developed standards and heuristic rules are the main expert data that provide the work of DIIPIS. They are set before starting the IPI detection system [21]. There is a possibility of their correction. For this purpose, the MCM is transferred into the system in the BCM or the HRCM, during which the benchmarks and HR can be changed by an expert or system operator. The purpose of fuzzy arithmetic operations module is to compare the current values of the parameters with the benchmarks and determine the HR, which coincides with the current situation, that is, decisions made regarding the existence or possibility of IPI emergence.

In PHIB, identifying parameters are measured and fasificated using the generalized Heming distance method, are compared with the benchmarks, defining the corresponding (the closest) terms of the current situation, and the current state identifier is generated on this basis. In DB, the current states identifier is compared with the sets of the HR, in the process the rule that matches the current identifier is searched. The probability of an IPI emergence is equated with the value of the linguistic identifier of the IPI's implementation of the rule that worked.

The purpose of the module for results representing is to reflect the results obtained in a system in the form understandable to the system operator. The result can be displayed in a linguistic form.

For DIIPIS input parameters are – E_{log} та N_{log}; output parameters – the degree of expert confidence in the decision to identify the fact of IPI realization. The maximum parameters values are determined by the expert depending on the field of application and type of impact, which is the reason of the current situation. In addition, identifying parameters can be described as linguistic variables, such as "low" (H), "medium" (C), "big" (B), etc.

DIIPIS research is performed in a full accordance with the stages of the IPI detection method - a set of identifying parameters is given; their current values are measured and phasified, which are then compared with the benchmarks; the identifiers of the current situation are checked for compliance with the given rules from the set of HR, the result is formed (decision on the possibility of implementation of the IPI is taken).

The experiment was conducted within 10 days (01.05.2017 - 11.05.2017). The participants of the experiment were competent representatives of state authorities, employees of special services of Ukraine, scientists in the field of information warfare. The purpose of the experiment is to detect and identify the information psychological impact on Ukrainian citizens in 2016.

During the examination, a number of information psychological impact methods were detected and identified (Table 11):

Table 11.

The name of the method	Number of detections
Manipulation	14
Psychological isolation	0
Suggestion	19
Persuasion	12
Propaganda	9
Misinformation	5
Reframing	1
Coercion	0
Poisoning	10

5. Conclusions

The results of the work are solving the actual scientific and practical task of developing the method and the system for detection and identification of information and psychological impact.

In the process of implementation, we obtained such significant results:

1. An analysis of modern theories of information confrontation was conducted. It is determined that the subject of information confrontation is widespread in scientific research, but there are no developments in establishing the fact of detection and identifying information psychological impact.
2. Benchmarks of identifying parameters and heuristic rules for identification of information psychological impact have been developed. The method was developed for the first time, and on it's basis a system of information psychological impact. The results obtained in the work can be used to develop countermeasures against the negative informational impacts in the context of information warfare in order to increase the information-psychological safety of citizens, society and the state, as well as the preliminary and current assessment of the congestion level of specific IPIs.
3. The method and system of detection and identification of information psychological impact have been experimentally investigated. The result was the detection and identification of IPI methods in the Ukrainian information space during 2016. Thus, the adequacy of the developed methods and models, the possibility of using the developed system in real conditions was confirmed. In the future it is necessary to continue the IPI study in terms of assessing the criticality of their impact and choosing specific effective countermeasures.

REFERENCES

1. GRIGA V., GNATYUK S., GIZUN A.: Information-psychological security of society as a way of nation preserving, Ukrainian Scientific Journal of Information Security, 21(2015)2, 179-190.
2. GIZUN A., GRIGA V.: Analysis of modern information-psychological impact theories in aspect of information confrontation, Ukrainian Scientific Journal of Information Security, 22(2016)3, 272-282.
3. SHIYAN A.: Methodology of complex security for the person and social groups against the negative information-psychological impact, Ukrainian Scientific Journal of Information Security, 22(2016)1, 94-98.
4. GRYCHUK R., DANIIK Y.: The basis of cybernetic security, ZNAEU, 2016, 636 p.
5. Van NIEKERK B., MAHARAJ M.S.: The Information Warfare Life Cycle Model, SA Journal of Information Management, 13(2011), 1–9.
6. PEW R.: Modeling Human and Organizational Behavior: Application to Military Simulations, National Academy Press, 1998, 418 p.
7. JORMAKKA J.: Modelling Information Warfare as a Game, Journal of Information Warfare, 4(2)(2005)12, 25 p.
8. JOHNSON S.: Toward a Functional Model of Information Warfare, Center for the Study of Intelligence, 8 p.

9. HUTCHINSON B.: Information Warfare: Using the Viable System Model as a framework to attack organizations, Australasian Journal of Information Systems, 9(2002)2, 10p.
10. ROSTORGUEV S.: Mathematical Models in Information Confrontation. Existential Mathematics", Center for strategic assessment and forecasts, 2014, 260 p.
11. MÄNNISTÖ I.: The strategic framing of foreign policy, Uppsala University, 36.
12. KORCHENKO A.: The formation method of linguistic standards created for the intrusion detection system", Ukrainian Information Security Research Journal, 16(2014)1, 5-12.
13. KARPINSKY M., KORCHENKO A., GIZUN A.: Integrated model of crisis situations presentation and formalized procedure for constructing reference identifiers, Legal, regulatory and metrological support of information security in Ukraine, (2015)29, 76 - 85.
14. KORCHENKO O.: Construction of information security systems on fuzzy sets: Theory and practical solutions, MK-Press, 2006, p. 320.
15. GIZUN A., GNATYUK V., SUPRUN O.: Formalized model for building heuristic rules for detecting incidents, Bulletin of Engineering Academy of Ukraine, (2015)1, 110-115.
16. KARPINSKY M., KORCHENKO A., GIZUN A.: Method of detecting incidents - potential crisis situations, Ukrainian Information Security Research Journal, 17(2015)2, 124-130.
17. GIZUN A., GNATYUK V., BALYK N., FALAT P.: Approaches to improve the activity of computer incident response teams, Proceedings of the 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2015, p. 442-447.
18. GIZUN A., VOLYANSKA V., GNATYUK V.: Etalon models of linguistic variables for information security intruders detection and identification, Ukrainian Scientific Journal of Information Security, 19(2013)1, 13-21.
19. KORCHENKO A., GIZUN A., VOLYANSKA V., GNATYUK S.: Method of detecting and identifying an offender in information and communication systems, Ukrainian Information Security Research Journal, Vol.15, №4, 387-393.
20. GIZUN A.: Computer complex for detection and evaluation of crisis situations in information sphere", Ukrainian Information Security Research Journal, 18(2016)1, 66-73.
21. KORCHENKO A., VOLYANSKA V., GIZUN A.: System of intruder detection and identification in information & communication networks, Ukrainian Scientific Journal of Information Security, 19(2013)3, 158-162.

Yuriy HULKA [1], Ruslan KOZAK [2]

Scientific supervisor: Nataliya ZAGORODNA [3]

OTWARTE ZAGADNIENIA DOTYCZĄCE BEZPIECZEŃSTWA INFORMACJI P2PP W DYSTRYBUCJI MULTIMEDIALNEJ

Streszczenie: Systemy P2P napotykają poważne problemy w zakresie łączenia funkcji bezpieczeństwa, prywatności i dostępności w celu rozpowszechniania legalnych treści. Omówiono otwarte problemy związane z dystrybucją informacji w sieci P2P. Przedstawiono koncepcje rozwiązania niektórych problemów bezpieczeństwa w dystrybucji multimediów.

Słowa kluczowe: zabezpieczenia multimedialne, sieć P2P, dystrybucja multimediów

OPEN ISSUES OF P2P INFORMATION SECURITY IN MULTIMEDIA DISTRIBUTION

Summary: P2P systems face serious problems in terms of combining security features, privacy and accessibility for the distribution of legitimate content. Open issues in contend distribution over P2P network were analyzed. The concepts to solve some problems of security in multimedia distribution have been described.

Keywords: multimedia security, P2P network, multimedia distribution

1. Introduction

Nowadays Internet network is transmitting very and very large amount of data. Among them many data are transferred inefficiently around the world. To improve this situation we may use first existed kind of network communication – Peer-to-peer (P2P). In early 2000's existed large P2P networks such as Gnutella, eDonkey2000, Napster or good known to everyone BitTorrent [1].

Peer-to-peer architecture is a commonly used computer networking architecture in which each workstation, or node, has the same capabilities and responsibilities in front

[1] Ternopil Ivan Pul'uj National Technical University, Cybersecurity Department, Bachelor specialty: Information system security, email: yhulka@ymail.com

[2] Ph.D., Assoc. Prof., Ternopil Ivan Pul'uj National Technical University, Cybersecurity department, email: ruslan.o.kozak@gmail.com

[3] Ph.D., Head of Cybersecurity department, Ternopil Ivan Pul'uj National Technical University, email: zagorodna.n @gmail.com

of others. It is often compared and contrasted to the classic client/server architecture, in which some computers are dedicated to serving others, which can lead to bad scalability and single point of failure (Figure 1). This kind of nature makes P2P architecture good in load balancing, fault tolerance, availability, content-based addressing and helps to reach high bandwidth due to the peers being close geographically.

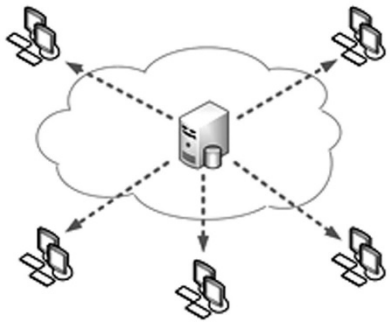


Figure 1. Client Server architecture with single point of failure

P2P networks have many applications, but the most common is for content distribution network (CDN) (Figure 2). CDN serves to ensure that when *file/video/ object/resource* are commonly used by many people, were always available, even when active user count is close to millions. The CDN providers have servers deployed around the globe, and it is its responsibility to decide how to best serve that file to the audience [2].

There are many different types of content delivery networks, depending on the network topology and business relationship. In the star topology CDNs, the content is inserted into the center of the network, and further distributed and replicated into caches of specialized serving servers residing close to the end users [3]. The content is then delivered to the end users from the serving servers.

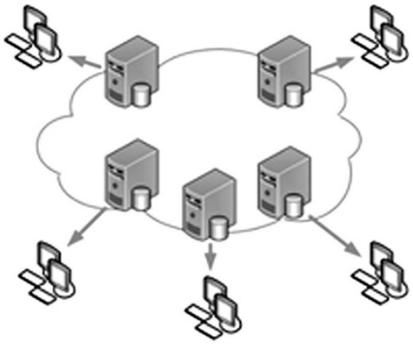


Figure 2. Scheme of content delivery network

Adding P2P to a CDN simply means that now instead of always serving the file directly from the CDN to the end users, the end users can share that file or blocks of

it between them in a peer-to-peer fashion. This reduces the load on the CDN and the bandwidth required on the CDN's side. Caching or broadcast technologies can be used for content distribution. The serving servers are typically located inside the networks of the access connection provider, such as an Internet Service Provider (ISP), and can act as a caching proxy or receive the content they host from, e.g., satellite broadcast.
In the P2P CDNs the content is delivered straight from and to the end user terminals that constitute the delivery network. The advantages of P2P CDNs include independency of central servers enhancing system robustness and network topology enabling more balanced network utilization than in the client-server architectures (Figure 3). The P2P technology-based content delivery systems have obvious advantages however, the pure P2P model is not an applicable architecture choice for a commercial CDN as such in case of copyright infringement.

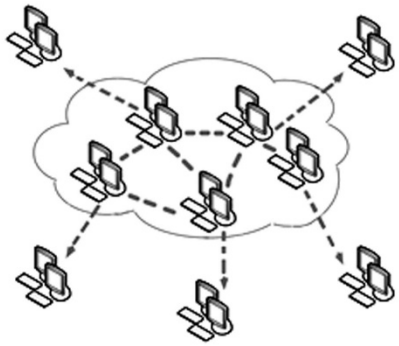


Figure 3. CDN using peer-to-peer network

Detecting a P2P DoS attack is easy; defending against it is difficult. An organization's perimeter defense devices would be overwhelmed by a large attack. Blocking the large number of source IP addresses is time consuming and would still slow the packet processing to a crawl.
Detecting a P2P DoS attack is easy; defending against it is difficult. An organization's perimeter defense devices would be overwhelmed by a large attack. Blocking the large number of source IP addresses is time consuming and would still slow the packet processing to a crawl.

2. Open issues in contend distribution over P2P network

Despite the good prospects the P2P has problems that do not make it good for easy use in modern content delivery networks, and are described as follows:

- Security (provide peers safety and protect an intellectual property);
- Privacy (save content provider and end user privacy, but prevent illegal re-distribution);
- Accessibility (ensuring the correct operation of the P2P network).

Security
P2P network for correct work needs to be installed on user's computer as software application which can lead us to a large number of attacks. This nature needs to make

copyright holders more trustworthy to end-user which can unfairly steal content with some software manipulation. On the other hand hackers also can modify software to be able to intercept user's confidential data and use for their own purposes. The most common security issues are following [4]:

- Eavesdropping (unauthorized node can connect to the network and monitor communications);
- Injection and Modification;
- DDoS Attack;
- Sybil Attack;
- Communication Jamming.

Detecting a copyright infringer is an extremely important task that requires suppliers to be conjunction with watermarking and digital fingerprinting technology, vendors and developers of the P2P CDN. The frequency of peer-to-peer (P2P) enabled denial-of-service (DoS) is quite big. Detecting a P2P DDoS attack is easy, but defending against it is very difficult. An organization that is attacked would be overwhelmed by a large attack. Blocking the large number of source IP addresses is time consuming and would still slow the packet processing to a crawl by blacklist.

One of these was an attack on Dyn DNS servers in Friday, October 21, 2016 which led to interruptions in the work of many sites in North America such as Facebook, Twitter and Amazon AWS. For example, a severe virus known as Antinny, affected the Japanese-based P2P content distribution system "Winny". This virus led to the disclosure of a large amount of U.S. military base security codes along with private documents of a police investigator [5].

Privacy

Open nature of P2P systems makes data privacy a major challenge. Many existing P2P content distribution systems with copyright protection mechanisms monitor activities of the users. These systems are more concerned about security of data rather than protection of user privacy. So there is a need for the development of systems in such a way that don't disclose any personal information of end users if the user does not violate any terms of service provision Few P2P systems have been proposed that address both security concerns of providers and users privacy concerns such as distributed fingerprinted content based multicast, digital rights management (DRM) [6].

Accessibility

Due to the small number of addresses in the IPv4 protocol functioning of pure P2P networks it is impossible due to NAT (network address translation). From the social side widespread use of peer network is hampered by such things as the use of P2P for the spread of illegal content, criminals can also use the network for their own purposes because the network allows for complete anonymity, and it can show P2P from the bad side both to users and to the copyright owners. Since there is no central authority in these systems that can authenticate and protect against malicious end users, it is up to the user to protect itself and be responsible for its own actions.

3. Ways to solve problems of content delivery P2P network

Privacy. Onion routing which used in Tor (anonymity network) is a distributed P2P mechanism that allows two users to communicate anonymously over the network. It

protects its communication against traffic analysis. The main aim of onion routing is to prevent intermediary nodes from knowing the source, destination and contents of the message [7].

Fingerprinting is a technique, outlined in the research by Electronic Frontier Foundation, of anonymously identifying [8]. For example a browser is queried for its agent string, screen color depth, language, installed plugins with supported mime types, time zone offset and other capabilities, such as local storage and session storage (Figure 4). Then these values are passed through a hashing function to produce a fingerprint that gives weak guarantees of uniqueness. No cookies are stored to identify a browser.

Figure 4. Structure of digital fingerprint

Security. *Encryptio*n is a basic content protection technique. Before distribution, the content is encrypted by the provider, and the decryption key is only available to users who have permission to access legitimate copies of the content [9].

Digital watermarking is the method of embedding data into digital multimedia content. This is used to verify the credibility of the content or to recognize the identity of the digital content's owner [2]. In audio and video can be used robust watermarks which include blending signal amplitude with large bandwidths and a short message length. Frequency domain capabilities and mixed-domain techniques, when added to signals, are believed to provide the right amount of robustness in order to guard against attacks on watermarks.

Digital rights management (DRM) is a systematic approach to copyright protection for digital media [10]. The purpose of DRM is to prevent unauthorized redistribution of digital media and restrict the ways consumers can copy content they've purchased. DRM products were developed in response to the rapid increase in online piracy of commercially marketed material, which proliferated through the widespread use of peer-to-peer file exchange programs. Typically DRM is implemented by embedding code that prevents copying, specifies a time period in which the content can be accessed or limits the number of devices the media can be installed on [11].

Accessibility. Hybrid solution between a static server infrastructure and dynamic volunteer P2P resources provides work to ensure the correct operation of the network which is blocked in some counties by ISP. But the transition to a new version of the IPv6 protocol will help to better disseminate this technology.

4. Conclusions

We have not yet fully understood the use of P2P networks to avoid nowadays problems. The concepts to solve some problems have been described. Much of the work has been done by using applications of watermarking, fingerprinting and DRM mechanisms. However, most of the research work involving fingerprinting protocols for copyright protection incurs high computational and communicational burdens due to the use of public-key encryption of the contents, secure multiparty protocols and other techniques. This brief review illustrates that P2P systems face serious challenges in terms of combining security, privacy and accessibility properties for legal content distribution.

REFERENCES

1. JOHNSEN J., KARLSEN L., BIRKELAND S.: Peer-to-peer networking with BitTorrent. Accessed at: http://web.cs.ucla.edu/classes/cs217/05BitTorrent.pdf.
2. LEVENT-LEVI T.: WebRTC P2P CDN: Where are the Use Cases?. Accessed at: https://bloggeek.me/webrtc-p2p-cdn-use-cases/.

A. QURESHI H. RIFÀ-POUS, MEGÍAS D.: Security, Privacy and Anonymity in Legal Distribution of Copyrighted MultimediaContent over Peer-to-Peer Networks: A Brief Overview. Accessed at: http://cv.uoc.edu/webapps/dspace_rei/bitstream/ 10609/38581/1/Qureshi_MINES2013_Security.pdf.

3. WARARKARA P., KAPILB N., REHANIB V., MEHRAB Y., BHATNAGAR Y.: Resolving Problems Based on Peer to Peer Network Security Issue's Accessed at: http://www.sciencedirect.com/science/article/pii/S1877050916001150.
4. INGRAM M.: 66,000 Names and Personal details leak on Peer-to-Peer, Accessed at: http://www.slyck.com/news.php?story=1169.
5. DOMINGO-FERRER J., MEGÍAS D.: Distributed multicast of fingerprinted content based on a rational Peer-to-Peer community, Comput. Commun., 2013, vol. 36, no. 5, pp. 542-550.
6. POUWELSE J.A., GARBACKI P., EPEMA D.H.J., SIPS H.J.: The bitt orrent p2p file-sharing system: measurements and analysis. Accessed at: http://www.divms.uiowa.edu/~ghosh/bittorrent1.pdf.
7. COURTRIGHT M., PASHUPATI K.: The Impact of Digital Fingerprinting and Identity Verification on Data Quality. Accessed at: http://www.journalofadvertisingresearch.com/ content/ 54/3/263.
8. PAKKALA D., LATVAKOSKI J. : Towards a Peer-to-Peer Extended Content Delivery Network. Accessed at: https://www.eurasip.org/Proceedings/Ext/IST05/papers/ 99.pdf.
9. LIU ., SAFAVI-NAINI R., SHEPPARD N.P.: Digital Rights Management for Content Distribution. Accessed at: http://crpit.com/confpapers/CRPITV21ALiu.pdf.
10. JAE-YOUN SUNG, JEONG-YEON JEONG, KI-SONG YOON: DRM Enabled P2P Architecture. Advanced Communication Technology, 2006. ICACT 2006. The 8th International Conference.

Igor IAKYMENKO[1], Stepan IVASIEV[2]

Scientific Supervisor: Mykhajlo KASIANCHUK[3]

TEORETYCZNE PODSTAWY BUDOWY PIĘCIOMODUŁOWEJ ZMODYFIKOWANEJ POSTACI DOSKONAŁEGO SYSTEMU KLAS RESZTKOWYCH

Streszczenie: W artykule przedstawiono teoretyczne podstawy budowy pięciomodułowego zmodyfikowanej postaci doskonałego systemu klas resztkowych. Wykorzystanie tej postaci zasadniczo zmniejsza złożoność obliczeń przy tłumaczeniu liczb z systemu klas resztkowych do systemu pozycyjnego obliczeń unikając wyszukiwania modułów odwrotnych i mnożąc je na liczby bazowe. Podano przykład konstrukcji pięciomodułowej zmodyfikowanej postaci doskonałego systemu klas resztkowych. Pokazano, że w zależności od wyboru systemu modułów bitów liczb, po których będą wykonywane operacje arytmetyczne, złożoność spada 2-3-krotnie, co jest szczególnie ważne w przypadku obliczeń z wielobitowymi liczbami. Zbudowano i przeanalizowano wykresy zależności modułów i możliwego zakresu obliczeń. Ustalono, że największy zakres obliczeń dla pierwszego modułu oraz ich ilości wystąpi wtedy, gdy wartości bezwzględne wszystkich kolejnych modułów na jednostkę są większe od iloczynu wartości bezwzględnych poprzednich modułów.

Słowa kluczowe: system klas resztkowych, zmodyfikowana doskonała postać, odwrotny element modułu, bity liczb, system modułów

THEORETICAL FOUNDATIONS FOR CREATING FIVE MODULAR MODIFIED PERFECT FORM OF THE SYSTEM OF RESIDUAL CLASSES

Summary: The theoretical foundations for creating five modular modified perfect form of the system of residual classes are presented in the article. The use of this form significantly reduces the computational complexity in transferring numbers from the system of residual classes into positional number system by avoiding procedures for searching inverse element by the module and in multiplying basis numbers. An example of creating five modular modified perfect form

[1] Ternopil National Economic University, Department of Computer Engineering, Ph.D., Associate Professor, iyakymenko@ukr.net

[2] Ternopil National Economic University, Department of Computer Engineering, Ph.D., Associate Professor, stepan.ivasiev@gmail.com

[3] Ternopil National Economic University, Department of Computer Engineering, Ph.D., Associate Professor, kasyanchuk@ukr.net

of the system of residual classes is presented.
It is shown that the bit of numbers, over which arithmetic operations are performed, is reduced by 2-3 times depending on the choice of the system of modules that is especially important in the calculation of multi-digit numbers. Graphs of dependence modules and the possible range of calculations are constructed and analyzed. It was established that the greatest range of calculations at given first module and their quantity will be when the absolute values of all following modules are on a unit larger than the product of the absolute values of the previous.

Keywords: system of residual classes, modified perfect form, inverse element by the module, bit of numbers, system of modules.

1. Formulation of the problem

Due to the significant increase in volumes of calculations and increasing bit capacity of numbers [1] that are used for calculations, shortcomings of the most common at this time binary number system (for example, its multi-digit, strictly sequential structure, availability of between digits hyphenation [2], etc.) which largely slow down performing arithmetic operations.
Parallelization of the process for information processing is the most promising way to improve performance of modern computational systems. Some position-independent number system, including the system of residual classes (SRC) [3], which is one of the alternatives to the binary representation of numbers have this property. It enables applying new approaches to organizing calculation when performing basic mathematical operations [4]. Although SRC has its drawbacks, which in particular include the complexity of dividing ([5] and comparing numbers [6], the need to determine the conditions of overflow of digit grid, but it can be successfully used for addition, subtraction, exponentiation, multiplication of integer multi-digit numbers [7], which is very important, especially in asymmetric cryptography (algorithms RSA, El Gamal [8], in an electronic digital signature, the use of elliptic curves [9], etc.), in developing new tools and algorithms for protected from interference coding [10], in improving reliability of data control [11], for large matrix calculations [12] and other tasks of applied and discrete mathematics.
Undoubted advantage of SRC is the ability to perform operations on numbers that are smaller than the selected module, the parallelization of the calculation process, which is the most promising way to improve the performance of computing systems and the absence of transfers between digits.

2. Analysis of recent research and publications

The fundamental basis of the system of residual classes (SRC) is number theory [13]. Any integer decimal number N_B of SRC is represented as a set ($b_1, b_2, ..., b_n$) of the least positive residue of dividing this number into fixed natural pairwise co-prime numbers $p_1, p_2, ..., p_n$ ($b_i = N \bmod p_i$), which are called the modules (n – number of module). An inequality $0 \le N < P-1$ should be performed, where $P = \prod_{i=1}^{n} p_i$ – number that determines the condition of the overflow of digit grid. Arithmetic operations in SRC

(except dividing and comparing numbers) occur independently in each module without transfers between digits.
Inverse transformation of SRC in the decimal number system is based on the use of Chinese Remainder Theorem and is rather cumbersome process that is another drawback of SRC, which kept its development and distribution:

$$N = \left(\sum_{i=1}^{k} b_i B_i \right) \bmod P \ , \tag{1}$$

where $B_i=M_i m_i$, $M_i = \frac{P}{p_i}$, $m_i = M_i^{-1} \bmod p_i = 1$.

Currently, the most common are three search methods for inverse element by module: an exhaustive search of possible options, the use of Euler's theorem and the extended Euclidean algorithm. All of them are characterized by considerable computational complexity; require significant consumption of time and computing resources when performing module operation, exponentiation, finding function of Euler, etc. Moreover, these operations must be performed on very large numbers, which may lead to an overflow of digit grid.

A perfect form (PF) of SRC, where the condition $M_i \bmod p_i = 1$ is performed thereby avoiding search procedure of inverse element and multiplying on it into (1), is described in [14]. The problem was solved in [15] and the conditions for analytical finding m_i were determined. However, in both cases, the value p_i rapidly increase, which is unacceptable when it is necessary to use modules of the same digits. The theoretical basis of the modified perfect form (MPF) of SRC was developed in [16], where the following equation is performed:

$$M_i \bmod p_i = \pm 1 \, , \tag{2}$$

that also eliminates the search operation of inverse element, and the proposed method for construction MPF of SRC from three modules on the example p_2-p_1=5. However, at present there are no universal methods for constructing a system with more modules that satisfy the conditions MPF of SRC.

3. Problem definition

Based on the foregoing, the aim of our work is the further development of theory and methods for analytical determination the sets of modules for MPF of SRC and determining conditions that allow constructing all the possible options for a given number of modules of MPF of SRC.

4. Statement of the main material

Let us write the expression (2) in the form of the system:

$$\begin{cases} M_1 \bmod p_1 = \pm 1 \\ \ldots \\ M_n \bmod p_n = \pm 1. \end{cases} \tag{3}$$

Multiplying each equation on the corresponding module, we'll get:

$$\begin{cases} P \bmod p_1^2 = \pm p_1 \\ \ldots \\ P \bmod p_n^2 = \pm p_n. \end{cases} \tag{4}$$

Solving the expression (4) by standard methods of number theory according to the Chinese Remainder Theorem, we'll have:

$$P = \left(\sum_{i=1}^{n} p_i M_i^2 m_i^2 \right) \bmod M \ , \tag{5}$$

where $M = \prod_{i=1}^{n} p_i^2 = P^2$.

Taking into account that PF of SRC m_i=±1, and reducing module, the left and right part (5) on their common divisor $P = \prod_{i=1}^{n} p_i$, we'll write (5) as follows:

$$\left(\sum_{i=1}^{n} M_i \right) \bmod P = \pm 1 \ . \tag{6}$$

Expression (6) is equivalent to the equality:

$$\sum_{i=1}^{n} M_i = kP \pm 1 \, , \tag{7}$$

where k=±1, ±2, ±3, ...

Dividing the left and right parts (7) on $P=p_1p_2p_3$, we'll get the final expression for finding a set of modules in PF of SRC:

$$\sum_{i=1}^{n} \frac{1}{p_i} = k \pm \frac{1}{\prod_{i=1}^{n} p_i} \, . \tag{8}$$

In contrast to the PF of SRC where all modules are positive and therefore k> 0, modules have different signs in MDF of SRC and for simplification of task you can take k=0, that is correspond to the largest range of calculations for a given number of modules. Thus, the equation (8) is in such form:

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \ldots + \frac{1}{p_{n-1}} + \frac{1}{p_n} = \pm \frac{1}{p_1 p_2 p_3 \cdots p_{n-1} p_n} \, . \tag{9}$$

It should be noted that the positive values of modules p_i correspond to the conditions m_i=1 and in the condition m_i=-1 are negative. In addition, the first two modules are strictly defined in PF of SRC (p_1=2, p_2=3) [23], they can be any in MPF of SRC. For the demonstration of method and simplifying the calculations we confine ourselves with the value of the first module p_1=3, the number of modules n = 5 and without reducing generality solutions we'll consider that

$$p_1=3<|p_2|<|p_3|<|p_4|<|p_5|. \quad (10)$$

The numerical calculations show that fulfillment of this condition for integer solutions (9) is possible only when p_2, $p_3<0$.

The surface appearance characterizing dependence of the module p_5 from p_3 and p_4 according to the expression $p_5=\frac{-1-3p_2p_3p_4}{p_2p_3p_4+3(p_3p_4+p_2p_4+p_2p_3)}$, which is obtained from (9), when $p_1=3$ and $p_2=-5$ is shown for example in Figure 1.

Then, considering last two modules p_4 and p_5 as unknown, from (9) we can obtain Diophantine equation of the second order for their search:

$$p_4p_5(p_2p_3+3(p_2+p_3))+3p_2p_3(p_4+p_5)=\pm1. \quad (11)$$

Let us introduce the notation:

$$p_{4,5}=\frac{a,b-3p_2p_3}{p_2p_3+3(p_2+p_3)}. \quad (12)$$

After substituting (12) into (11) and corresponding mathematical transformations we can obtain an expression for integer solution (12):

$$\pm(p_2p_3+3(p_2+p_3))+(3p_2p_3)^2=ab. \quad (13)$$

This means that the left part (13) should be factorized, based on what parameters *a and b* are defined. In addition, modules p_4 and p_5 must be integers. Therefore, from (12) the next follows:

$$(a,b-3p_2p_3)\bmod(|p_2p_3+3(p_2+p_3)|)=0. \quad (14)$$

Expressions (13) and (14) determine the conditions for finding five modules MPF of SRC, two of which are unknown.

For p_2=-4, p_3=-7 expressions (12)-(14) are transformed as follows: $p_{4,5}=\frac{a,b-84}{-5}$; $(a,b-84)\bmod 5=0$; $ab=\pm5+7056=\begin{cases}7051=11\cdot641\\7061=23\cdot307.\end{cases}$

All possible position-independent integer values *a* and *b* that are defined by factorization of *a* and *b*, and cases where there are sets of MPF of SRC modules and a range of appropriate calculations are shown in Table 1.

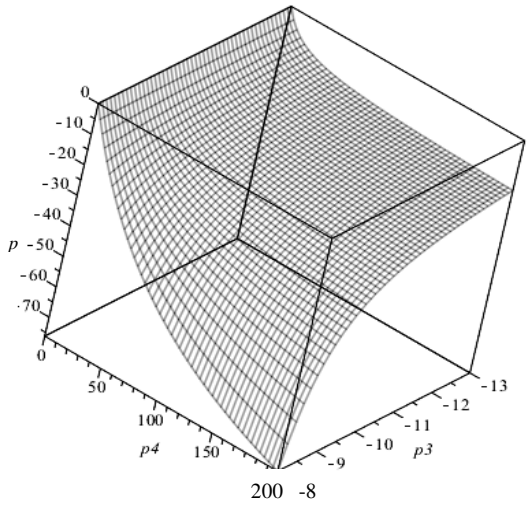


Figure 1. The surface appearance characterizing dependence of the module p_5 from p_3 and p_4 for p_2=-5

Table 1. Possible variants of systems with five modules of MPF of SRC for p_1=3, p_2=-4, p_3=-7 (in brackets - digits in the binary system)

№	p_1, p_2	*ab*	*a*	*b*	p_4	p_5	*P*
1	p_1=3(2) p_2=-4(3) p_3=-7(3)	7051	1	7051	does not exist		
2			-1	-7051	17(5)	1427(11)	2037756(21)
3			11	641	does not exist		
4			-11	-641	19(5)	145(8)	231420(18)
5		7061	1	7061	does not exist		
6			-1	-7061	17(5)	1429(11)	2040612(21)
7			23	23	does not exist		
8			-23	-23	does not exist		

It can be seen in Table 1 that the modules p_4 and p_5 acquire positive values and order of numbers, which arithmetic operations are performed on, is reduced by about half. In addition, in five out of eight possible cases, which are formed during factorization, there are no integer module sets. The value $p_4 = 17$ corresponds to two values of the module p_5, each of which differs by one from the product of the previous four modules.

Obviously, that the most finding sets will be when the modules p_1, p_2, p_3 form themselves MPF of SRC, as in this case $p_1p_2 + p_2p_3 + p_1p_3 = \pm 1$, and condition (14) is always performed.

The number of different options will be determined by the number of multipliers at the factorization of left part (13). Such cases are appropriate to consider more details.

For p_2=-4, p_3=-11 we can obtain from the expressions (12)-(13): $p_{4,5}=\frac{a,b-132}{-1}$; For p_2=-4, p_3=-11 we can obtain from the expressions (12)-(13): $ab=\pm1+17424=\begin{cases}17423=7\cdot19\cdot131\\17425=5\cdot5\cdot17\cdot41.\end{cases}$. All possible options of modules and the range of calculations are presented in Table 2.

The analysis of Table 2 shows that values p_4 acquire only positive values and the sign p_5 is opposite to the sign *a and b*. Bit of numbers which arithmetic operations will be performed on are reduced by 2-2.5 times. Line 7, wherein p_4, $p_5 = \pm 1$, shows that this set of three modules $p_1 = 3$, $p_2 = -4$, $p_3 = -11$ forms of MPF of SRC. The values p_4=131 and p_4=133 correspond to two values of the module p_5, absolute values of which differ by one from the product of previous four modules.

Table 2. Possible options of systems from five modules for MPF of SRC for p_1=3, p_2=-4, p_3=-11 (in brackets – the bit in the binary system)

№	p_1, p_2, p_3	*ab*	*a*	*b*	p_4	p_5	*P*
1	3 (2), -4 (3), -11 (4)	17423	1	17423	131 (8)	-17291 (15)	298995972 (29)
2			-1	-17423	133 (8)	17555 (15)	308195580 (29)
3			7	2489	125 (7)	-2357 (12)	38890500 (26)
4			-7	-2489	139 (8)	2621 (12)	48090108 (26)
5			19	917	113 (7)	-785 (10)	11709060 (25)
6			-19	-917	151 (8)	1049 (11)	20908668 (25)
7			131	133	1 (1)	-1(1)	132 (8)
8			-131	-133	263 (9)	265 (9)	9199740 (24)
9		17425	1	17425	131 (8)	-17293 (15)	299030556 (29)
10			-1	-17425	133 (8)	17557 (15)	308230692 (29)
11			5	3485	127 (7)	-3353 (12)	56209692 (26)
12			-5	-3485	137 (8)	3617 (12)	65409828 (26)
13			17	1025	115 (7)	-893 (10)	13555740 (24)
14			-17	-1025	149 (8)	1157 (11)	22755876 (25)
15			25	697	107 (7)	-565 (10)	7980060 (23)
16			-25	-697	157 (8)	829 (10)	17180196 (25)
17			41	425	91 (7)	-293 (9)	3519516 (22)
18			-41	-425	173 (8)	557 (10)	12719652 (21)
19			85	205	47 (6)	-73 (7)	452892 (19)
20			-85	-205	217 (8)	337 (9)	9653028 (23)

A set of p_1=3, p_2=-4, p_3=-13 also forms MPF of SRC. That is why from the expression (12)-(13) we can obtain the followings: $p_{4,5} = \frac{a,b-156}{1}$; $ab = \pm 1 + 24336 = \begin{cases} 24335 = 5 \cdot 31 \cdot 157 \\ 24337 - \text{prime number.} \end{cases}$ All possible options of modules, the range of possible computation and corresponding bit are presented in Table 3. From the Table 3 follows that the module p_4 takes only negative values and the sign p_5 coincides with the sign a and b. Bit of numbers, over which arithmetic operations are performed, is reduced in 2-2.5 times. Line 7 shows that this set of three modules form MPF of SRC. The value p_4=-155 and p_5=-157 corresponds to two values of module p_5, that is different from the product of the absolute values of the previous four modules per unit.

In the latter two cases for conducting further researches for division of the absolute values of modules they must be renumbered in ascending order $|p_4|$, which are presented in Tables 4, 5.

Table 3. Possible options of systems from five modules for MPF RNS for p_1=3, p_2=-4, p_3=-13 (in brackets – the bit in the binary system).

№	p_1, p_2, p_3	ab	a	b	p_4	p_5	P
1	3 (2), -4 (3), -13 (4)	24335	1	24335	-155 (8)	24179 (15)	584648220 (30)
2			-1	-24335	-157 (8)	-24491 (15)	599833572 (30)
3			5	4867	-151 (8)	4711 (13)	110972316 (27)
4			-5	-4867	-161 (8)	-5023 (13)	126157668 (27)
5			31	785	-125 (7)	629 (10)	12265500 (24)
6			-31	-785	-187 (8)	-941 (10)	27450852 (25)
7			155	157	-1 (1)	1(1)	156 (8)
8			-155	-157	-311 (9)	-313 (9)	15185508 (24)
9		24337	1	24337	-155 (8)	24181 (15)	584696580 (30)
10			-1	-24337	-157 (8)	-24493 (15)	599882556 (30)

Table 4. Ordering modules in ascending $|p_4|$ for p_1=3, p_2=-4, p_3=-11

№	1	2	3	4	5	6	7	8	9	10
p_4	1	47	91	107	113	115	125	127	131	131
p_5	1	73	293	565	785	893	2357	3353	17291	17293
№	11	12	13	14	15	16	17	18	19	20
p_4	133	133	137	139	149	151	157	173	217	263
p_5	17555	17557	3617	2621	1157	1049	829	557	337	265

Table 5. Ordering modules in ascending $|p_4|$ for p_1=3, p_2=-4, p_3=-13

№	1	2	3	4	5	6	7	8	9	10
p_4	1	125	151	155	155	157	157	161	187	311
p_5	1	629	4711	24179	24181	24492	24493	5023	941	313

The nature of changes the values of modules p_4 and p_5 is shown in Figure 2 depending on the number of the module according to Tables 4, 5 in the logarithmic scale with the basis 2, indicating the bit of received modules in the binary system.

As seen in Figure 2, in both cases the value $|p_4|$, is growing relatively slowly. At the same time, the graph for $|p_5|$ growing much more rapidly and reaches to flat maximum in the middle of numbering range of modules, and then falling to value $|p_4|$.

A set of modules p_1=3, p_2=-5, p_3=-7 also forms MPF of SRC, so expressions (12)-(13) are written in such a way: $p_{4,5} = \frac{a,b-105}{-1} = 105 - a,b$ and $ab = \pm 1 + 11025 = \begin{cases} 11024 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 13 \cdot 53 \\ 11026 = 2 \cdot 37 \cdot 149. \end{cases}$ All possible options of systems from five modules for MPF of SRC for p_1=3, p_2=-5, p_3=-7 are presented in Table 6.

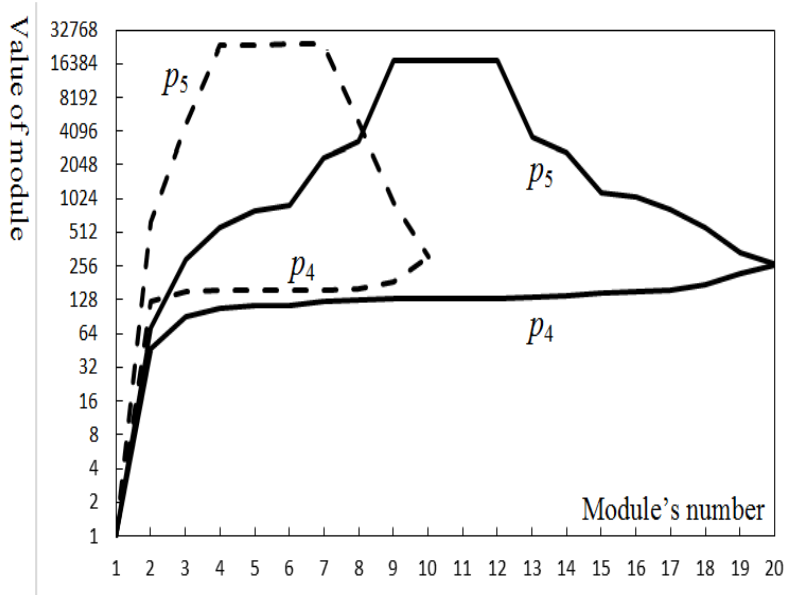


Figure 2. The nature of changes the values of modules p_4 and p_5 for p_1=3, p_2=-4, p_3=-11 (solid line) and p_3=-13 (dotted line) depending on the number of the module according to Tables 4, 5

Numerical experiment results presented in Table 6 show that the module p_4 acquires positive values and the sign p_4 and p_5 is opposite to the sign *a* and *b*. Bit of numbers, over which arithmetic operations will be performed, is reduced in 2-3 times. Line 17, wherein p_4, p_5=±1, shows that this set of three modules p_1=3, p_2=-5, p_3=-7 forms MPF

of SRC. The value p_4 = 103, 104, 106, 107 corresponds to two values of module p_5, which is caused by the presence of joint multipliers 1 and 2 in both cases of the decomposition of product *ab*.

Modules p_1=3, p_2=-5, p_3=-8 also form the MPF of SRC, so expressions (12)-(13) lead to the following results: $p_{4,5} = a,b - 120$ and $ab = \pm 1 + 14400 = \begin{cases} 14399 = 7 \cdot 11 \cdot 11 \cdot 17 \\ 14401 - \text{prime number.} \end{cases}$

All possible options of systems from five modules for MPF of SRC when p_1=3, p_2=-5, p_3=-8 are presented in Table 7.

From the Table 7 you can see that the module p_4 acquires only negative values and the sign p_5 coincides with the sign *a* and *b*. Bit of numbers, over which arithmetic operations are performed, is reduced in 2-3 times.

Line 7 shows that the set p_1=3, p_2=-5, p_3=-8 forms MPF of SRC. The value p_4=-119 p_5=-121 corresponds to two values of the module p_5, which differ per unit from the product of absolute values of the previous four modules.

Table 6. Possible options of systems from five modules for MPF of SRC for p_1=3, p_2=-5, p_3=-7 (in brackets – the bit in the binary system)

№	p_1, p_2, p_3	*ab*	*a*	*b*	p_4	p_5	*P*
1	3 (2), -5 (3), -7 (3)	11024 (14)	1	11024	104 (7)	-10919 (14)	119235480 (27)
2			-1	-11024	106 (7)	11129 (14)	123865770 (27)
3			2	5512	103 (7)	-5407 (13)	58476705 (26)
4			-2	-5512	107 (7)	5617 (13)	63106995 (26)
5			4	2756	101 (7)	-2651 (12)	28113855 (25)
6			-4	-2756	109 (7)	2861 (12)	32744145 (25)
7			8	1378	97 (7)	-1273(11)	12965505 (24)
8			-8	-1378	113 (7)	1483 (11)	17595795 (25)
9			13	848	92 (7)	-743 (10)	7177380 (23)
10			-13	-848	118 (7)	953 (10)	11807670 (24)
11			16	689	89 (7)	-584 (10)	5457480 (23)
12			-16	-689	121 (7)	794 (10)	10087770 (24)
13			26	424	79 (7)	-319 (9)	2646105 (22)
14			-26	-424	131 (8)	529 (10)	7276395 (23)
15			52	212	53 (6)	-107 (7)	595455 (20)
16			-52	-212	157 (8)	317 (9)	5225745 (23)
17			104	106	1 (1)	-1 (1)	105 (7)
18			-104	-106	209 (8)	211 (8)	4630395 (23)
19		11026 (14)	1	11026	104 (7)	-10921 (14)	119257320 (27)
20			-1	-11026	106 (7)	11131 (14)	123888030 (27)
21			2	5513	103 (7)	-5408 (13)	58487520 (26)
22			-2	-5513	107 (7)	5618 (13)	63118230 (26)
23			37	298	68 (7)	-193 (8)	1378020 (21)

24			-37	-298	142 (8)	403 (9)	6008730 (23)
25			74	149	31 (5)	-44 (6)	143220 (18)
26			-74	-149	179 (8)	254 (8)	4773930 (23)

As in the case p_1=3, p_2=-4, ordering modules in ascending the absolute value of the module p_4 is presented in Tables 8 and 9.

The character of changes of the values of modules p_4 and p_5 in depending on the number of module according to Tables 8, 9 in a logarithmical scale with the base 2, indicating digit capacity of received modules in the binary system, is presented in Figure 3. The nature of the relevant graphs that is presented in Figures 2 and 3 are similar, but maximum flat in Figure 3 has more values.

Ordering the value of the range of calculations according to tables 4, 5, 8, 9 in ascending absolute value module p_4 is presented in Table 10.

Table 7. Possible options of systems from five modules for MPF of SRC for p_1=3, p_2=-5, p_3=-8 (in brackets – the bit in the binary system).

№	p_1, p_2, p_3	*ab*	*a*	*b*	p_4	p_5	*P*
1	3 (2),	14399	1	14399	-119 (7)	14279 (14)	203904120 (28)
2	-5 (3),	(14)	-1	-14399	-121 (7)	-14519 (14)	210815880 (28)
3	-8 (3)		7	2057	-113 (7)	1937 (11)	26265720 (25)
4			-7	-2057	-127 (7)	-2177 (12)	33177480 (25)
5			11	1309	-109 (7)	1189 (11)	15552120 (24)
6			-11	-1309	-131 (8)	-1429 (11)	22463880 (25)
7			17	847	-103 (7)	727(10)	8985720 (24)
8			-17	-847	-137 (8)	-967 (10)	15897480 (24)
9			77	187	-43 (6)	67 (7)	345720 (19)
10			-77	-187	-197 (8)	-307 (9)	7257480 (23)
11			119	121	-1 (1)	1 (1)	120 (7)
12			-119	-121	-239 (8)	-241 (8)	6911880 (23)
13		14401	1	14401	-119 (7)	14281 (14)	203932680 (28)
14		(14)	-1	-14401	-121 (7)	-14521 (14)	210844920 (28)

Table 8. Ordering modules in ascending $|p_4|$ for p_1=3, p_2=-5, p_3=-7

№	1	2	3	4	5	6	7	8	9
p_4	1	31	53	68	79	89	92	97	101
p_5	1	44	107	193	319	584	743	1273	2651
№	10	11	12	13	14	15	16	17	18
p_4	103	103	104	104	106	106	107	107	109
p_5	5407	5408	10919	10921	11129	11131	5617	5618	2861
№	19	20	21	22	23	24	25	26	
p_4	113	118	121	131	142	157	179	209	
p_5	1483	953	794	529	403	317	254	211	

Table 9. Ordering modules in ascending $|p_4|$ for $p_1=3$, $p_2=-5$, $p_3=-8$

№	1	2	3	4	5	6	7
p_4	1	43	103	109	113	119	119
p_5	1	67	727	1189	1937	14279	14281
№	8	9	10	11	12	13	14
p_4	121	121	127	131	137	197	239
p_5	14519	14521	2177	1429	967	307	241

The graph of dependency of the range of calculations according to the number in Table 10 for various values p_2, p_3 is shown in Figure 4. Similarly to graphs for p_5 in Figures 3, 4 dependency of the range of calculations P initially increases sharply; in the middle number range has a flat maximum that for fixed p_2 is placed higher when larger $|p_3|$ The graphics fall slowly at further increase of the number.

Table 10. Ordering the value of the range of calculations P in ascending absolute value p_4 for different values p_2, p_3 and $p_1=3$

№	P (p_2=-4, p_3=-11)	P (p_2=-4, p_3=-13)	P (p_2=-5, p_3=-7)	P (p_2=-5, p_3=-8)
1	132	156	105	120
2	452892	12265500	143220	345720
3	3519516	110972316	595455	8985720
4	7980060	584648220	1378020	15552120
5	11709060	584696580	2646105	26265720
6	13555740	599833572	5457480	203904120
7	38890500	599882556	7177380	203932680
8	56209692	126157668	12965505	210815880
9	298995972	27450852	28113855	210844920
10	299030556	15185508	58476705	33177480
11	308195580		58487520	22463880
12	308230692		119235480	15897480
13	65409828		119257320	7257480
14	48090108		123865770	6911880
15	22755876		123888030	
16	20908668		63106995	
17	17180196		63118230	
18	12719652		32744145	
19	9653028		17595795	
20	9199740		11807670	
21			10087770	
22			7276395	
23			6008730	
24			5225745	
25			4773930	
26			4630395	

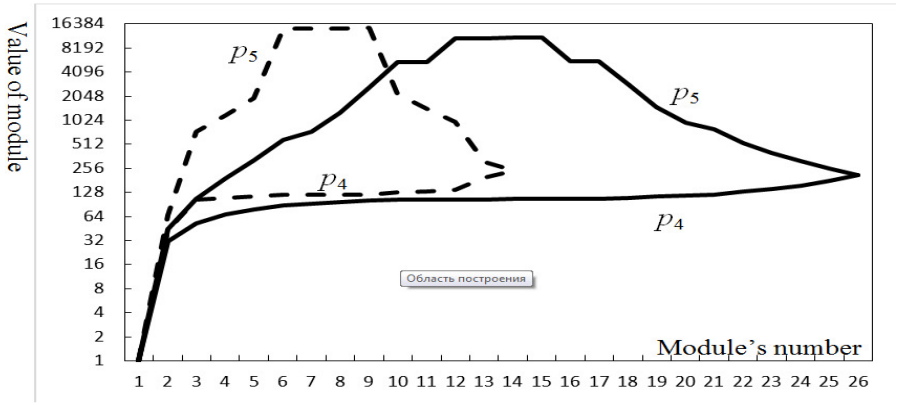


Figure 3. The variation of the values of the modules p_4 and p_5 for p_1=3, p_2=-5, p_3=-7 (solid line) and p_3=-8 (dotted line) depending on the number of module according to Tables 7, 8

The other possible sets of modules for different values p_2, p_3, that form of MPF of SRC, are shown in Table 11 for p_1=3 according to similar numerous studies.

Table 11. The other possible sets of modules, that form of MPF of SRC for p_1=3.

№	p_2, p_3	p_4	p_5	P
1	-4 (3), -17 (5)	-41 (6)	-8363 (14)	69948132 (27)
2		-41 (6)	-8365 (14)	69964860 (27)
3	-5 (3), -11 (4)	-28 (5)	-149 (8)	688380 (20)
4		-23 (5)	949 (10)	3601455 (22)
5		-19 (5)	98 (7)	307230 (19)
6		-17 (5)	61 (6)	171105 (18)
7		-13 (4)	29 (5)	62205 (16)
8	-7 (3), -10 (4)	-11 (4)	2309 (12)	5333790 (23)
9		-11 (4)	2311 (12)	5338410 (23)

It was demonstrated in Table 11 that the number of sets of five modules is significantly reduced if the first three of them do not form MPF of SRC. This is due to the need to fulfill the condition (14). Bit of numbers, over which arithmetic operations will be performed, is reduced in 2-3 times. The largest range of calculations for a given number of modules will be when the absolute value of each subsequent module is by one bigger than the product of the absolute values of the previous ones.

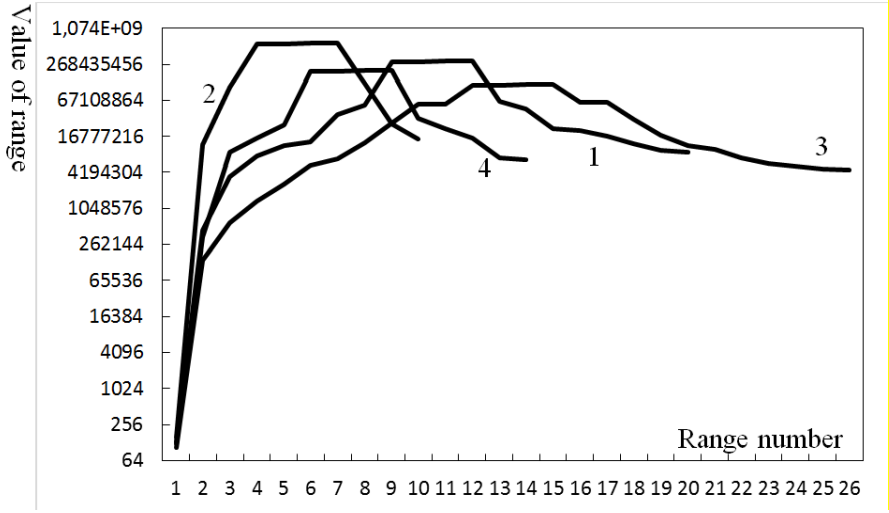


Figure 4. The graph of dependency of the range of calculations according to the number in Table 9 for various values p_2, p_3(1 - p_2=-4, p_3=-11, 2- p_2=-4, p_3=-13, 3 - p_2=-5, p_3=-7, 4 - p_2=-5, p_3=-8)

5. Conclusions

An analysis of the process of performing arithmetic operations on multi-digit numbers in computation systems established that the most promising way to improve the performance of information processing is the ability to parallelization, which is inherent in the system of residual classes. However, one of its major drawbacks is the complexity of reverse transfer to the decimal number system. The paper shows that the use of a modified perfect form greatly simplifies this procedure by avoiding the search operation of inverse element by module and multiplying on basic numbers.
An example of construction of five modular modified perfect form of the system of residual classes is presented. It is shown that depending on the choice of the system of modules the bit of numbers, over which arithmetic operations will performed, is reduced in 2-3 times, which is especially important for use of multi-digit numbers. The graphs of dependence modules and the possible range of calculations are constructed and analyzed. It was established that the largest range of calculations for a given number of modules will be obtained when the absolute value of each subsequent module is by one bigger than the product of the absolute values of the previous ones.

REFERENCES

1. MIRZAEI O., YAGHOOBI M., IRANI H.: A new image encryption method, Issue parallel sub-image encryption with hyper chaos. Nonlinear Dynamics, (2012), 67(1), 557-566.

2. PARHAMI B.: Computer Arithmetic Algorithms and Hardware Architectures. Oxford University Press, New York 2010.
3. OMONDI A., PREMKUMAR B.: Residue Number Systems. Theory and Implementation. Imperial College Press, London 2007.
4. DHANABAL R., SARAT KUMAR SAHOO, BARATHI V., NAAMATHEERTHAM R. SAMHITHA NEETHU ACHA CHERIAN, PRETTY MARIAM JACOB: Implementation of Floating Point Mac Using Residue Number System. Journal of Theoretical and Applied Information Technology, (2014) 62(2), 458-463.
5. YANG J., CHANG C., CHEN A.: High-speed division algorithm in residue number system using parity. International Journal of Computer Mathematics, (2004) 81(6), 775-780.
6. KRASNOBAYEV V.A., YANKO A.S., KOSHMAN S.A.: A Method for Arithmetic Comparison of Data Represented in a Residue Number System. Cybernetics and Systems Analysis, (2016) 52(1), 145-150.
7. KOZACZKO D., IVASIEV S., YAKYMENKO I., KASIANCHUK M.: Vector Module Exponential in the Remaining Classes System. Proceedings of the 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2015), Warsaw, Poland, (2015), 161-163.
8. KOVALCHUK A., BORZOV Y., PELESHKO D.: A blend of algorithms RSA and bit, additive-difference operations and algorithms in El-Gamal images. Journal of Global Research in Computer Science, (2013), 1-7.
9. AHMED A., EL-LATIF A., XIAMU NIU: A hybrid chaotic system and cyclic elliptic curve for image encryption. AEU-International Journal of Electronics and Communications, (2013) 67(2), 136-143.
10. YATSKIV V., YATSKIV N., JUN S., SACHENKO A., ZHENGBING H.: The Use of Modified Correction Code Based on Residue Number System in WSN. Proceedings of the 7-th2013 IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS'2013), Berlin, Germany, (2013), 513-516.
11. KRASNOBAYEV V., KOSHMAN S., MAVRINA M.: A Method of Increasing the Reliability of Verification of Data Represented in a Residue Number System. Cybernetics and Systems Analysis, (2014) 50(6), 969-976.
12. YAKYMENKO I., KASYANCHUK M., NYKOLAJCHUK Y.: Matrix algorithms of processing of the information flow in computer systems based on theoretical and numerical Krestenson's basis. Proceedings of the X–th International Conference "Modern Problems of Radio Engineering, Telecommunications and Computer Science" (TCSET–2010), L'viv–Slavske, Ukraine, (2010), 241.
13. STEIN W.: Elementary Number Theory: Primes, Congruences, and Secrets. A Computational Approach. Springer Science+Business Media, NewYork (2009).

14. KASIANCHUK M., YAKYMENKO I., PAZDRIY I., ZASTAVNYY O.: Algorithms of findings of perfect shape modules of remaining classes system. XIII International Conference "The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM-2015)", Polyana-Svalyava (Zakarpattya), Ukraine, (2015), 168-171.
15. NYKOLAYCHUK Y., KASIANCHUK M., YAKYMENKO I.: Theoretical Foundations for the Analytical Computation of Coefficients of Basic Numbers of Krestenson's Transformation. Cybernetics and Systems Analysis, (2014) 50(5), 649-654.
16. NYKOLAYCHUK, Y., KASIANCHUK M., YAKYMENKO I.: Theoretical Foundations of the Modified Perfect Form of Residue Number System. Cybernetics and Systems Analysis, (2016) 52(2), 219-223.

Mariia IVASHCHENKO[1], Anna STORIZHKO[2]

Scientific Supervisor: Ivan PARKHOMENKO[3]

GENERAL MODEL OF STEGANOSYSTEM AND TYPES OF ATTACKS ON STEGANOSYSTEM

Summary: One of the popular ways of data security is steganography. Steganography is the art and science of hiding data into information. The task of steganalysis is very important today. To protect steganosystem it is necessarily to know the model of its structure, main properties and possible attacks on it to prevent them.

Keywords: attack, container, key, steganalysis, steganography, steganosystem

OGÓLNY MODEL SYSTEMU STEGANOGRAFICZNEGO ORAZ TYPY ATAKÓW NA SYSTEMY STENOGRAFICZNE

Streszczenie: Jednym z popularnych sposobów bezpiecznego przesyłania danych jest steganografia. Steganografia jest sztuką oraz nauką o ukrywaniu danych w przesyłanych informacjach. Zatem, cele i zadania analizy steganograficznej są szczególnie ważne obecnie. W celu ochrony systemu steganograficznego jest koniecznym, aby znać model jego struktury, główne własności oraz możliwe ataki żeby im przeciwdziałać.

Słowa kluczowe: atak, pojemnik, klucz, steganografia, system steganograficzny

Introduction

The construction of steganographic techniques attracts many professionals engaged the development of new technologies to ensure high reliability of information systems. In general, the problem of steganography is in embedding data for the hidden transmission, and opposite to it is steganalysis task, that the identifying hidden information is one of the basic problems in the theory of reliability and security of information technology. There is also the opposite problem to steganography - steganalysis. The task of steganalysis is to identify the fact of transmission of the

[1] Taras Shevchenko National University of Kyiv, the Faculty of Information Technology, Information security management, iv-@ukr.net
[2] Taras Shevchenko National University of Kyiv, the Faculty of Information Technology, Information security management, annastor12345@ukr.net
[3] Ph.D. Docent of computerized security systems, Taras Shevchenko National University of Kyiv, the Faculty of Information Technology, parkh08@ukr.net

secret message. We can say that steganography and steganalysis – are two areas of science that are developing in parallel. Steganalysis method can be developed for existing steganography method, which usually restricts the output circuit of embedding information in a container. Steganalysis is widely used in the field of information security and, in particular to combat the illicit transmission of information.

Steganographic system or steganosystem is a set of tools and techniques used to build the covert data transmission channel.

The building of steganosystem should take into account the following provisions:

- the enemy has a full understanding of the steganographic system and details of its implementation. The only information that remains unknown to a potential enemy, is the key by which only the holder can establish the fact of presence and contens of the hidden message;
- if the enemy knows about the existence of a hidden message, it should not allow him to extract similar message from other data as long as the key is kept secret;
- a potential enemy must be deprived of any technical and other advantages in identifying or disclosing the content of secret messages.

Steganography allows the user to hide a message inside a container.

Empty container is a container without built-in messages; the filled container or stego-container that contains inline information.

Embedded (hidden) message is a message that is embedded in the container.

Steganographic channel is a channel that transfers stego.

Stegano-key is a private key needed for hiding information. Depending on how many levels of protection are in steganosystem may be one or more stegano-key.

By analogy with the cryptography type, the key of steganosystem can be divided into two types:

- with the private key;
- public key.

In steganosystem with the private key one key, which must be defined or prior to the exchange secret messages or transmitted on a secure channel, is used.

In steganosystem with the public key for embedding and extracting of a message different keys, which differ in such way that with the help of calculations it is impossible to derive one key from another, are used. Therefore, one key (public) can be transmitted freely over unsecured communication channel. In addition, this scheme works well in case of the mutual distrust of the sender and recipient.

Every steganosystem must meet the following requirements:

- Properties of the container must be modified so that it was impossible to detect a change during visual inspection. This requirement determines the quality of hiding embedded message: to ensure the smooth passage of stegano-message through channel when it should not attract the attention of the attacker in any way.
- Stegano-message should be resistant to distortions, including harmful. In the process of transferring images (sound or other container) various transformations may undergo: decrease or increase, be converted to another format, etc. In addition, it can be compressed, including using compression algorithms with data loss.

- To preserve the integrity of embedded messages the code with error correction must be used.
- To improve the reliability of the embedded message it must be duplicated.

Attacks on steganosystem

Steganosystem considered to be broken if the infringer could prove the existence of a hidden message in the intercepted container.
It is assumed that the offender can carry out any types of attacks and has unlimited computing power. If he fails to confirm the hypothesis that in the container there is a hidden message, steganographic system is considered to be stable.
In most cases there are several stages of breaking steganographic systems:

- Detection of the presence of hidden information.
- Retrieving of the hidden message.
- Modification of the hidden information.
- Prohibition on any transmission of information, including hidden pieces.

The first two stages are passive attacks on steganosystem, and the last are active (or malicious) attacks. There are the following types of attacks on steganosystems (similar to cryptanalysis) :

- The attack based on the filled container. In this case, the offender has at its disposal one or more filled containers. Offender`s task is to identify the existence of the steganochannel (main task), and also to extract data or to define the key. Knowing the key, the intruder has the ability to analyze other steganopodes.
- The attack based on the embedded message. This type of attack is more typical for systems of intellectual property protection when the digital watermarking, for example, is a well-known company logo. The objective of the analysis is to obtain the key. If the filled container corresponded to the hidden message is unknown, the problem is practically insoluble.
- The attack based on the selected hidden message. In this case, the offender may propose to transfer your messages and analyze the resulting containers.
- Adaptive attack on the basis of the selected message. This attack is a special case of the previous one. The offender has the ability to choose messages for the adaptive imposition of transmission, depending on the results of the analysis of the previous containers-results.
- The attack based on the selected filled container. This type of attack is more typical for systems of the digital watermarking. A steganoanalyst has a detector of filled containers in the form of "black box" and several of such containers. While analyzing detected hidden messages, the intruder tries to open the key.

In addition, the offender may be able to apply three more attacks that has no analogues in cryptanalysis:

- The attack based on the known empty container. If the last one is known by offender, he can always establish the presence of steganochannel by comparing empty container with container, which is suspect on the presence

of hidden data. Despite the triviality of this case, the information-theoretical justification is provided in a number of publications.

More interesting is the scenario where a container is known only approximately, comprises some error (e.g., some added noise). In this case there is the possibility of building a sustainable steganosystem.

• Attack based on the selected empty container. In this case, the offender can cause the usage of the proposed container. The latter, for example, can have a significant homogeneous regions (solid image), and then ensuring the secrecy of the embedding would not be an easy thing.

• Attack based on a known mathematical model of the container or its part. In this case the attacker tries to determine the difference between the suspicious message and the known model. For example, we can assume that the bits in the middle of the particular image area are correlated. Then the lack of such correlation may signal the presence of a hidden message. The task of the person who embeds the message is not to disturb the statistics of the container. The sender and attacker can have at their disposal various models of signals, then the battle would be winned by one who has more effective model.

The main purpose of the attack on steganographic system is similar to the purpose of attacks on the cryptosystems, with the only difference that dramatically increases the importance of an active (malicious) attacks. Any container can be replaced in order to remove or destroy the hidden message, regardless of whether it exists there or not. Identification of existing hidden data limits time at the removal stage, that is required to handle only those containers that comprise hidden information.

Even under optimal conditions for attack the task of extracting the hidden message from the container can be very difficult. Argueing definitely about the existence of hidden information is possible only after its release explicitly. Sometimes the purpose of steganoanalysis not finding an algorithm at all, but searching, for example, specific steganokey, which is used to select bits in a container in stereoperception .

REFERENCES

1. BARSUKOV V.S., ROMANTSOV A.P.: Computer steganography: yesterday, today and tomorrow. Security technologies of the XXI century. Spetsial'naia tekhnika - Special technology, 1998, 116-132.
2. KHOROSHKO V., AZAROV A. SHELEST M., YAREMCHUK Y.: Basics of computer steganography: Textbook for students. - Vinnytsya: VGTU, 2003.
3. MAO B.: Modern Cryptography: Theory and Practice, Wenbo Mao.: "Williams" Publishing House, 2005.
4. PFITZMANN B.: Computer Based Steganography: How It Works And Why Therefore Any Restriction On Cryptography Are Nonsense, At Best. – Springer: Lecture Notes in Computer Science 1996, 7-21.
5. GRIBUNIN V. OKOV I. TURINTSEV I. Digital steganography: Solon-Press, 2002.

Łukasz JUROSZEK[1]

Opiekun naukowy: Stanisław ZAWIŚLAK[2]

WIZUALIZACJA GRAFU PRZY POMOCY APLIKACJI PRZEGLĄDARKOWEJ

Streszczenie: W pracy omówiono algorytmy oraz program komuterowy do generowania losowego grafu oraz jego kolorowania. Zastosowano losowanie o zmiennym prawdopodobieństwie przy pomocy pseudo-losowego generatora liczb losowych. Kolorowanie wierzchołków grafu zostało zrealizowane przez algorytm quasi-zachłanny na podstawie ciągu liczb. Spójność grafu sprawdzano przy użyciu przeszukania w głąb. Wyniki generowania zostały zaprezentowane w formie graficznej oraz macierzy sąsiedztwa. Jedną z opcji programu jest możliwość utworzenia nowego wierzchołku lub krawędzi na wygenerowanym grafie. Użytkownik ma możliwość wybrania prawdopodobieństwa na otrzymanie krawędzi oraz liczbę wygenerowanych wierzchołków.

Słowa kluczowe: losowanie o zmiennym prawdopodobieństwie, generowanie grafu, przeszukiwanie w głąb, kolorowanie quasi-zachłanne, generator liczb psełdolosowych

WEB BASED GRAPH VISUALIZATION APPLICATION

Summary: In the paper, the problem of graph visualization is considered. The algorithms and a computer program that generate random graph and colors its nodes are described. The generation numbers by changeable probability where used. The pseudo-random numbers generator was utilized. Colors of nodes are created by quasi-greedy algorithm based on string of numbers. For checking its connectivity Depth First Search algorithm was implemented in program. Result of generating are displayed for user in two forms, graph itself and adjacency matrix. There are options to add node or edge. User can change probability and number of nodes that will be generated.

Keywords: drawing numbers of different probability, graph generation, depth first search, quasi-greedy coloring, pseudo-random numbers generator

1. Introduction

Graph theory is a branch of mathematics which was originated in 1736 by Leonhard Euler [1,2]. The algorithmic approach to graph theory is extremely important due to development of networks. Nowadays, there are a lot of ways to generate a particular

[1] University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, lukjuroszek@gmail.com

[2] Assoc. Professor [Dr hab. inż.], University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, szawislak@ath.bielsko.pl

graph by programing algebraic structures. However, it is hard to implement visualization in desktop application. Another thing is generating a random number in the range of 0 to 1 by specific probability. Relatively easy to implement, but without our problem consisting in generation of numbers by changeable probability values. The probability increases via a set increment until the graph is connected, disconnected graph is considered as incorrect. Graph drawing is a widely discussed problem having his own books [3,4,10] as well as annual world-wide conference [8]. Coloring of graph nodes [9] by quasi-greedy algorithm based on string of numbers is also incorporated in the prepared application.

2. Problem formulation

In the paper, the visualization of graph problem is considered. One can prepare its own application or incorporate the standard procedure published in net as a so called open source. Nowadays, a lot of visualization libraries for desktop application exist, but they don't work in web based application. Web application doesn't need an installing process.

Graph G (V, E) consists of two sets: V – set of vertices (non-empty) and E – set of edges. Vertices are drawn usually as dots, circles or ellipses and edges as straight lines or bows. If E is empty than graph consists of separate vertices, only. Graph is named connected if for each pair of its vertices, the path exists – which connects these chosen vertices. Connected graphs are considered.

Pseudo-Random Number Generation (PRNG) is an algorithm for generating a sequence of the numbers by computer program. PRNG get approximately random numbers, because it is completely determined by an initial value, so it can generate similar number in some situation.

Adjacency matrix will be generated by method that get a number [0,1] based on probability e.g. enter [1] will have $0{,}3 * 100\% = 30\%$ *chance to gets, and enter* [0] *will have* $1 - 0{,}3 = 0{,}7 * 100\% = 70\%$.

The goal of the consideration is to use the above solutions and to draw a connected graph - for user – in a graphical window of the programme. Moreover, a user will have still ability to add edge and nodes via mouse-driven options.

3. Algorithm

There was utilized the simplified generation with probability algorithm. The advantages of this approach are: greater independence from pseudo-random number generator, better control of generating numbers. The generation with probability may be presented as a strip of paper deviated to segments with fixed length; length is a measure of probability. Bigger probability, bigger chance to get a certain result.

Quasi-greedy coloring solves approximately minimal color problem, approximately, because the number of colors received by this method in not optimal. It would be add that there is not available an algorithm which solves problem of coloring for an arbitrary graph. The order of searching should start by default from first node consecutively to last node in a graph. There are cases when the order is not

optimal and a solution to this problem will be the reshuffling of the strings of nodes to reduce the number of colors to minimum.

Depth-first search (DFS) [5] is a recursive algorithm that uses the idea of backtracking. It involves exhaustive searches of all the nodes by going ahead, if possible, in other cases by backtracking. Backtrack means that when you are moving forward and there are no more nodes along the current edge, you move backwards on the same edge to find nodes to traverse. All the nodes will be visited on the current path till all the unvisited nodes have been traversed and after that the next edge will be selected. The result of this review is a sequence of nodes e.g. 0, 1, 3, 4, 2. If a sequence doesn't contain all the nodes of a graph it means that a graph is not connected. Therefore search algorithms are used for connectivity detection.

4. Program description

Computer program was written in HTML, CSS, JavaScript language and visualization library "visjs"[6].

4.1. Application specification

An Application may be used for the visualization of both random and regular, plain graphs. It will tell a user if a graph is connected, what was used to generate; how many nodes and colors were used. Our application also allows to display results in an adjacency matrix representing the considered graph.

A user doesn't have to know the interface to use the application. There is option to add or edit edges and nodes when modifying an existing one. Nodes should use some pseudo-physics in order to not overlap themselves. The Number of colors, nodes, probability and the consistence of a graph must be displayed in program.

Adjacency matrix should be visible on screen. Coloring process should change node colors and color codes should be visible in nodes (hex color code). Eliminate of random number generation problem by using generation with specific probability for getting 1 in adjacency matrix.

4.2. How to use the application

Application doesn't need to get any data from user, the first run will automatically generate default graph with 5 nodes and edges based on the probability of 0,5. A user will be able to move nodes by clicking on specific node and dragging. Nodes have some pseudo-physics (library has implemented a solver called " hierarchical Repulsion", so distance between nodes will adjust itself).

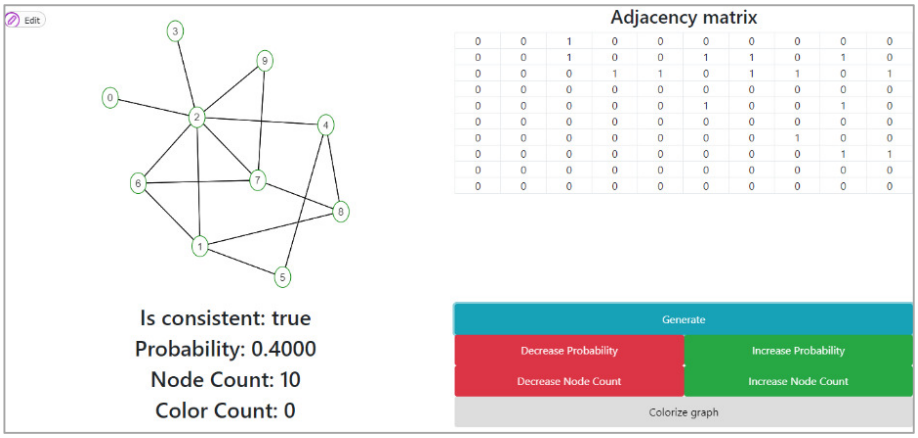


Figure 1. Application main window

In the upper-left corner of the programme window, there is a button named "Edit" (Fig.2) that has option for editing graph:

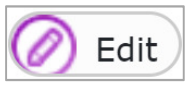


Figure 2. Graph Edit button

- Add Node (Fig. 3) – allows user to create a new node on screen,

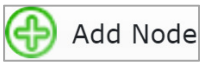


Figure 3. Graph Add Node button

- Add Edge (Fig. 4) – adds a new edge from node to node, by clicking first node and drag to second one, is option to add edge for one node (loop),

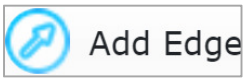


Figure 4. Add Edge button

- Delete Selected (node) (Fig. 5) – deletes selected node, this option is only visible if node is selected,

Figure 5. Delete button (for edge, node)

- Delete Selected (edge) – deletes selected edge, this option is only visible if edge is selected,
- Edit Edge - his option is only visible if edge is selected, user can change edge start and edge ends into another node.

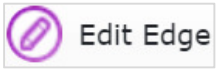


Figure 6. Edit Edge button

On right side of menu user will see adjacency matrix (Fig. 7).

Adjacency matrix

0	0	1	0	0	0	0	0	0	0
0	0	1	0	0	1	1	0	1	0
0	0	0	1	1	0	1	1	0	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Figure 7. Generated Adjacency matrix

Rows and Columns of this matrix represents nodes, first row has index [0] (in programing array start from zero and zero is the first index array), so Node [0] has an edge to Node [2] and Node [2] has edges with Nodes [3,4,6,7,9]. As it is shown in Figure 8. Generated Graph with low probability from Adjacency matrix. As we see, on its diagonal are only zeros, under diagonal too, that means is simple graph. Is way to create directed graph from this matrix, just by mirroring values from values above diagonal.

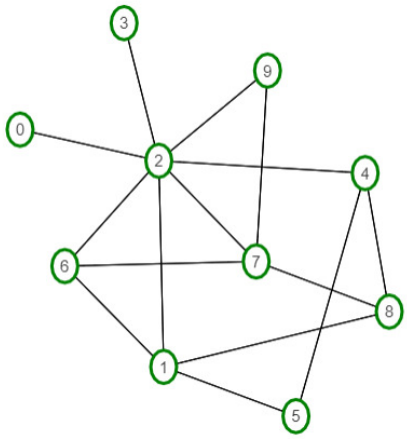


Figure 8. Generated Graph with low probability from Adjacency matrix

Generated graph uses all available space on left side of the programme window, by default number of nodes are placed in circles, distance between all nodes are similar, edges are straight, but they can be changed to round one (bows) in option "roundness". Options for viewing graph are set to default, but advanced user can change it in script option. This Options are hiden for normal user to make application. Graph window can be manipulated by standard option e.g. click on free area and drag will move graph, scrolling is responsible for zoom.

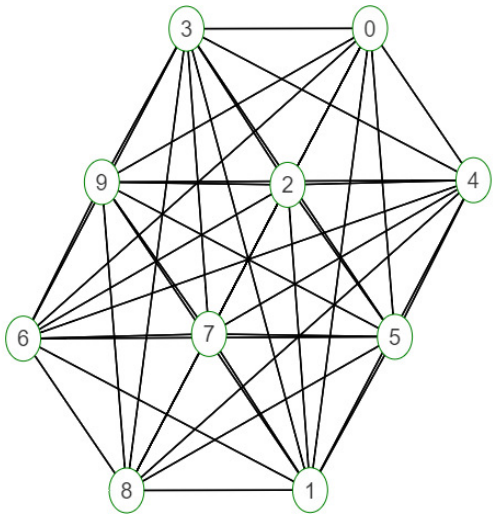


Figure 9. Generated graph with high probability

In bottom-left corner of window are displayed information and parameters about graph:

- Is connected – result of checking connected of graph after its generation, have two values: [true, false],
- Probability – a parameter for generating [0,1] in adjacency matrix, this value represent probability for [1], minimum: 0, maximum 1.0,
- Node Count – a number of nodes (max: 100) that will be generated,
- Color Count – number of color used to colorize graph, it changes only before clicking "Colorize graph" button

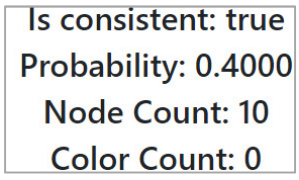


Figure 10. Information and parameters of generated Graph

In the bottom-right corner of window are displayed control buttons:

- Generate – generate graph by parameters
- Decrease/Increase Probability – change value of Probability by 0,1 (+/-)
- Decrease/Increase Node Count – change value of Node Count by 1 (+/-)
- Colorize graph – colorize graph by quasi-greedy algorithm, change graph labels in view.

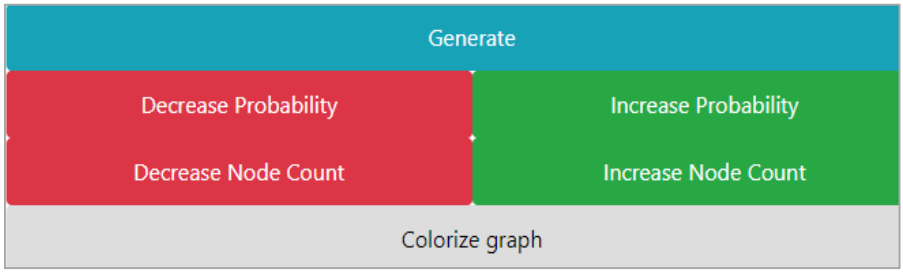


Figure 11. Control buttons for generating and coloring Graph

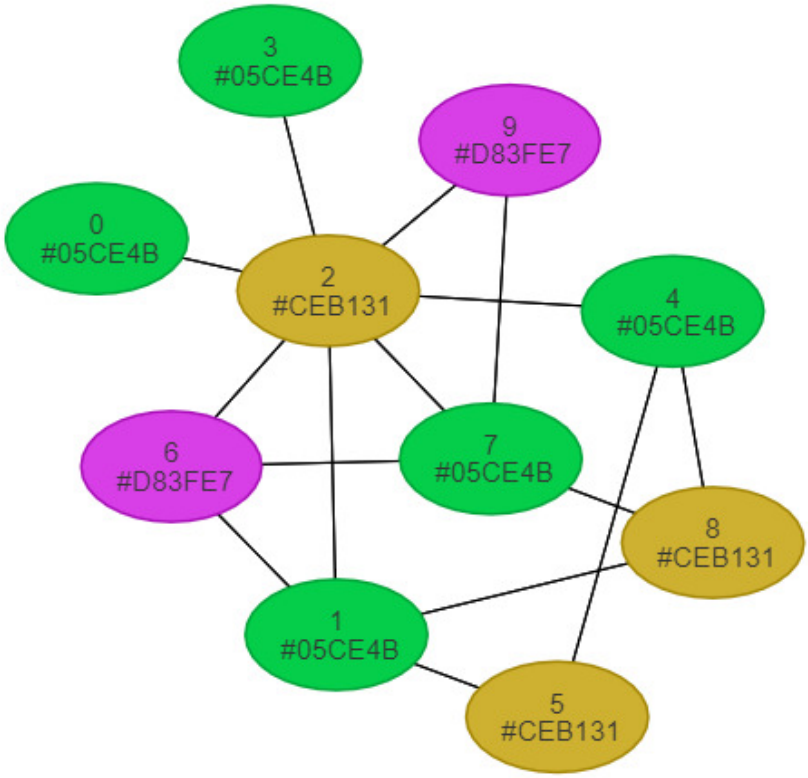


Figure 12. Result of clicking Colorize graph button

Result of graph coloring is not optimal, because quasi-greedy algorithm was used. However, like it was mentioned above, there is not available any perfect algorithm for perform coloring like e.g. for spanning tree or for the shortest path problems.

Labels of nodes now have other color and information about color, e.g. #05CE4B is shade of green, #D83FE7 is shade of pink. As shown in Figure 12. Result of clicking Colorize graph button.

5. Summary

The discussed application, thanks to generations algorithms, allows for solving visualization problem with web based library. User decide what to generate, which node or edge should be deleted, so user knows when graph is connected and after this, coloring procedure can be applied to the existing graph, user can check if he is able to colorize in better way. The application allows also for coloring of graph vertices. In this case the pictogram used for vertex displaying changes from a circle into an ellipse which makes enough room for entering the code of color. The programme could be used in didactics of graph theory related subjects. It gives also the glimpse how incorporate open source codes into our own programme.

REFERENCES

1. WILSON R.J.: Introductory graph theory, (Polish translation: „Wprowadzenie do teorii grafów"), PWN, Warszawa (2017).
2. WOJCIECHOWSKI J., PIEŃKOSZ K.: Graphs and networks (in Polish: Grafy i sieci), PWN, Warszawa (2013).
3. NISHIZEKI T., RAHMAN S.: Planar graph drawing, Lecture Notes Series on Computing: Volume 12, Word Scientific, Singapre (2004).
4. DI BATISTA et al.: Algorithms for visualisation of graphs, Prentice Hall, New York (1999).
5. HackerEarth, web service: Depth First Search https://www.hackerearth.com/practice/algorithms/graphs/depth-first-search.[3]
6. A dynamic browser based visualization library, http://visjs.org.
7. http://mrvar.fdv.uni-lj.si/pajek/
8. https://gd2017.ccis.northeastern.edu/
9. KUBALE M. (editor): Graph colorings, Contemporary Mathematics, Volume: 352, American Mathematical Society (2004) 208 pages.
10. JUNGER M., MUTZEL P.: Graph drawing software, Springer Verlag, Heidelberg, 2004.

[3] Cited Web Pages were opened on: 30 Oct. 2017.

Taras KAVKA[1], Ivan OPIRSKYY[2]

Opiekun naukowy: Ivan OPIRSKYY[2]

ANALYSIS OF THE MAIN SECURITY RISKS OF WIRELESS

Summary: In this work is carried out the analysis of the main security risks of wireless networks such as: risk concerning non-stationary communication, risks of information leaks from wire networks, risks concerning peculiarities of wireless networks functioning, risks concerning new threats and attacks, risks concerning networks vulnerabilities and forced downtime. The carried out analysis has allowed to represent a problem of protection of wireless communication systems and identify the most critical risks.

Key words: risks, threat, attack, access point, wireless network, attacker's tools

ANALIZA GŁÓWNYCH ZAGROŻEŃ BEZPIECZEŃSTWA W SIECIACH BEZPRZEWODOWYCH

Streszczenie: W pracy przeprowadzono analizę głównych zagrożeń bezpieczeństwa sieci bezprzewodowych, takich jak: ryzyko związane z komunikacją niestacjonarną, ryzyko wycieku informacji z sieci przewodowych, zagrożenia związane z funkcjonowaniem sieci bezprzewodowych, zagrożenia związane z nowymi zagrożeniami i atakami , zagrożenia związane z luką w sieci i wymuszone przestoje. Przeprowadzona analiza pozwala na zdefiniowanie najważniejszych problemów ochrony systemów komunikacji bezprzewodowej oraz identyfikowania najbardziej krytycznych zagrożeń.

Słowa kluczowe: zagrożenia, atak, punkt dostępu, sieć bezprzewodowa, narzędzia atakujące

1. Introduction

Protection of information in modern conditions becomes more and more difficult problem, that are caused by different reasons, the main of them are: massive distribution of computers; complicating of encrypting technologies; a necessity of protection not only State and military secret but also industrial, commerce and financial secrets; spreading possibilities of unauthorized actions over the information. Especially is actual an issue considering channels of wireless communication, that are considered now as perspective mean of receiving and sending information, because

[1] student of Department of Information Protection.
[2] Ph.D., Lviv Polytechnic National University, a docent of Department of Information Protection

of flexible network organization, especially in places where wire networks cannot be organized in general, and also due to minimum time spends, that currently are the most important requirements for information transfer systems. But on the other hand, they are the most vulnerable.
The fact, that anyone can connect to wireless networks, requires from creators of these networks serious approach for ensuring sufficient security level. That's why, the analysis of main risks of security of wireless networks is actual task.

1.1 Main part.

Wireless technologies, that work without physical and logical limitations of its wire analogs, undergo large risks for network infrastructure and users. To understand how to ensure safe functioning of wireless networks, let's examine them in detail.

1.1.1. The risks concerning authorized access

The "aliens" are called devices, which provide unauthorized access to corporate network and often bypassing the protection mechanisms, specified by corporate security policy. The most frequently these are willfully created access points. The statistics all over the world, points at aliens as a reason of the majority of hacks of corporate networks. Accessibility and cheapness of Wi-Fi devices have led to a case, that in Ukraine for instance, almost every network with more than 50 users, has faced with such phenomenon.
Except access points, in a role of alien can be home router with Wi-Fi support, software based access point Soft AP, a laptop with turned on simultaneously wire and wireless interfaces, a scanner, a projector etc.
Risks concerning non-stationary connection.
Wireless devices are not "connected" with cable to a socket and can change access points to network directly during the work. For instance, can happen "casual associations", when a laptop with Windows XP or just not correctly configured wireless client automatically associates and connects a user to the nearest wireless network. Such mechanism allows hackers to "switch victims over themselves". In addition, if a user is connected simultaneously to wire network, he becomes an easy entrance point, in other words – a classical alien [2].
Many users of laptops, equipped with Wi-Fi and wire interfaces and are not satisfied with quality of wire network (it works slowly, an administrator has filtered certain URLs, or Skype/Viber protocol), prefer to switch to nearest access zones. Or the operating system does this automatically for them in case of wire network malfunction. There is no use to say, that in such case all attempts of IT department to ensure network security will be futile.
The ad-hoc networks – are peer-to-peer connections between wireless devices without access point - quickly allow to send a file to a colleague or print a necessary document on printer with Wi-Fi adapter. However, this method of network organization doesn't support the majority of necessary methods of ensuring security, providing hackers easy way to hack computers.

1.1.2. Risks concerning networks vulnerabilities and forced downtime

Some network devices can be more vulnerable than others: not correctly configured, they use weak encrypting keys or methods of authentication with well-known

vulnerabilities. And it is clear, that the hackers attack them first of all. The reports of analysts confirm, that more than 70 % of successful hacks of wireless networks have happened in result of not correct configuration of access points or client software.

Not correctly configured access points (AP).

Only one not correctly configured AP (including an alien) can be a reason of hack of corporate network. Settings by default of the majority of AP do not include authentication or encrypting or use static keys, written in manual and that's why are well known. In addition, a combination with low price of these devices, this factor complicates very much a task of monitoring the integrity of configured network infrastructure and a level of its protection.

Not correctly configured wireless clients.

This category has bigger threat, than not correctly configured access points. In fact, these devices "come and go" from the enterprise, and often they are not configured especially with purpose of minimization of wireless risks or often are configured with default settings (that cannot be recognized as safe). Such devices provide priceless help for hackers, providing an easy access point for network scanning and spreading in it malware [6, 18].

A hack of encrypting algorithms.

Special tools are available for hackers to hack networks, that are based on encrypting standards. These tools are widely represented in the Internet, and their usage doesn't require special skills. They exploit the vulnerabilities of algorithm, collecting the traffic statistic passively in wireless network till the moment, when the received data will not be enough for recovery the encrypting key. In case of using the latest generation of hack tools, that use special methods of traffic injections, a term "till the moment" fluctuates between 15 minutes and 15 seconds. Recently, have been revealed first, currently not serious, vulnerabilities in TKIP, that allow to decrypt and send in protected network not big packets.

Risks concerning new threats and attacks

Wireless technologies have generated new methods of realization of old threats, and also some new, hitherto not possible in wire network. In all cases to struggle with attacker has become much harder, because it is not possible to track his location, and to isolate him from the network.

Exploration.

The majority of traditional attacks start from exploration, as a result of which, the hacker specifies the further steps of their exploitation. For wireless exploration are used as well as wireless scan tools (NetStumbler, Wellenreiter, embedded JC client), and also packages collection and analytical tools, because many of leading WLAN packages are not encrypted. At the same time, it is very hard to distinguish a situation, which collects information, from ordinary one, that tries to obtain authorized access to the network or a try of casual association.

It is universally recognized, that methods, that relate to Security through Obscurity class, are not sufficient, because the attacker observes wireless network in certain radio channel anyway, and he has to wait for first authorized connection, because in this process ESSID is sent in not encrypted view. After this, this security measure just loses its sense.

Identity Theft.

The identity theft of authorized user is a serious danger in any network, and not only wireless one. However, in last case is harder to identify the user authenticity.

Of course, there are SSIDs, that can be filtered by MAC addresses, but both of them are transferred in open view, and it is not difficult to fake both of them.
There is an erroneous idea, that user Identity Theft is possible only in case of MAC-authentication or by using static keys, and scheme on base of 802.1x, such as LEAP, is absolutely safe. Unfortunately, it is not true, and even now, is available a hack tool, for instance for LEAP. Other schemes, let say EAP-TLS and PEAP, are more reliable, but they don't guarantee endurance in case of complex attack, that uses several factors simultaneously.
Denial of Service, DoS. A task of this attack is violation of quality indicators of functioning of network services, or compete elimination of access possibility to them for authorized users. To achieve this goal, the network can be filled up with "trash" packets (with wrong checksum), sent from legitimate address. In case of wireless network, to track a source of such attack without special tools, is impossible. In addition, it is possible to organize the DoS attack on physical level, by launching quite powerful generator of obstacles in necessary frequency range. An example of DoS attack is shown on picture no. 1.

Figure 1. The example of DoS attack

Special tools of attacker. The toolkit of attacks on wireless networks is widely accessible in the Internet and constantly is supplemented with new means. The main types of attacker's toolkit are:

- Exploration tools – scanning networks and defining their parameters, traffic collection and analyzing (Kismet, NetStumbler, AirMagnet, Ethereal, Wireshark with special module, THC-RUT, Airmon-NG);
- Encrypting hack tools (AirCrack, WEPWedgie, WEPCrack, WepAttack, AirSnort, Airmon-NG);
- The hack tools of authentication mechanisms for their bypassing or obtaining the credentials (ASLEAP, THC-LEAPCracker, Airmon-NG) on picture no 2;
- The tools of DoS attacks organization (WLANjack, hunter_killer);
- Vulnerabilities scanner (Nessus, xSpider);
- The tools of wireless connections manipulation (HotSpotter, SoftAP, AirSnarf);

• Traditional toolkit (SMAC, IRPAS, Ettercap, Cain & Abel, DSNIFF, IKEcrack). This list can be expanded.

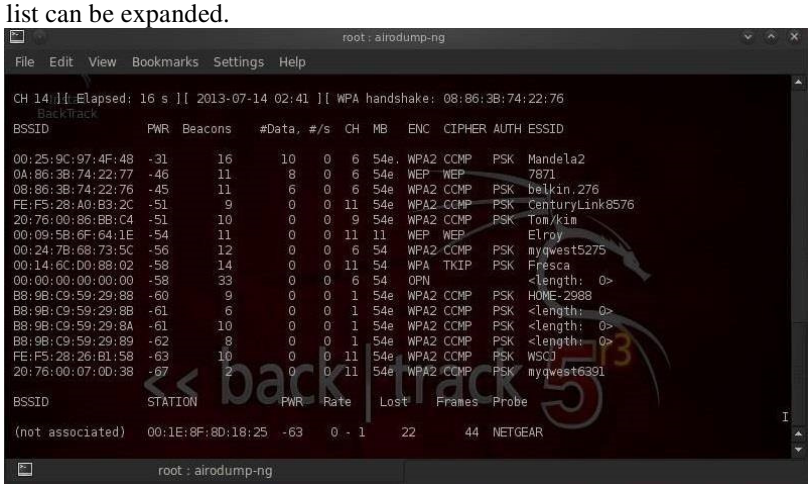


Figure 2. An example of a hacking mechanism for authenticating utilities airmon-NG

1.1.3. Risks of information leaks from wire networks

Almost all wireless networks at certain moments are connected to wire ones. Correspondingly, any wireless AP can be used as a base for attack. But it is not all: some errors in their configuration in combination with configuration errors of wire networks can open the ways for information leaks. The widest example – the APs, that work in bridge mode (Layer 2 Bridge), they are connected in flat network (or a network with VLAN segmentation violations) and spread wide broadcasting packages from wire segment, ARP, DHCP requests, STP frames etc., some of this data can be useful for Man-in-The-Middle attacks organization, different Poisoning and DoS attacks, and just exploration tasks.

Another wide spread scenario is based on peculiarities of 802.11 protocols realization. In case, when on one AP are configured several ESSIDs, wide broadcasting traffic will be spread over all ESSIDs at once. As a result, if on certain point are configured the protected network and public access zone, the hacker, connected to the last one, can, for instance, violate the functioning of DHCP or ARP protocols in protected network. It can be fixed by turning on Multi-BSSID mode, or other its name - Virtual AP, which is supported by almost all vendors of Enterprise class equipment (and not all from Consumer class), it worth knowing.

1.1.4. Risks concerning peculiarities of wireless networks functioning

Some peculiarities of wireless networks functioning generate additional problems, that can influence on their accessibility, efficiency, security and cost of operation. For complex solving these problems is necessary special toolkit of support and operation, special mechanisms of administration and monitoring, not realized in traditional toolkit of wireless network management.

Activity in not working time.

It is possible to connect to wireless network from any place in coverage area at any time. Due to this, many organizations limit the accessibility of wireless networks in their offices by only working hours (even by turning off the AP). By taking into consideration all that has been said, we can suppose that any wireless activity in network during not working time, has to be considered as suspicious and has to be investigated.

Speed.

The APs, that allow the connections with low speed, have larger coverage zone. Thus, they provide additional possibility of remote hack. If in office network, where everybody works with the speed 54/100 Mbps, occasionally a connection appears with 10 or 20 Mbps, this can be a signal, that someone tries to break through into the network from outside.

Obstacles.

The quality of wireless network functioning depends on many factors. The brightest example are obstacles, that reduces the bandwidth of network and the amount of supported clients, even to complete network inaccessibility. The source of obstacles can be a device, that emits a signal with enough power in the same frequency range that the AP does. On the other hand, the hackers can use the obstacles for DoS attacks organization on network.

Except obstacles, there are other aspects: an aerial creates AP malfunction that can create problems, as on physical and on channel level, leading to deterioration of maintenance quality of other network clients.

2. Conclusion

The carried out analysis has permitted us to define the main risks of wireless network security, that can be divided into the following groups: the risks concerning authorized access, risk concerning non-stationary connection, risks concerning networks vulnerabilities and forced downtime, risks concerning new threats and attacks, risks of information leaks from wire networks, risks concerning peculiarities of wireless networks functioning. This analysis has shown, that the most dangerous are risks concerning authorized access and risks concerning new threats and attacks.

REFERENCES:

1. IVAN OPIRSKYY, The peculiarities of forecasting procedure of unauthorized access. Naukovo-practychniy zhurnal “Zakhyst informatii” [Scientific and practical magazine “Information protection”], 2014, no. special edition, pp. 74-80 (in Ukrainian).
2. DUDUKEVICH V.B., OPIRSKYY I.R., The analysis of models of information protection in State informational networks. Systemy obrobky informatsii [Information processing systems], Kharkiv, 2016, no. 4 (141), pp. 86-90 (in Ukrainian).
3. HORDEYCHYK S.V., DUBROVIN V.V. Bezpeka bezdrotovykh merezh [Wireless networks security] Moscow, Haryacha liniya - telekom Publ., 2008.
4. PAVLOV I.M., KHOROSHKO V.O. Proektuvanya complexnykh system zakhystu informathii [The designing of complex information protection systems] Kiyv, VITI-DUIKT Publ., 2011.

Nataliya KLYMUK[1], Nataliya KRAVETS[2]

Scientific supervisor: Vasyl MARTSENYUK[3]

AN APPROACH FOR DEVELOPMENT OF MEDICAL INFORMATION SYSTEM BASED ON MICROSERVICES ARCHITECTURE

Summary: The work is devoted to presentation of microservices approach for development of medical information system. Implementation is based on usage of Spring, Spring Boot and Spring Cloud frameworks.

Keywords: medical information system, microservices, Spring

OPRACOWANIE MEDYCZNEGO SYSTEMU INFORMACYJNEGO NA PODSTAWIE ARCHITEKTURY MIKROSERWISÓW

Streszczenie: Celem artykułu jest przedstawienie wykorzystania mikroserwisów do opracowania medycznego systemu informacyjnego. Realizacja jest oparta na użyciu szablonów Spring, Spring Boot, Spring Cloud.

Słowa kluczowe: medyczny system informacyjny, mikroserwisy, Spring

1. Introduction

A variety of approaches and techniques are used when developing information systems for medical practice and research [1].
Primarily they are based on so called monolithic architecture. In such case you are developing a server-side enterprise application. It must support a variety of different clients including desktop browsers, mobile browsers and native mobile applications [2].
In turn Microservice is a paradigm that serves for organization and usage of distributed services that can have different proprietors. The basic ideas of this architectural approach was stated by Martin Fowler in 2014 in [3].
Microservices allow large systems to be built up from a number of collaborating components. It does at the process level what a lot of frameworks (e.g., Spring) has always done at the component level: loosely coupled processes instead of loosely coupled components.

[1] Ternopil State Medical University
[2] Ternopil State Medical University
[3] Prof., Dr hab., Akademia Techniczno-Humanistyczna w Bielsku-Białej, Wydział Budowy Maszyn i Informatyki, email: vmartsenyuk@ath.bielsko.pl

Much technologies and the protocols that are created to realization of business processes [4] and their support are included in the content of widely understoodable microservices architechture. Those technologies together create powerfull instruments for

- implementation of business processes,
- opening for access to the client for different type of services through a network,
- seeking out of services (UDDI Universal Description, Discovery of and of Integration);
- uses of services through a client,
- determinations of business processes with help of languages of determination of flow of problems and creations of complex services.

Mentioned above possibilities of microservices technology and in particular arrangement of services in greater processes are as background of application of microservices to the construction of the system of medical information services. However at the market there is a shortage of such solutions now.

Composition of services in microservice architecture results in combining of separate services (WS, Web Service) in a structure named by a process, that describes the algorithm of implementation of series of services. In order to do it, possession of the detailed information on motion of process is needed (before it will be determined).

In case of emergency medicine service the detailed determination of tasks is not possible. Reason is unpredictability of motion of incident and also factors that influence on flow of possible treatment scheme:

- number and state of health of patients;
- variety of services that belong to many health establishments;
- dynamics of patient state can change in the process of implementation of treatment;
- accessability (primarily distance) to healthcare establishments for patient.

We can divide those factors into two groups:

1) information on state of patient previously known. It should be known from the point of view of healthcare, e.g. patient passport data, health assessment, previous histories of diseases and so on,
2) exceptional/change of certain factors. It can be instant change of well-known or unknown previously factor, e.g.: fracture, blooding, emergency etc.

It all causes that the detailed setting by default, how procedure of emergency medicine should look like, is impossible. Moreover, settings are not possible usually, how must be conducted healthcare in situation, when this situation will arise up already. The reason of that is a typical shortage of sufficient data during undertaking of action, since at the beginning of healthcare as a rule we don't have possibility to deliver exact data [5].

The **objective** of this paper is to present the way of application of Microservices architecture for the problem of development of medical information system (MIS).

Potential problems, that touch medicine and idea of construction of the system that basing on the paradigm of Microservice forms separate healthcare services into one coordinated healthcare process, will be presented in the paper. The offered solution means to determine process dynamically without detailed definition from point of view of the whole action. When basing the behaviors on typical charts, the system executes certain introductory steps. During their execution, additional data on the

basis of the system for determination of further actions are gotten. As far as progress of medicine gets each time more detailed data dealing with certain case and necessary actions, a process is being specified in the process of implementation.

2. Materials and methods

2.1. Conception of solution of problem of emergency medicine

Consider that on a certain street in city accident has happened. In accident a dozen of car passengers damaged. Many people feel damages and need emergency medical assistance.

From the point of view of rescue services a problem is difficult for the solution, because information about a case is usually inexact on this stage: it is unknown, how many people are injured and how their damages are serious, how many coaches are needed on the place of accident, how many places in a hospital needed to be prepared to give a help for injured.

We need fast and well-coordinated rescue action. Human factor can lead to errors under such circumstances, that can influence on saving lifes of injured ones. In this situation it is the best to dispose of the certain system able to manage a rescue action, liquidating the errors related to the emotions, by work in stress and necessity of the rapid reacting.

Possibility of support of rescue services is covenient here through the computer system that has an access to all well-known information about a case, that is capable efficiently to plan actions on the basis of all well-known data, is able dynamically to adjust actions to the changable conditions and also to coordinate the actions of elements of different services.

2.2. Requirements for information system

In ideal case system that would solve the above-mentioned problems that arise in rescue medical service should be able:

- to give independence to the different elements that are his constituents;
- to do possible a suggestion of its own services in the system;
- to do possible composition of component services into more complicated one such as a complex service for injured person.

Main principle is opening for access of actions of medical services as network services of web-service at application of paradigm of microservices.

The system implements its task through a construction of a skeleton of process on the initial stage. Then as far as the flow of additional data will grow gradually, it allows corresponding adaptation of actions to the queries for concrete case.

2.3. Basic services of emergency medicine

For the purpose of development of the system we need to determine the basic complete set of services that will be able to serve to creation of arbitrary rescue action. It seems that determining a complete set of services in emergency medicine is impossible, taking into account a fact, that together with development of medicine there will appear new services that should be taken into account during composition. The reasoning requires to determine certain standard of description of services, that enables development of new ones in future instead of attempts of determination of

complete base that is necessary for system running. Examples of services that will be token into account when developing rescue action are [6]:
Arrival of ambulance to the place of incident (ArrivalToPlaceOfIncident) is simple service that means arrival of ambulances to the given place of incident. It includes arrival to the place of case; implementation of assessment of injured person; delivering primary care; preparing report including data of injured persons in details.
Hospitalization of injured persons (HospitalizationOfPatient) is service that consists of the acceptance of patient to the hospital and the delivering corresponding help to him.
Transportation of patient (TransportationOfPatient) is service implying the transportation of injured patient to the hospital.

3. Results

In this part of the article there will be presented an example of problem solution of emergency medicine action using basic set of services mentioned above (Fig.1).

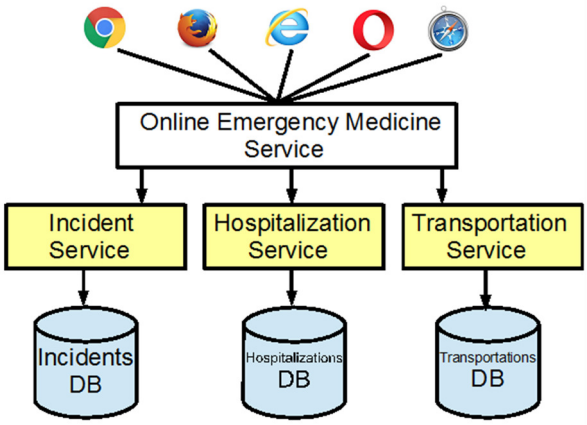


Figure 1. Problem solution of emergency medicine action using basic set of microservices

Let's consider that a travelling accident happened and casual witness is calling on a line. Operator of line inputs the system data concerning the case, after that the system begins processing the case immediately.
We imagine an online emergency medicine service with separate microservices for arrival-to-place-of-case, hospitalization-of-patient and transportation-of-patient:
Inevitably there are a number of moving parts that we have to setup and configure to build such a system. How to get them working together is not obvious - we need to have good familiarity with Spring Boot since Spring Cloud leverages it heavily, several Netflix or other OSS projects are required and, of course, there is some Spring configuration "magic" [7].

3.1. Service Registration

When you have multiple processes working together they need to find each other. If you have ever used Java's RMI mechanism you may recall that it relied on a central

registry so that RMI processes could find each other. Microservices has the same requirement.
The developers at Netflix had this problem when building their systems and created a registration server called Eureka ("I have found it" in Greek). Fortunately for us, they made their discovery server open-source and Spring has incorporated into Spring Cloud, making it even easier to run up a Eureka server. Here is the complete discovery-server application:

```
@SpringBootApplication
@EnableEurekaServer
public class ServiceRegistrationServer {

  public static void main(String[] args) {
    // Tell Boot to look for registration-server.yml
    System.setProperty("spring.config.name",  "registration-
server");
    SpringApplication.run(ServiceRegistrationServer.class,
args);
  }
}
```

Spring Cloud is built on Spring Boot and utilizes parent and starter POMs. The important parts of the POM are:

```
    <parent>
        <groupId>org.springframework.cloud</groupId>
        <artifactId>spring-cloud-starter-parent</artifactId>
        <version>_Brixton_.RELEASE</version>   <!-- Name of
release train -->
    </parent>
    <dependencies>
        <dependency>
            <!-- Setup Spring Boot -->
            <groupId>org.springframework.boot</groupId>
            <artifactId>spring-boot-starter</artifactId>
        </dependency>
        <dependency>
            <!-- Setup Spring MVC & REST, use Embedded Tomcat
-->
            <groupId>org.springframework.boot</groupId>
            <artifactId>spring-boot-starter-web</artifactId>
        </dependency>
        <dependency>
            <!-- Spring Cloud starter -->
            <groupId>org.springframework.cloud</groupId>
            <artifactId>spring-cloud-starter</artifactId>
        </dependency>
        <dependency>
            <!-- Eureka for service registration -->
            <groupId>org.springframework.cloud</groupId>
            <artifactId>spring-cloud-starter-eureka-
server</artifactId>
        </dependency>
    </dependencies>
```

By default Spring Boot applications look for an application.properties or application.yml file for configuration. By setting the spring.config.name property we

can tell Spring Boot to look for a different file - useful if you have multiple Spring Boot applications in the same project.
This application looks for registration-server.properties or registration-server.yml. Here is the relevant configuration from registration-server.yml:

```
# Configure this Discovery Server
eureka:
  instance:
    hostname: localhost
  client:  # Not a client, don't register with yourself
    registerWithEureka: false
    fetchRegistry: false

server:
  port: 1111    # HTTP (Tomcat) port
```

By default Eureka runs on port 8761, but here we will use port 1111 instead. Also by including the registration code in my process I might be a server or a client. The configuration specifies that I am not a client and stops the server process trying to register with itself.
Try running the *RegistrationServer* now. You can open the Eureka dashboard here: http://localhost:1111 and the section showing Applications will be empty.
From now on we will refer to the *discovery-server* since it could be Eureka or Consul.

3.2. Creating a Microservice: Arrival-to-Place-of-Incident-Service

A microservice is a stand-alone process that handles a well-defined requirement.

When configuring applications with Spring we emphasize Loose Coupling and Tight Cohesion, These are not new concepts (Larry Constantine is credited with first defining these in the late 1960s [7]) but now we are applying them, not to interacting components (Spring Beans), but to interacting processes.
In this example, we have a simple Arrival-to-Place-of-Incident (Incident) microservice that uses Spring Data to implement a JPA IncidentRepository and Spring REST to provide a RESTful interface to incident information (Fig. 2). In most respects this is a straightforward Spring Boot application.

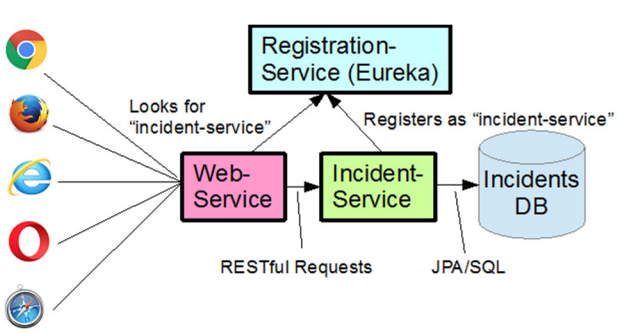


Figure 2. Incident microservice that uses Spring Data to implement a JPA IncidentRepository and Spring REST to provide a RESTful interface to incident information

What makes it special is that it registers itself with the *discovery-server* at start-up. Here is the Spring Boot startup class:

```
@EnableAutoConfiguration
@EnableDiscoveryClient
@Import(IncidentsWebApplication.class)
public class IncidentsServer {
    @Autowired
    IncidentRepository incidentRepository;
    public static void main(String[] args) {
        // Will configure using incidents-server.yml
        System.setProperty("spring.config.name", "incidents-
server");
        SpringApplication.run(IncidentsServer.class, args);
    }
}
```

The annotations do the work:

1. @EnableAutoConfiguration - defines this as a Spring Boot application.
2. @EnableDiscoveryClient - this enables service registration and discovery. In this case, this process registers itself with the *discovery-server* service using its application name.
3. @Import(AccountsWebApplication.class) - this Java Configuration class sets up everything else.

What makes this a microservice is the registration with the *discovery-server* via @EnableDiscoveryClient and its YML configuration completes the setup:

```
# Spring properties
spring:
  application:
     name: accounts-service
# Discovery Server Access
eureka:
  client:
    serviceUrl:
      defaultZone: http://localhost:1111/eureka/
# HTTP Server
server:
  port: 2222   # HTTP (Tomcat) port
```

Note that this file

1. Sets the application name as incidents-service. This service registers under this name and can also be accessed by this name.
2. Specifies a custom port to listen on (2222). All our processes are using Tomcat, they can't all listen on port 8080.
3. The URL of the Eureka Service process

Run the *IncidentsService* application now and let it finish initializing. Refresh the dashboard http://localhost:1111 and you should see the INCIDENTS-SERVICE listed under Applications. Registration takes up to 30 seconds (by default) so be patient - check the log output from *RegistrationService*

For more detail, go here: http://localhost:1111/eureka/apps/ and you should see something like this:

```
<applications>
    <versions__delta>1</versions__delta>
    <apps__hashcode>UP_1_</apps__hashcode>
    <application>
```

```
        <name>INCIDENTS-SERVICE</name>
        <instance>
            <hostName>autgchapmp1m1.corp.emc.com</hostName>
            <app>INCIDENTS-SERVICE</app>
            <ipAddr>172.16.84.1</ipAddr><status>UP</status>
            <overriddenstatus>UNKNOWN</overriddenstatus>
            <port enabled="true">3344</port>
            <securePort enabled="false">443</securePort>
            ...
        </instance>
    </application>
</applications>
```

Alternatively go to http://localhost:1111/eureka/apps/INCIDENTS-SERVICE and see just the details for Incidents*Service* - if it's not registered you will get a 404.

4. Conclusions

In the work an innovative approach to construct business process is presented. In opposite to typical application where the process is completely determined before its starting in this case process is not completely determined at the moment when its running is started. It is implemented using service that can be presented as handler for another process. That service composes process based on data obtained and runs it.
There is some lack of solutions for the problem presented. Although Microservice approach offers tools leading to development of system supporting emergency medicine. An advantage of the system offered is its usage based on mechanism of market. All services are searched through Internet. Moreover any institution can add its own service and in turn to join to the system.
It is not entirely known mechanism of integration of processes in emergency medicine. One of the most promising possibilities is application of Micriservices. Methods of search of appropriate services will be object of future research. In such case an application of ontological descriptions can be solution of the problem. Also the future investigations should be dealing with document formats for exchanging by the system elements, identification of basic services and taking into account economic factors when selecting services for patient

REFERENCES

4. SEMENETS A.S., MARTSENYUK V. P.: On the CDSS platform development for the open-source MIS OpenEMR, Med. Informatics Eng., (2015)3, 22–40.
5. Web page: http://microservices.io/patterns/monolithic.html
6. Web page: https://martinfowler.com/articles/microservices.html
7. Web page: https://www.infoq.com/articles/soa-healthcare
8. Web page: https://www.pcpcc.org/initiative/primary-care-information-project-pcip
9. WĄCHOCKI G.: Zastosowanie SOA do celów konstrukcji systemu wspomagającego ratownictwo medyczne, *Automatyka* 13/2 (2009), 653-661
10. Web page: https://spring.io/blog/2015/07/14/microservices-with-spring

Yevgeniy KOSYUK[1]

Scientific Supervisor: Liudmyla TEREIKOVSKA[2]

METODY TEORII PRZEKSZTAŁCEŃ FALKOWYCH W PROBLEMATYCE PROGNOZOWANIA OBCIĄŻENIA SERWERA INTERNETOWEGO

Streszczenie: W pracy opisano jest problem prognozowania obciążenia serwera internetowego. Wskazane są niedociągnięcia istniejących metod jego rozwiązania. Perspektywy wykorzystania metod teorii przekształceń falkowych są opisane w celu prognozowania parametrów determinujących obciążenie serwera internetowego.

Słowa kluczowe: Internet-serwer, parametry obciążenia, transformaty falkowe

METHODS OF THE THEORY OF WAVELET TRANSFORMATION IN THE PROBLEM OF FORECASTING OF INTERNET-SERVER LOAD

Summary: The problem of predicting Internet-server load. Disadvantages of existing methods of its solution. Outlines the prospects of methods wavelet transformation theory for predicting the parameters defining of Internet-server load.

Keywords: Internet-server, load the parameters of the wavelet transform

1. Formulation of the problem

Due to intensive growth of distributed network systems, the introduction of such systems in the various sectors of human activity the problem of ensuring the reliability of their operation is becoming more important and urgent [7]. The structure of almost all modern networked computer systems include one or more servers that integrate with the global Internet computer network. Typical functions of Internet-servers are providing services: WWW, FTP and e-mail.

[1] National Aviation University of Ukraine, Department of Information Security: Information security systems, Post-graduate student, yevgeniy_kosyuk @gmail.com
[2]Candidate of Technical Sciences, Kiev National University of Construction and Architecture, Department of Cybernetic Security and Computer Engineering, Associate Professor, terejkowski@ukr.net

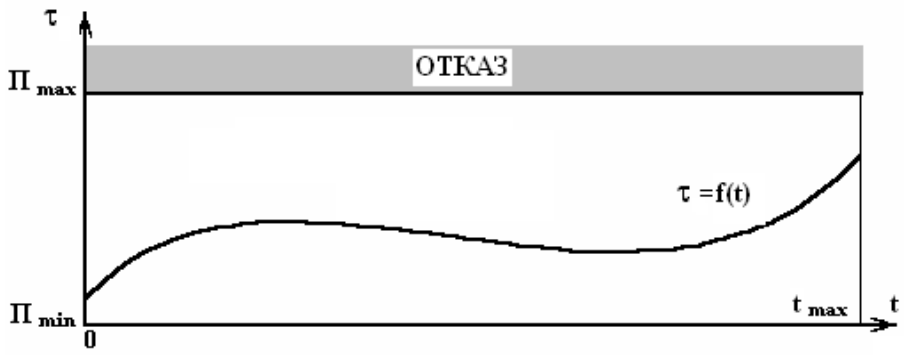
Practical experience indicates repeated violations performance network systems due to malfunction of the software Internet-server, as a result of a server failure, and as a result of a successful attack by hackers. For example, there are cases of failures net banking payment systems due to failures of the web server. This points to the need for research aimed at ensuring the required level of reliability of the software Internet-server, which is a common problem of this article.

2. Analysis of recent achievements and publications relied upon by the author

In general the term secure Internet-server software reliably realize its ability to perform certain functions at predetermined conditions during the predetermined time period. The problem of ensuring the required level of reliability of Internet-servers, the subject of many scientific and practical work, the analysis of which indicates that one of the main directions of its solutions is ensuring a balance between computing power of Internet-server and the expected load. It is assumed that the Internet-server load can be estimated by the value of one or more parameters controlled in the operation.

These parameters are called the defining parameters. It is also believed that as long as their values are within the set limits, the Internet-server is healthy. In the opposite case a failure occurs due to inability to perform the required amount of computational operations.

The range of values of characteristic parameters between the upper and lower predetermined region within a predetermined Internet-server functionality. Simplified illustration of this approach is rice. 1, which shows a graph of the change of one parameter in the operation efficiency.



τ – defining parameter Pmax, Pmin - upper and lower limit of the area efficiency, τ = F (t) - a function determining the parameter changes, t - the operation of Internet-server, tmax - limiting the operation of Internet-server

Figure 1. Graph determining parameter

In [5] pointed out that the task of calculating the optimal mode of maintenance work should be based on the values of the prediction model is determined by a parameter. Note that in the theoretical work in the field of information protection analogue

prediction model are normal patterns of behavior that apply in the systems detect network attacks.
Currently, there are quite a lot of the forecast model is determined by a parameter, however, the most widely used models based on statistical analysis of the data recorded in the operation of computer systems. Despite the great potential of these models, their use for predicting the technical state of Internet-server processing complexity difficult source statistics. By solving this problem would be to use the methods of wavelet transform theory, which has already proven to be effective in such cases.

3. The wording of Article purposes

To evaluate the possibility of using methods of wavelet transform theory to construct forecasting models of Internet-server load.

4. Statement of the main material

Obviously, the prospects of methods wavelet transformation theory to load forecasting of Internet-server can be estimated based on a ratio of capacity of these methods to the types of statistical data processing tasks in the formation of the corresponding prediction models.
In general, the technical system load forecasting model should take into account the trend of the process, and the frequency of the local features of the studied process.
The term trend understand the basic trend of the dynamics of the test process (permanent decline or rise). The trend shows the development process regardless of periodic oscillations and local features. As a rule, the trend predicted by the functional dependence of the expectation of the parameters characterizing the process from time to time. The definition of this function is now considered a trivial task that is divided into two stages.
In the first stage, based on the physics of the process, the approximating function is selected. In the second step the parameters are calculated approximation function. used mainly linear ($y = k\,x + b$), a parabolic ($y = k\,x^2 + b$) or exponential ($y = \exp(k\,x) + b$) function. Calculation of the parameters of these functions can be implemented by the method of least squares.
Under frequency (self-similar) process repeatability understand it at regular intervals. These intervals are called periods or lags. Complexity of analysis is periodic potential multiperiodichal and unsteadiness of the test process. In this case, it refers to non-stationary emergence and existence of periodic components. Fig. 2 shows an example of complex graphs three periodichal function $Y = \sin(t) + 2\sin(2t) + 4\sin(7t)$ and its components on the interval $t \in [0.360]$.
In this example, the periodicity of the function is independent of time.

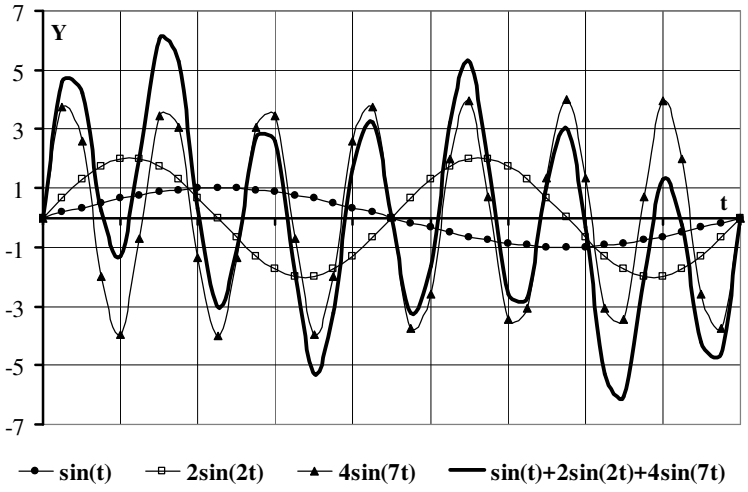


Figure 2. Schedule three periodical function and its components

Fig. 3 shows an example graph two periodical function $Y = \sin(t) + \sin(7t)$ and its components on the interval $t \in [0.360]$. In this example, a periodic component sin ($7t$) occurs only in part of the range of existence two periodical function. In consequence of this feature is dependent on the frequency of the time.

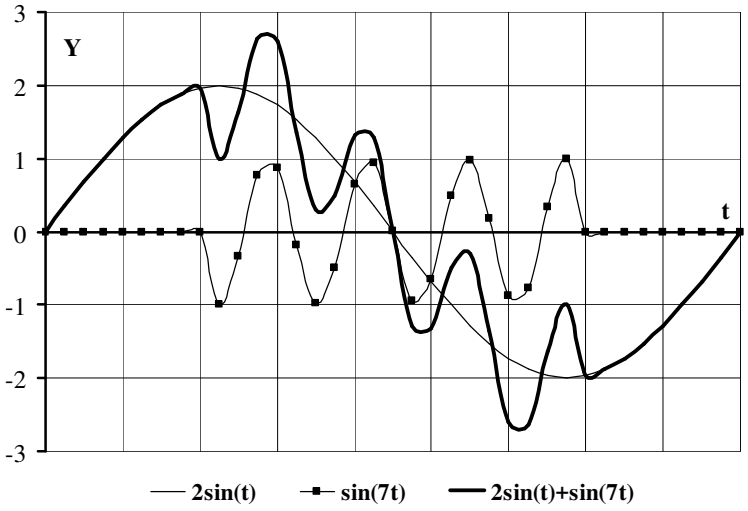


Figure 3. Schedule stationary two periodical function and its components

Analysis Fig. 2, 3 confirms the complexity of calculating the periodic components and process load change indicates the need for specialized techniques for such calculation. Another important factor to be taken into account when drawing up the load forecast are local features, ie sharp, abrupt changes in its characteristics. A corresponding

example of the function shown in Fig. 4. In Fig. 4 shows a graph of the one-periodic function Y = 2sin (t) in the interval $t \in [0.360]$ the local feature points in the vicinity of t = 120.

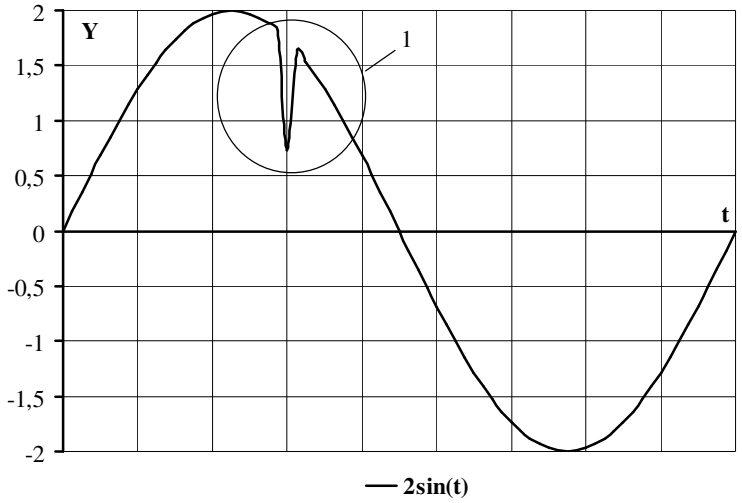


Figure 4. Schedule the one-periodic function having a local feature

Local features can be both random and systematic. Competent analysis of local features allows you to recover information on the dynamics of the technical state of "calm", a stable environment and a more reliable analytical information.

Another important task of drawing up a load forecast model Internet-server It is the preliminary processing of statistical data in order to isolate them from the noise components.

Consider the typical dependence of the change of parameters characterizing the Internet-server load [2, 3, 6]. In practice, different characteristics are used to assess the load, including the parameters of network activity and user settings references to resources Internet-servers. In this paper, as a load parameter used amount of TCP / IP packets received of Internet-server per unit time and the number of viewing Web pages. Fig. 5 and 6 are graphs amount TCP / IP packets received by the Web server when the frequency of registrations 1 minute and 1 hour. Fig. 7 is a graph of viewing a web page, with a recording frequency of 1 hour. Fig. 5, 6 are built according to [3], Figure 7 - Based on data www.finance.ua informational website dedicated to financial transactions.

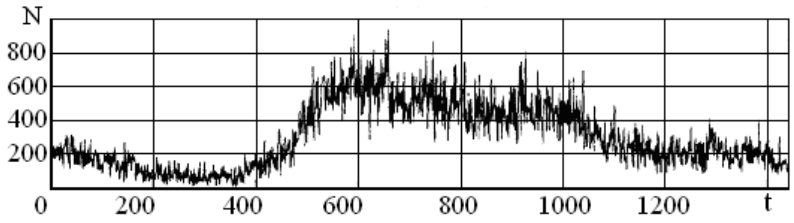


Figure 5. Schedule a TCP/IP packet number at a frequency of 1 minute registrations

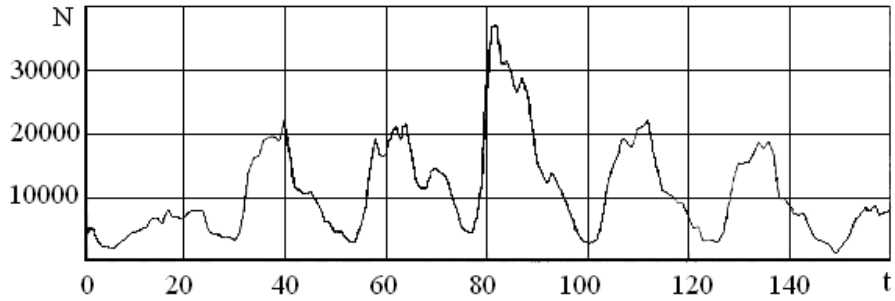


Figure 6. Schedule a TCP / IP packet number at a frequency of 1 hour registrations

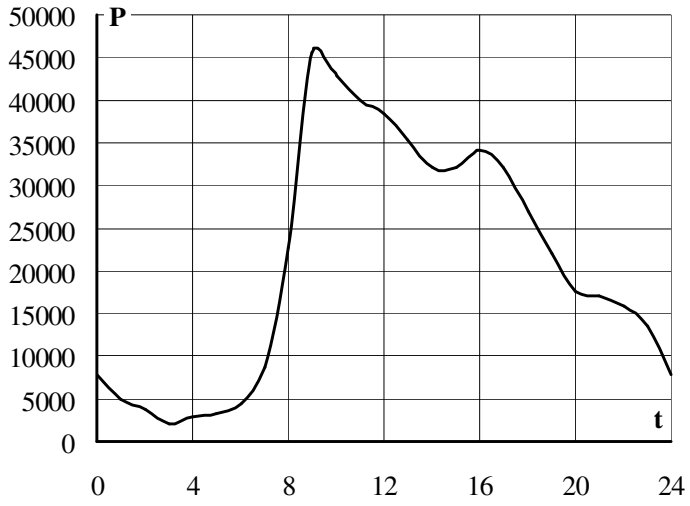


Figure 7. Schedule the number of Web page views

Obvious multiperiodichal dynamics load parameters, the dependence of some periodic components on the time and the need for preliminary isolation of the noise source statistics. In this case, in Fig. 5-7 there is a pronounced trend that can be explained in a short interval of observations and stable conditions for the functioning of Internet-server in these examples. Given sufficiently proven methodology for calculating the trend, we can conclude that its definition will mainly depend on the representative statistics. Definition of the main tasks of processing of statistical data

in the development of Internet-server load forecasting model allows us to pass to the analysis of possibilities of methods wavelet transformation theory [1, 4]. This theory is an extension of the spectral analysis, represented by the classical Fourier transform. Formally, integral wavelet transform of the function f (t)∈ *L2 (R)* written as

$$W(a,b) = |a|^{0,5} \int_{\infty}^{\infty} f(t)\psi\left(\frac{t\,b}{a}\right)\mathrm{dt}, \tag{1}$$

where ψ - the base wavelet, * - complex conjugation procedure, *a* - scale wavelet, *b* - wavelet shift, *a, b∈R, a≠0.*

Wavelets - a generic name of a family of functions well localized in a small neighborhood of a point and a drastically decreasing to zero as one moves away from it in the time and frequency domain. Area of wavelet is 0.In the family all functions are obtained from it by a single base changes (displacement field localization in time) and scaling (displacement field frequency localization). EXAMPLE wavelet shown in Fig.8.

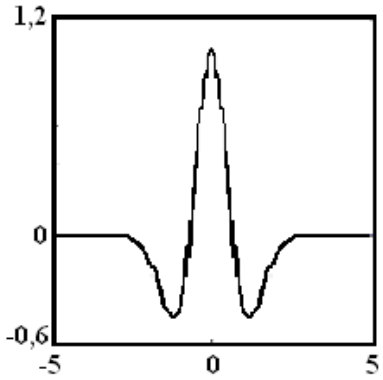


Figure 8. Gauss Wavelet

When carrying out a wavelet analysis of the classical concept of frequency spectral analysis replaced scale, and to block all wavelet time axis shift introduced functions over time. The wavelet transformation is considered analyzed temporary function in terms vibration localized in time and frequency. Distinguish discrete and continuous wavelet transform. Discrete used for transformation and coding of signals and continuous - for signal analysis. The meaning of the continuous wavelet transform is calculating scalar product (Value showing a degree of "similarity" of the two patterns of) the test data with different shifts of a wavelet at various scales. The result is a set of coefficients, showing how the process behavior at a given point like "behavior" wavelet on this scale. The closer the form of the analyzed dependence in the vicinity of the point of sight to the wavelet, the greater the absolute value of appropriate factor. Relative to the classical methods of spectral analysis advantage of wavelet analysis is the ability to determine the non-stationary signals not only the overall frequency response, but also information about specific local coordinates, which manifest themselves or another group of frequency components, or on which there are rapid changes in the frequency components of the signal. Additionally, one-dimensional

wavelet transform signals provides two-dimensional scanning, with the frequency and the coordinate are considered as independent variables, which allows signal analysis in two dimensions.

5. Conclusions

1. Basic methods of use prospects wavelet transformation theory for predicting the Internet-server load associated with the release of noise statistics and calculating the amount, duration and timing of periodic components dynamics performance parameters. Through the use of these methods may improve the accuracy of prediction models Internet-server load parameters that largely determine the reliability and security distributed network of computer systems.
2. Prospects for further studies in this direction are to develop on the basis of wavelet transformation theory model filtering and statistics calculation model periodic components load parameters dynamics Internet-servers. It is also necessary to develop tools and methods of application of these models.

REFERENCES

1. ASTAFIEVA N.M.: Wavelet analysis: basic theory and application examples. Successes of physical sciences, **166**(1996)11, 1145-1170.
2. DOVLAD O. A.: Doslidzhennya ta rozrobka modeli procesu ataky ta trafiku lokalnoï merezhi. Zahyst Informacji, (2009) 1, 83-86. (in Ukrainian)
3. MENASCO D., Virgilio A.: Performance Web-services. Analysis, evaluation and planning, transl. from English. DiaSoftYup, SPb., 2003.
4. DEACONS V.P.: Wavelets. From theory to practice. Solon-P, M., 2002.
5. TEREYKOVSKY I. A.: Kontseptsiya viznachennya optimal control mode zahischenosti software zabezpechennya Komp'yuterniy systems. Legal, Regulatory and Metrological Support Information Security System in Ukraine, (2006)1 Key infrastructure (12), 88-96. (in Ukrainian)
6. TEREYKOVSKAYA L. A.: Prospects for the use of methods wavelet transformation theory to predict technical condition of Internet-server. - Naukova-tehnichny zbirnik "Upravlinnya rozvitkom folding systems" Kiïvskogo natsionalnogo universitetu budivnitstva i arhitekturi. (2010)4 100-103. (in Ukrainian)
7. TEREYKOVSKAYA L. A.: Application of the discrete wavelet transform to determine the time-frequency operational load parameters Web-server remote education. - Naukova-tehnichny zbirnik "Upravlinnya rozvitkom convertible systems" Kiïvskogo natsionalnogo universitetu budivnitstva i arhitekturi, (2011)5, 124-127. (in Ukrainian)

Volodymyr KOVALOK [1], Andrii SEMENETS [2]

Supervisor: Vasyl MARTSENYUK [3]

ON CDSS PLATFORM DIALOG'S COMPONENT CODE REFACTORING FOR USAGE WITH THE OPEN-SOURCE MIS OPENMRS

Summary: The significance of Medical Information Systems (MIS) for medical practice is emphasized. The wide use of Electronic Medical Records (EMR) software is displayed. Benefits of the open-source MIS usage are shown. Effectiveness of the Clinical Decision Support System (CDSS) application in the medical decision making process is emphasized. The open-source MIS OpenMRS developer tools and software APIs are reviewed. The results of code refactoring of the dialog subsystem of the CDSS platform which is made as a module for the open-source MIS OpenMRS are presented. The structure of the information model of the CDSS subsystem database was updated according to the MIS OpenMRS requirements and Liquibase framework guidelines. The Model - View - Controller (MVC) based approach to the CDSS dialog subsystem architecture was re-implemented with Java programming language using Spring and Hibernate frameworks. The MIS OpenMRS Encounter portlet form for the CDSS dialog subsystem integration is developed as an extension. The OpenMRS administrative forms for the CDSS platform are created. The data exchanging formats and methods for the interaction of the OpenMRS CDSS dialog subsystem module and the GAE Decision Tree service are implemented with the help of AJAX technology through the jQuery library.

Keywords: medical information systems, electronic medical records, decision support systems, decision tree, open-source software, MIS, EMR, OpenMRS, CDSS, Java, Spring, Hibernate, Google App Engine.

REFRAKTORING KODU KOMPONENTU DIALOGOWGO OPROGRAMOWANIA DLA SZPITALI – DOPASOWANIA APLIKACJI DO POTRZEB UŻYTKOWNIKA

Streszczenie: W pracy podkreślono znaczenie medycznych systemów informatycznych (MSI) w praktyce leczniczej. Omówiono oprogramowanie do analizy elektronicznych zapisów medycznych (EZM). Przedstawiono zalety systemów MSI o otwartym kodzie. Ponadto, omówiono system wspierania decyzji klinicznych (SWPDK). Między innymi, rozpatrywany

[1] Ternopil State Medical University, engineer at Information Technology department, kovalok@tdmu.edu.ua

[2] Technical Science. Ph.D., Ternopil State Medical University, associate professor of Medical Informatics department, semteacher@tdmu.edu.ua

[3] Prof. DSc, University of Bielsko-Biala, professor of Department of Computer Science and Automatics, vmartsenyuk@ath.bielsko.pl

jest system o kodzie otwartym MSI OpenMRS oraz odpowiednie API. Przedstawiono wyniki refraktoringu kodu podsystemu dialogowego platformy SWPDK. System bazy danych systemu informatycznego SWPDK został dostosowany do wymagań SMI OpenMRS oraz poleceń szablonu Liquibase. Podejście typu Model-Widok—Sterownik (MWS) zostało wdrożone z użyciem języka programowania Java oraz szablonów Spring oraz Hibernatus. Wprowadzono wiele szczegółowych rozwiązań informatycznych – jak na przykład: stworzono formy administracyjne do platformy SWPDK, natomiast formaty wymiany danych oraz metody interakcji pomiędzy: podsystemem dialogowym SWPDK OpenMRS a serwisem GAE Decision Tree są realizowane z pomocą technologii AJAX przez bibliotekę jQuery.

Słowa kluczowe: medyczne systemy informacyjne, elektroniczne zapisy medyczne, systemy wspierania podejmowania decyzji, drzewo decyzyjne, oprogramowanie o otwartym kodzie zrodlowym, MSI, EZM, OpenMRS, SWPDK, Java, Spring, Hibernate, Google App Engine.

Introduction

The importance of wide application of the Medical Information Systems (MIS) as a key element of informatization of healthcare, especially in Ukraine, is shown in [1, 2]. The development of information technologies makes it possible to improve the quality of medical care by providing medical personnel with hardware and software tools for the efficient processing of clinical information [2, 3]. A conceptual direction of modern information technologies adoption in hospitals pass through patient's Electronic Medical Record (EMR) formation and support [1, 2, 3].

1. On the Decision Support Systems application - as part of an open-source MIS usage

An overview of approaches of implementation into as well as brief list of the leading MIS developers is given in [2]. MIS global market has stable positive dynamics as it is shown in [4]. A few high-quality MIS has been created by Ukrainian software development companies too, for example, "Doctor Elex" (http://www.doctor.eleks.com), "EMSiMED" (http://www.mcmed.ua), etc. In fact, all they are commercial software with a high cost [1].

An open-source-based software solutions for healthcare has been actively developing for the last decade along with the commercial software applications [3, 5, 6]. Most widely used open-source MIS EMR are WorldVistA (http://worldvista.org/), OpenEMR (http://www.open-emr.org/) and OpenMRS (http://openmrs.org/) [3, 7]. Advantages of such MIS software are shown in [2, 3]. Prospects for open-source and free MIS software usage in developing countries, or countries with financial problems has been considered by F. Aminpour, F. Fritz, C. J. Reynolds and others [3, 5, 7]. The approaches to implementing open-source MIS, especially OpenEMR, OpenMRS and OpenDental, in Ukraine healthcare system has been studied as well as methods of integrating these MIS EMR with other MIS software has been developed by authors of this work during last few years [2, 8, 9].

Clinical Decision Support Systems (CDSS) regular usage in physician's practice is strongly recommended for improving of the quality of care. This thesis was confirmed in [10, 11, 12]. Advantages of CDSS usage in healthcare systems of the developing countries was shown in [13]. The importance of integration of different types of MIS,

and MIS EMR with CDSS especially, is provided in [14]. The CDSS theoretical approaches as well as software applications has been developed by TSMU Medical Informatics Department staff for last few years [15 - 19].
Approaches of the CDSS usage in obstetrics for early detection of pathologies of miscarriage of pregnancy are analyzed in [20, 21, 22]. A prototype of such CDSS has been developed by Semenets AV, Zhilyayev MM and Heryak SM in 2013 [23]. The effectiveness of proposed algorithm was confirmed by experimental exploitation of this CDSS prototype in the Ternopil state perinatal center "Mother and Child" during 2013-2015 that is proved in [24]. As result, the fully functional CDSS application for miscarriage pathology diagnostic has been developed by authors in form of an information module (plugin) for free- and open-source MIS OpenEMR [25, 26].
Purpose of this work is to represent authors' experience on code refactoring of the plugin, which implements dialog component of custom CDSS platform, for usage with free- and open-source MIS OpenMRS.

2. The CDSS platform dialog component's code refactoring implementation

The alternative method of the decision making process, based on the algorithm for induction of "decision trees", was proposed by VP Martsenyuk as result of preceding investigations described in [16-19]. Finally, given decision-making diagnostic algorithm was implemented with Java programming language as a web-service for the Google App Engine platform. A web-service training database has been deployed to Google Datastore service, which is a form of no-SQL data warehouse [25, 26]. This approach provide flexible way to integrate above GAE Decision Tree service with third-party MIS EMR by developing appropriate dialog components (modules, plugins) as well as administrative tools (Fig. 1). Therefore the feasibility of CDSS dialog component's plugin [25, 26] code refactoring for usage with free- and open-source MIS OpenMRS is obvious.

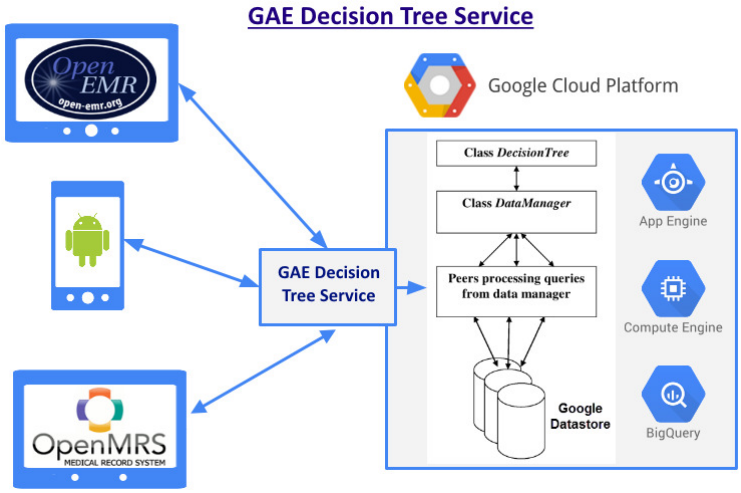


Figure. 1 Integration of the GAE Decision Tree CDSS web service with arbitrary EMR MIS

2.1. The OpenMRS add-ons (modules) development capabilities

OpenMRS is a free- and open source software platform dedicated to develop and integrate MIS EMR solutions (https://github.com/openmrs/). This MIS is focused on EMR automation of primary health care institutions like ambulances and small clinics. Several academics and non-governmental organizations, including the Institute Regenstrief (http://regenstrief.org/) and In Health Partners (http://pih.org/), are responsible to support and maintain OpenMRS core code. There are dozens of implementation [27] are registered, mainly in Africa and Asia (https://atlas.openmrs.org/).

The OpenMRS core is written in Java programming language using Spring and Hibernate frameworks. An MySQL RDBMS is used as data storage. There are tree main way to perform OpenMRS customization and adoption process:

1. The visual interactive editor for managing templates of patient registration forms and their components - Concepts, Form Metadata and Form Schema – Form Administrator (https://wiki.openmrs.org/display/docs/Administering+Forms).
2. The tool for integration of forms, developed by InfoPath (http://www.infopathdev.com/) – InfoPath Administrator (https://wiki.openmrs.org/display/docs/InfoPath+Notes).
3. Set of programming interfaces (API) for creating custom modules using Java programming language (https://wiki.openmrs.org/display/docs/API and https://wiki.openmrs.org/display/docs/Modules).

Tools 1 and 2 are easy-to-use and do not require knowledge of programming languages. However, they do not have features which are required to implementation of given CDSS. Therefore, OpenMRS Modules API has been selected to develop a module that implements features of the dialog component of CDSS platform. Corresponded module architecture is shown on Fig. 2.

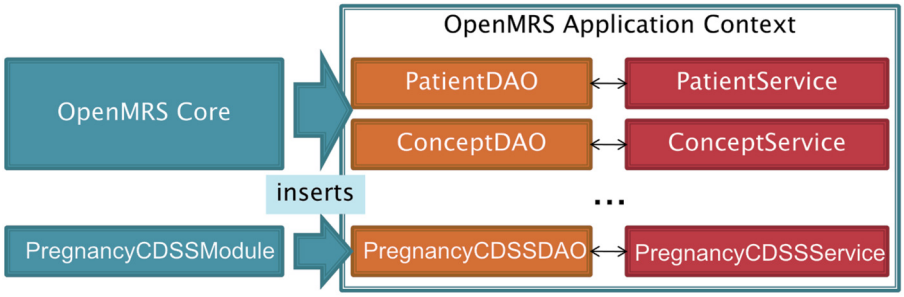


Figure 2. Software architecture of Pregnancy CDSS module for OpenMRS that implements the dialog component of the CDSS platform

2.2. Adaption of the information model of dialog component of the CDSS module

The external representations of the information model (IM) of CDSS dialog component, as well as the necessary data structures, are described in [25, 26]. The internal representation of information model has been adapted according to OpenMRS database requirements for the custom modules (https://wiki.openmrs.org/display/docs/Data+Model):

- a mechanism of IM key concepts identification by the universal identifier (UUID) values assignment has been introduced (https://wiki.openmrs.org/display/docs/UUIDs);
- some tables key field data types has been adopted according to OpenMRS coding guidelines (https://wiki.openmrs.org/display/docs/Conventions);
- module's database tables installation procedure according Liquibase technology (http://www.liquibase.org) description has been developed and set of special XML files has been formed.

Data structures for the recorded patient's data representation has been developed as the following Java-classes according to general (MVC, Model - View - Controller) approach adoption with the Spring framework usage (Fig. 3):

- SymptCategoryModel.java - represent symptom's categories;
- SymptomModel.java - represent symptom's description;
- SymptomOptionModel.java – represent possible symptom's values;
- DiseasesSymptOptModel.java – represent information about probability of a certain diagnosis depending on the given symptom's value;
- PatientExamModel.java – represent general Patient questionnaire data model;
- PatientSymptomByExamModel.java – represent each patient's questionnaire submission.

The Java Hibernate framework should be used within OpenMRS to implement database management operations according coding guidelines (https://wiki.openmrs.org/display/docs/For+Module+Developers. Therefore, necessary service classes has been developed as shown on the Fig. 3.

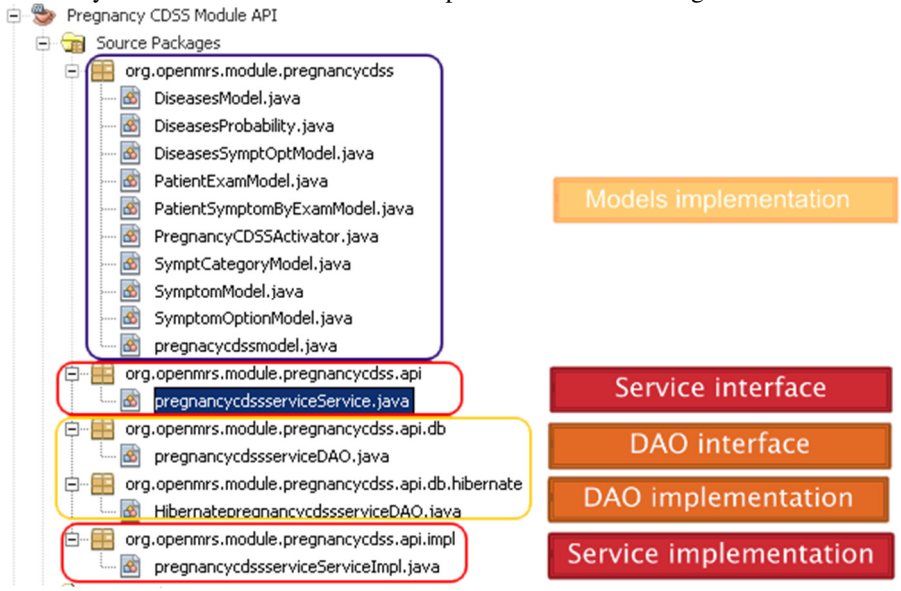


Figure 3. Structure of Java classes of Models and Services for patient data management in the Pregnancy CDSS module for OpenMRS MIS

2.3. CDSS dialog component User Interface development

Most of modern web- technologies could be used for UI development of OpenMRS custom modules, including HTML 5, CSS 3, AJAX (JQuery usage is recommended). According to above, set of flexible forms and reports has been developed to effectively implement necessary Pregnancy CDSS module User Interface views according to IM external representations as it was shown in [25] and MVC paradigm (Fig. 7). Those views are included especially:

– patientExamForm.jsp – the patient's survey main form (Fig. 4);
– encounterPatientExamData.jsp – the portlet which represent pregnancy miscarriage pathology diagnostic data, provided by Pregnancy CDSS module, inside OpenMRS patient encounter form (Fig. 5);
– patientExamForm2Print.jsp – the survey report with patient's answers and diagnostic conclusion;
– series of forms under OpenMRS Administration section for the CDSS platform dialog component content management, settings adjustment and configuration customization (Fig. 6).

Figure 4. A patient's survey form View in the Pregnancy CDSS module for OpenMRS MIS

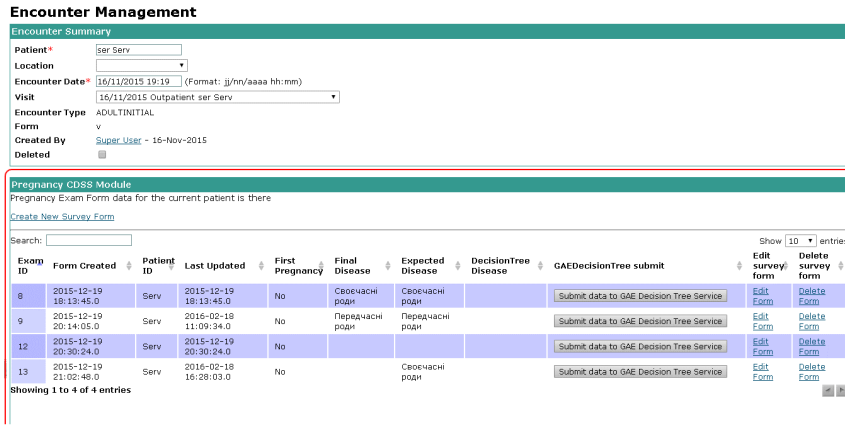


Figure 5. Representation of pregnancy miscarriage pathology examination summary, provided by Pregnancy CDSS module, inside OpenMRS patient encounter form

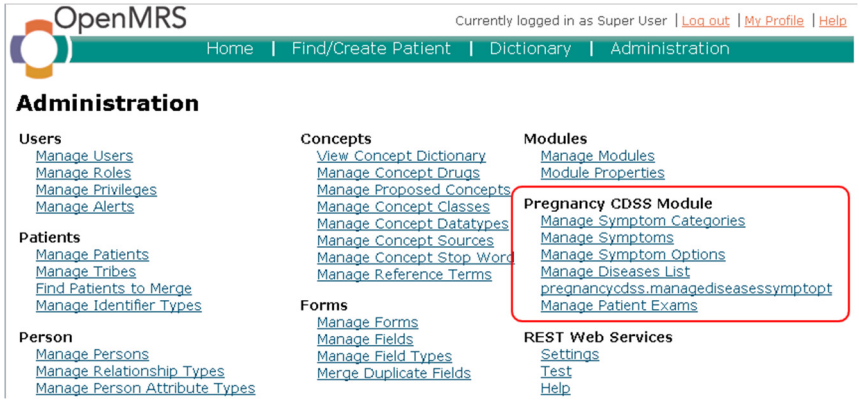


Figure 6. Pregnancy CDSS module customization forms under OpenMRS Administration section

Main decision-making algorithm are based on results of research obtained in [19, 23, 25, 26]. This algorithm as well as common module's management activities has been implemented in form of Java servlets, according to general MVC approach (Fig. 7):

– EncounterPatientExamDataPortletController.java – portlet controller to manage module data representation within OpenMRS patient's encounter form;
– PatientExamFormController.java – patient's survey form controller;
– GAEDecisionTreeController.java – provides interaction of the Pregnancy CDSS module with GAEDecisionTree diagnostic web-service;
– PregnancyCDSSManageController.java - provides Pregnancy CDSS module administrative features and customization capabilities.

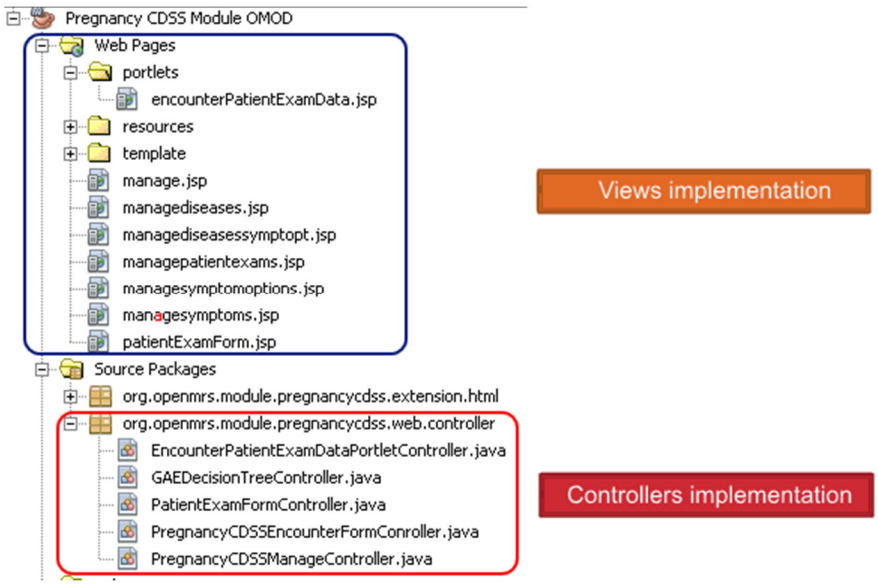


Figure 7. Structure of Java classes of views and controllers in the Pregnancy CDSS module for OpenMRS MIS

The presented CDSS platform dialog's component and provided GAE Decision Tree web-service interaction procedure has been developed according to recommendations how to cross-site data request being performed (http://www.gwtproject.org/doc/latest/tutorial/Xsite.html#design). The following methods of the GAEDecisionTreeController.java controller are responsible for:

- getPatientDataJson2 – handles GET-type of HTTP request and returns data for the selected survey form as a JSON object;
- getAllPatientDataJson – handles GET-type of HTTP request and returns data for all survey forms, where final diagnosis is given, as a JSON object. It is used for the training dataset formation during GAE Decision Tree web-service education stage (http://decisiontree-1013.appspot.com);
- setGAEDecision – handles POST-type of HTTP request and store GAE Decision Tree diagnostic output in Pregnancy CDSS module database for appropriate patient's record.

Practically, Querying service GAE Decision Tree service has been queried directly from view (portlet encounterPatientExamData.jsp) with AJAX technology using jQuery library via the following code sniplet (listing 1):

- gaeDecisionTreeSubmitFunction – retrieves a survey form data by asynchronous calling of the getPatientDataJson2 method of the GAEDecisionTreeController.java servlet;
- submitData2GAE – submits a survey form data to the GAE Decision Tree service via asynchronous request;
- setDecisionTreeResponceFunction – receives a diagnostic conclusion provided by GAE Decision Tree service and redirect it to the GAEDecisionTreeController.java servlet by asynchronous calling of the setGAEDecision method.

A training dataset deployment to the GAE Decision Tree service has been implemented in the same way within the managepatientexams.jsp view in OpenMRS administrative panel of the Pregnancy CDSS module.

Listing 1. Implementing of asynchronous interaction of the OpenMRS Pregnancy CDSS module with the GAE Decision Tree web-service

```
<script type="text/javascript">
  function submitData2GAE(formData){
    jQuery.ajax({
      type : 'GET',
      url              :              'http://decisiontree-
1013.appspot.com/patientdata',
        data : formData,
        dataType : 'json',
        success : function(response) {
          var mystr = JSON.stringify(response);
          setGAEDecision (response);
        },
        error : function(e) {
          alert('Error: ' + e);
        }
    });
  };
```

```
  function
gaeDecisionTreeSubmitFunction(examId,encounterId,patientId){
    jQuery.ajax({
      type : 'GET',
      url                                                    :
'${pageContext.request.contextPath}/module/pregnancycdss/gAE
DecisionTree/single.json',
      data  :  'examId='  +  examId  +  '&encounterId='  +
encounterId + '&patientId=' + patientId,
      dataType : 'json',
      success : function(response) {
         submitData2GAE(response);
      },
      error : function(e) {
        alert('Error: ' + e);
      }
    });
  };
  function setGAEDecision(GAEresponse){
    jQuery.ajax({
      type : 'POST',
      url                                                    :
'${pageContext.request.contextPath}/module/pregnancycdss/gAE
DecisionTree/setdisease.json',
      data : gAEresponse =' + GAEresponse,
      dataType : 'json',
      success : function(response) {
         alert('Sucessfully saved!');
      },
      error : function(e) {
        alert('Error: ' + e);
      }
    });
  };
</script>
```

The Pregnancy CDSS module installation process has been performed according general OpenMRS administration guide (https://wiki.openmrs.org/display/docs/Administering+Modules):

1. Download the Pregnancy CDSS module compiled file (pregnancycdss-1.xx-SNAPSHOT.omod) from author's GitHub repository (https://github.com/semteacher/pregnacy_cdss).
2. Log in to OpenMRS as administrator. Go to MIS module administration page (*Administration →Manage Modules*).
3. Press *Add or Upgrade Module* button. In "Add Module" dialog click *Choose File* in the *Add Module* section. Specify downloaded module file location and click *OK* than *Upload*.
4. After installation will complete – new "Pregnancy CDSS Module" section will appears in OpenMRS patient Encounter form (Fig. 4).

Conclusion

Effectiveness of the Clinical Decision Support System (CDSS) application in the medical decision making process has been signed. An opportunities provided by CDDS in diagnostics of miscarriage pathologies with aim to prevent of preterm birth has been shown as a result of trial evaluation of the CDSS prototype in Ternopil regional perinatal center "Mother and Child".
An approach to the decision making process which is based on the decision tree algorithm has been recommended. The implementation of the given above approach as separate web-service based on the Google App Engine (GAE) capabilities has been provided.
The results of code refactoring of the dialog subsystem of the CDSS platform which is made as module for the open-source MIS OpenMRS has been presented. The Model-View-Controller (MVC) based approach to the CDSS dialog subsystem architecture has been implemented with Java programming language using Spring and Hibernate frameworks. The OpenMRS Encounter portlet form for the CDSS dialog subsystem integration has been developed as a module. The data exchanging formats and methods to establish interaction between OpenMRS newly-developed Pregnancy CDSS module and GAE Decision Tree service are developed with AJAX technology via jQuery library.
The prospects for the further research is to extend web-service core decision tree algorithm capabilities to support different types of diagnostic problems. Such achievements will allow to more comprehensive end more effective utilize of patient's health data which are collected within both supported MIS – OpenEMR and OpenMRS.

REFERENCES

1. AVRAMENKO V.I., KACHMAR V.O. (2011). Creation of the new ways for development of information technologies in medicine for ukrainian health care grounded at worlwide approaches. Ukrainian of Telemedicine and Medical Telematics, (9, № 2), 124–133. Retrieved from http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILA=&2_S21STR=Ujtm_2011_9_2_3 [in Ukrainian]
2. SEMENETS A. V.: On organizational and methodological approaches of the EMR-systems implementation in Public Health of Ukraine. Medical Informatics and Engineering, 23(2013)3, 35–42. doi:10.11603/mie.1996-1960.2013.3.1742 [in Ukrainian]
3. AMINPOUR F., SADOUGHI F., AHAMDI, M. (2014). Utilization of open source electronic health record around the world: A systematic review. Journal of Research in Medical Sciences : The Official Journal of Isfahan University of Medical Sciences, 19(1), 57–64. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3963324&tool=pmcentrez&rendertype=abstract
4. Global Healthcare IT Market Analysis And Segment Forecasts To 2020 - Healthcare IT Industry, Outlook, Size, Application, Product, Share, Growth Prospects, Key Opportunities, Dynamics, Trends, Analysis, Healthcare IT Report

- Grand View Research Inc. (n.d.). Retrieved May 1, 2015, from http://www.grandviewresearch.com/industry-analysis/healthcare-it-market
5. REYNOLDS C. J., WYATT, J. C.: Open source, open standards, and health care information systems. Journal of Medical Internet Research, 13(2011)1, e24. doi:10.2196/jmir.1521
6. List of open-source healthcare software - Wikipedia, the free encyclopedia. Retrieved from: http://en.wikipedia.org/wiki/List_of_open-source_healthcare_software#Electronic_health_or_medical_record.
7. FRITZ F., TILAHUN B., DUGAS M.: Success criteria for electronic medical record implementations in low-resource settings: a systematic review. Journal of the American Medical Informatics Association, 22(2015)2, 479–488. doi:10.1093/jamia/ocu038
8. SEMENETS A. V.: About expirience of the patient data migration during the open source emr-system implementation. Medical Informatics and Engineering, 25(2014)1, 28–37. doi:10.11603/mie.1996-1960.2014.1.3756 [in Ukrainian]
9. SEMENETS A. V., KUZIV, N. I.: About experience of the dr.suni1800 intraoral radiovisiograph and opendental emr-system integration. Medical Informatics and Engineering, 26(2014)2, 43–50. doi:10.11603/mie.1996-1960.2014.2.3841.
10. ROSHANOV P. S., FERNANDES N., WILCZYNSKI J. M., HEMENS B. J., YOU J. J., HANDLER S. M., HAYNES R. B.: Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ (Clinical Research Ed.), 346(2013)feb14_1, f657. doi:10.1136/bmj.f657
11. BRIGHT T. J., WONG A., DHURJATI R., BRISTOW E., BASTIAN L., COEYTAUX R.R., LOBACH D.: Effect of clinical decision-support systems: a systematic review. Annals of Internal Medicine, 157(2012)1, 29–43. doi:10.7326/0003-4819-157-1-201207030-00450
12. JASPERS M.W.M., SMEULERS M., VERMEULEN H., PEute L.W.: Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. Journal of the American Medical Informatics Association : JAMIA, 18(2011)3, 327–34. doi:10.1136/amiajnl-2011-000094
13. ESMAEILZADEH, P., SAMBASIVAN, M., KUMAR, N., & NEZAKATI, H. (2015). Adoption of clinical decision support systems in a developing country: Antecedents and outcomes of physician's threat to perceived professional autonomy. International Journal of Medical Informatics. doi:10.1016/j.ijmedinf.2015.03.007
14. GOLDSPIEL B. R., FLEGEL W. A., DIPATRIZIO G., SISSUNG T., ADAMS S.D., PENZAK S.R., MCKEEBY J.W.: Integrating pharmacogenetic information and clinical decision support into the electronic health record. Journal of the American Medical Informatics Association : JAMIA, 21(2014)3, 522–8. doi:10.1136/amiajnl-2013-001873
15. MARTSENYUK V.P., SEMENETS A.V. Medical Informatics. Instrumental and Expert systems. Ternopil: Ukrmedknyha. (2004) [in Ukrainian]
16. BORYS R.M., MARTSENYUK V. P.: Classification algorithm polytrauma by induction of decision trees. Medical Informatics and Engineering, (2013)2, 12-17. doi: 10.11603/mie.1996-1960.2013.2.1693 [in Ukrainian]
17. MARTSENYUK V.P., ANDRUSHCHAK I. Y.: Development of clinical exert system based on rules with help of method of sequential covering. Proceedings

[Petro Mohyla Black Sea State University complex "Kyiv-Mohyla Academy"], 237(2014) 225, 5-10. [in Ukrainian]
18. MARTSENYUK V. P., STAKHANSKA O.O.: About clinical expert system based on rules using data mining technology. Medical Informatics and Engineering, (2014)1, 24-27. doi: 10.11603/mie.1996-1960.2014.1.3788
19. MARTSENYUK V.P., ANDRUSHCHAK I. Y., GVOZDETSKA, I. S.: Qualitative Analysis of the Antineoplastic Immunity System on the Basis of a Decision Tree. Cybernetics and Systems Analysis, 51(2015)3, 461–470. doi:10.1007/s10559-015-9737-6
20. HASMIK MARTIROSYAN, MONIQUE FRIZE, DAPHNE E. ONG, JEFF GILCHRIST E. B.: A Decision-Support System for Expecting Mothers and Obstetricians. In 6th European Conference of the International Federation for Medical and Biological Engineering MBEC 2014, 7-11 September 2014, Dubrovnik, Croatia. Springer International Publishing (Vol. 45, pp. 703–706).
21. EDELMAN E. A., LIN B. K., DOKSUM T., DROHAN B., EDELSON, V., DOLAN S. M., SCOTT J.: Evaluation of a novel electronic genetic screening and clinical decision support tool in prenatal clinical settings. Maternal and Child Health Journal, 18(2014)5, 1233–45. doi:10.1007/s10995-013-1358-y
22. PAHL C., ZARE M., MEHRBAKHSH N., BORGES M. A. DE F., WEINGAERTNER D., DETSCHEW V., IBRAHIM O.: Role of OpenEHR as an Open Source Solution for the Regional Modelling of Patient Data in Obstetrics. (2015) Journal of Biomedical Informatics. doi:10.1016/j.jbi.2015.04.004
23. SEMENETS A. V., ZHYLIAYEV M. M., HERYAK S. M. (2013). The computer program “Information system of the decision support ”Pregnancy". Ukraine, patent N 51256.
24. ZHYLIAYEV M. M., HERYAK S. M. Effectiveness of computer screening system for diagnosing and prediction of preterm delivery. Medical Newsletter of Vyatka, (2014) 2, 18-22. Retrieved from http://cyberleninka.ru/article/n/effektivnost-primeneniya-kompyuternoy-skriningovoy-sistemy-dlya-diagnostiki-i-prognozirovaniya-prezhdevremennyh-rodov [in Russian]
25. SEMENETS A. V., MARTSENYUK, V. P.: On the CDSS platform development for the open-source MIS OpenEMR. Medical Informatics and Engineering, (2015)3, 22–40. doi:10.11603/mie.1996-1960.2015.3.4999 [in Ukrainian]
26. MARTSENYUK V., SEMENETS A.: System elektronicznych zapisów medycznych dla wspomagania decyzji z wykorzystaniem Google Application Engine (GAE). Studia Ekonomiczne, (2016), 308, 157–172. Retrieved from http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-b7e3b3be-367a-4df6-b7d8-97f5cb8684be?q=bwmeta1.element.cejsh-9348705d-114a-444d-963b-6854d1bbcccb;11&qt=CHILDREN-STATELESS [in Polish]
27. MOHAMMED-RAJPUT N. A., SMITH D. C., MAMLIN B., BIONDICH P., DOEBBELING B. N. (2011, October 1). OpenMRS, A Global Medical Records System Collaborative: Factors Influencing Successful Implementation. Retrieved from https://scholarworks.iupui.edu/handle/1805/6663

Oleksandra KUCHVARA[1]

Scientific supervisor: Vasyl MARTSENYUK [2]

ON CONCEPTUAL MODEL OF INFORMATION SYSTEM FOR EPIDEMIOLOGICAL RESEARCH

Summary: Web-integrated software environment for the modeling of epidemiological disease development is proposed in this paper. We used mathematical models of population dynamics which is developed to the description of different virus strains interaction. Program implementation is based on Java-technology.

Keywords: mathematical epidemiology, optimal vaccination, Java-technology.

MODEL KONCEPTUALNY SYSTEMU INFORMACYJNEGO BADAŃ EPIDEMIOLOGICZNYCH

Streszczenie: W artykule jest zaoferowano środowisko modelowania rozwoju choroby epidemiologicznej. W tym celu korzystaliśmy z matematycznych modeli dynamiky populacyjnej, którzy są opracowane dla opisywania interakcji różnych szczepów wirusów. Implementacja programu jest na podstawie technologii Java.

Słowa kluczowe: epidemiologia matematyczna, wakcynacja optymalna, technologia Java.

1.Introduction

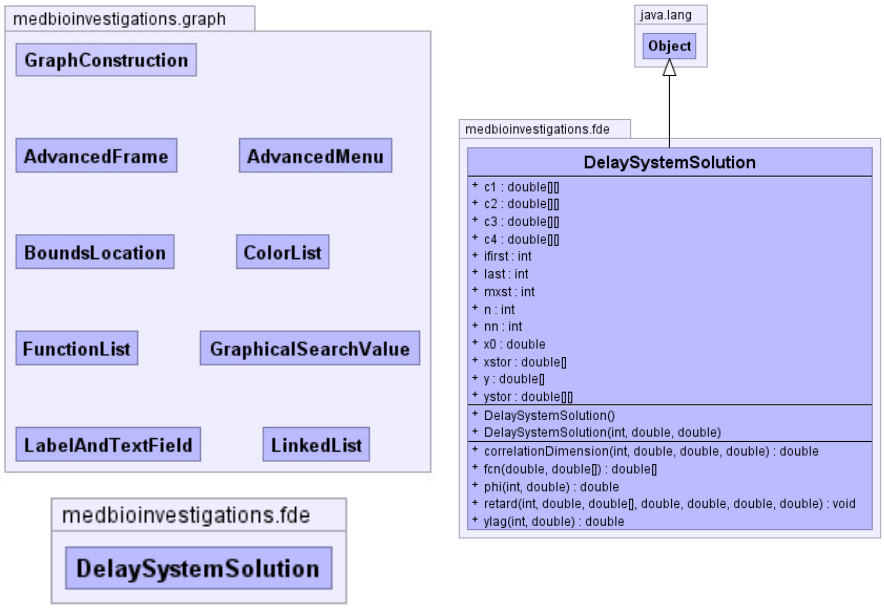
A flu and infectious diseases remain one of the most actual medical and social problems [1]. The special value is acquired by advanced scientific researches dealing with analysis and prognosis of probable scenarios of development of epidemics. Prognostication of origin and extension of epidemics is needed for the most effective realization of immunoprophylactic events, and in particular, it is important both for realization of educational actinity and for the organizational planning of events.

[1] Mgr., Ternopil State Medical University, an assistant professor at the department of medical informatics, kuchvara@tdmu.edu.ua
[2] Prof. D.Sc., University of Bielsko-Biala, a professor of department of computer science and automatics

The aim of this article is on the basis of analysis of principles and approachesto computer simulation and on the basis of existent software that is used in epidemiology researches, to offer the web-integrated software environment that will implement basic mathematical models [2].

2. The conceptual approach for the software environment of support of epidemiological system researches

For providing of functioning of the system is reasonable to develop Decision Support System (DSS) for epidemiological system researches. In a functioning of the DSS three groups of users participate: epidemiologists which are scientists in field of medical epidemiology (workers of departments and laboratories of infectious diseases, microbiology and other); state epidemiologic service; system analysts which are developers of compartmental epidemiological models [3]. The generalized diagram of classes of the system is presented in Figure 1.



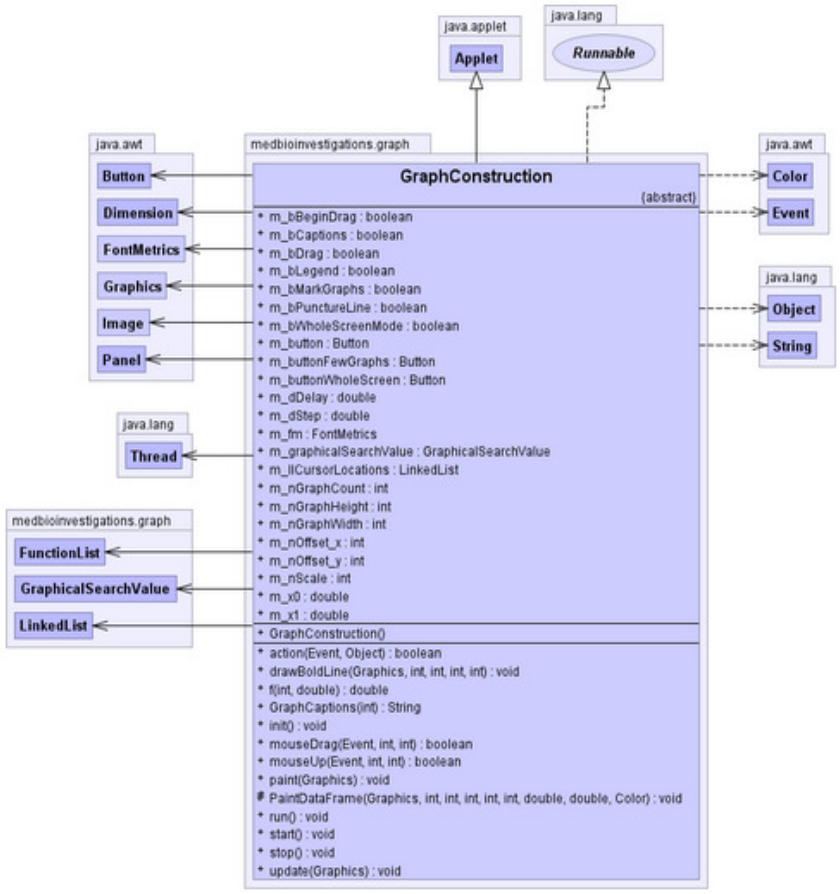


Figure 1. UML-models of java-classes of project for client application "System analyst"

A productional model is chosen as basis of base of knowledge of DSS. The functions of the base of knowledge are: choice of mathematical model (in accordance with data which are transfered by epidemiologists and queries of epidemiological service) and inferencing (in accordance with results of modeling). It is implemented in the environment of server of applications Apache Tomcat. The results of the system can be displayed as videograms (in the character or graphic mode) and as machine-reports (the printed reports after every stage of decision-making or fundamental analysis after the certain stages which are determined by the user).

When developing DSS of epidemiologic system researches, it is necessary to take into account a requirement of simultaneous access of many users to information. For the solution of problems of support of the clustered access to data and effective use of resources of the system, client-server architecture is chosen, that contains the following components in the composition:

Server of databases which manages the databases of project. Database Management System (DBMS) solves a problem of the reception/ sending of data to workstations (WS).

Client applications are WS "System analyst", "Epidemiologist" and "Epidemiological service", which allow to enter primary data, implement queries to the database and base of knowledge, calculation of input coefficients of models; they calculate operations of the design of processes and implement a "conclusion based on knowledge" and generation of managing influences.

The basic goal of application of the client-server architecture is attempt to bring down requirements to the client computers, that it could be possible to use the underproductive workstations (even diskless network terminals). Hence, considerable part of software resources, which are related to the management of data, are carried on the separate server of databases. It is the server that co-operates directly with databases, realizing mediation between them and client applications.

For the purpose of implementation of client-server architecture within DSS of epidemiologic system researches JDBC technology was chosen. In JDBC the model of work of user is used without connection with the source of data. Applications are connected to the database only on the small interval of time. Connection is set on only if a client from a remote computer queries data on a server. Since a server prepared the necessary set of data, formed and sent them to the client as a web-page, connection of application with the server is set off at once, and the client looks over obtained information without connection with the server.

In this model we differ three levels:

- the level of data;
- the level of business-logic;
- the level of application.

A level of data is a base level on that the given (for example, tables of database of MySQL) are located. At this level physical storage of information and manipulation are provided with data at the level of initial tables (selection, sorting, inserting, removing, updating and others like that).

A level of business-logic is a set of objects that determine, with what database it is required to set connection and what actions are needed to be peformed with information that is stored in it. For a connection with databases the object DataManager is used. For storage of commands that make actions with data, the objects-peers Peer are used. And, finally, if the process of selection of information from a database was executed, the object ResultSet is used for storage of results of selection.

A level of application is a set of objects that allow to keep and display data on the computer of end-user. For storage of information the object ResultSet is used, and for the displaying of data there is a large enough set of visual custom controls, for example RIA-components.

Classes and algorithms that provide implementation of certain methods of classes and software of DSS of epidemiologic system researches are implemented as a server application in Java-language in the integrated developer environment Netbeans and DBMS MySQL. DSS has an iterative interface. A window contains all needed in accordance with current requirements elements of Windows-oriented applications, in particular, program menu, facilities of management child windows, toolbar and statusbar.

An example of presentation of calculation of optimal vaccination is shown on Figure 2.

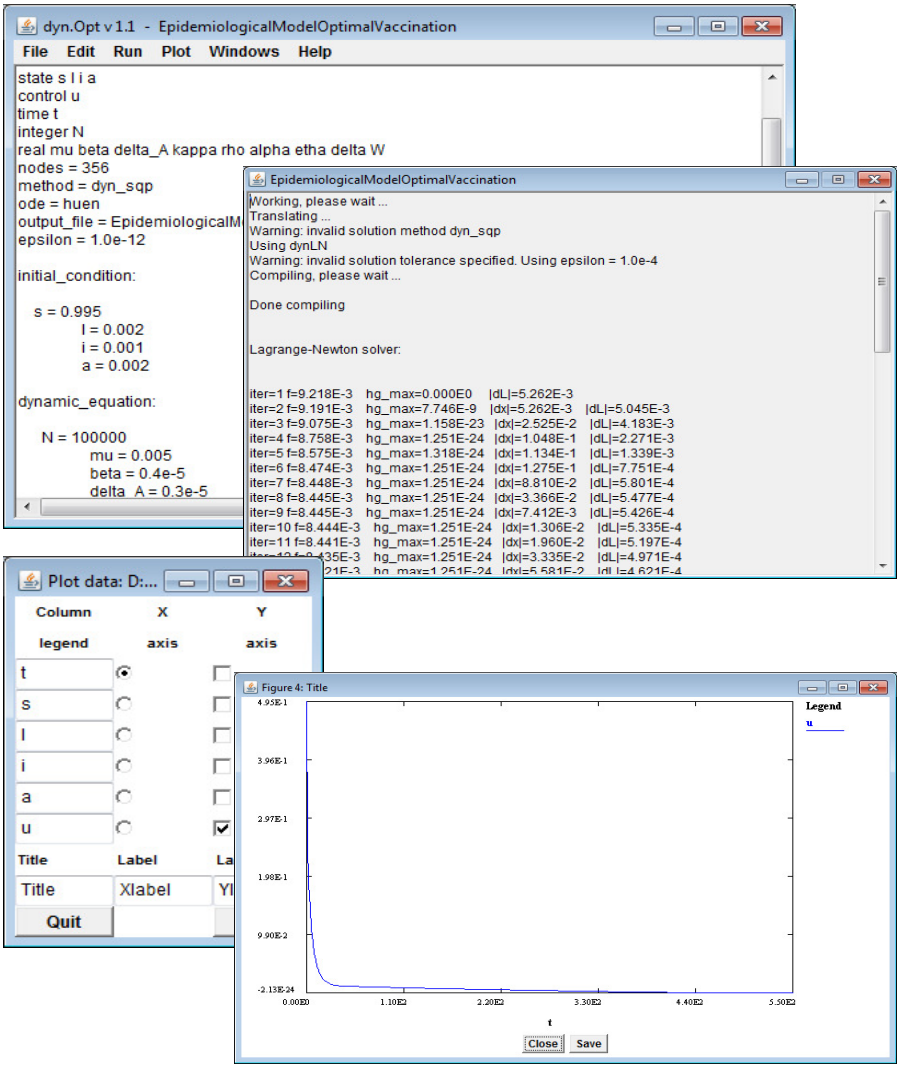


Figure 2.- example of present ation of calculation of optimal vaccination [4]

Results of system analysis of process of support of decision making in epidemiological system researches, analysis of structure of the system and project researches, which are executed with application of modern CASE - tools, allow to carry out development and implementation of software for DSS of problems of mathematical epidemiology [4]. DSS uses the modern algorithms of processing of data and contains a database for maintenance of input, internal and output streams of data and base of knowledge, that gives an opportunity to estimate objectively the state of process of distribution of epidemiological disease and choose optimal decisions in relation to a prophylaxis and treatment [5]. Use of the client-server architecture allows

to provide effective access to data, to implement safety facilities and promote efficiency of processing of data.
As an IDE we have chosen Netbeans. This is due to the features of universality, speed of development and openness. Besides the freely available libraries of AJAX and RIA- components are used in a project - both for work with databases and for a graphic interface.

3. Conclusions

Thus, the software environment of support of epidemiological system researches, which in fact is DSS is presented. A project is implemented on the basis of technologies Java, that does this software Web- integrated one.
A software environment provides determination of thresholdings of base numbers of reproduction taking into account the type of compartmental epidemiological model, and also development of optimal scheme of prophylactics.

REFERENCES

1. BAILEY N. T. J.: The Mathematical Theory of Infectious Diseases, Oxford University Press, Incorporated. – 1987. – P. 430.
2. HETHCOTE H. W.: The Mathematics of Infectious Diseases, SIAM Review. 42(2000)4, 599–653.
3. MARTSENYUK V. P., KUCHVARA O.M. et al.: Information and statistical approach for the modeling of distribution of infectious disease using example of epidemy of acute respiratory diseases during October-November 2009 in Ternopil region, Infectious diseases, (2009)4, 50–59 [in Ukrainian].
4. NAKONECHNYI O.H., MARTSENYUK V.P., KUCHVARA O.M.: Models of population dynamics for problems of mathematical epidemiology of acute respiratory diseases, Cybernetics and Computers (2010)159, 45–64 [in Russian]
5. MARTSENYUK V.P., ANDRUSHCHAK I.Y., KUCHVARA A.M.: On Conditions of Asymptotic Stability in SIR-Models of Mathematical Epidemiology, Journal of Automation and Information Sciences, 43(2011)12, 59-68. DOI: 10.1615/JAutomatInfScien.v43.i12.70

Hanna KUZNETSOVA[1], Ivan KOPYCHENKO[2]

Supervisor: Nadija KAZAKOVA[3]

GŁÓWNE ATAKI NA PROTOKOŁY KRYPTOGRAFII KWANTOWEJ Z CIĄGŁYMI ZMIENNYMI

Streszczenie: W tym artykule zbadano podstawowe aspekty niezawodności służące do oceny bezpieczeństwa metodą rozkładu kluczy kwantowych przy użyciu zmiennych ciągłych opartych na modulacji Gaussa, stanach Gaussa i pomiarach Gaussa. Omówiono analizę indywidualnych, zbiorowych i sekwencyjnych ataków.

Słowa kluczowe: kryptografia kwantowa, zmienne ciągłe, spójne ataki, zbiorowe ataki, ataki indywidualne, dystrybucja kluczy kwantowych, protokół kwantowy

MAIN ATTACKS ON THE QUANTUM CRYPTOGRAPHY PROTOCOL WITH CONTINUOUS VARIABLES

Summary: The main reliability evidence was considered in this paper to evaluate the safety of quantum key distribution using continuous variables based on Gaussian modulation, Gaussian states, and Gaussian measurements. The analysis of individual, collective and coherent attacks has been discussed.

Keywords: quantum cryptography, continuous variables, coherent attacks, collective attacks, individual attacks, quantum key distribution, quantum protocol.

1. Introduction

One of the most important areas of ensuring the information confidentiality in open communication networks was and remains its protection by cryptographic methods. In the last two decades, the new direction of cryptographic information security - quantum cryptography - is rapidly developing.

[1] Graduate student, Odesa national academy of telecommunications n.a. O.S. Popov, the department of information security and data transfer, kuznetsova__anna@hotmail.com

[2] Head of technical information protection department, Municipal Enterprise Regional Information and analytical center, i.kopychenko@odessa.gov.ua

[3] Prof. D.Sc., Odesa State Academy of Technical Regulation and Quality, a head of department of computer, information and measurement technologies, kaz2003@ukr.net

Quantum cryptography is one of the most advanced direction of quantum informatics, both from the theoretical and the practical point of view. And it includes methods for protecting confidential information that have corresponding counterparts in classical cryptography.

The purpose of the quantum key distribution protocol for two partners, traditionally called Alice and Bob, is a secret casual string contract (keys). This secret key should not be known to the eavesdropper (Eve), which is supposed to have access to much more advanced technology than Alice and Bob. If Eve had unlimited resources and was able to do everything but not violate the laws of quantum physics, then they talk about unconditional security.

Protocol BB84 was developed taking into account single states of photons. Coherent states are much easier to produce, i.e. they are used only as approximations of individual photon states. In addition to the fact that coherent states generally contain more than one photon, they also form a family of states that are not orthogonal, in contrast to $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$. Back in 1992, Bennett [14] showed that the use of two nonorthogonal quantum states is a sufficient condition for the quantum distribution of keys. This result opens up the possibility for a quantum distribution of keys with continuous variables. Although compressed and coherent states can very well contain more than one photon, their nonorthogonality is sufficient to implement the quantum distribution of keys. Many of the first proposals for the quantum distribution of keys with continuous variables used mixed fringes of light, showing the EPR correlation [1, 4]. These states are difficult to produce, however, some other protocols use more traditional training and measurement procedures, where Alice prepares the state and sends it to Bob without preserving the intricate subsystems. Cerf and his colleagues proposed in [1] to use continuous Gaussian key elements, hence the continuous modulation of compressed states. This protocol looks like the equivalent of BB84, using Gaussian key elements instead of binary units and zeros. However, the production of compressed states is quite difficult. Using coherent states is much easier to produce. In spite of this, coherent states are separate cases of compressed states.

2. Quantum key distribution protocols with using continuous variables

Quantum key distribution protocols can be divided into two main categories: preparation and measurement (P&M) and (EB) based on the entanglement schemes. The P&M protocol works like this: Alice is preparing quantum systems (usually light pulses) in some states and sends them to Bob in a quantum channel that is supposed to be controlled by Eve. After Bob meets the quantum systems he has received, Alice and Bob share the corrupted classical information from which they release the secret key through classical communication by a public authenticated channel. Of course, Eve was supposedly interacting with her own discretion, with quantum systems on their way from Alice to Bob. She also overheard all messages transmitted on the classic channel.

In the EB protocol, Alice and Bob initially share the entanglement state (which even may be prepared by Eve) and perform measurements on their side. Everything else is identical to the P&M scheme. Since Alice measurements can be considered as "training," these protocols are indeed equivalent to the P&M schemes [1]. Although protocols EB are harder to implement experimentally, they are easier for theoretical

study. Its not only because of the symmetry between Alice and Bob, but also because the "monogamy" of entanglement allows you to study the attack of Eve in general.

The analysis of the main types of attacks on quantum cryptography protocols with continuous variables is the main purpose of this work.

The classic connection between Alice and Bob allows them to sift the secret key from their correlated data. This is usually divided into three stages:

1 - channel estimation: Alice and Bob publish a random sample of their measurements and compare them to evaluate the characteristics of the quantum channel (and derive possible Eve's actions from it);

2 - matching: they use error correction methods to correct bugs during transmission and agree common bits of lines, partly known to Eve;

3 - enhancement of protection: they use technology based on the hash function for extraction, the common line, the secret key unknown to Eve.

When referring to continuous variables, the general description of the quantum key distribution remains valid. But, in addition, quantum key distribution protocols with continuous variables can be understood in a limited or more general form.

3. Main attacks on quantum cryptography protocols with continuous variables

The physical attack of Eve allows her to hold (Gaussian or non-Gaussian) "purifications" of the density matrix ρ_{AB}. This is not enough on its own to give the information about secret keys. Eve must do some measurements to get this information. These measurements can be divided into three categories for power increase, namely:

1) individual attacks;

2) collective attacks;

3) coherent attacks.

In individual attacks, Eve creates one auxiliary impulse that interacts with each pulse individually and performs measurements on it. This measurement can not depend on the classical connection between Alice and Bob (except for the selection of bases). Since the result of this measurement is classical information, then Eve's information measured by the magnitude of the Shannon (classical) mutual information. The type of such attacks is "finite-size attack", where the interaction involves several pulses.

In collective attacks, the interaction with auxiliary impulses remains individual (or at least finite), but they are stored in quantum memory and measured only after Alice and Bob have agreed to execute the key allocation stage. At this moment, a complex collective measurement is performed on quantum memory. Information received by Eve, using this strategy, is calculated using von Neumann's entropy instead of the Shannon entropy, which leads to Holevo's information. This strategy potentially gives Eve more information than individual attacks.

By definition, coherent attacks are the most powerful attacks, "allowed" quantum mechanics: Eve interacts globally with all impulses and then performs global measurement with a delay. This global interaction makes any statistical assumption difficult, since Alice, Bob and Eve are sharing a single multidimensional quantum system. However, collective attacks that are currently optimal among a limited class

of explicit attacks are likely to be fully optimal for all coherent attacks [2], but there is still no clear proof for this in the case of protocols with continuous variables. To distinguish the secret key from their correlated data, Alice and Bob must comply (error correction) and enhance their privacy. Matching, as a step of a protocol with continuous variables, is slightly different from the protocol with discrete states. The total bit string can be removed from the continuous data of Alice and Bob. Then there is a growing confidentiality that allows them to filter out the bits, known to Eve. Everything they need to know to apply privacy enhancements is the upper bound for I_E info Eve. Once I_E becomes known, they can get a secret key of at least a length $(A:B)-I_E$, where $I(A:B)$ is the mutual information between Alice and Bob.

The value of I_E usually depends on the strategy of Eve, but also on the direction of the classical information flow: if the classical one-way communication and directed from Alice to Bob to correct errors, then this means that the Alice data is a secret key, so that I_E is the amount of information which Eve received from Alice's data. This is known as Direct Reconciliation. For obvious reasons for symmetry, such a strategy can not be successful when the physical channel with losses is more than 50%. The symmetry between Bob and Eve must be broken, which can only be done with feedback, that is, with some classical bond that runs from Bob to Alice. This can be done through feedback: this is a Reverse Reconciliation script, where the secret key is determined based on Bob's data. In this case, I_E represents the amount of information accumulated by Eve on the basis of Bob's data.

Thus, in the work the individual, collective and coherent attacks on quantum cryptography protocols with continuous variables are analyzed, the methods of coordination and strengthening of confidentiality for such protocols are considered.

The publication contains the results of research conducted with the grant support of the State Fund for Fundamental Research under the competition project F73 / 49-2017.

REFERENCES

1. CERF N.J.: Security of quantum key distribution using d-level systems, N.J. Cerf, M. Bourennane, A. Karlsson, N. Gisin, Physical Review Letters 88(2002)12, 127902.
2. KILIN S.Ya.: Quantum cryptography: ideas and practice, Kilin S.Ya, Khoroshko D.B., Nizovtsev A.P. – Minsk: "Belarusian science ", 2008. – 392 p.
3. KORCHENKO O.: Modern quantum technologies of information security against cyber – terrorist attacks, O. Korchenko, Ye. Vasiliu, S. Gnatyuk, Aviation: Research Journal of Vilnius Gediminas Technical University, 14(2010)2, 58–69.
4. VASILIU E.V.: Non-coherent attack on the ping-pong protocol with completely entangled pairs of qutrits, Eugene V. Vasiliu, Quantum Information Processing, 10(2011)2, 189–202.

Maciej KOBIAŁKA[1]

Opiekun naukowy: Szymon WĄSOWICZ[2]

SZYFROWANIE ORAZ METODYKA ATAKÓW NA STRONY W SIECI WEB

Streszczenie: Mechanizmy kryptograficzne są powszechnie stosowane w dziedzinie bezpieczeństwa systemów komputerowych. Są bardzo uniwersalnym narzędziem do osiągnięcia poufności, integralności lub autentyczności, są również wykorzystywane w procedurach uwierzytelniania i do ochrony danych. Mechanizmy kryptograficzne niewątpliwie należą do najważniejszych mechanizmów bezpieczeństwa. Artykuł zawiera elementarne koncepcje pola kryptograficznego i prezentuje podstawowe koncepcje algorytmów kryptograficznych. Głównym celem tej pracy naukowej jest wykazanie, jak działają techniki kryptograficzne. Kolejny moduł przedstawia podstawowe zastosowania mechanizmów kryptograficznych w informatyce, ich wady i przedstawia najbardziej popularne metody ataku na strony internetowe.

Słowa kluczowe: Kryptografia,MD5,Szyfry,Szyfrowanie na liczbach pierwszych

ENCRYPTION AND METHODS OF ATTACKS ON WEB SITES

Summary: The cryptography mechanisms are commonly used in the field of computer system security. They are very universal tool to achieve confidentiality, integrity or authenticity, they are also used in authentication procedures and to protect data. The cryptography mechanisms undoubtedly belongs to the most important security mechanisms. The paper feature an elementary concepts of the cryptography field and presents a basic conceptions of cryptographic algorithms. The main purpose of this scientific work is to demonstrate how cryptographic techniques works. Next module will present the basic uses of cryptography mechanisms in computer science, their disadvantages and depict the most popular methods of attacking web pages.

Keywords: Cryptography, MD5, Ciphers, Encryption on the first digits

1. Wiadomości wprowadzające

Szyfrowali wszyscy: starożytni Grecy, Spartanie, Rzymianie. Używając hieroglifów, łącząc kilka alfabetów, pisząc na kawałkach skór, czy budując do tego celu specjalne

[1] Student, Akademia Techniczno-Humanistyczna w Bielsku-Białej

[2] dr hab. prof. ATH, Akademia Techniczno-Humanistyczna w Bielsku-Białej, Katedra Matematyki, swasowicz@ath.bielsko.pl

maszyny. Dziś tę pracę wykonują zaawansowane komputery, ale kiedyś ludzie musieli sobie radzić inaczej.

Szyfrowanie pomaga chronić dane przechowywane na komputerze lub te, które przesyłane są przez Internet lub inne sieci komputerowe. Nowoczesne algorytmy odgrywają bardzo ważną rolę w bezpieczeństwie systemów IT i komunikacji. Tajne informacje są niebywale cenne, dlatego każdy chce się zabezpieczyć przed ich utratą. Wiele jest firm, które na systemy bezpieczeństwa przeznaczają duże sumy pieniędzy. Jak jednak stosowano szyfrowanie, kiedy nie było komputerów?

2. Szyfr Cezara

Jednym z pierwszych szyfrów, a również prawdopodobnie najbardziej znanym, jest szyfr Cezara. Opiera się on na prostej zasadzie przesunięcia liter w wyrazach o określoną liczbę 'pól', jednak jest bardzo łatwy do złamania przez osoby, które znają sposób jego działania. Został on opisany 150 lat po Cezarze przez Swetoniusza.

Sposób działania
Szyfr ten polega na zwykłym przesunięciu alfabetu. Jeżeli zaszyfrowana ma zostać litera A, a alfabet przesunięty został o 3 litery, to w jej miejsce otrzymujemy literę D (tak jak poniżej). Szyfr ten obejmuję zatem 25 potencjalnych ustawień w zależności od przesunięcia jakie zastosujemy. Nie jest on jednak zbyt bezpieczny - jeżeli podejrzewane jest jego użycie, wystarczy sprawdzić wszystkie 25 ustawień.

Wpisany tekst: Akademia Techniczno-Humanistyczna w Bielsku-Białej
Zaszyfrowany tekst: Nxnqrzvn Grpuavpmab-Uhznavfglpman j Ovryfxh-Ovnłrw

Obecnie najczęściej używaną wersją Szyfru Cezara jest tzw. ROT13, który został wykorzystany do szyfrowania tekstu w linijce powyżej. Polega on na przesunięciu znaków o 13, co w przypadku standardowego alfabetu łacińskiego pozwala na odszyfrowywanie go przy użyciu dokładnie tej samej metody – ROT13(ROT13(x))=x. Szyfr ten nie koduje znaków specjalnych, cyfr ani znaków spoza alfabetu łacińskiego (m. in. znaków diakrytycznych - ą, ę, ź ...)

3. Szyfrowanie za pomocą liczb pierwszych

Liczby pierwsze są absolutnie zadziwiające mimo swej prostoty. Rzadko sobie uświadamiamy, że mamy z nimi do czynienia wszędzie, a szczególnie jeśli dotyczy do ochrony danych - kart bankowych, systemów bezpieczeństwa komputerów czy ochrony prywatności rozmów mailowych czy telefonicznych. Jeden z nowoczesnych systemów kryptograficznych KRYPTOSYSTEM RSA opiera się właśnie na operacjach z wykorzystaniem liczb pierwszych. Można powiedzieć, że dzięki liczbom pierwszym możemy czuć się bezpieczni.

Liczby pierwsze

Liczby pierwsze to takie liczby naturalne, które mają dokładnie dwa dzielniki, np.: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, , 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, … i tak dalej w nieskończoność.

Jak już wyżej wspomniałem liczb pierwszych jest nieskończenie wiele. Stwierdził to już grecki matematyk Euklides około trzysta lat przed naszą erą. Dowód tego faktu przedstawić można następująco:Przyjmijmy, że istnieje największa liczba pierwsza (nazwijmy ją **G**). Następnie utwórzmy iloczyn wszystkich liczb pierwszych, które są od niej mniejsze. Otrzymany wynik pomnóżmy przez **G** i do otrzymanego iloczynu dodajmy **1**. Rezultat nazwijmy**Y**. Liczba ta z pewnością jest większa od **G** (ponieważ **G** zostało pomnożone przez liczby całkowite oraz dodaliśmy **1**). Nie jest również podzielna przez żadną z liczb pierwszych mniejszych od **G** i przez samo **G** (dlatego, że wszystkie mniejsze od niej liczby pierwsze łącznie z **G**, dają przy dzieleniu przez **Y** resztę **1**). Wynika stąd, że liczba **Y** jest liczbą pierwszą lub musi być podzielna przez jakąś liczbę pierwszą większą od **G**, czyli w obu przypadkach istnieje liczba pierwsza większa od **G**. Widać tu, że jeśli weźmiemy jakakolwiek liczbę pierwszą, to zawsze możemy wskazać od niej większą, co pokazuje, że liczb tych jest nieskończenie i wiele.

Mam nadzieję, że ten krótki dowód nie odstraszył Cię od czytania dalszej części (w końcu niewielu normalnych ludzi lubi dowody matematyczne).Opiszę teraz, w jaki sposób można otrzymywać liczby pierwsze. Metodę wyznaczania liczb pierwszych odkrył około 250 lat p.n.e. Eratostenes z Cyreny, dlatego metodę tą nazywamy **sitem Eratostenesa**. Wypisuje się kolejno liczby naturalne od 2 do n (w naszym przykładzie do 100). Liczba 2 jest pierwsza, wiec ją zostawiamy wykreślając jednocześnie wszystkie jej wielokrotności: 4, 6, 8, 10, 12, 14, 16 itd. Kolejna liczba to 3. Ją również zostawiamy i wkreślamy te jej wielokrotności, które jeszcze zostały: 9, 15, 21, 27, 33 itd. Liczbę 4 pomijamy (bo została skreślona), przechodzimy do liczby 5 i znowu kreślimy. Czynności te powtarzamy, aż pozbędziemy się wszystkich liczb złożonych (tych, które są wielokrotnościami innych). To co zostało po naszym skreślaniu to liczby pierwsze.

2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19
~~20~~	~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~
29	~~30~~	31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37
~~38~~	~~39~~	~~40~~	41	~~42~~	43	~~44~~	~~45~~	~~46~~
47	~~48~~	~~49~~	~~50~~	51	~~52~~	53	~~54~~	~~55~~
~~56~~	57	~~58~~	59	~~60~~	61	~~62~~	~~63~~	~~64~~
~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~	71	~~72~~	73
~~74~~	~~75~~	~~76~~	77	~~78~~	79	~~80~~	~~81~~	~~82~~
83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~	91
~~92~~	93	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~

Początkowo matematycy uważali liczby pierwsze za obiekty ciekawe, lecz zupełnie bezużyteczne. Okazuje się jednak, że mają one ogromne znaczenie w kryptologii.

Szyfrowanie

Szyfrowanie opisane w poprzednim rozdziale ma pewna wadę. Jest nią konieczność przekazania klucza, co niesie za sobą ryzyko jego przechwycenia. Jednak pod koniec

lat 70-tych dwudziestego wieku opracowano metodę szyfrowania, która ten problem eliminuje. Od początkowych liter nazwisk jej twórców (Rivest, Shamir, Adelman) nazwano ją **RSA**. Polega ona na tym, że wybiera się trzy liczby (tzw. **magiczne**) N, E i D. Dwie z nich N i E służą do zaszyfrowania informacji, natomiast rozszyfrowanie możliwe jest **tylko** przy pomocy N i D. Osoba, która będzie odbierać zaszyfrowane informacje może liczby N i E rozgłosić całemu światu, natomiast liczbę D zamknąć w sejfie i wyciągać ją z niego, tylko gdy zajdzie potrzeba. Aby jednak zrozumieć tą metodę musisz poznać **dzielenie modulo** oraz sposób tworzenia **liczb magicznych**.
Dzielenie modulo to takie, w którym wynikiem jest reszta z dzielenia np.:
5 mod 2 = **1**, bo 5 = 2·2 + **1** (w piątce mieszczą się dwie pełne dwójki i reszta wynosi 1);
30 mod 4 = **2**, bo 30 =7·4 + **2** (w trzydziestce mieści się siedem czwórek, a reszta wynosi 2);
78 mod 5 = **3**, bo 78 = 15·5 + **3**;
13 mod 7 = **6**, bo 13 = 1·7 + **6;**
39 mod 3 = **0,** bo 39 = 13·3 + **0.**
Dzielenie modulo ma pewną cechę, której używa się do uproszczenia obliczeń np.: 78mod5= **3,** (78·78) mod 5 = 605284 mod 5 = 4 oraz (3·3) mod 5 = 9 mod 5 = 4. Wynika stąd, że (78·78) mod 5 = (3·3) mod 5.
Do obliczeń reszty można użyć również windowsowego kalkulatora naukowego, który posiada funkcję **MOD.**
Teraz postaraj się skupić, gdyż nastąpi opis tworzenia trójki **liczb magicznych**.

1. Bierzemy dwie liczby pierwsze np. p=11 i q=17 (w praktyce używa się liczb pierwszych składających się z więcej niż stu cyfr).
2. Mnożymy przez siebie liczby p i q, otrzymując magiczną liczbę **N**= 11·17=187 (duży klucz).
3. Obliczamy pomocniczą liczbę z, która pomoże nam określić małe klucze **E** i **D**. Otrzymujemy ją zmniejszając p i q o 1, a następnie mnożąc otrzymane wyniki. z =(p-1)·(q-1) = 10·16 = 160.
4. Wybieramy liczbę **E**. Nie może ona posiadać z liczbą z żadnego wspólnego dzielnika. Najwygodniej jest wybrać jakąś liczbę pierwszą mniejszą od z i sprawdzić przez dzielenie, czy jest ona dzielnikiem z. Jeśli tak, szukamy dalej, jeśli nie, liczba ta może być wykorzystana jako **E**. W naszym przykładzie wybrałem **E**=37
5. Wyznaczenie liczby **D** trochę trudniejsze. Jeśli nie zrozumiesz tego za pierwszym razem, nie przejmuj się. Spróbuj przeczytać opis kilka razy, najlepiej robiąc obliczenia na kartce. Liczba **D** powinna po pomnożeniu przez **E**, a następnie po podzieleniu przez z dać resztę 1. (**D**·**E**) mod z = 1 inaczej **D**·**E** = x·z +1, gdzie x pewna liczba naturalna.
   - Bierzemy **E** i z, następnie dzielimy większą z nich przez mniejszą i określamy resztę. z mod **E** = 160 mod 37 = **12**, bo 160 = 4·37 + **12** (z = 4·**E** +**12** inaczej **12**= z - 4·**E**).
   - Liczbę **E** dzielę modulo przez otrzymaną resztę. **E** mod 12 = 37 mod 12 = **1**, bo 37 = 3·12 +**1** (**E** =3 ·12 + **1** inaczej **1** = **E** - 3·12);
   - Gdy nie otrzymamy reszty 1, to przedostatnią resztę dzielę przez ostatnią , itd., aż otrzymam resztę **1**.

- W ostatnim równaniu za 12 podstawiam z - 4·**E**, zatem 1 = **E** -3·(z - 4·**E**) = **E** - 3·z +12·**E** = 13**E** - 3·z, otrzymaliśmy zatem 1 = 13·**E** - 3·z, po przekształceniu mamy: 13**E** =3·z +1.
- Porównajmy teraz zapisy 13·**E** =3·z +1. i **D**·**E** = x·z +1. Widać tu, że **D** = 13.

6. Nasze liczby magiczne to: **N** = 187, **E** = 37 i **D** = 13.

Po przebrnięciu przez proces tworzenia liczb magicznych, można przystąpić do szyfrowania. Aby zrozumieć ten sposób szyfrowania, zaczniemy od szyfrowania pojedynczej litery np. *m*. Literę zamieniamy na liczbę, zamianę przeprowadzić można wiele sposobów, my podstawimy numer jaki *m* ma w alfabecie.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Literze *m* odpowiada liczba 13.Liczbę należy podnieść do potęgi **E**, czyli do 37, a wynik podzielić modulo przez **N,** czyli u nas 187 i otrzymamy liczbę zaszyfrowaną. Jednak 13^{37} to liczba składającą się 41 cyfr. Na szczęście, nie jest nam ona potrzebna w całości. Gdy liczby są tak duże robimy to tak:

- mnożymy 13· 13 = 169;
- szukamy reszty z dzielenia przez **N**=187 i otrzymujemy 169 mod 187 = **169,** bo 169 = 0·187+ **169**;
- $13^4=13^{2}\cdot13^2$, stąd 13^4 mod 187 = ($13^{2}\cdot13^2$) mod 187 = (169·169) mod 187 = 28561 mod 187 = **137**;
  liczbę 137 wyznaczyłem następująco:
  - 28561 : 187=152,732.. (152 pełne liczby 187);
  - 152·187=28424;
  - 28561-28424= **137** (czyli 28561 = 157·187+**137**).
- $13^8=13^{4}\cdot13^4$, stąd 13^8 mod 187 = ($13^{4}\cdot13^4$) mod 187 = (137·137) mod 187 = 18769 mod 187 = **69**;
- $13^{16}=13^{8}\cdot13^8$, stąd 13^{16} mod 187 = ($13^{8}\cdot13^8$) mod 187 = (69·69) mod 187 = 4761 mod 187 = **86**;
- $13^{32}=13^{16}\cdot13^{16}$, stąd 13^{32} mod 187 = ($13^{16}\cdot13^{16}$) mod 187 = (86·86) mod 187 = 7396 mod 187 = **103**;
- $13^{36}=13^{32}\cdot13^4$, stąd 13^{36} mod 187 = ($13^{32}\cdot13^4$) mod 187 = (103·137) mod 187 = 14111 mod 187 = **86**;
- 13 mod 187 = **13**;
- $13^{37}=13^{36}\cdot13^1$, stąd13^{37} mod 187 = ($13^{36}\cdot13^1$) mod 187 = (86·13) mod 187 = 1118 mod 187 = **183**;
- **liczba tekstu tajnego wynosi 183**.

Rozszyfrowywanie przebiega analogicznie - liczbę tekstu tajnego podnosimy do potęgi **D** i wyznaczamy resztę z dzielenia przez **N**. Ponieważ znowu otrzymujemy duże liczby (183^{13}= 2581452668046920778582615126 63), warto dla uproszczenia zrobić to w sposób opisany powyżej:

- 183 mod 187 =183 ;
- 183^2 mod 187=183^2 mod 187 = 33489 mod 187 = 16;

- $183^4 \bmod 187 = 16^2 \bmod 187 = 256 \bmod 187 = 69$;
- $183^8 \bmod 187 = 69^2 \bmod 187 = 4761 \bmod 187 = 86$;
- $183^{12} \bmod 187 = (183^8 \cdot 183^4) \bmod 187 = (86 \cdot 69) \bmod 187 = 5934 \bmod 187 = 137$;
- $183^{13} \bmod 187 = (183^{12} \cdot 183^1) \bmod 187 = (137 \cdot 183) \bmod 187 = 25071 \bmod 187 = \mathbf{13}$;
- otrzymaliśmy liczbę tekstu jawnego.

Szyfrowanie po jednej literce nie jest jednak przy dłuższych tekstach zalecane, gdyż taki szyfr jest łatwy do złamania. Spróbujemy teraz zaszyfrować tekst "o*to ja*". Najpierw zamieniamy tekst na liczby, a w odstępy między wyrazami wstawiamy dwa zera. Dokonujemy zamiany

t o j a

20 15 00 10 01

Nasz tekst wygląda teraz tak: **2015001001**

Zapisujemy go w blokach np. czterocyfrowych: **2015 0010 0100** (aby ostatni blok miał cztery cyfry, na końcu dopisałem dwa zera). Będziemy szyfrować każdy blok oddzielnie, potrzebujemy jednak innych liczb magicznych - tak by **N** było pięciocyfrowe.

- Weźmy p=113 i q=101;
- **N**=113·101=11413;
- z=(p-1)(q-1)=112·100=11200;
- niech np. **E**= 43;
- szukam **D**:
  - Bierzemy **E** i z, następnie dzielimy większą z nich przez mniejszą i określamy resztę.
    z mod **E** = 11200 mod 43= **20**, bo 11200 = 260·43 + **20** (z = 260·**E** +**20** inaczej **20**= z - 260·**E**).
  - Liczbę **E** dzielę modulo przez otrzymaną resztę.
    E mod 20 = 43 mod 20 = **3**, bo 43 = 2·20 +**3** (**E** =2 ·20 + **3** inaczej **3** = **E** - 2·20);
  - Przedostatnią resztę dzielę przez ostatnią , itd., aż otrzymam resztę **1**.
    20 mod 3 = 2, bo 20 = 3·18 +**2** (20 =6 ·3 + **2** inaczej **2** = 20 -6·3);
    3 mod 2 = 1, bo 3 = 1·2 +**1** (3 =1 ·2 + **1** inaczej **1** = 3 -2);
  - W ostatnim równaniu za 2 podstawiam 20 -6·3, zatem
    1 = 3 **-**(20 - 6·3) = 3 **-**20 **+**6·3 = -20+7·3;
  - za 3 podstawiam **E** - 2·20, zatem
    1 = -20 **+**7·(**E** - 2·20) = **-**20 **+**7·**E -** 14·20 = 7·**E** -15·20;
  - za 20 podstawiam z - 260·**E**, zatem
    1 = 7·**E** - 15·(z - 260·**E**) = 7·**E** -13·z **+**3900·**E** = 3907·**E** - 13·z;
    otrzymaliśmy zatem 1 = 3919·**E - 1**3·z, po przekształceniu mamy: 3907·**E =1**3·z +1.
  - zatem **D** = 3907;

Nasze liczby magiczne to: **E=43, D=3907, N= 11413**

Teraz szyfrujemy bloki. Liczbę z każdego bloku należy podnieść do potęgi **E**, czyli do **43**, a wynik podzielić modulo przez **N,** czyli u nas **11413** i otrzymamy liczbę

tekstu tajnego. Do obliczeń reszt użyj metody opisanej wyżej lub kalkulatora naukowego z Windows.

- pierwszy blok - 2015
  $2015^{43} \bmod 11413 = 336$;
- drugi blok - 0010
  $10^{43} \bmod 11413 = 3323$;
- trzeci blok - 0100
  $100^{43} \bmod 11413 = 5958$.

Nasz tekst po zaszyfrowaniu ma postać: **0336 3323 5958** (jeśli w bloku jest mniej cyfr, dopisujemy na początku zera).
Aby tekst rozszyfrować, liczbę z każdego bloku tekstu tajnego podnosimy do potęgi **D** i szukamy reszty z dzielenia przez **N**.

- pierwszy blok - 0336
  $336^{3907} \bmod 11413 = 2015$;
- drugi blok - 3323
  $3323^{3907} \bmod 11413 = 10$;
- trzeci blok - 5958
  $5958^{3907} \bmod 11413 = 100$.

Po rozszyfrowaniu znów mamy: **2015 0010 0100.**
Przykłady, jakie tu podałem były w miarę proste, gdyż miały na celu pokazanie działania metody. Dlatego też szyfr, którego używaliśmy jest łatwy do złamania. Liczby **N** i **E** mogą być podawane do publicznej wiadomości (tajna jest tylko **D**) i jeśli ktoś rozpisze **N** jako iloczyn liczb pierwszych (czyli znajdzie p i q), to znając **E** łatwo wyznaczy **D**. Liczby **N**, których używałem w przykładach (187 i 11413) można dość szybko rozłożyć na czynniki. Aby uniemożliwić złamanie szyfru, do tworzenia liczb magicznych używa się liczb pierwszych p i q, które po pomnożeniu dają wynik ponad stucyfrowy (im większy tym lepiej). Rozłożenie takiej liczby na czynniki pierwsze, to nawet dla nowoczesnych komputerów praca na wiele miesięcy, a nawet lat. Ocenia się, że rozłożenie np. liczby pięćsetcyfrowej wymaga czterdziestocyfrowej liczby operacji.
Metoda **RSA,** znalazła również zastosowanie w bankach, do szyfrowania numeru PIN.

3. Algorytm RSA

Geneza
Opracowany w 1977 roku przez Ronald'a L. Rivest'a, Adi'a Shamir'a i Leonard'a Adleman'a niesymetryczny system szyfrowania danych, nazwany od pierwszych liter ich nazwisk. Jego zasadniczą cechą jest użycie dwóch kluczy - publicznego (do kodowania informacji) i prywatnego (do jej odczytywania). Algorytm RSA może być używany w środowisku narażonym na wiele nadużyć. Trudność jego złamania polega na złożoności rozkładu dużych liczb na czynniki pierwsze.
Co to jest klucz publiczny/prywatny ?
- Klucz publiczny
Umożliwia szyfrowanie danych, ale w żaden sposób nie ułatwia ich odczytywania, dzięki czemu można go udostępniać publicznie.

- Klucz prywatny
Służy do odczytywania zakodowanych kluczem publicznym danych. Klucz ten nie może być udostępniany publicznie.

Generowanie kluczy i szyfrowanie:
Aby wygenerować parę kluczy (publiczny i prywatny) należy wykonać kilka operacji:
1) Wybranie dwóch liczb pierwszych p i q, tak aby miały one jak najbliższą sobie długość w bitach przy możliwie najbardziej różniącej je wartości.
2) Wyznaczenie wartości n=p*q.
3) Wyznaczenie wartości funkcji Eulera dla n: φ(n)=(p-1)*(q-1).
4) Wybranie takiej liczby e z przedziału 1 < e < φ(n), aby była ona względnie pierwsza z φ(n).
5) Znalezienie liczby d, której różnica z odwrotnością liczby e jest podzielna przez φ(n): d ≡ e^(-1) mod(φ(n)) → d·e ≡ 1 (mod φ(n))
Klucz publiczny definiowany jest jako para liczb (n, e), a klucz prywatny jako (n, d).
Aby zaszyfrować wiadomość należy zamienić ją na takie liczby naturalne t, że 0 < t < n, a następnie każdą z nich zakodować w taki sposób, że: c ≡ t^e mod(n).
Zaszyfrowana wiadomość składać się będzie z kolejnych bloków c. Taki szyfrogram możemy przekształcić na tekst jawny za pomocą danej funkcji: t ≡ c^d mod(n).

Trudność w rozszyfrowaniu:
Do złamania szyfru RSA potrzebne jest rozbicie klucza publicznego na dwie liczby pierwsze będące jego dzielnikami (ich znajomość pozwala na rozszyfrowanie danych przy użyciu klucza publicznego i prywatnego). Nie istnieje jednak wzór pozwalający na wyznaczenie dzielników liczb, a więc ich szukanie polega na testowaniu podzielności kolejnych liczb.

Z rozważań na temat liczb pierwszych wynika, że jeden z dwóch dzielników pierwszych musi znajdować się poniżej wartości √n, a drugi powyżej niej, zatem aby znaleźć jeden z nich musimy wyliczyć pierwiastek z rozkładanej liczby i testować liczby nieparzyste mniejsze od niego. Statystycznie pierwiastek ten powinien znajdować się w górnej połówce przedziału <2,√n>.
A więc dla 128-bitowego klucza pierwiastek jest liczbą 64-bitową (liczb tych jest 2^64). Możemy odrzucić wszystkie liczby parzyste, a więc pozostaje nam 2^64/2=2^63 możliwości. Poszukiwania możemy ograniczyć również do górnej połówki, więc znów zmniejszamy ilość możliwości do 2^63/2=2^62.
Zakładając, że posługujemy się komputerem o mocy obliczeniowej umożliwiającej sprawdzanie 1 000 000 000 liczb w ciągu 1 sekundy, szukanie tej właściwej mogłoby w najgorszym wypadu potrwać nawet 146 lat (2^62 / 10^9=4611686018 sekund=76861433 minut=1281023 godzin=53375 dni=146 lat).
Czas ten można wydłużyć do niewyobrażalnych wartości korzystając np. z klucza o długości 1024 bitów zamiast 128 (w tym wypadku czas ten mógłby wynieść nawet 1.06*10^146 lat!)

Przykład liczbowy:
1) Wybieramy p=11 i q=19.
2) Wyznaczamy wartość n=11*19=209.
3) Wyznaczamy wartość funkcji Eulera φ(n)=(11-1)*(19-1)=180.

4) Wybieramy e=13, względnie pierwsze z φ(n)=180.
5) Szukamy liczby d, korzystając ze wzoru d*13 mod180=1. d=277

Klucz publiczny to (209, 13), a klucz prywatny to (209, 277).

Chcąc zaszyfrować wiadomość przekonwertowaną na liczby 127 101 72 korzystamy ze wzoru i otrzymujemy c1 ≡ 127^13 mod(209) ≡ 205, c2 ≡ 101^13 mod(209) ≡ 118, c3 ≡ 72^13 mod(209) ≡ 29.
Otrzymany szyfrogram: 205 118 29.

Aby odszyfrować otrzymany ciąg liczb należy zastosować drugi ze wzorów, wtedy: t1 ≡ 205^277 mod(209) ≡ 127, t2 ≡ 118^227 mod(209) ≡ 101, t3 ≡ 29^277 mod(209) ≡ 72.
Otrzymana wiadomość: 127 101 72.

4. Szyfrowanie metoda MD5

Geneza powstania
MD5 jest jednym z serii algorytmów zaprojektowanych przez profesora Rona Rivesta z MIT (Rivest, 1994). Poprzednikiem był algorytm MD4, który w wyniku analizy przeprowadzonej przez Hansa Dobbertina okazał się zbyt mało bezpieczny. Jego bezpiecznym następcą był MD5 opracowany w 1991.W 1996 Dobbertin zaprezentował analizę kolizji algorytmu MD5. Chociaż nie był to jeszcze pełny atak na funkcję skrótu to jednak wystarczył, aby specjaliści w dziedzinie kryptografii zaczęli stosować silniejsze odpowiedniki, takie jak SHA-1 lub RIPEMD-160. W marcu 2004 powstał rozproszony projekt nazywany MD5CRK (ang.) . Twórcą projektu był Jean-Luc Cooke i jego współpracownicy. Miał on na celu wykazanie, że możliwe jest wyznaczenie wiadomości różnej od zadanej, która ma taką samą wartość skrótu. Do tego celu wykorzystano sieć Internet oraz dużą liczbę komputerów biorących udział w projekcie. Projekt wykazał, że dysponując bardzo dużą mocą obliczeniową możliwe jest podrabianie generowanych podpisów.Dopiero prace badawcze chińskich naukowców Xiaoyun Wang, Dengguo Fen, Xuejia Lai i Hongbo Yu w pełni wykazały słabość algorytmu. 17 sierpnia 2004 opublikowali oni analityczny algorytm ataku, dzięki któremu do podrobienia podpisu wystarczyła godzina działania klastrowego komputera IBM P690.W marcu 2005 Arjen Lenstra, Xiaoyun Wang i Benne de Weger zaprezentowali metodę umożliwiającą znalezienie kolizji dla algorytmu MD5 i przeprowadzenie ataku polegającego na wysłaniu dwóch różnych wiadomości chronionych tym samym podpisem cyfrowym. Kilka dni później Vlastimil Klima opublikował algorytm, który potrafił znaleźć kolizję w ciągu minuty, używając metody nazwanej tunneling.Pod koniec 2008 roku niezależni kalifornijscy specjaliści ds. bezpieczeństwa, we współpracy z ekspertami z Centrum voor Wiskunde en Informatica, Technische Universiteit Eindhoven oraz Ecole Polytechnique Fédérale de Lausanne odkryli lukę w MD5 umożliwiającą podrobienie dowolnego certyfikatu SSL w taki sposób, że zostanie on zaakceptowany przez wszystkie popularne przeglądarki internetowe. Do podrobienia certyfikatu wystarczyła moc obliczeniowa 200 konsol do gier PlayStation 3.

Od lat 90. MD5 nie jest uważany za bezpieczny do większości zastosowań i w jego miejsce zaleca się stosowanie algorytmów z rodziny SHA-2 lub SHA-3.

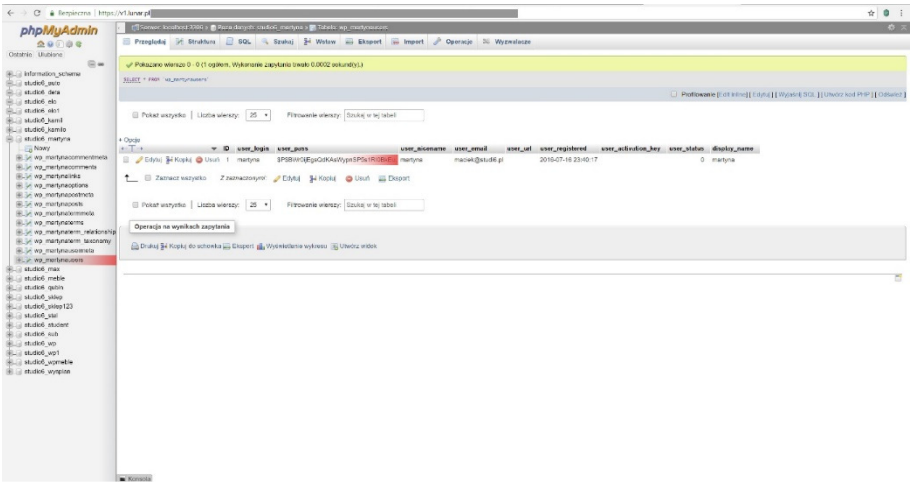
Sposób działania:
Algorytm działania MD5 można zapisać w 6 punktach:
1. Doklejenie do wiadomości wejściowej bitu o wartości 1.
2. Doklejenie do wiadomości tylu zer, aby składała się ona z bloków o długości 512 bitów i ostatniego, niepełnego 448-bitowego.
3. Doklejenie do wiadomości 64-bitowego licznika oznaczającego rozmiar wiadomości (zaczynając od najmniej znaczącego bitu) - otrzymujemy wiadomość składającą się z pełnych, 512 bitowych fragmentów.
4. Ustawienie stanu początkowego '0123456789abcdeffedcba9876543210'.
5. Uruchomienie na każdym bloku funkcji zmieniającej stan.
6. Po przetworzeniu ostatniego bloku uzyskanie stanu jako obliczonego skrótu wiadomości.
Funkcja zmiany stanu składa się z 4 cykli - 64 kroków. Stan jest rozumiany jako 4 liczby 32-bitowe. W każdym kroku do jednej z tych liczb dodawany jest jeden z 16 32-bitowych fragmentów bloku wejściowego, pewna stała zależna od numeru kroku oraz pewna prosta funkcja boolowska 3 pozostałych liczb. Następnie liczba ta jest obracana (przesuwana cyklicznie z najstarszymi bitami wsuwanymi w najmłodsze pozycje) o liczbę bitów zależną od kroku, oraz jest dodawana do niej jedna z pozostałych liczb.

Zastosowanie kodowania md5 w naszym życiu:
Wszystkie hasła których używamy w portalach internetowych np.: faceboook,poczta oraz różnego rodzaju sklepach internetowych ,są zapisywane w postaci hasła zaszyfrowanego metodą md5 ponieważ hasło np.: Admin1 które wyjadę się bardzo proste do złamanie metodą bute force, staje się o wiele bardziej skomplikowane jeśli zapiszemy to hasło w bazie za pomocą md5 ,wtedy hasło Admin1 ma postać 2e33a9b0b06aa0a01ede70995674ee23 , obliczając czas na złamanie tego o to hasła metodą brute force,

Hasła:	**Metoda**	**Czas do złamania**
Admin1	Brf	200 milisekund
2e33a9b0b06aa0a01ede70995674ee23	Brf	$10x10^{54}$ lat

*Brf - Atak brute force – technika łamania haseł lub kluczy kryptograficznych polegająca na sprawdzeniu wszystkich możliwych kombinacji

Dlaczego nasze hasła są haszowane MD5 ??
Zapewne każdy zastanawiał się kiedyś czy administrator strony np. Facebooka ma dostęp do naszych haseł? Każdy szanujący się portal albo sklep internetowy ma dostęp do naszego hasła ponieważ jest ono zapisane w bazie SQL .W rzeczywistości nikt nie może znać twojego hasła ,ponieważ nie było by to zgodne z polityką prywatności. Dlatego hasło użytkownika jest widoczne tylko i wyłącznie w formie zahaszowanej w postaci szyfru MD5 (rys.1).



Rysunek 1. Ochrona hasła w przeglądarce internetowej

Administrator widzi login oraz wszystkie informacje lecz hasło jest zahaszowane szyfrem MD5.

Zagrożenia związane z MD5

Podczas 25 konferencji Chaos Communication Congress (25C3) w berlińskim Centrum Kongresowym, został przedstawiony problem związany z podatnością w funkcji haszującej MD5 na kolizje do stworzenia fałszywych certyfikatów SSL. Specjaliści ds. bezpieczeństwa wykorzystując lukę, stworzyli w strukturze PKI - używanego do wystawiania certyfikatów dla bezpiecznych stron internetowych, podrobiony certyfikat CA zaufany we wszystkich przeglądarkach internetowych. Atak wykorzystuję tzw. "kolizję" w funkcji haszującej MD5, czyli możliwość posiadania przez dwa różne pliki tej samej sumy kontrolnej. Podatność ta odkryta została już w 2004 roku przez naukowców z Chin.
Pozwala to na podszycie się pod jakąkolwiek stronę protokołu HTTPS. Niczego nie świadomy
użytkownik będzie w posiadaniu fałszywych certyfikatów i będzie mógł z łatwością być przekierowywany do fałszywych stron HTTPS, np. stron bankowych legitymujących się jako autentyczne. Zalecane jest zaprzestanie wykorzystywania MD5 i przejście na algorytm SHA-1 lub
SHA-2.

Co to jest kolizja i na czym ona polega?
Kolizja funkcji skrótu H to taka para różnych wiadomości m1, m2, że mają one taką samą wartość skrótu, tj. H(m1) = H(m2). Ponieważ funkcja skrótu zwraca skończenie wiele wartości, a przestrzeń argumentów jest nieskończona (w przypadku funkcji akceptujących dowolnie długie argumenty), lub przynajmniej znacznie większa od przestrzeni wyników, dla każdej funkcji skrótu istnieją kolizje.

Jeśli funkcja skrótu zwraca k bitów, to zgodnie z paradoksem dnia urodzin sprawdzenie wśród zbioru losowo wybranych wiadomości rozmiaru rzędu $2^{k/2}$ prawdopodobnie istnieje jakaś kolizja. Jest to zasada działania ataku urodzinowego*.

Najprostszy sposób, czyli pamiętanie wszystkich dotychczas sprawdzonych skrótów, wymaga bardzo dużo pamięci, istnieją jednak algorytmy "bez pamięciowy" o szybkości gorszej tylko o czynnik stały.

Tak więc znalezienie kolizji 128-bitowej funkcji skrótu (takiej jak MD5) jest zadaniem o trudności porównywalnej ze znalezieniem klucza 64-bitowego szyfru symetrycznego. Nie jest to zadanie trywialne, aczkolwiek znajduje się w zasięgu możliwości współczesnego sprzętu i sieci rozproszonych. Znajdowanie kolizji funkcji 160-bitowych (SHA1, RIPEMD-160) jest równie trudne jak łamanie 80-bitowego szyfru symetrycznego, i jest obecnie uważane za zbyt trudne.

Liczby te dotyczą tylko sytuacji, w której funkcją skrótu posługujemy się jako "czarną skrzynką", tzn. nie korzystamy z wiedzy o jej strukturze. Wykorzystując słabości struktury możemy często znajdować kolizje o wiele szybciej (np. dla MD4 kolizje można znaleźć w czasie rzeczywistym).

Znajdowanie przeciw obrazu czy drugiego przeciw obrazu jest równoważne atakowi na wszystkie bity funkcji skrótu, dlatego też 128-bitowe MD5 jest uważane za bezpieczne, jeśli zależy nam tylko na tych właściwościach, choć nie chroni przed kolizjami.

* Paradoks dnia urodzin ma znaczenie w kryptografii i jest podstawą działania tzw. ataku urodzinowego. Załóżmy, że badamy funkcję haszującą H, która zwraca kod o M bitach, czyli daje 2^M możliwych odpowiedzi. Szukamy kolizji, czyli dwóch takich wiadomości W_1 i W_2, że $H_{(W1)} = H_{(W2)}$. Każdy kwantyl rozkładu liczby prób n potrzebnych do znalezienia kolizji wśród $K=2^M$ kodów, spełnia zależność (5), gdzie $1 - p$ to rząd kwantyla. Średni czas łamania funkcji haszującej rośnie więc w przybliżeniu proporcjonalnie do pierwiastka liczby wszystkich możliwych odpowiedzi tej funkcji.

LITERATURA

1. Serwis internetowy: http://www.cert.gov.pl/cer/wiadomosci/zagrozenia-i-podatnosc/104,Niebezpieczna-kolizja-w-funkcji-skrotu-MD5.html
2. Serwis internetowy: http://matematykainnegowymiaru.pl/open/lekcje.php?mode=pokaz&id=71
3. Serwis internetowy: https://pl.wikipedia.org
4. Serwis internetowy: http://matematyka.studio6.pl/
5. KIPPENHAHN R.: Tajemne przekazy, Prószyński i S-ka Warszawa 2000.
6. MENEZES A., OORSCHOT P., VANSTONE S.: Kryptografia stosowana, WNT, Warszawa.
7. KARBOWSKI M.: Podstawy kryptografii. Wydanie III (Ebook), Helion.

Viktor MOLITSKYI[1], Nazariy YUZVIN[2]

Supervisors: Andriy LUTSKIV[3]

UŻYCIE DEEPLEARNING4J DO WERYFIKACJI DYNAMICZNEGO PODPISU

Streszczenie: W artykule opisano użycie Deeplearning4j do budowania systemu weryfikacji dynamicznego podpisu opartego na algorytmach uczenia głębokiego. Zostały wdrożone i prowadzone badania wykorzystania sieci neuronowej do weryfikacji dynamicznego podpisu z użyciem Deeplearning4j.

Słowa kluczowe: weryfikacja dynamicznego podpisu, uczenie głębokie, sieć neuronowa, biometryczny system uwierzytelniania

USING DEEPLEARNING4J FOR ONLINE SIGNATURE VERIFICATION

Summary: This paper describes using Deeplearning4j for build online signature verification system based on deep learning algorithms. Implemented and conducted research feedforward neural network for online signature verification using Deeplearning4j.

Keywords: online signature verification, deep learning, feedforward neural network, biometric authentication system

1. Introduction

Nowadays, biometric technologies are increasingly used to provide information security and control access right to secure resources. They allow unambiguous identification of the user and his authority over a particular resource. Unlike traditional authentication methods (passwords, cards, various electronic keys), biometric features are very difficult to counterfeit and can not be lost, stolen or transferred to another person.

One of the most common biometric methods of authentication is handwritten signature verification, that employ various specifications of a signature.

There are two types of signature verification:

- offline (static);

[1] Ternopil Ivan Pul'uj National Technical University, Department of computer systems and networks, molitskyi@gmail.com

[2] Ternopil National Economic University, nazariy.yuzvin@gmail.com

[3] Assoc.Prof, PhD, Ternopil Ivan Pul'uj National Technical University, Department of computer systems and networks, l.andriy@gmail.com

- online (dynamic).

In the offline verification, we have the shape of the signature by capturing or scanning them from papers and the system must extract features from the picture of the signature. Therefore, in offline verification system, input data contains x-y coordinates of signatures. However, in the online signature verification, the system uses devices for capturing additional information while the user is signing [1].

Online signature verification systems perform better than offline systems because more dimensions of information are available.

2. Analysis of recent research

Handwritten signature verification is an active research area because of the long-term and widespread use of signatures for personal authentication.

Many previous works in online signature verification use Dynamic Time Warping [2] [3] or Hidden Markov Models [1] [4]. DTW is an effective templatebased method for online signature verification in which only small amounts of data is available. HMMs can be regarded as a soft version of DTW and outperforms DTW when enough training signatures are available [1]. Recent works in field of signature verification use deep learning algorithms: feedforward neural networks, recurrent neural networks, convolutional neural network and their modifications[5][6][7].

3. Problem statement

Neural networks show good result in online signature verification and nowadays exist many frameworks for build different kind of neural networks. The goal of this paper is select most suitable deep learning framework that work on JVM for building online signature verification system.

In the process of research were considered TensorFlow and Deeplearning4j. TensorFlow is one of the most popular frameworks but it Java API is currently experimental and is not covered by TensorFlow API stability guarantees, that is why was chosen Deeplearning4j.

Deeplearning4j is a deep learning programming library written for Java and the Java virtual machine (JVM) and a computing framework with wide support for deep learning algorithms.

4. Test data

For training models and checking results of working system was used dataset “SigComp11” [8] that contains simultaneously acquired online and offline samples.

Training set consist signatures of 10 reference writers and skilled forgeries of these signatures. Total number of online signature samples in training set are 449. Test set consist signatures of 54 reference writers and skilled forgeries of these signatures and contain 1907 online signature.

One online signature is range of points that contain three parameters: x-coordinate, y-coordinate and pressure. This points was captured with sampling rate 200 Hz using WACOM Intuos3 A3 Wide USB Pen Tablet.

5. Experiments

The experiments were conducted on data which are described in the previous section and consist of three steps.
First step is preprocessing of input data which include the following operations [9]:

- determination of signature trend and movement of every signature realization to the origin (0,0);
- signature duration scaling.

Second step is training feedforward neural network (Fig. 1) which is build using Deepleadning4j.

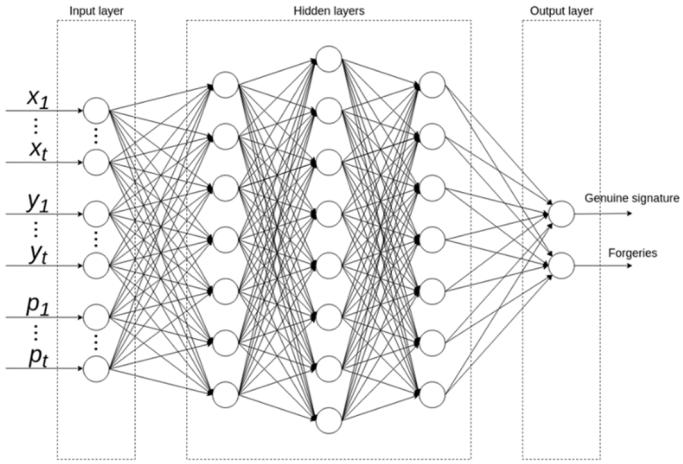


Figure 1. Base schema of used neural network

Last step is evaluation of work trained at previous step neural network.
During the experiment were built and evaluate neural networks with different number of hidden layers and different numbers of neurons on each of them. As activation function at hidden layers was used Relu and for output layer was used Softmax. Results of experiments is present in table 1.

Table 1. Results of experiments

Layers	Epoch	Accurency
6000-8000-4000-2	200	82.31%
6000-8000-4000-2	500	83.56%
6000-8000-4000-2	1000	83.88%
6000-8000-9000-7000-3000-500-2	200	80.9%
6000-8000-9000-7000-3000-500-2	500	92.34%
6000-8000-9000-7000-3000-500-2	1000	93.16%

6. Conclusion and future work

This paper describes the process of building a handwritten signature verification system based on an artificial feedforward neural network using Deeplearning4j. For testing accuracy of working system was used dataset "SigComp11" which containes signatures of 10 reference writers and skilled forgeries of these signatures. As it can be inferred from experimental results signature verification system based on neural network show good results.
As a future work for improve accurancy may use other more complex architectures of artificial neural networks such as LSTM.

REFERENCE

1. FIERREZ-AGUILAR J., NANNI L., LOPEZ-PEALBA J., ORTEGA-GARCIA J., MALTONI D.: An On-Line Signature Verification System Based on Fusion of Local and Global Information. Springer Berlin Heidelberg, 2005.
2. KHOLMATOV A., YANIKOGLU B.: Identity authentication using improved online signature verification method, Pattern Recognition Letters 26(2005)15, 2400–2408.
3. SHARMA A., SUNDARAM S.: An enhanced contextual dtw based system for online signature verification using vector quantization, Pattern Recognition Letters, 84(2016), 22–28.
4. FIERREZ J., ORTEGA-GARCIA J., RAMOS D., GONZALEZ-RODRIGUEZ J.: Hmm-based on-line signature verification: Feature extraction and signature modeling, Pattern Recognition Letters, 28(2007)16, 2325– 2334.
5. HEINEN M. R.: Handwritten Signature Authentication using Artificial Neural Networks. IEEE International Conference on Neural Networks - Conference Proceedings. 10.1109/IJCNN.2006.247206., Heinen Milton Roberto, Osorio Fernando 2006, 5012 - 5019.
6. SONGXUAN LAI, LIANWEN JIN, WEIXIN YANG: Online Signature Verification using Recurrent Neural Network and Length-normalized Path Signature Descriptor arXiv:1705.06849v1 [cs.CV] 19 May 2017
7. KSHITIJ S. Off-line Handwritten Signature Verification using Artificial Neural Network Classifier, Kshitij Sisodia, S. Mahesh Anand, International Journal of Recent Trends in Engineering, 2(2009)2, ACADEMY PUBLISHER ACEEE, 205-207.
8. LIWICKI M., BLUMENSTEIN M., ELISA VAN DEN HEUVEL, BERGER C.E.H., STOEL R.D., FOUND B., CHEN X., MALIK M.I.: SigComp11: Signature Verification Competition for On- and Offline Skilled Forgeries, Proc. 11th Int. Conference on Document Analysis and Recognition, 2011
9. ЛУЦКІВ А.М.: Математичне моделювання і обробка динамічно введеного підпису для задачі аутентифікації особи у інформаційних системах: Дис. кандидата техн.наук: 01.05.02; - Захищена 03.06.2008; Затв. 03.12.2008. - Тернопіль, 2008.

Elena NYEMKOVA[1], Taras KOSTYRKO[2]

Opiekun naukowy: Vyacheslav CHAPLYGA[3]

METODA PROGNOZOWANIA STOCHASTYCZNYCH SZEREGÓW CZASOWYCH O ZMIENNEJ DYSPERSJI

Streszczenie: Metoda jest poświęcona prognozowaniu szeregów czasowych w przypadku, kiedy funkcja autokorelacji nieznacznie zmienia się w pewnym przedziale czasowym. Stwierdzono, że metoda daje dobre wyniki dla modeli sekwencji stochastycznej, kursów walut oraz sekwencji białego szumu.

Słowa kluczowe: autokorelacja, stochastyczny szereg czasowy, stacjonarność, kursy walut

FORECASTING METHOD FOR STOCHASTIC TIME SERIES WITH VARYING DISPERSION

Summary: The article is devoted to the method of predicting stochastic time series based on the autocorrelation function, which does not significantly change for a certain time range. The method showed good results for the model stochastic sequence, the cross-rate of currencies, the sequences of real noises.

Keywords: autocorrelation, stochastic time series, stationarity, currencies cross-rate

1. The formulation of the problem of forecasting stochastic time series

Time series studies draw the attention of researchers from different fields of science. Above all there are the task of forecasting weather, prediction of financial markets, investigating the cryptographic stability of pseudo-random sequence generators, and many other tasks.

Time series study is a passive method of studying complex systems. It is applied to systems, the simulation of which is almost impossible because there are a large

[1] Lviv Polytechnic National University, Faculty of Computer Technologies, Automation and Metrology, Information Technologies Security department: cyberlbi12@gmail.com

[2] Banking University, Phd and Doctoral studies, specialty: mathematical methods, models and information technologies in economics: taraskostyrka@gmail.com

[3] Banking University, Lviv Educational-scientific Institute, Faculty of Finance, Economics and Accounting, Economics department: 4vyach@gmail.com

number of subsystems and vague interactions between them. As a rule, such systems are social and economic systems or natural systems.

Two goals of time series research are the identification of a complex system and the prediction of the behavior of a complex system. Identification and prediction of the system's behavior mean to find certain invariant characteristics, which are independent from time. In this paper, the prediction of the behavior of complex systems is considered.

Researchers distinguish the following components of time series when they analyze them. There are firstly, the trend line; secondly, periodic changes in the series; thirdly, the stochastic component. The trend is determined by several methods. These methods are technical analysis, moving average, volume indicator and others. Spectral analysis is used to determine the periodic oscillations of the time series. Also, artificial neural networks are used to predict the behavior of time series. ARCH-model, the method of local approximation, linear homodynamical models are used to predict stochastic changes in time series. Despite a huge amount of research, forecasting with a good accuracy of an arbitrary time series has not been done to date. The reason for this is in a wide variety of complex systems and in the openness of these systems. Nevertheless, for certain systems and conditions, forecasting can be performed and the accuracy of the forecast can be increased.

The passive method is based on the analysis of the system's own stochastic signals. Suppose the observed variable, a series of N numbers, is present. These are the values of some measured dynamic variable $x(t)$ with a constant step τ in time, $t_i = t_0+(i-1)\tau$: $x_i=x(t_i),\ i=1,...,N$. The main requirement for study is the following. The invariant characteristics of the initial system and those obtained from the time series must coincide. These characteristics can be determined from the experiment without knowing all the dynamic variables of the system.

The task of this study is to find the invariant characteristics of complex system on the basis of the time series data $x(t_i)$ without carrying out external influences on the complex system and without computing the spectral characteristics.

2. Invariant characteristics of stationary time series

A stochastic process is called stationary, if its basic properties are unchanged in time. The stationary process (stationary series) is characterized by the following four properties: 1) the mathematical expectation (*Mean)* of a stationary series is the constant; 2) the variance of the stationary series *(Var)* is a constant quantity; 3) the autocovariance of the stationary series is a constant, it depends only from the lag value; 4) The autocorrelation coefficient of a stationary series with the lag value is the constant. If these four properties are satisfied, then the process is stationary. Checking on the time series stationarity can be carried out by calculating the above-mentioned four characteristics for multiple subsequences studied series.

Let's consider a few time series. All calculations are carried out in the Mathcad program. The first sequence is a pseudo-random number generator, which is built into the Mathcad program, *{rnd(1)-0,5}*. Ideally, this sequence mimics white noise; the mathematical expectation should be equal zero. The second sequence is taken from the archive of the cross-rate of the USD dollar to the Ukrainian hryvnia. This sequence was converted to test for randomness, namely: a large constant component has been

allocated and the stochastic sequence was normalized to the maximum range, as the sequence has been shifted with respect to zero. These transformations are linear and do not affect to statistical averaging. The third sequence is the noise signal of the computer's audio card; the amplitude of the noise signal is recorded using a program Oscillometr. These three sequences are of sufficient length to allow subsequences to be made. Subsequences must have a sufficient number of samples for statistical averaging.
For each of the three sequences, two sequences were identified. The shift of the second subsequence relative to the first is equal to one sample. Thus, the subsequences are almost the same. The number of samples is equal to one hundred. The results of the calculations are given in Table 1. The average value *Mean1* and dispersion *Var1* are for the first subsequence and the *Mean2* and *Var2* are for the second subsequence.

Table 1. Mean values and variances for subsequences

Sequence	*Mean 1*	*Mean 2*	*Var 1*	*Var 2*
rnd(1)-0,5	0,0059	0,0100	0,081	0,078
normalized cross-rate	-0,059	-0,056	0,064	0,065
real noise of sound card	1,08	1,12	1,743	1,670

The calculation results show significant non-stationary of the first and third sequence, the second sequence is also non-stationary.
For non-stationary sequences, an autocorrelation function is often used to determine the characteristic frequencies. For sequences similar to white noise, the autocorrelation function was not previously used by other researchers. As shown by the authors' research, the autocorrelation function can be used to study noise-like time series due to the unique shape for each complex system.

3. Invariant characteristics of stochastic time series with varying dispersion

The time realization of a series of dynamic variables $x_j(t_i)$ of each system j will be different. Calculations show that for noise-like time series the autocorrelation functions change insignificantly, they retain own form. The autocorrelation functions for the time series under study are presented in Figures 1 to 3. The third subsequence was used to determine how quickly the autocorrelation functions are change. The shift for the third subsequence is 10 samples for all figures.
The function *lcorr(x,x)* was used to calculate the autocorrelation. The result represents 100 values for each subsequence. The procedure of linear interpolation *linterp(i,y(i),x)* was applied to the values of autocorrelation for convenience of comparison.
Three characteristic features are observed. Firstly, the form of the autocorrelation function for different subsequences of each sequence under study remains practically constant. Secondly, there are samples for which the autocorrelation function of different subsequences is the same with great accuracy. For example, samples number 25 and number 28 are in the first figure, the sample number 35 is in the third figure. In the second figure, there is the sample number 50. Thirdly, the autocorrelation function for the cross-rate of currencies is similar to the autocorrelation of flicker noise, which was investigated earlier.

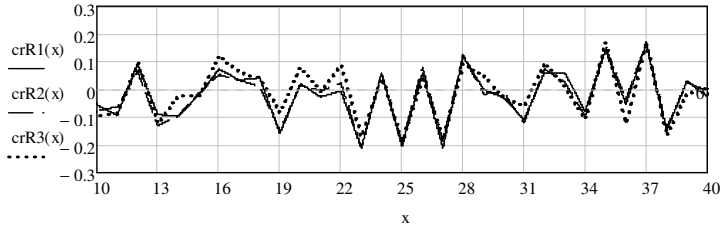


Figure 1. Autocorrelation functions of subsequences of a pseudo-random sequence {rnd(1)-0.5}

Similar graphs are obtained for cross-rates of other currencies.

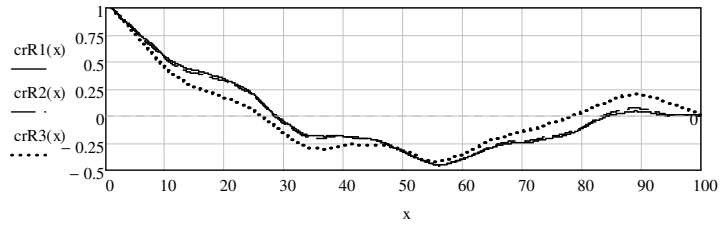


Figure 2. Autocorrelation function of currency cross-rate (USD vs UAH)

At first glance, the autocorrelation function for the noise signal of the audio card is similar to one for pseudo-random sequence *rnd(x)*. Their difference is not critical for this study.

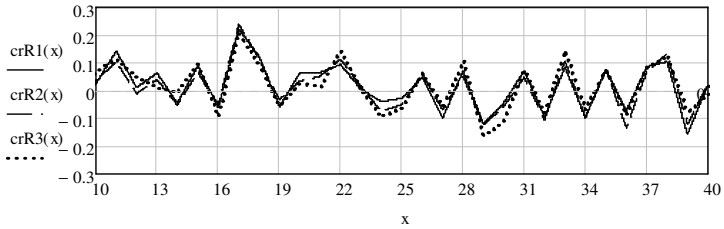


Figure 3. Autocorrelation functions of subsequences of the audio card's signal noise

It is possible with great accuracy to perform prediction of the sample number 101 for the first subsequence using the first and second characteristic features. The next method is proposed for this prediction. The sample with number 101 for the first subsequence is the sample number 100 for the second subsequence. The value of this sample $R2_{100} = y$ can take any value from the range of possible values: from *-0,5* to *+0,5* in increments of *0,01* (for example), y_i*=-0,5+0,01(i-1), i=1..101*. For each value y_i, the autocorrelation functions $A3(y_i)$ are calculated and the values $A3(y_i)_l$ for lag l are selected. Each of them is compared with the value of the autocorrelation function $crR1_l$ with the lag l from the first subsequence. The value y_i, which is the solution of

the equation $A3(y_i)_l = crR1_l$, determines one of the hundred numbers i and y_i. It should be noted that $crR1_l \approx crR2_l$. The graphical solution of the equation $A3(y_i)_l = crR1_l$ is shown in Figure 4. The function *root(A3(y)-crR1$_l$)* was used to determine y_i analytically.

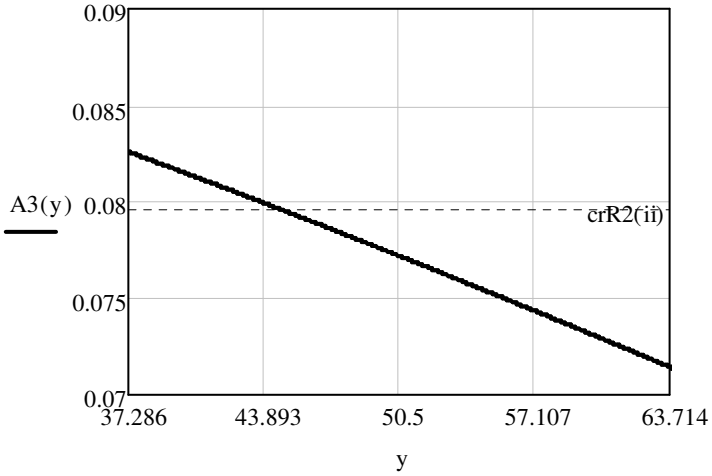


Figure 4. Graphical solution of the equation to determine the predicted value of the sample

The forecast was made for a pseudo-random sequence *{rnd(1)-0,5}*. The first forecast value coincided with the true one with great accuracy. In general, the calculations were carried out for the next 10 samples; all the calculated values practically coincided with the true ones. The results of the forecast are presented in Table 2. For clarity, the coincident numerals of the predicted samples and true samples are underlined.

Table 2. Results of forecasting the values of 10 samples of the sequence {rnd(1)-0,5}

Number i	*Predicted r_i*	*True r*	*δ*
1	-0,0633244	-0,0633169	$-7{,}5\cdot10^{-6}$
2	0,0778651	0,0778666	$-1{,}5\cdot10^{-6}$
3	0,1287013	0,1286670	$34{,}3\cdot10^{-6}$
4	0,0041477	0,0041493	$-1{,}5\cdot10^{-6}$
5	0,1957697	0,1957680	$1{,}7\cdot10^{-6}$
6	-0,3100367	-0,3100483	$11{,}6\cdot10^{-6}$
7	-0,3216296	-0,3216249	$-4{,}7\cdot10^{-6}$
8	-0,0425355	-0.0425416	$6{,}0\cdot10^{-6}$
9	-0,4024815	-0,4024773	$-4{,}2\cdot10^{-6}$
10	-0,4056022	-0,4055958	$-6{,}4\cdot10^{-6}$

It is important to note that the magnitude of the error *δ=Predicted r_i – True r* does not increase when prediction number gets up.

The prediction of the sequence *R1={rnd(1)-0,5}* with the help of the built-in function of the Mathcad *predict* shows the absolute inapplicability of this function to the considered sequence. For example, using the function *predict(R1,99,1)* gives a value of *-0,138* instead of *-0,063*. The peculiarity of the described method is that not all

autocorrelation coefficients are used, but only those that are almost identical for neighboring subsequences.

4. Conclusion

The proposed method for predicting time series has demonstrated good results for nonstationary stochastic sequences. The method works in the case when autocorrelation functions change little for neighboring subsequences, although autocorrelation functions can differ greatly for time-separated subsequences. Closed systems have this property. Actually, good results were obtained for closed systems such as the algorithm for generating pseudo-random numbers and the noise signal of the computer's audio card.
The prediction situation with the help of this method is different if several essentially different processes operate in a complex open system, for example, for flicker noise or a cross-rate of currencies. For such systems, prediction is possible in a certain time range, when the nature of the time series does not change significantly.
The method can be used to test the cryptographic resistance of stream sequences.

LITERATURA

1. BREDIKHIN A., LOSKUTOV A.: To the problem of financial time series analysis. III. ARCH-models in the Russian financial market. Surveys of Applied and Industrial Math., **11**(2004)3, 468-486.
2. BREDIKHIN A., LOSKUTOV A., SEDYKH A.: To the problem of financial time series analysis. I. Linear homodynamical models. Moscow Univ. Phys. Bull., **1**(2000), 10-12.
3. DYVAK M., KASATKINA N., PUKAS A., PADLETSKA N.: Spectral analysis of information signal in the task of identification the recurrent laryngeal nerve during thyroid surgery, Proc. 13th International Workshop Computational Problems of Electrical Engineering, Grubow 2012, p.55.
4. LOSKUTOV A., KOTLYAROV O.: Local approximation: A new method of forecasting of economic indexes. Currency Stag, **11**(2008), 8-13.
5. PETROVICH V.N.: Identification of parameters of mathematical models of dynamic control system. Artificial Intelligent, **4**(2011), 343 - 349.
6. Internet service: Archive of dollar courses to grivne *https://kurs.com.ua/ valyuty/*, 01.09.2017.

Tamara OLESHKO [1], Nadiia IVANCHENKO[2]

SEMANTYCZNO-RAMKOWE MODELE W ZAPEWNIENIU EKONOMICZNEGO BEZPIECZEŃSTWA PRZEDSIĘBIORSTWA

Streszczenie: W artykule opisano tworzenie semantyczno-ramkowych modeli wiedzy z zakresu technicznego i technologicznego potencjału przedsiębiorstwa dla zapewnienia jego bezpieczeństwa ekonomicznego. Omówiono różne systemy semantyczno-ramkowe zapewnienia bezpieczeństwa ekonomicznego przedsiębiorstwa.

Słowa kluczowe: model semantyczny, reprezentacja wiedzy, model, ramka, bezpieczeństwo ekonomiczne, wielowymiarowa wirtualna rzeczywistość

SEMANTIC- FRAME MODEL OF TECHNICAL AND TECHNOLOGICAL POTENTIAL OF THE ECONOMIC SAFETY OF THE ENTERPRISE

Summary: The article is represented to construction of semantic-frame model of knowledge's of technical-technological potential of enterprise economic safety. It is offered to use semantic-frame technologies for the design of technical-technological potential of economic safety. For presentation of this constituent used frame.

Keywords: semantic model, knowledges representation, model, frame, economic safety, multidimensional informational variable adaptive reality

1. Formulation of the problem

The frame modelling method responds to the tasks of the real study and has the following advantages:

- the semantic frame model is universal and describes various aspects, detailed information of economical safety of the enterprise, systems and strategies, structure and business process.

[1] National Aviation University, Department of Economic Cybernetics, Doctor of Engineering, Professor, Head of Department of economic cybernetics, ti_oleshko@ukr.net
[2] National Aviation University, Department of Economic Cybernetics, PhD, Docent of Department of economic cybernetics, ivan730@ukr.net

- the model is applicable at different levels, from the upper, middle level describing the basic categories of economic security up to designing information of the system.
- integrates the concept of various aspects of the economic security of the enterprise and systemize the conceptual provisions.
- the model is user friendly, understandable and able t correct business and IT specialists with the capability to present the categories of the economic security of the enterprise.
- certain aspects of the economic security the system can describe them in deferent languages by selecting it from the introduction in the frame model.
- the structure of the semantic frame model allows to create services and applications on different level of structure.
- semantic frame models are suitable for translating into other languages descriptions in various systems such as UML or XML

2. Analysis of recent research and publications

The problems of model of knowledge and economic safety of enterprises draw attention of many foreign and national scholars, such as O. Varlamov, G. Ivanchenko, A. Kleshchev, I. Artemjeva and others.

3. Problem definition

The permanent change of external environment where people, enterprises, organizations and countries work, adapt to and survive is the feature of the present. In the conditions of competition the market lot of enterprises is determined in a great deal by the speed and exactness of reaction on external environment changes, therefore it requires application of new unconventional conceptions, techniques and tools in the management.

Development of computer technologies enabled planning of knowledge bases (KB), which will organize activity in subject industries of economic safety of the enterprise (ESE). Technical and technological potential is a component of ESE. It is closely associated with financial, innovative and other constituents and, at the same time, possesses its own specific features. Therefore, the solution of intellectual tasks in the industry of technical and technological potential requires consideration of large volumes of information.

The semantic and ontological model of knowledge is necessary for KB creation. Ontology of ESE is a formal specification of concepts and indexes of technical and technological of ESE potential, which describes different properties and attributes of concepts (slots) and limitations encapsulated in slots. Ontology together with the set of individual copies of classes forms the semantic model of KB.

Visual models like the ontological ones possess the special cognitive force. Visualization of ontology allows specialists on *knowledge engineering* to directly design, formulate and explain both the nature and structure of economic processes.

4. Statement of the main material

The principal reason of cyclic crises of the economic systems is general production funds wearing-out. This rule operates both for separate subject of the system and a separate economic player. In this connection the estimation of the technical state and motion of capital assets becomes an indispensable condition. It is also needed for planning and creation of the enterprises' reserve for depreciation, which is formed with the purpose of financial resources accumulation necessary for the basic facilities renewal.

Technical and technological potential of ESE of a separate economic player depends on the technical and technological level of production and determines the level of products, its competitiveness and expenditures. This potential is foremost estimated by the level of technical perfection of labour facilities and technology of production.

The technical and technological potential of ESE consists of a few successive stages:

1) market analysis of technologies in production of goods similar to the type of a certain enterprise or a system designer organization: collection and analysis of information about the features of technological processes on enterprises that make similar products; analysis of scientific and technical information on new developments in a certain industry, and also technologies which are able to carry out intervention in the industrial technological market; forming of KB technical and technological potential of ESE.
2) analysis of particular technological processes and discovery of internal reserves for the improvement of the used technologies. The calculation of indicators is to be made.
3) estimation of prospects of the market development of enterprises' products and prognostication of possible specific of necessary technological processes for the issue of competitive innovative commodities.
4) design of a technological strategy of an enterprise development using the accumulated knowledge of KB technical and technological potential of ESE, in order:
   - to find out perspective commodities;
   - to plan the complex of technologies for production of perspective commodities positions;
   - to finance technological development of an enterprise through the charges optimization program of technological development;
   - to design the general plan of an enterprise's technological development;
   - to make the plan of own corporate resources in accordance with the plan of an enterprise's technological development.
5) operative realization of enterprise's technological development plans in the process of production and economic activity.
6) analysis of results of the practical used actions for providing of technical and technological potential of ESE on the basis of the special card of efficiency calculations.

In addition to the above mentioned indicators, technical and technological potential of ESE is also characterized by the following indicators:

- the level of technologies progressiveness ;
- the level of progressive products;
- the level of technological potential.

Model creation. It is expedient to use a network model for knowledge representation of ESE.

Semantics determines the sense of signs and relations between characters and objects, which they determine.

Semantic network of ESE is a graphic system of denotations for knowledge representation in the templates of linked knots and arcs. In other words: a semantic network is the oriented graph with its tops denoting concept, and arcs denoting relations between them.

Network models of ESE formally can be set as $H = <I, C_1, C_2, ..., Cn\ G>$, where I is the number of informative units; $C_1, C_2, ..., Cn$ are the number of types of links between informative units; G is the reflections between informative units that belong to I, links out of a set of links types.

Declarative graphic representation is general for all semantic networks of ESE. It can be used for representation of knowledge or creation of the automated systems of decision-making representation on the basis of knowledge.

Tops may represent: concepts, events, properties. The marks of tops have a reference character and represent some entities identifiers. The marks of arcs mark the elements of multitude of relations.

The classification of objects types and selection of some fundamental types of links between objects are important when using the semantic network for knowledge representation. Regardless of the features of an environment being designed, it is possible to assume that any more or less difficult model represents some generalized, particular and aggregate objects.

The particular object of ESE is the specially selected single essence.

The aggregate object of ESE is the object of a problem environment made of other objects which are its components. Both the generalized and the particular object can be aggregated there.

Links between the objects are determined in the described typification of problem environment objects.

Generic links can exist between two generalized objects. The use of inheritance provides the effective method of knowledge simplification and volume reduction of information needed to be memorized for a particular knot. It considerably enables both acceleration of knowledge processing and information retrieval by means of the requests of general nature.

Link «is a representative» and can exist between generalized and particular objects. It takes place in that case, when a particular object belongs to the class with the proper generalized object.

For the creation of a semantic and ontological as well as a frame model of technical and technological potential of ESE the ontology editor *Protégé* was used, which allows to design ontology opening out the hierarchical structure of abstract and concrete classes and slots. The results of the creation of semantic and ontological and frame models of technical and technological potential of ESE are represented in pictures 1, 2.

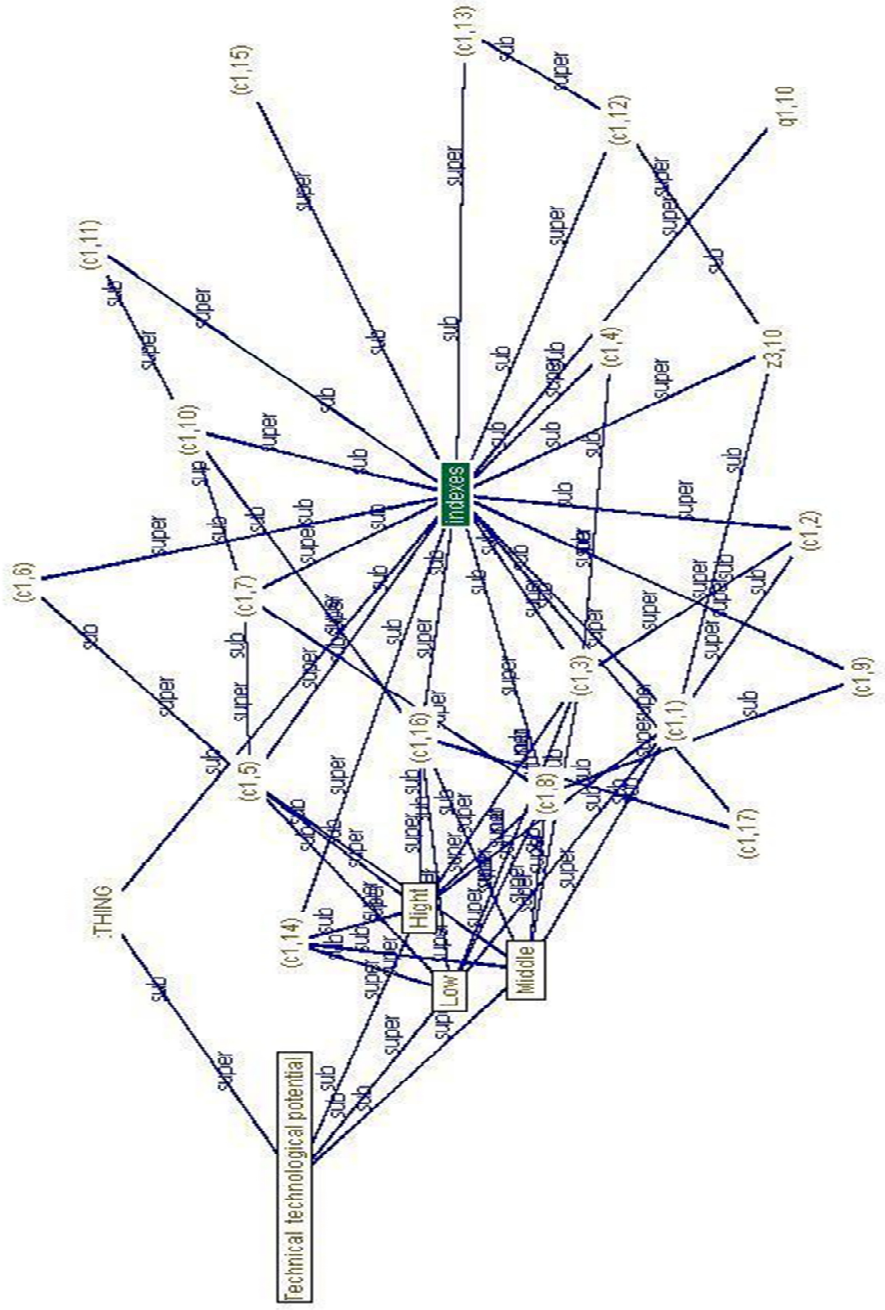


Figure 1. Semantic model of technical and technological potential of ESE

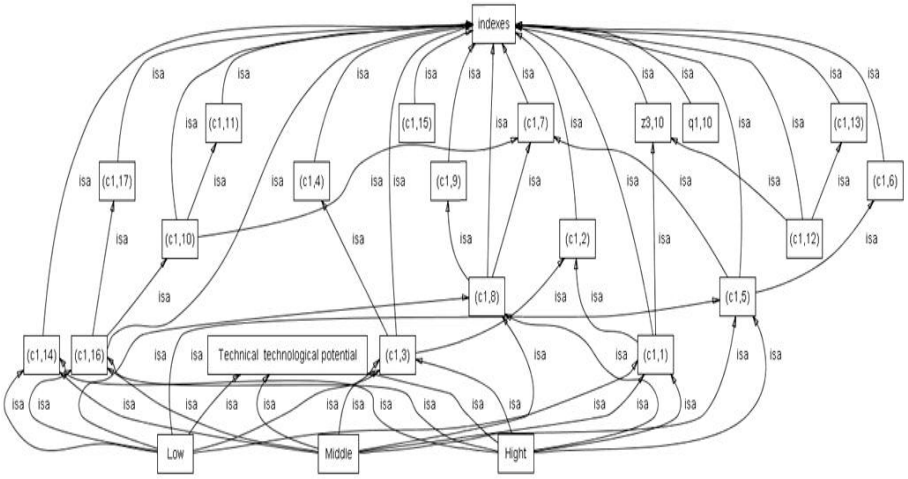


Figure 2. Frame model of technical and technological potential of ESE

5. Conclusions

Thus, it is foreseen to use semantic-ontological models as conceptual facilities helping to design and create semantic-ontological technical-technological potential of enterprises economic safety which gives possibility of development: reliable semantic base in determination of content of technical-technological potential, general logical model of rules, consisting of dictionary and set of assertions in logical language, providing a basis for communication between KB and computers agents with the purpose of creation the management information system of enterprises economic safety.

REFERENCES

1. KLESHCHEV A.S., ARTEMJEVA I.L.: A structure of domain ontologies and their mathematical models. Proceeding of The Pacific Asian Conference on Intelligent systems 2001 (PAIS 2001), Korea Intelligent Information Systems Society, 2001, 410–420.
2. ZATSERKLIANYI M.M., MELNYKOV O.F.: Fundamentals of Economic Security : Tutorial. KNT, Kyiv 2009.
3. IVANCHENKO N., IVANCHENKO G.: Semantic networks of financial potential of economic safety of enterprise. EIIC 2012: Proceedings in Electronic International Interdisciplinary Conference, 3-7 September 2012, 238-241.
4. IVANCHENKO N.: MIVAR technologies modeling of technical - technological potential of enterprise. Actual problems of the economy, (2014)151, 505-511.

Volodymyr POGORELOV[1], Oleh TEREIKOVSKYI[2]

Scientific Supervisor: Ihor TEREIKOVSKYI[3]

ROZPOZNAWANIE CYBERATAKÓW PRZY UŻYCIU SIECI NEURONOWEJ Z RADIALNYMI FUNKCJAMI BAZOWYMI

Streszczenie: Praca poświęcona jest określeniu wykonalności wykorzystania klasycznej sieci neuronowej z radialnymi funkcjami bazowymi do rozwiązywania problemów rozpoznawania ataków cybernetycznych w celu zakłócenia funkcjonowania oprogramowania systemów i sieci komputerowych. Przyjęto metodologię budowy sieci, przeprowadzono weryfikację i analizę ograniczeń obliczeniowych. Uzasadniono możliwe obszary zastosowania sieci.

Słowa kluczowe: rozpoznawanie cyberataku, ochrona informacji, sieć neuronowa

CYBERATTACK RECOGNITION WITH RADIAL BASIS FUNCTION NEURAL NETWORK

Summary: The article is devoted to determining the feasibility of using a classical neural network using radial basis functions for cyberattack recognition aimed at disrupting the computer system and network software operation. The methodology of network construction is considered, verified and computational constraints analysed. Possible areas of network application are substantiated.

Keywords: cyberattack recognition, information protection, neural network

1. Formulation of the problem

Over the past few years, interest in the application of neural networks (NN) in the means of technical control and diagnostics has increased significantly. NN are mainly used as a control element in the technical systems state recognition blocks. Efficiency

[1] National Technical University of Ukraine, Igor Sikorsky Kyiv Polytechnic Institute, Department of System Programming and Specialized Computer Systems: Computer Systems and Components, Doctoral Student, volodymyr.pogorelov@gmail.com

[2] National Technical University of Ukraine, Igor Sikorsky Kyiv Polytechnic Institute, Department of Computer Engineering: Computer engineering, Student, terejkowski@ukr.net

[3] Doctor of Technical Sciences, Assoc., National Technical University of Ukraine, Igor Sikorsky Kyiv Polytechnic Institute, Proffesor of the Department of System Programming and Specialized Computer Systems, terejkowski@ukr.net

of the application has been proved to largely depend on the NN's computational capabilities, which in turn are determined by its architecture. Radial basis function (RBF) system-based architectures are considered promising [1,2]. Obviously, networks of this kind have certain prospects in computer system and network parameter diagnostics used to recognize cyberattacks aimed at disrupting the software operation. Defining these prospects is the main topic of this article. The problematic is directly related to such important scientific and practical tasks as ensuring reliability of distributed computer system and network operation.

2. Analysis of recent research and publications used as sources

The application of RBF NN assumes that to increase the likelihood of linear division of images into classes, it is necessary to place these images in a space of high dimension in some non-linear way [3]. The simplest RBF form is a tree-layer NN: input, hidden and output. A simplified RBF scheme with one neuron in the input layer is shown in Fig.1. The input layer is tasked with the distribution of input data to the neurons in the NN's hidden layer. The hidden layer includes neurons with a radially symmetric activation function. Each of the hidden neurons is intended to store a separate reference image that corresponds to a separate class. Often, the number of neurons in the hidden layer is greater than the number of input neurons.

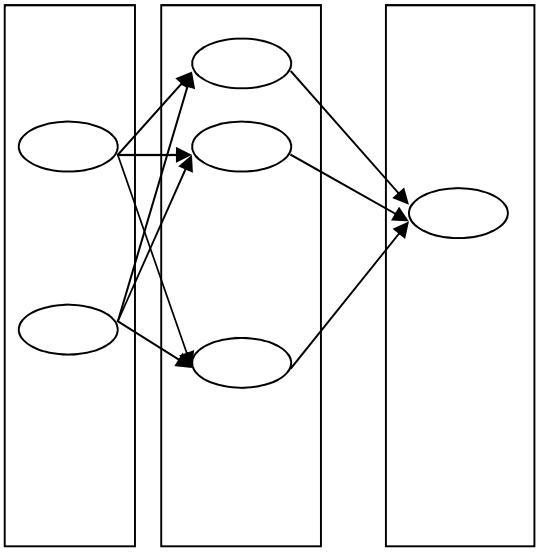


Figure 1. Simplified RBF scheme

For a j-th neuron in a hidden layer, the total input signal (net) from a certain input vector (x) is calculated as a Euclidean norm:

$$net_j = \sqrt{\sum_{i=1}^{N} (x_i - w_{ij})^2} \ , \tag{1}$$

where x_i – is the i-th component of the input vector, x, w_{ij} – is the weight coefficient of the j-th hidden neuron with the i-th input neuron, N – is the number of input neurons.
As a function of activation for neurons in the hidden layer, it is typical to use the Gaussian function:

$$\varphi_j(net) = \exp\left(-\frac{1}{2\sigma}\sum_{i=1}^{N}(c_i - x_i)^2\right), \tag{2}$$

where $\varphi_j(net)$ –is the activation function of the j-th neuron of the intermediate layer, net – is the total input signal, x – is the output vector, c – is the Gauss function centre, σ – is the radius of the Gaussian function.

After a non-linear transformation, the signals from the hidden layer neurons fall into the output layer of the neurons that have linear activation functions. The calculation of the total input signal for any neuron of the output layer is carried out accordingly (1). Note that the set of values of the activity of all hidden neurons is defined by the vector on which the input vector is displayed:

$$\varphi(x) = |\varphi_1(x), \varphi_2(x), ..., \varphi_M(x)|, \tag{3}$$

where x - is the input vector, $\varphi(x)$ - is the output vector, $\varphi_i(x)$ - is the component of the output vector associated with the i-th hidden neuron, M - is the number of hidden neurons.
Since the activation function of the hidden layer neurons is non-linear, for the simulation of any input information, there is only one intermediate layer with a sufficiently large number of neurons. The total number of synaptic connections (Z_i) of the RBF network can be calculated as follows:

$$Z_\Sigma = Z_1 + Z_2, \tag{4}$$

$$Z_1 = N \times M, \tag{5}$$

$$Z_2 = M \times K, \tag{6}$$

where Z_1 – is the number of synaptic connections of hidden neurons, Z_2 – is the number of synaptic connections of the output neurons.

RBF is trained in stages. During the first stage, the number of neurons in the hidden layer and the coefficients (centre and radius of the Gaussian function) for the activation functions of the hidden layer neurons are calculated. To calculate the Gaussian function centre, it is recommended to apply the “K-mean” method or the Kohonen network learning method “the winner takes everything” [3]. The next stage of training is the calculation of the radii of Gaussian functions. To do this, you can apply the “K of the closest neighbours” method. After calculating the parameters of the Gaussian function, which represent the weight coefficients of the hidden layer neurons, it is necessary to determine the weight coefficients of the neurons of the output layer. In [3], the definition is proposed to be implemented using the “teaching with a teacher” method according to the Widrow-Hoff rule:

$$\Delta w_j = \eta \times net_j \times \delta_j , \tag{7}$$

where Δw_j - is the correction of weight coefficients of the j-th neuron of the output layer, η - is the standard of training, δ_j - is the output signal error of the j-th output neuron, net_j - is the cumulative input signal of j-th output neuron.

In its turn, the error of the output signal for the j-th neuron is calculated as follows:

$$\delta_j = w_j^f - w_j^o , \tag{8}$$

where w_{fj} - is the actual output, w_{oj} - is the expected output of the j-th output neuron.

In many cases, in the calculations (7,8) it is assumed that all the connections of the output neurons require the same magnitude of correction of weight coefficients. Therefore, for RBF with one output neuron (7, 8) it is possible to rewrite it as follows:

$$\Delta w_j = \frac{\eta \times net \times \delta}{M}, \tag{9}$$

$$\delta = w^f - w^j , \tag{10}$$

where M is the number of hidden neurons, w_f - is the actual output, w_o - is the expected output of the network, net - is the total input signal of the output neuron.

The application of (7–10) indicates an iterative training process. In this case, the author did not succeed in finding the theoretical dependence of the optimal number of training iterations on the parameters of RBF. At the same time, the number of iterations (i), hidden (M), input (N), and output neurons (K) directly affect the duration of training (T), which is one of the main characteristics of NN:

$$T \approx i \times K \times N \times M . \tag{11}$$

After training, it is recommended to check the quality of RBF recognition on test samples that are not included in the training. If the quality is unsatisfactory then the weight coefficients are corrected, first for the hidden and then for the output neuron layer. The performed analysis of the mathematical support and functional features of the RBF network allows us to proceed to the definitions of:

- the advantages and disadvantages of its application in the tasks of information protection in relation to other types of NN;
- specific areas of application for solving information security problems.

3. Formulating the purpose of the article

Evaluate the possibility of using an RBF network to recognize cyberattacks aimed at disrupting the operation of the computer system and network software.

4. Presentation of the main research material

The method is based on the procedure of spectral subtraction of signals, the theoretical basis of which is the assertion that the sum of spectrum of the voice signal and noise is equal to the sum of the spectra of this signal and noise.
At the first stage of the research, numerical experiments were carried out to verify the RBF model (1-10) and to determine the optimal number of training iterations, which directly affects the computational efficiency of the network. Experiments were carried out using two programs created by the author. In the first series of experiments, the approximation of the function was carried out $y = 0{,}5x + 2x^2 - x^3$. The choice of the function is due to it being presented as an RBF application example in [3]. As in [3], an RBF network and study samples were used, with the following parameters: Gaussian function radius 0,5, training standard 0,1, hidden neuron count = 9, input and output neuron count = 1, training examples with x [-1,1] count = 30, the Gaussian function centres are at points 0.88889, 0.66667, 0.44444, 0.22222, 0, 0.222222, 0.444444, 0.666667, 0.888889. A fragment of the calculated RBF output indicators is shown in Table 1.

Table 1. RBF output values for different number of training iterations

№.	Iterations			Actual function value
	1	100	1,000	
1	-0.8365	1.1255	2.0503	2.5
2	-0.9034	1.8632	1.5626	1.392
3	-0.8655	1.4190	0.9246	0.636
4	-0.6997	0.4050	0.3194	0.184
5	-0.4083	-0.0140	-0.0752	-0.012

The analysis of data in Table 1 indicates that both the maximum and the average relative error of the of RBF output stabilizes with the number of training iterations more than 100. Increasing the number of training iterations does not reduce the magnitude of these errors, although it significantly affects the values of the weight coefficients of the output neuron. To increase the visibility of the obtained results in Fig. 2 a graph of the functions of approximated RBF, which trained for 1000 iterations and the actual function graph is shown $y = 0{,}5x + 2x^2 - x^3$.

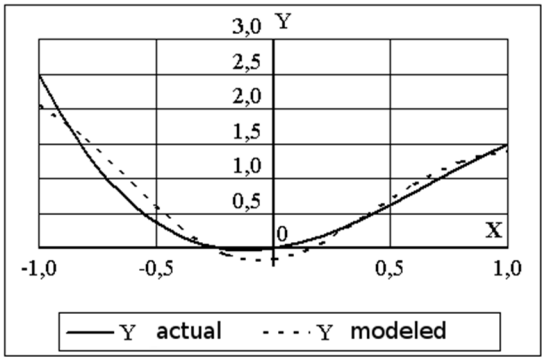


Figure 2. Actual and simulated RBF function graph

One can note the proximity of simulated and actual charts, which is a confirmation of the accuracy of the model. In addition, experiments were performed to approximate linear and quadratic RBF functions dependent on two and three variables. The training of the model was conducted in the range of training examples from 10 to 100 with the hidden neuron count from 10 to 20.

The model showed sufficient accuracy of approximation of linear functions after 100 training iterations, quadratic functions—after 1000 iterations. In this case, the relative average error of approximation was within 1%, and the relative maximum error was within 3%. Thus, the optimal number of training iterations directly depends on the number of input parameters and the type of simulated function. In this case, the computational complexity of the training iteration is proportional to the number of hidden neurons. For this reason, in complex cases, the training of RBF may require a significant amount of computation, which contradicts the known theoretical conclusions [3].

The results of experiments and conclusions [2, 3] make it possible to indicate some advantages of the RBF network in comparison with the multi-layered perceptron selected as a comparative basis for the application of other NN types in the tasks of recognition of cyberattacks. First, RBF allows you to simulate an arbitrary function using only one intermediate layer, which, to some extent, simplifies the architecture of the network.

Secondly, training of the intermediate and output layer of neurons of the RBF can be carried out with the help of sufficiently tested methods of linear simulating. Another advantage of RBF is a simple program implementation, which is achieved through simpler training methods. At the same time, the results of comparing the computational capabilities of RBF and the multi-layered perceptron are given in [1, 2], indicating that for the simulation of complex functions, the RBF network needs a slightly larger number of neurons.

This is because in the process of approximating data at any point in the perceptron all hidden neurons are involved, and in RBF only the closest ones are involved. Therefore, for RBF, the number of neurons required for the approximation of a function with a given accuracy increases exponentially with the increase in the dimension of the input signal. As a result, the program implementation of the RBF will make the classification longer, spend more resources, but will learn more quickly than the software implementation of the multilayer perceptron. Thus, the application

of RBF in the tasks of information security should consist in conducting an operational analysis of information in the process of which the question of the speed of the rough definition of classes prevail over the objectives of accuracy of classification.
In traditional spheres of application, such an analysis is often carried out with the aim of rough estimation of the architecture of the perceptron, which will more accurately solve a similar problem. Also, the corresponding goal can be addressed to the RBF in the field of information security.
It can also be concluded that the RBF network is expedient to use in those means of recognition of cyberattacks that are based on the analysis of the set of interrelated discrete parameters. These security features include network cyberattack detection, vulnerability detection systems, antiviruses and anti-keyloggers. Possible input parameters of the RBF network for these protective devices are given in Table 2.

Table 2. RBF network input parameters in the cyberattack recognition solutions

Name of protection solution	Input parameters
Network attack recognition system	Parameters of network requests and events in the computer system: user input/output, process count, file access, intervals for requests to objects of a computer system.
Vulnerability detection system	Computer system settings: user count, user privileges, computer system object access parameters, open port count and nomenclature, running network services, DCOM/COM + administrative settings.
Antiviruses, anti-keyloggers	Parameters of events in the computer system: running program and process count and range, access to file processes, network service access attempts, executable file modification attempts, operating system API access. Parameters of program code signatures responsible for self-reproduction and destructive actions: file object access, false situation interception operators, network resource use functionality.

In the author's view, the main applicability restrictions for the RBF network in the field of protection are:

- insufficiently studied possibilities in the field of generalization and formation of new knowledge.
- the impossibility of independent study in the process of practical exploitation.

In practice, these restrictions can negatively affect the ability to diagnose new types of attacks or unknown vulnerabilities. To solve this problem, it is necessary to develop a methodology for the formation of a qualitative primary training sample and to

conduct research in the direction of developing such an RBF characteristic as a generalization of such input information. In addition, a study should be conducted in the direction of combined use of RBF with other types of NN.

5. Conclusion

- The general preconditions for using the RBF network are: simplicity of structure, which leads to ease of program realization and high speed of training.
- The general restrictions of the RBF network include: limited computing capabilities in comparison with multilayer perceptron, many empirical parameters used in the training of hidden layer and poor extrapolation of results outside the field of study data. Therefore, the training sample must be presented with practically the entire range of possible input data.
- The application of RBF is expedient in the tasks of recognition of cyberattacks, if necessary, to carry out rapid operational analysis of data to further use the results in more powerful systems. For example, with the help of RBF laboratory analysis of computer virus signatures, it is possible to approximately determine the characteristics of a multi-layer perceptron suitable for use in an antivirus detection unit.

REFERENCES

1. TEREYKOVSKY I.A.: Application of neural networks in the recognition of macroviruses. Legal, Regulatory and Metrological Support Information Security System in Ukraine, (2006) Issue 2 (13), 176-183.
2. TEREYKOVSKAYA L., PETROV O., ALEKSANDER M.: Prospects of neural networks in business models. TransComp 2015. November 30–December 3, 2015, Zakopane, Poland, 1539–1545.
3. KORCHENKO A., TEREYKOVSKYI I., KARPINSKI M., TYNYMBAYEV S.: Neural network models, methods and security options assessment tools Internet-oriented information systems. Nash Format, Kiev 2016. (in Russian)
4. TEREYKOVSKY I.: Neural networks in computer information protection solutions. PolygraphConsulting, Kiev, 2007.
5. TEREYKOVSKY I. A.: Application of radial basis function network to problems of software security diagnostics. Scientific and Technical Collection "Management of the development of complex systems" of Kyiv National University of Civil Engineering and Architecture, (2010), Issue 3, 111-114.
6. KARPINSKI M.: Information Security. Measurements, Automation and Monitoring, Warsaw, 2012. (in Polish)
7. PETROV O., BOROWIK B., KARPINSKYY M., KORCHENKO O., LAKHNO V.: Immune and defensive corporate systems with intellectual identification of threats. Slaska Oficyna Drukarska, Pszczyna 2016.

Artem POLOZHENTSEV[1], Andriy FESENKO[2]

Opiekun naukowy: Viktor GNATYUK[3]

METODA OCENY EFEKTYWNOŚCI CSIRT

Streszczenie: W tej pracy opracowano metodę oceny skuteczności funkcjonowania CSIRT, realizowaną w następujących etapach: określanie wyników CSIRT, określanie KPI, budowanie panelu wskaźników. Opracowana metoda może być wykorzystana do monitorowania, zarządzania, analizowania i zwiększania skuteczności CSIRT.

Słowa kluczowe: CSIRT, KPI, macierz korelacji, efektywności

METHOD FOR CSIRT PERFORMANCE EVALUATION

Summary: The article develops a method for evaluating the effectiveness of the CSIRT, which is implemented in the following stages: determining the performance of the CSIRT, defining the KPI, building a panel of indicators. The developed method can be used to monitor, manage, analyze and enhance the effectiveness of the CSIRT.

Keywords: CSIRT, KPI, correlation matrix, efficiency

Introduction

Now, the information security of persons, societies and countries is one of the main components of national security in general because information and communication technologies are widely used in all areas.

The problem of information security is not only actual, but also global. Information security incidents become more complex and often. Usually, the response to cyber incident directed at CSIRT (Computer Security Incident Response Team) which every year receive more and more assignments and challenges. It becomes necessary to evaluate and analyze the work of CSIRT. This index is most important to informational security of some organization or country. Periodic (monthly, quarterly, etc.) evaluation of CSIRT`s work authorize strong and weak departments, groups, some employees for improving their work in future and highlight some trends based

[1] National Aviation University, Kyiv, Ukraine
[2] National Aviation University, Kyiv, Ukraine
[3] National Aviation University, Kyiv, Ukraine, viktorgnatyuk@ukr.net

on statistical data. The analysis showed that CSIRT performance evaluation not given enough attention, and this could adversely affect the level of information security.
After analyzing, the existing methods for evaluating staff or unit discovered that none of the methods is universal. Everyone has advantages and disadvantages. In addition, in order to achieve the maximum result in the evaluation it is possible to use several methods simultaneously. Moreover should take into account the specifics of the organization, staff or unit is estimated. The chosen methods should meet to the structure of the enterprise, the nature of the activities of staff, the objectives of evaluation, to be simple and understandable; include both qualitative and quantitative indicators. Based on this, has developed a method that combines the advantages of known techniques to minimize gaps and takes into account the specifics of the CSIRT. The developed method consists of three steps: determining the performance of the CSIRT, determining the key performance indicators of the CSIRT, building a panel of indicators and visualizing the dependence of KPI and E.

Stage 1 - Determining the performance of the CSIRT

When a CSIRT is functioning, the information about Cyber incidents is recorded to the database (DB). Among the basic indicators of the functioning of CSIRT [1, 3], which have quantitative values should be allocated the following (table. 1).

Table 1. CSIRT Performance Indicators

Mark	Name
E	Efficiency
LRI	level of resolving the incident
INAI	Incorrect number appointments of the incident
DRI	Duration of resolving the incident
ECS	Evaluation customer satisfaction
PRI	The priority of the incident
DIR	Duration of the incident registration
CII	Information provided about the incident

For the implementation of this stage, use a set of performance indicators CSIRT $\boldsymbol{PI}$:

$$\boldsymbol{PI} = \{ \overset{p}{\underset{q=1}{}} \boldsymbol{PI}_q \} = \{ \boldsymbol{PI}_1, \boldsymbol{PI}_2, ..., \boldsymbol{PI}_p \}, \quad (1)$$

where $\boldsymbol{PI}_q \subseteq \boldsymbol{PI}$, $(q = \overline{1, p})$, p – the number of performance indicators of the CSIRT.

Stage 2 – Determination of Key Performance Indicators for CSIRT.

To determine the key performance indicators from the set of CSIRT performance indicators was used the multiple correlation-regression analysis process [4], which includes the following steps:
Step 1. Selection of all possible factors, which affect on the indicator (or process) that being investigated. Each factor determines numerical characteristics if some factors can`t be quantitatively or qualitatively determined or statistics are not available to them, they will removed from further consideration.

Step 2. Choosing a regressive or multi-factor model, that is finding an analytical expression that describes the link between factor factors with the resultant (function selection):

$$\hat{Y} = f(x_1, x_2, x_3, \ldots, x_d), \tag{2}$$

where $\hat{Y}$ – resultant variable function; $x_1, x_2, x_3, ..., x_d$ – factors signs.

An important problem is the choice of an analytical form for a function that links factors with a resultant feature-function. This function have to show real connections between the studied parameters and factors. It is important to note that the empirical justification of the type of function using the graphic analysis of the connections for multi-tasking models is unsuitable. Given that, any function of many variables by logarithms or replacement of variables can be reduced to a linear form then in practice the multiple regression equations are given linearly:

$$\hat{Y} = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_d x_d, \tag{3}$$

where $a_0, a_1, ..., a_d$ – the parameters of the equation are to be measured.

If for every factor and for a productive feature known d values $y_h, x_{1h}, x_{2h}, \ldots, x_{dh}$, at $h = 1, 2, \ldots, m$ then using the standard procedure of the least squares method to evaluate the parameters a system of linear algebraic equations will be obtained.

$$\begin{cases} a_0 m + a_1 \sum_{j=1}^{m} x_{1j} + a_2 \sum_{j=1}^{m} x_{2j} + \ldots + a_d \sum_{j=1}^{m} x_{dj} = \sum_{j=1}^{m} y_j; \\ a_0 \sum_{j=1}^{m} x_{1j} + a_1 \sum_{j=1}^{m} x_{1j}^2 + a_2 \sum_{j=1}^{m} x_{1j} x_{2j} + \ldots + a_d \sum_{j=1}^{m} x_{1j} x_{dj} = \sum_{j=1}^{m} x_{1j} y_j; \\ \ldots \\ a_0 \sum_{j=1}^{m} x_{dj} + a_1 \sum_{j=1}^{m} x_{dj} x_{1j} + a_2 \sum_{j=1}^{m} x_{dj} x_{2j} + \ldots + a_d \sum_{j=1}^{m} x_{dj}^2 = \sum_{j=1}^{m} x_{dj} y_j. \end{cases} \tag{4}$$

The obtained system $d+1$ of equations with $d+1$ unknowns $a_0, a_1, \ldots, a_d$ can be solved by methods of linear algebra. For many equations would be best to use the method of choice Gauss main element. Since the matrix of the system of linear equations is symmetric, it is always a solution, and the only one. If the number of equations is small, then can be successfully used the inverse matrix method to solve the problem.

Step 3. Activity check of received model. To do this need to calculate:

– remnants of the model as the differences between the observed and estimated values:

$$u_h = y_h - \hat{y}h = y_h - (a_0 + a_1 x_{1h} + a_2 x_{2h} + \ldots + a_d x_{dh}), \; h = 1, 2, \ldots, m; \tag{5}$$

– relative error of the residues and its average value:

$$\delta_h = \frac{u_h}{y_h} \cdot 100\%, \; \delta = \frac{\sum_{h=1}^{m} \delta_h}{m}; \tag{6}$$

– RMS error variance disturbances:

$$\sigma_u = \sqrt{\frac{\sum_{h=1}^{m} u_h^2}{m - d - 1}}; \tag{7}$$

– determination factor:

$$R^2 = 1 - \frac{\sum_{h=1}^{m} u_h^2}{\sum_{h=1}^{m} (y_h - \bar{y})^2} \text{ or } R^2 = 1 - \frac{\sum_{h=1}^{m} (y_h - \hat{y})^2}{\sum_{h=1}^{m} (y_h - \bar{y})^2}; \quad (8)$$

– the coefficient of multiple correlation, which is the main indicator of the correlation density of a generalized indicator with factors:

$$R = \sqrt{1 - \frac{\sum_{h=1}^{m} (y_h - \hat{y}_h)^2}{\sum_{h=1}^{m} (y_h - \bar{y})^2}}. \quad (9)$$

All values of the coefficient of correlation R belong to the interval from -1 to 1. The sign of the coefficient shows the "direction" of the connection: the positive value indicates a "direct" connection, the negative value - about the "reverse" connection, and the value "0" - the absence of linear correlation communication. With R=1 or R=-1 system has functional link between the signs. The multiplicity of the correlation coefficient is the main characteristic of the tightness of the link between the resultant sign and the combination of factors.

Step 4. Checking the statistical significance of the results. Testing is carried out using Fisher statistics with d and $(m - d - 1)$ degrees of freedom:

$$F = \frac{\frac{\sum_{h=1}^{m} (\hat{y}_h - \bar{y})^2}{d}}{\frac{\sum_{h=1}^{m} (y_h - \hat{y}_h)^2}{m - d - 1}} \text{ or } F = \frac{R^2}{1 - R^2} \cdot \frac{m - d - 1}{d}, \quad (10)$$

where d – the number of factors included in the model; m – total number; $\hat{y}_h$ – estimated value of the dependent variable at *h*-th observation; $\bar{y}$ – the average value of the dependent variable; y_h – the value of the dependent variable at *h*-th observation; R – coefficient of multiple correlation.

According to Fisher's tables critical value F_{kp} at d and $(m - d - 1)$ degrees of freedom. If $F > F_{kp}$, it is means about adequacy of the constructed model. If the model is not adequate then it is necessary to return to the stage of constructing the model and possibly introduce additional factors or switch to a nonlinear model.

Step 5. Check significance of regression coefficients. Testing is carried out using t-statistics that parameters for multivariate regression is:

$$t_h = \frac{a_h}{\sigma_{a_h}^2}, \quad (11)$$

where σ_{a_h} – standard deviation assessment of *h* parameter.

If the value of t_h exceeds the critical value, which is based on the tables of the t-criterion of the Student, then the corresponding parameter is statistically significant and has a significant impact on the aggregate indicator.

Step 6. Calculate the elasticity factor. Differences in the units of measurement of factors are eliminated by using partial elasticity factors, which are given by the ratio:

$$\varepsilon_h = \frac{\partial \hat{y}}{\partial x_h} \cdot \frac{\overline{x}_h}{\overline{y}}, \tag{12}$$

where x_h – average value of *h*-th parameter; $\overline{y}$ – the average value of effective signs. Partial elasticity coefficient indicates the percentage change in average productive sign of a change of 1% factor for fixed values of other parameters.

Step 7. Determination of confidence intervals for regression parameters.

Confidence interval at reliability level (1-α) is an interval with randomly defined limits with confidence level (1-α) Overstate the true value of the coefficient of the regression equation a_h and has the following form:

$$(a_h - t_{\alpha/2,z}\sigma_{ah}^2; a_h + t_{\alpha/2,z}\sigma_{ah}^2), \tag{13}$$

where $t_{\alpha/2,z}$ – Student`s statistics with $z = m - d - 1$ degrees of freedom and levels of significance α; σ_{ah}^2 – average square deviation of estimation parameter a_h .

Suppose system has *s* random variables $x_1, x_2, \ldots, x_r, \ldots, x_s$ (investigated parameters) represented by samples by *v* values $x_r = \{x_{r1}, x_{r2}, \ldots, x_{rz}, \ldots, x_{rv}\}$ For each pair of random variables x_r and x_w the equation can estimate the value of the empirical coefficient of linear correlation r_{rw} . The obtained coefficients are written into the matrix size $s \times s$:

$$\begin{pmatrix} 1 & r_{12} & \ldots & r_{1w} & \ldots & r_{1s} \\ r_{21} & 1 & \ldots & r_{2w} & \ldots & r_{2s} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{r1} & r_{r2} & \ldots & 1 & \ldots & r_{rs} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{s1} & r_{s2} & \ldots & r_{sw} & \ldots & 1 \end{pmatrix}. \tag{14}$$

All correlation coefficient r belong to the interval from -1 to 1. The sign of the coefficient shows the "direction" of the connection: the positive value indicates a "direct" connection, the negative value - about the "reverse" connection, and the value «0» - the absence of linear correlation communication. With R=1 or R=–1 system has functional link between the signs. The multiplicity of the correlation coefficient is the main characteristic of the tightness of the link between the resultant sign and the combination of factors. [5].

So, using the above calculation procedure of multiple regression analysis it is possible to evaluate the degree of influence on the researched result indicator PI_1 each of the factors introduced into the model PI_2, PI_3,..., PI_p and identify a set of key performance indicators:

$$KPI = \{\bigcup_{aw=1}^{av} KPI_{aw}\} = \{KPI_1, KPI_2, ..., KPI_{av}\}, \tag{15}$$

where $KPI_{aw} \subseteq KPI$, $(aw = \overline{1, av})$, av – number of KPI.

Stage 3 – Construct the indicators panel and visualize the KPI and E dependencies.

The next stage of this developing method is constricting the indicators panel, which will help with monitoring and CSIRT performance management. The indicators panel is tool for visualization and information analysis about business processes and their effectiveness. The data displayed on the panel indicators usually looks in the ***KPI*** form. Panel indicator system may be part of a corporate information system or act as a standalone application [6]. Using the indicator panel will present the data in a convenient form – diagrams, charts and data charts. For each organization, depending on its operational, planning and strategic goals, this panel is made individually.

So, the method for assessing the effectiveness of the CSIRT is developed in this paper, which is due to determining the performance of the CSIRT, the allocation of key performance indicators among of them, using a multi-factor correlation-regression analysis in construction of indicators panel and visualization of KPI and E dependencies gives an opportunity to audit the activities of the CSIRT and other centers of information and telecommunication systems maintenance. This method and the tools based on it will be useful to the incident response centers managers for monitoring, analyzing, assessing and managing the effectiveness of the CSIRT. Since the method is universal and can be applied to any company or government agency, in order to increase both the level of information security and the efficiency of the employee, department and organization.

REFERENCES

1. JAN VAN BON. IT Service Management / translate from English by «IT Expert», 2003.
2. GNATYUK V.: Analyze definitions of the term «incident» and its interpretation in cyberspace, V. Gnatyuk, Information security. №3 (19). – 2013. 175-180.
3. GNATYUK V. Basic indicators of the effectiveness of teams responding to cyber incidents, V. Kinzeryavyy, V. Gnatyuk, Information security. – № 20, №2. – 2014. 193-196.
4. MARMOZA A.: Theory of statistic, A. Marmoza, The textbook for university students. – K. 2013. – P. 333-397.
5. SIZOVA T.M.: Statistics: Tutorial. – SPb .: SPb GUITMO, 2005. – 80 p.
6. Performance Dashboards, Wayne W. Eckerson; Translate from English. Al′pyna Business Books, 2007.

Anna ROMANOVA[1]

Opiekun naukowy: Sergiy TOLIUPA[2]

PERSPECTIVE STEGANOGRAPHIC SOLUTIONS AND THEIR APPLICATION

Summary: Steganography is widely used to hide secret data. However, not all of its methods receive equal attention. Steganographical solutions, that lack either theoretical basis or practical implementation are analyzed. Ways of their possible practical application are suggested.

Keywords: steganography, cryptography, steganoanalysis, key, concealed information

PERSPEKTYWICZNE ROZWIĄZANIA STEGANOGRAFICZNE ORAZ ICH ZASTOSOWANIA

Streszczenie: Steganografia jest szeroko stosowana do ukrywania sekretnych/tajnych danych. Jednakże, nie wszystkie tego typu metody cieszą się tym samym zainteresowaniem. W pracy analizowano rozwiązania steganograficzne, którym albo brakuje podstaw teoretycznych albo nie są implementowane praktycznie. Zaproponowano sposoby ich możliwych praktycznych zastosowań.

Słowa kluczowe: steganografia, kryptografia, analiza steganograficzna, klucz, informacja usunięta

1. Introduction

Nowadays, as the role of information technologies in our lives is growing and growing with every passing minute, there is no question about the fact that information security measures are an issue of the highest importance. It would not be wrong to say that there is no ‘zero-day threat’ anymore; it is more like ‘a threat of a zero second’. Every modern technology automatically causes several new vulnerabilities and threats to come to existence, which, on its part, makes security specialists work their best to counter such efforts. Thus, there are quite a lot of information security solutions at hand already.

[1] Taras Shevchenko National University of Kyiv, the Faculty of Information Technology, Information Security Management, email: anitraromanova@gmail.com

[2] Full prof, Taras Shevchenko National University of Kyiv, the Faculty of Information Technology, tolupa@i.ua

Cryptography is doubtlessly one of the most effective, developed and approbated methods to be used when it comes to the protection of information resources. Nevertheless, it might be more effective to hide the communication channel itself instead of making unreadable the information within it. The practice of concealing data within text- or media-file is called steganography and has its roots deep in the history of the humankind.

A considerable number of steganographic methods are well-known and implemented in various steganosystems and applications. There are some methods of information concealment, though, that receive the attention not so much. The reasons of such lack of popularity can differ depending on specific solutions: their complexity for one, low cost-effectiveness of their realization for another. In any instance, they are either poorly described or are not widely used regardless of their perspectiveness.

The purpose of this article is to conduct an analysis and suggest possible practical use of steganographic solutions that are known and can be applied to a variety of information security systems, and yet lack either theoretical basis or practical application.

2. Steganography as a means of hiding information

2.1 Basic terminology

Steganography is an art and science of storing and transferring secret messages within covert channels that are based on and created inside open channels in such a way that the cover data is perceived as if not having any embedded messages for its unplanned recipients.

The main concepts are:

- Container b (also: carrier) is open data used to conceal secret information;
- Message m (also: payload) is secret information to be concealed;
- Key k is secret information that is known only to a legitimate user and defines a specific concealment algorithm;
- Empty container c (also: unmodified container) is a container devoid of any secret data; it is a sequence of l_c-long elements;
- Modified container s (also: package, steganogram) is the one that contains a secret message;
- Steganographic algorithm means two transforms, a direct $F: M \times B \times K \rightarrow B$ and an inverse one $F^{-1}: B \times K \rightarrow$ □;
- Steganographic system (also: steganosystem) is a totality of messages, secret keys, containers and transforms that connect them [1, 3].

Most steganography methods are based on two key principles:

- Human senses cannot distinguish slight changes in colour, shape and sound perception;
- Consequently, there are files that do not demand absolute preciseness and therefore can be modified without losing their functional value.

As a result, said methods imply allocation of insignificant fragments of the container and replacement of the information within them with information that needs to be hidden.

Finally, the process of encoded steganogram detection is called *steganoanalysis*.

2.2 Popular steganographic solutions

In this section the brief overview of widely used steganographic solutions is presented. Mostly, steganography uses the data concealment within digital images and audio files, less so video files and text. Electronic communications may also include hiding data inside of a transport layer (program or protocol) [4].
Starting with non-digital methods, physical steganography technics cannot be omitted. They have been developing for centuries and include, for example, blinking one's eyes in Morse code to spell a secret message [5].
Another example is adding tiny yellow dots to each page while printing a document. They are not detectable by the bare eye and contain the model, serial number and timestamps. This information cannot be obtained from a computer file and is embedded in a printout using dot-matrix code. The technology is used by many brand color laser printers, such as Xerox and Hewlett-Packard for traceability reasons [6].
Methods of embedding data within an image container are [2, 3]:

- Least Significant Bit method (LSB) (Sequential Insertion) is the most popular steganographic method. The least siginficant bit of each pixel is in fact a noise. If it is changed, the difference in the image will not be noticed by a human eye. Thus, these bits can be replaced with the bits of a secret message.
- LSB Psuedo Random Insertion. In contrast to the previous method, in which every changed data bit follows the next, this method uses pseudo random distribution of the secret message bits through the container. Thus, the interval between two bits is pseudo-randomly defined, which complicates both visual and statistical attacks, as well as extraction of all the hidden bits.
- Block hiding method. The container is split into disjoint blocks; for each of them a parity bit is calculated. One secret bit is concealed within one block. If the parity bit does not equal the respective secret bit, then one of the LSB in the block is inverted, so that the parity and the secret bits are the same.
- Palette permutation. Any colour palette consists of pairs of indexes. Each pixel of the image corresponds to a certain index in the table. The sequence of colours in the palette is not important, so it is possible to conceal a covert message by changing this sequence.
- Koch-Zhao (Relative DCT (Discrete Cosine Transform) values change method). Initial image is split into blocks of 8x8 pixels. As the result of applying DCT to every block a table of DCT coefficients is formed. Every secret bit is hidden in a separate block. Frequencies quantization causes some rate of distortion in the image, which is still not noticeable by the human eye.
- Benham-Memon-Yeo-Yeung method. Optimized version of the previous method. Firstly, only the most suitable blocks are used. Secondly, three DCT coefficients are selected instead of two, which decreases destortion in the container.
- Fridrich method implies a cascade embedment in low- and high-frequency DCT coefficients.
- Spread-Spectrum method consits of three possible variants:
    - The used frequency band is much wider than needed;
    - Spectrum is expanded by using a special independed (also: code) signal. The signal energy is distributed through all frequency bands, which makes the signal noise immune;

  - Restoration of the initial information is carried out by comparing the received signal and a synchronized copy of the code signal.
- Embedding pictures within video-files [5].

Audio steganography [3]:

- LSB-method for audio-files is the same as for images, but working with the audio-file format. It causes considerable destortions in the container.
- Phase coding method implies the substitution of the initial sound segment phase with the reference phase, which is the data to be concealed. Phases of adjacent segments are agreed to preserve the difference phase between them.
- Echo-signal use. Data is embedded in the container by injecting an echo-signal in it. Three echo-signal parameters are changed: initial amplitude, attenuation and shear rate. The echo-signal is perceived only as an additional resonance [7].

Linguistic steganography [3]:

- Random interval methods. Changing the number of spaces in the end of the text string does not cause significant changes in the meaning of the sentence. What is more, an average reader is unlikely to detect insignificant space modifications:
  - Changing the interval between sentences. One or two additional spaces are added after the sentence.
  - Changing the number of spaces in the end of text lines. Spaces are added according to the secret bit to be hidden. Two spaces encode one bit a line, four spaces – two bits etcetera.
  - Changing the number of spaces between words in a flattened text. When the text is width aligned, spaces between words are not of the same length and some of them can be used to hide data.
- Making the text of the same colour as the background [5];
- Using similarly looking Unicode and ASCII characters [4, 8];
- Using non-printable Unicode characters [8];
- Creating a pattern of deliberate errors and/or marked corrections [4].

Some other methods:

- Converting a file so that it has the statistical characteristics of another one [4];
- Format steganography;
- Blog-steganography. Secret data is added as commentary pin boards on social networks [5].

Finally, there are different software applications that use the methods of steganographic concealment mentioned above:

- Using LSB-method: OutGuess, JSTEG, JPHS, Hide-and-Seek, Steganos, Steghide, DC-Stegano;
- Using the palette permutation: Gifshuffle;
- JPEG format: OutGuess, JSTEG, JPHS;
- GIF format: Gifshuffle, Hide-and-Seek;
- BMP format: Steganos, Steghide;
- PCX format: DC-Stegano;
- LSB-method in audio-files: Invisible secrets, Hide4PGP, Steghide, StegoWav, Steghan, S-Tools;
- Using parity of quantization of frequency coefficients: MP3Stego;
- Using incorrect frames in a compressed stream: UnderMP3Cover [9].

3. Perspective steganographic solutions and their application

3.1. Internet of Things and cyber-physical systems

A cyber-physical system is a mechanism that is controlled or monitored by computer-based algorithms, tightly integrated with the Internet and its users. Examples of CPS are autonomous automobile systems, medical monitoring, smart grids, automatic pilot avionics etc [10].
The Internet of Things (IoT) is the network of physical devices, vehicles and other items embedded with electronics, sensors, software and network connectivity, which enable them to collect and exchange data. It is more or less an instance of a class of cyber-physical systems [11].
The network steganography uses communication protocols' control elements and their functionality to hide information inside [12]. The modifications can be carried out either over a single network protocol (applied to the Protocol Data Unit, the time relations between PDUs or both) or to several protocols at the same time (inter-protocol steganography). Such network steganography methods can be applied to the systems mentioned above, too. The IoT is believed to be a phenomenon that will expand its influence greatly within the next few years. As a perspective network instance it requires thorough attention of steganography specialists. Information circulates within it the same or the fairly similar way as in any other system. Thus, optimal and the most suitable methods of hiding data in communication protocols should be developed specifically for the IoT.
What is more, as the items within the IoT possess a vast variety of sensors and software, they can be used to conceal data in. For example, covert messages can be stored in unused registers of the CPS/IoT components or in the states of their actuators [12].

3.2. The use of stream containers

As mentioned above, by the type of access to the data one can distinguish fixed and stream containers [3]. All the methods mentioned in Chapter 2.3 use the first ones to conceal information in. Such a container is a constant pre-defined sequence of bits that are displayed before a steganographer all at once. To the contrary, a *stream container* is a sequence of bits that are continuously changing, as in a phone conversation. A message is embedded in real time so that the final size of the container is never known beforehand. The intervals between the embedded bits are generated by a pseudorandom sequence (PRS) generator and uniformly distributed between readouts [3].
There is hardly a couple of scientific works devoted to this type of steganography, let alone examples of its real-life practical implementation. Despite any reasons, it can be successfully applied as an efficient means of information security. There is a number of solutions for encrypted secure real-time communication. However, what if we could, for example, make a confidential phone conversation not only indecipherable but also seem to be an innocent chat? A stenographic telephone set-top box could be a solution. The same concerns video-conferences. An extraneous observer would only see an average conversation not having any access to the real audio, video or any other embedded data.

The unpopularity of the stream-container steganography can be explained by defining major issues concerning its use. First and foremost, it is never known whether the size of the container will be enough to conceal the whole message as the length of the first (and likely of the latter, as well) is undefined. The same property creates and advantage as one carrier file can be capacious enough to contain several messages. In any case, the secret data has to be somehow synchronized with the container, thus one of the biggest questions is how to define the beginning and the end of the embedded sequence within the container. The problem becomes more serious concerning video communication. The solution would be of extreme complexity, as we would need to synchronize the image-image stream (both open and covert), the sound-sound stream and image and sound respectively.

The solution may lie in using special built-in libraries. They would consist of structured groups of words of the same length, which would in ideal case possess pronunciation similarities. Such groups should then be grouped in semantic dictionaries, so that they would form simple, but logically and semantically structured sentences. The linguistic means for this are well-developed and are similar to those of forming synonymic dictionaries and machine translation applications. The words and sentences could then be synchronized with the container using synchronization bits, package headers and/or other means of dividing encapsulated data; the covert message can be embedded after them and be synchronized using the initial properties of the container.

The possible situations with video communication would be more complex. If only the content of a given conversation is confidential, then the issue is just to steganographically encrypt the sound and synchronize it with the real video image. On the other hand, if the identities of conversation participants are also a secret, then other methods should be provided. It is not necessary for a steganographic solution to be all-purpose. It is possible to design a system consisting of a cryptographic and a steganographic modules and providing different scenarios according to the situation.

The biggest remaining problem is a significant delay which is unacceptable in real-time conversations. Then again, there are numerous solutions in cryptography in this field, that can be adapted to the task.

3.3. Semantic and syntactic methods

These two classes of methods belong to steganography with text containers. Instead of using digital format features, they work with the language itself. It is an advantage in comparison to the first type. As an average reader may not be aware of the covert message existing in an open text, a text editor may automatically change the number of spaces or conduct other actions that would ruin embedded data [3]. In fact, any reformatting will lead to the same result.

Syntactic and semantic methods, though, do not use the presentation of text, but work with the text itself. The first type uses the fact that in most languages there are some optional rules of punctuation and grammar forms. Any given language sticks to specific rules, but is still not so solid of a structure, which presents a great number of linguistic possibilities. For example, in the Ukrainian language a colon and a dash can replace each other in some cases. This can be used to encode bit of secret message: “0” for one punctuation mark and “1” for the other one. A more complicated method

could be using grammatical similarities in different sentence constructions, such as changing the sequence of some words.
An example of semantic steganography is using the table of synonyms to encode the secret bits. If there are two of them, say, 'however' and 'but', then again one of them can mean "0" and the other "1". If there are more synonyms, possibly context ones, then 4 words can encode 2 bits, 6 words 3 bits of information etcetera. The average data transfer speed when using these methods is several bits per kilobyte [3].
The main problems with such linguistic methods are obvious. First of all, they are very language-dependent. Secondly, they require large amounts of initial text as a container, which is not exactly effective. Finally, even if some punctuation rules are ambiguous, their deliberate and controversial usage can be detected by a censor/editor.
It could be wise to suggest the usage of more complex methods of language-based steganography. Every language can be analyzed to create special tables of syntactic correspondence. For example, for the English language and other Germanic languages the use of active and passive voice, as well as of complex object and complex subject is optional. Sentences can be easily and naturally transformed using equivalent constructions that most likely will not raise suspicion. The advantages of such solution are numerous. One of them is high resistance to various attacks (they are here similar to one-time pads). Another is that the concealment capacity is much higher than that of basic semantic methods. The only question is an algorithm of selecting initial text material. It is likely the best option is to create special libraries of texts, sorted by the topic. This way many fields of interest may be covered so that the covert message is not detected.
A creation of a multi-purpose linguistic steganography complex is suggested. It will doubtlessly require linguistic work of high quality and profoundness. An optimal approach is to be found to, if possible, reduce the language-dependency of each solution. In other case, such an application will have to be designed according to a separate language or, at least, a group of languages with the same paradigm. Thus, the task at hand is to group the languages within each family by the similar tendencies in grammar usage. The next step is to create tables of correspondence for grammatical constructions and stylistic expressions that can be interchanged. Finally, text material libraries are required to provide unobtrusive containers with as much options described in the tables as possible.

3.4. Biochemical Steganography

Most modern steganographic systems use only digital containers, such as files of various nature, binary sequences etc [9]. However, there are other fields of interest for steganography, as the environment can provide a considerable variety of non-digital containers.
We are surrounded by billions of organisms, every cell of which contain DNA-molecules. They are the central repository of information in the cell [13]. Biological computing and quantum computing are believed to be the two most promising technologies under development right now [14]. And as cryptography now mostly works with factorization problems, which makes the messages subjects to attack by quantum computers, biochemistry presents the whole new sphere of potential information security solutions.

DNA-steganography is a process of camouflaging a DNA-encoded message within the enourmous complexity of genomic DNAs [13]. Due to the DNA-code variability among different species, an organism, selected at hazard, possesses random DNA-code. Such a characteristic makes these molecules potentially good containers. Another doubtless advantage is their tiny size, as huge amounts of information can be encoded within a container that cannot even be seen by a human eye without proper amplification. DNA is also a quite solid structure and one highly resistant to possible biochemical attacks.

Nevertheless, despite obvious perspectives of using DNA as a means of biochemical steganography, most of the attention has been received by DNA-cryptography so far. A DNA-molecule is a sequence of four nucleotides – Adenine (A), Cytosine (C), Guanine (G) and Thymine (T), that are grouped in triplexes, so called codones, and form two anti-parallel strands [9, 14]. Complementary DNA strands can self-assemble by forming hydrogen bonds between bases (base pairing) of each strand specifically with A bonding only to T and G only bonding to C [10]. Four different bases mean 4n possible different n-meres that encode genetic information through a number of aminoacids [6, 8]. Another macromolecule to be potentially used is RNA. It is similar to a DNA-molecule with an exception that the base structure is a different 5-atom sugar – ribose, and uracil (U) corresponds to thymine [9].

So, how do we encode information within a biomolecular structure? A message can be encrypted in a DNA strand, every symbol being encoded by a codon defined in the specially designed table (the technic resembles the use of one-time-pads). For example, 'A' may be encoded by a CGA sequence, 'B' by CCA and so on [14]. The secret message is then presented as a sequence of codones. Some of the aminoacids are presented: alanine – GCT, GCC, GCA, GCG; asparagine – GAT, GAC; fenilalanine – TTT, TTC [9].

Then the strand is flanked by polymerase chain reaction (PCR) primer sequences and hidden by mixing it within many other additional "distracter" DNA strands [10, 9]. Polymerase Chain Reaction (PCR) is a process, during which PCR primers become complimentary to the F and R primer "keys" in Secret Message DNA [13]. They are then hidden in a microdot [14]. Knowing the secret key and the primer sequences, a user can extract the strand using known DNA separation methods (hybridization with the complements of the "secret key" strands might be placed in solid support on magnetic beads or on a prepared surface; may be combined with amplification steps and/or PCR [15]) and read the message.

A problem with such approach is that the probability profile of aminoacids in nature is not the same as that of a secret message [9]. The secret "tags" have to be indistinguishable from "distracter" DNA strands and the entropy has to be as in any DNA-molecule – between 1.2 and 2 [15]. This creates the need to use models of real DNA-molecules along with some other solutions. One of the enhancement technics suggested recently is the use of sequencing (determining the sequence of nucleotides in a DNA-fragment). There are a lot of sequenced genomes provided in open arrays already. Some of them are 55 genomes of bacteries, a yeast genome and those of other standart laboratory objects [6]. Another techinc is to construct the "distracter" strands so that their distribution mimics the plaintext source distribution. One of the easiest possible ways to do so is to synthesize a DNA-molecule that depends only on the plaintext and the secret key [9, 15]. The compression of the plaintext is also possible. If the resulting distribution

of the plaintext approximates a universal distribution, then a random distracter sequence may suffice to provide security needed [15]. The use of a substitute random combination of sequenced genomes (from exotic organisms, for example) may be an enhancement solution, as well [13]

There is a work [16] devoted to the use of run-length encoding (RLE) systems in biochemical steganography, though it is stated, that their application in practice still provides more questiones than answers.

Given everything stated above, any possible attack on a DNA-based steganographihc system would not be successful if it is purely computational [13]. The ways of resisting biochemical attacks, though, are an important question to pay attention to in the future development of DNA-steganography.

Current DNA-steganography technology is still in a period of laboratory exploration and focused on experiments [14]. A possible explanation of the lack of expected activity in the field is that it is a multidisciplinary area which demands knowledge in both biology and cryptography and so requires researchers from both areas to work in a new cooperative way [14]. Possible spheres to implement these technics in are negotiable instrument anti-forgery, personnel identity and access control, anti-theft marking and product authentication [17]. All of them are instruments of securing business profit and are thus attractive for service and production. Only a few examples of using DNA-steganography as a sort of watermark are know. In 2000, during the Olympics in Sydney the Australian Olympic Committee used the DNA based tracking technology to protect Sydney Olympic licensed merchandise from counterfeiting [17].

4. Conclusion

A variety of steganographic solutions was analyzed. Among them such methods were selected that are not either widely used in software applications or lack attention in general. Nevertheless, their perspective usage was discussed. Taking into account all the fact mentioned above, the next directions of development of steganography are suggested:

1) Steganography in cyber-physical systems and the Internet of Things in particular;
2) The use of stream containers;
3) Semantic and syntactic methods;
4)Biochemical steganography practical application.

Information is surely becoming an asset of the highest value. Seeing as the cyberspace is more of a battlefield for different forces continuously confronting each other, it is obvious that information security sphere requires the best solutions possible. Steganography has proven to be an effective means of secret data concealment ensured with centuries of practical use. And, just as any other science, it is in the state of constant development. Being aware of perspective ways to use its methods for our cause, we get access to numerous up-to-date possibilities of providing information security of the highest level.

REFERENCE

1. ЗОРИН Е.Л., ЧИЧВАРИН Н.В.: Стеганография в САПР. Учебное пособие. МГТУ им. Н.Э. Баумана, Москва (pdf).

2. ALEXANDRE MIGUEL FERREIRA: An Overview on Hiding and Detecting Stego-data in Video Streams. University of Amsterdam, System & Network Engineering – Research Project II, March 23 2015.
3. KONAKHOVICH G. F., PUZYRENKO A. YU.: Computer steganography. Theory and practice with Mathcad (Rus). MK-Press Kyiv, Ukraine 2006.
4. FRIDRICH, JESSICA, M. GOLJAN, D. SOUKAL: Searching for the Stego Key. Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VI 2004 (pdf): *http://www.ws.binghamton.edu/fridrich/Research/Keysearch_SPIE.pdf.*
5. CHRISTOPHER LEAGUE: An overview of digital steganography, particularly within images, for the computationally curious. Long Island University 2015: *https://www.youtube.com/watch?v=-7FBPgQDX5o.*
6. Secret Code in Color Printers Lets Government Track You; Tiny Dots Show Where and When You Made Your Print. Electronic Frontier Foundation October 2005: *https://www.eff.org/press/archives/2005/10/16.*
7. Echo Data Hiding (html): http://www.slidefinder.net/a/audio_steganography_echo_data_hiding/24367218
8. AKBAS E. ALI: A New Text Steganography Method by Using Non-Printing Unicode Characters. Eng& Tech. Journal, 28 (1) 2010 (pdf): *http://www.uotechnology.edu.iq/tec_magaz/volume282010/No.1.2010/researches/Text%287%29.pdf.*
9. АГРАНОВСКИЙ А.В., БАЛАКИН А.В., ГРИБУНИН В.Г., САПОЖНИКОВ С.А.: Стеганография, цифровые водяные знаки и стеганоанализ. Москва: Вузовская книга 2009.
10. Cyber-Physical system: *https://en.wikipedia.org/wiki/Cyber-physical_system.*
11. Internet of Things: *https://en.wikipedia.org/wiki/Internet_of_things.*
12. STEVEN J. MURDOCH, STEPHEN LEWIS: Embedding Covert Channels into TCP/IP. University of Cambridge, Computer Laboratory (pdf): *http://www.cl.cam.ac.uk/users/fsjm217, srl32g/.*
13. CARTER BANCROFT, PH.D.: DNA-Based Technologies: Computation, Steganography, Nanotechnology. Talk at Material Science and Engineering, Stony Brook University, April 2011.
14. ADITIT SHARMA: Security and Information Hiding based on DNA Steganography. International Journal of Computer Science and Mobile Computing, Vol. 5, March 2016: *www.ijcsmc.com.*
15. ASHISH GEHANI, THOMAS H. LABEAN, JOHN H. REIF: DNA-based Cryptography. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Volume 54, 2000 (pdf).
16. TOMONORI KAWANO: Run-length encoding graphic rules, biochemically editable designs and steganographical numeric data embedment for DNA-based cryptographical coding system. Commun Inteqr Biol. March 2013: *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3609851/.*
17. WENDELL M. SMITH: DNA Steganography for Security Marking. 5th World Product and Image Security Convention, PISEC '03, Czech Republic. Technology Transfer Group: *http://www.polestarltd.com/ttg/isspeeches/pisec03/index.html.*

Mykola ROMANYUKOV[1]

Scientific Supervisor: Vladimir KONONOVICH[2]

OGÓLNY MODEL OCENY SKUTECZNOŚCI OCHRONY SYSTEMÓW INFORMACYJNYCH

Streszczenie: Analizowane są dobrze znane modele optymalnych wydatków na ochronę systemu informacyjnego i określane są typowe funkcje dystrybucji zabezpieczeń obiektów informacji. Optymalne wydatki są ustalane z uwzględnieniem czynnika ludzkiego, ponieważ interakcja człowieka ze środowiskiem ma charakter nieliniowy, a reakcja człowieka jest proporcjonalna do logarytmu podrażnienia. Ustalono, że wybór takich funkcji dla konkretnego systemu informacyjnego jest specyficzny i wymaga w każdym przypadku obszernego badania eksperymentalnego.

Słowa kluczowe: optymalizacja kosztów, nieupoważniony dostęp, ochrona systemu informacyjnego

THE GENERAL MODEL FOR EVALUATING THE EFFECTIVENESS OF PROTECTION OF INFORMATION SYSTEMS

Summary: The analysis model best known costs to protect the information system and identified the typical distribution function of the safety of objects of information activity. The optimal ones are determined taking into account the human factor, since human interaction with the environment has a nonlinear character and the human response is proportional to the logarithm of irritation. It is determined that the choice of such functions for a particular information system is specific and requires a comprehensive experimental study in each case.

Keywords: cost optimality, unauthorized access, protection of the information system.

1. Formulation of the problem

Analyze known models of optimal expenses for the protection of the information system and determine the typical functions of distribution of security objects of information activities. Determine the best one considering the human factor as human

[1] Graduate student of informatics and security management information system Odessa national polytechnic university, nikolay.romanyukov@gmail.com

[2] Associate professor of informatics and security management information system Odessa national polytechnic university, vl_kononovich@ukr.net

interaction with the environment and non-linear response proportional to the logarithm of human irritation.

2. Analysis of recent research and publications

The classic work of managing information systems protection and finding economically justified security is the book [1]. However, in this paper, the dependence of the components of total costs on the construction of an effective security system and the dependence of the amount of damages for unauthorized access from the achieved level of security are not sufficiently researched. So the cost of compensation for violations when unauthorized access is considered linearly dependent on the level of security, which is quite doubtful.
A qualitative model of optimization of the organization's information system security system, based on the principle of «cost-efficiency minimization», described in his book Petrenko [1]. According to this model, security costs consist of expenses for precautionary measures, for monitoring the situation and restoring losses in the unauthorized access (Figure 1).

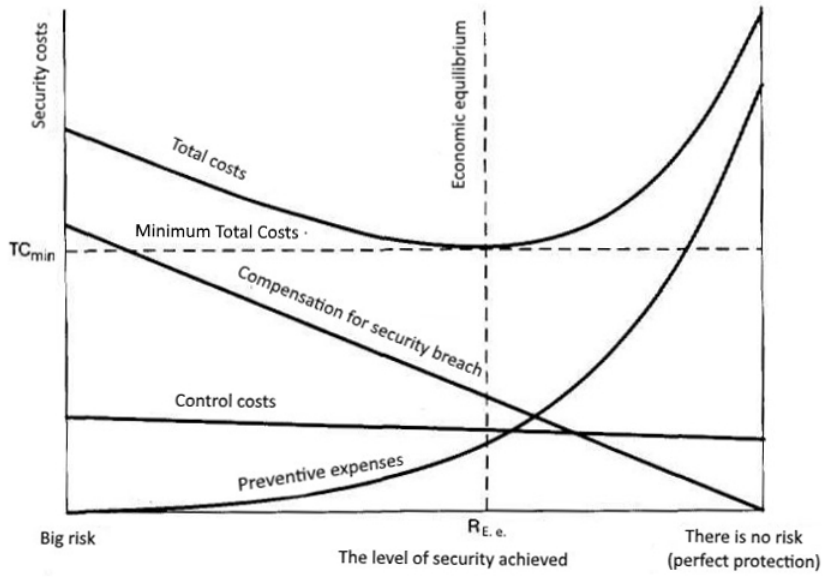


Figure 1. The link between security costs and a sufficient level of security

With the change in the level of security of the information environment, the values of total costs change, and accordingly their amount - the total cost of security. The following graph, which corresponds to the model of optimal costs for protecting the information environment of the enterprise, are constructed taking into account the following assumptions: firstly, when performing work on prevention of violations of the security policy, the work that gives the greatest effect on the protection of the information environment is first performed, and secondly the dependence and intensity of attacks does not change over time, while the economic balance does not change.

3. Statement of the main material

Apply this qualitative model for assessing the effectiveness of the mechanism of protection of information systems. To do this, we will carry out and substantiate the formalization of the constituent dependencies for the most typical cases. We will consider the dependence on the relative security level x, which varies from 0 to 1. This corresponds $x=0$ to an unprotected system, $x=1$ corresponding to a fairly high level of security of the system, $x<1$ means that absolutely protected systems in nature do not exist. The amount of component costs $y(x)$ is also relative and varies in the range from 0 to 1. In determining the ratio of security costs, a certain measurement base can be used. A typical measurement base may be the amount of resources and information environment of the organization.

In [2], it is shown that in the realization of threats the most important factor is the human factor. Human interaction with the environment has a non-linear character. Since the human response is proportional to the logarithm of irritation, then the human perception of the level of threats or the level of costs is also nonlinear. Therefore, mathematical formulas for typical dependencies should be sought among stable statistical distributions and dependencies.

The classical distribution is normal, which follows from the law of large numbers. In the theory of probabilities there exists a central limit theorem (Lyapunov's theorem) concerning the boundary laws of the distribution of the sum of random variables [3]. Thus, a candidate to formalize the relationship between the normalized value of damages for unauthorized access and security levels may be normal (Gaussian) distribution at the value of mathematical expectation *m=0*:

$$y_1 = \varphi(x) = \frac{k}{\sigma\sqrt{2x}} \cdot \exp(-\frac{x^2}{2\sigma^2}); \; x \in [0,1], \tag{1}$$

where k - the normalizing factor, σ - the mean square deviation.

Despite the widespread use of a normal law in science and technology, this distribution does not fit our dependencies. First, the direct proportional relationship between the magnitude of the damage with the non-level of security is more suitable for description than the Gauss curve. Secondly, this curve can not reflect the influence of the human factor due to its non-Gaussian distribution.

The second candidate for the formalization of the relationship between the normalized amount of damage from unauthorized access and the level of security may be an exponential function:

$$y_2(a, x) = \exp(-ax); \; x \in [0,1], \tag{2}$$

where $a>0$ is the parameter.

The exponential function has a remarkable property: the rate of its change is proportional to the value at this point:

$$\frac{dy_2(a, x)}{dx} = -a\exp(-ax). \tag{3}$$

In our case (Figure 2), the exponential function shows that, with increasing protection, losses decrease (curve 1). When x=1, $y(x) > 0$, that corresponds to the real situation. Therefore, at a high level of protection, there may be losses that require some financial cost to compensate. Exponential function is left for further research.

Recently, the theory of non-Gaussian processes, which may be the best candidates for the formalization of the considered dependencies, has become a significant development. Stationary distributions of values of variables are not basically Gaussian. In the case of large values of the variable, we have a hyperbolic distribution of Tsipfa-Pareto [4]:

$$y_3(\alpha, x) = \frac{c}{x^{1+\alpha}}, 0 < x_0 \le x \le \gamma; 0 < \alpha < \infty, \quad (4)$$

where, $y_3(x)$ - frequency, α - Zipf distribution index, which determines its form, the smaller α, the more long-tailed this distribution, c - the parameter that provides the condition for the normalization of the relative sample size x_0 - the minimum, γ- the maximum value x. The distribution of Tsipfa-Pareto in probability theory is a two-parameter family of absolutely continuous distributions that are power-generating. If the random variable is such that its distribution is given by equality $F_x(x) = P(X < x) = 1 - (\frac{x_m}{x})^k$; $x \in [x_m, \infty]; \forall x \ge x_m; x_m > 0; k > 0$. Then we can say that x has a distribution of Tsipfa-Pareto with parameters x_m and k. In our approach x - the random value relative to the level of security of the information system, accepts the value of the smaller $x \in (0,1]: X = 0$, corresponds to the unprotected system, X=1 corresponds to a high level of security of the system. Mathematical expectation of distribution $m = \frac{kx_m}{k-1}$, if $k > 1$; $x_m\sqrt{2}$ median, mod x_m variance $\left(\frac{x_m}{k-1}\right)^2 \frac{k}{k-2}$, if $k > 2$, probability density $\frac{kx_m^k}{x^{k+1}}$. For the distribution of Tsipfa-Pareto characterized by the lack of mathematical expectation: $m \to \infty$ at $x_m \to \infty$.

On the basis of the Gnidenko-Deblen theorem, it can be argued that non-Gaussian distributions are Cipsoff distributions with α<2. The use of Tsiph distribution to formalize the relationship between the normalized amount of damage and the level of protection is appropriate for the following reasons. Tsipf distribution is characterized by a sharp decline (steeper than the exponential) and a long tail (curve 2). A steep downturn corresponds to real data on the magnitude of losses at significant risks of unauthorized access, since they are more sensitive to the magnitude of risks, and smaller than in the case of formalization by an exponential function.

Indicative and parabolic functions can be used to formalize the relationship between the normalized cost of the precautionary measures and the level of protection [4]. The indicator function can be represented as an exponent:

$$y_4(b, x) = \exp[b(x-1)], x \in [0,1], \quad (5)$$

where b is the parameter. Function (5) is convenient for formalization because for $x = 0, y_4(b,0) = \frac{1}{\exp b} > 0$. However, it can not be used for our case (curve 3). From the graph, it is clear that even at a level of protection $x = 0{,}5$, the necessary costs for preventive measures far exceed the optimally acceptable $(y_4(b, 0{,}5) > 0{,}6)$.

The second candidate for the formalization of the relationship between the normalized amount of expenses for preventive measures and the level of protection can be a parabolic function:

$$y_5(b, x) = bx^2]; \quad x \in [0,1], \tag{6}$$

where b is the parameter. Parabola (curve 4), as the most simple nonlinear function, can be successfully used to formalize the relationship between the normalized cost of the precautionary measures and the level of protection of any information system.

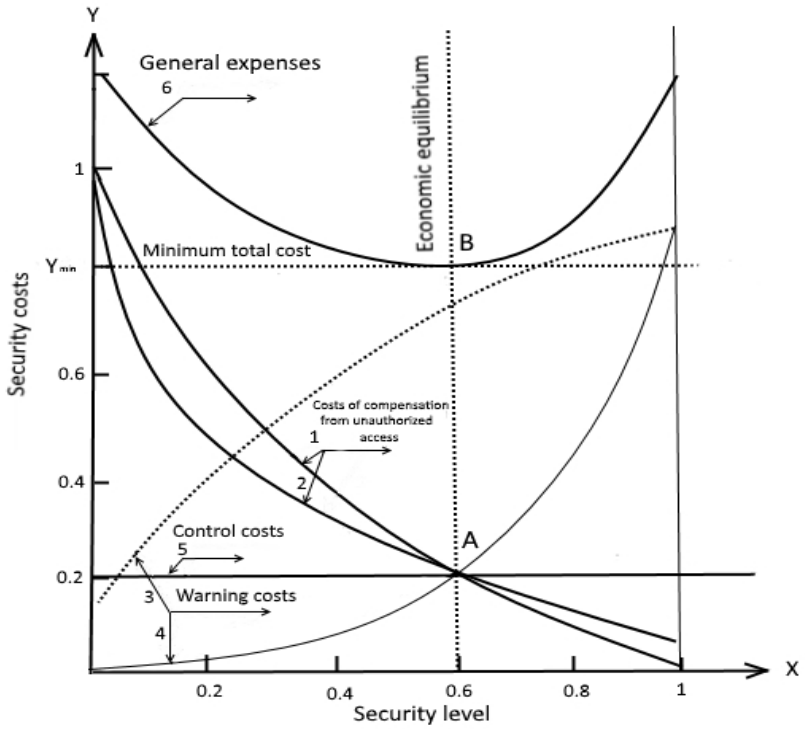


Figure 2. The general model of cost optimization for the protection of the information system.

The cost of monitoring the technical state of the information system is unchanged $y_6(x) = const$ (straight 5). Analyzing the previous costs of previous protection measures, compensating for the consequences of unauthorized access and controlling the functioning of the information system, we obtain the optimal total cost of protecting the information system (curve 6). To formulate the dependence of total

costs on the level of security of the information system, taking into account the considered component costs, you can use the analytical expression:

$$y_з(\beta, x) = \frac{d}{\sqrt{x} + e^{-\beta x}}, \; x \in [0,1], \tag{7}$$

where, d and b - parameters; $d \in [0,6; 1,3]$; $\beta \in [2,2; 4,5]$. The general model of the optimal cost of protecting the information system is presented in Figure 2.

Point A crossing cost curves 1, 2, 4, 5, corresponds to the level of economic equilibrium, ie component costs take the same values. The sum of the component costs for protection thus reaches the minimum value $y_з = 0,8$ (point B). The achieved security level corresponds to the value $x=0,6$. In this case $d = 0,7; \beta = 4$. With a decrease in the level of security, an increase in overall costs is observed mainly due to increased costs for possible unauthorized access to the information system. With a rise in the level of security, we also have an increase in overall costs, mainly due to increased costs for preventive measures.

An analysis of the general cost-optimization model for the protection of the information system allows us to set some confidence interval for the minimum total costs relative to point B. This interval can be chosen at the level of relative total costs $y_з = 0,82$. If we use the dependence established by us (7), then the optimal level of total costs corresponds to the level of security $x \in [0,4; 0,7]$.

4. Conclusions

An analysis of the optimal cost model for the protection of the information system shows that for the final choice of the typical functions of the distribution of security costs, a set of statistical data on these costs is required. The choice of such functions for a particular information system is specific and requires a comprehensive experimental study in each case.

REFERENCES

1. PETRENKO S.A., SIMONOV S.V.: Information Risk Management. Economically justified safety. Ai Ti Company; DMK Press Moscow, Information technology for engineers 2004. (in Russian).
2. KONONOVICH V.G., KONONOVICH I.V., STAIKUTSA S.V., CVILIY O.O.: DEFINITION OF IDENTITY IN THE SYSTEM OF SOCIAL AND INFORMATION SECURITY. MODERN PROTECTION OF INFORMATION, (2015)1, 19-27.
3. BARKOVSKY V.V., BARKOVSKAYA N.V., LOPATIN O.K.: Probability theory and mathematical statistics. CNL, Kiev 2010.
4. KILIN A.E., KONONOVICH V.G., KONONOVICH I.V., BEZZUBENKO V.V.: Model for evaluating the effectiveness of software code protection against unauthorized changes. Informatics and mathematical methods in modeling, 5(2015)3, 289-297.

Yanina SHESTAK[1], Stanislav MAHULA[2]

Opiekun naukowy: Vira VIALKOVA [3]

MATEMATYCZNY MODEL OCHRONY DANYCH PRZED CYBER-ATAKAMI W ROZPROSZONYCH SYSTEMACH INFORMACYJNO-TELEKOMUNIKACYJNYCH

Streszczenie: W ostatnich latach stało się oczywiste, że bardzo ważnym jest stworzenie/opracowanie skutecznego systemu zapewniającego cyber-bezpieczeństwo. Zatem, celem jest opracowanie systemu, który umożliwia ochronę informacji w jakimkolwiek rozproszonym systemie informacyjnym oraz telekomunikacyjnym – przed nowoczesnymi cyber-atakami. Ważność takiego narzędzia jest szczególnie podkreślona ze względu na ciągły rozwój systemów typu DITS ('distributed information and telecommunication systems', czyli systemów rozproszonych). W systemach tego typu wykonywane są różne działania polegające np. na gromadzeniu, przechowywaniu oraz przetwarzaniu ogromnych zasobów różnorodnych informacji [1]. Bezpieczne przechowywanie tychże informacji – z kolei – pozwala na wykonywanie dalsze czynności bez dodatkowego ryzyka oraz strat finansowych oraz strat typu intelektualnego.

Słowa kluczowe: rozproszone systemy informacyjne i telekomunikacyjne, model matematyczny, bezpieczeństwo informacji

THE MATHEMATICAL MODEL FOR PROTECTION OF DISTRIBUTED INFORMATION-TELECOMMUNICATION SYSTEMS FROM CYBER ATTACKS

Summary: It has become apparent in recent years, that the problem of creating an efficient cybersecurity system, which enables to protect the information of any complex distributed information and telecommunication systems from modern cyber-attacks is important. The importance of such measure in providing progressive development of DITS (distributed information and telecommunication systems) is particularly valuable because all aspects of RITS activity are directed at storing, accumulating and processing of a big amount of diverse information [1]. Reliable storage of this information allows maintaining further activities without additional financial and intellectual losses.

[1] Cybersecurity Ph.D, Taras Shevchenko National University of Kyiv, Department of Cybersecurity and Information Protection, lucenko.y@ukr.net

[2] Taras Shevchenko National University of Kyiv, Department of Cybersecurity and Information Protection, specialty: Information Security Management, Department Assistant, mahulchik@gmail.com

[3] Engineering Science As. Prof., Taras Shevchenko National University of Kyiv, Department of Cybersecurity and Information Protection, veravialkova@gmail.com

Keywords: distributed information and telecommunication systems, mathematical model, information security

1. Formulation of the problem

Admittedly, the protection methods depend on: the specific place, the scope of the object, the information that is being processed and the specific customers' needs. However, eventually, all data objects are meant to be protected from cyberattacks using various means and approaches.
The mathematical model for the protection (Fig. 1) is built based on various steps of criminal's attacks. These steps, in turn, have five main steps. The first stage, which can be described as the process of collecting information about the victim, that is, a detailed analysis of the situation takes place, IP-addresses, hosting services, security measures investigated etc. The third one is about when the attack occurred, the attacker broke the network and is looking for documents (files, folders, etc.) that he needs. This stage is special for a consistent effect on all segments of the network. An intruder penetrates into all segments gradually, by punching (destroying) them.
Following that, comes the fourth stage, when in the system, the attacker breaks (or partly) all its segments, installing malware. The final step is data transmission to remote computers, which controlled by a hacker. This stage is executed in order to have further obtained information used for its purposes (destruction, sale, copying, modification, etc.).

Figure 1. - Mathematical model of protection of DITS from cyber attacks.

Information security system (IC) of complex distributed information and telecommunication systems can be represented as a hierarchy of links to five levels.

The first level of the protection hierarchy stands for individual means of protection that parry specific methods of implementing the IS threats in the certain structural component of the local environment function part (finite systems, means of transmission, local communication channels).
The second level of the security system provides the resistance to a specific way of implementing the IS threat by all defense means of the subsystem for information environment..
The third level ensures safety to complex distributed information and telecommunication systems from all methods of implementing a threat in the functional part component of the local environment.
Then, on the fourth level, those systems are protected in the functional part component of the local environment from all threats to information security.

The fifth level security system ensures the security of complex distributed information and telecommunication systems provides security among all components of the information environment from the existing security threats in general.
This model is based on the principle of the importance of smoothly and evenly weight components of complex distributed information and telecommunication systems from all possible ways of realization of threats of violation of information security is equal to the maximum risk of a breach of information security among the components of complex distributed information and telecommunication systems in this segment, i.e. valuation is the most vulnerable component.
From a mathematical point of view, the model of DITS is a graph in the form:

$$DITS = (M_{CN} \ U \ M_{DW} \ U \ M_{SE}) \quad (1)$$

where DITS - distributed information-telecommunication system;
M_{CN} – a plurality of compute nodes DITS;
M_{DW}– a plurality of data warehouses;
M_{SE}– a plurality of switching network elements and exchange of physical data channels.

The model of the resource request is described by the expression:

$$RR = M_{VM} \ U \ M_{EL}, M_{VD}) \quad (2)$$

where RR – resource request;
M_{VM}– a plurality of virtual machines used by applications;
M_{EL} – a plurality of elements;
M_{VD}– a plurality of virtual data transfer channels between virtual machines and elements in the query.
The purpose of the resource request is formed as following:

$$A : RR \rightarrow \text{Data center} = \{M_{VM} \rightarrow M_{CN}, M_{EL} \rightarrow M_{DW}, M_{VD} \rightarrow M_{SE}\} \quad (2)$$

Thus, for the actual work required is the following:

- Each computing node must have a performance and total memory, which corresponds to the total composition of all the virtual machines thereto;
- Each virtual channel can be mapped onto a physical channel, provided that the total mass of the reflected virtual channels on a physical channel, is less than the nominal bandwidth of the transmission channel data;
- Each virtual channel can pass through the switching element, provided that a plurality of virtual channels passing through the switching element boundaries than the total bandwidth of this element (bytes/s);
- Each element of the total information space can be placed in the data warehouse, provided that each element and its type coincides with the type of data store in a shared set of conservation elements and does not exceed the amount of memory.

To minimize the information loss, by reducing risks and reducing possible external and internal influences, it is proposed to use mechanisms of optimization of virtual machine placement on physical servers, that is, to apply the principle of minimal fillings of servers: (4)

$$\min \sum_{i=1}^{n} y_i \tag{4}$$

$$y_i = \begin{cases} 1 - \text{the server has one virtual machine,} \\ 0 - \text{the server is not involved} \end{cases}$$

The practical implementation of this approach allows to minimize the level of both external and internal influences, as well as reduce the cost of maintaining common servers, in the case of identical technical characteristics.

The formation of the initial data analysis is built on the structure of individual DITS, based on the analysis carried out the calculation of the security metrics for individual resources:

$$Q_{pr}^{C} = \frac{\mu_{pr}^{C}}{I_{br}^{C} + \mu_{rec}^{C}}; \qquad Q_{pr}^{I} = \frac{\mu_{pr}^{I}}{I_{br}^{I} + \mu_{rec}^{I}} \qquad Q_{pr}^{A} = \frac{\mu_{pr}^{A}}{I_{br}^{A} + \mu_{rec}^{A}} \tag{5}$$

where: Q_{pr}^{C}, Q_{pr}^{I}, Q_{pr}^{A} – coefficients of protection of resources against threats of confidentiality, integrity and availability, respectively;

I_{br}^{C}, I_{br}^{I}, I_{br}^{A} – the intensity of breaches of confidentiality, integrity, and availability of resources, respectively.

μ_{rec}^{C}, μ_{rec}^{I}, μ_{rec}^{A} – the intensity of recovery of security for confidentiality, integrity, and availability of resources, respectively.

Where μ_{rec}^{C}, μ_{rec}^{I}, μ_{rec}^{A} – are calculated as

$$\mu_{pr}^{C} = \frac{L_{thr}^{C}}{M_{inv}^{C}}, \qquad \mu_{pr}^{I} = \frac{L_{thr}^{I}}{M_{inv}^{I}}, \qquad \mu_{pr}^{A} = \frac{L_{thr}^{A}}{M_{inv}^{A}} \tag{6}$$

where, L_{thr}^{C}, L_{thr}^{I}, L_{thr}^{A} – the number of new threats that have emerged over time Δt; M_{inv}^{C}, M_{inv}^{I}, M_{inv}^{A} – the number of invented threats of interference during the Δt.

Generalized type:

$$Q_{pr}^{C,I,A} = \frac{\mu_{pr}^{C,I,A}}{I_{br}^{C,I,A} + \mu_{rec}^{C,I,A}} \tag{7}$$

As set out in the preceding paragraphs, the model for assessing the security of complex distributed information and telecommunication systems based on the count of security on the types of threats to information security [2].

Assessment of the level of security is carried out by defining a set of possible threats through a particular remedy, given the separate functions of all software and hardware-software means of protection, as well as the severity of threats to complex distributed information-telecommunication system [3].

Based on the conducted analysis of methods and ways to ensure its protection and taking into account the models of sustainability and security of DITS is given in the

first section of this dissertation research, a mathematical model for assessing the security of its turns:

$$Q=\sum\{p_{con}+p_{\text{int}}+p_{av}+S_k\}/4p_i z_k \quad (8)$$

where: p_{con} – confidentiality;

p_{int} – integrity of information;

p_{av} –availability information;

S_k – information security;

p_i – the weighting factor of the threat;

z_k – a success rate of protection.

CONCLUSION

The level of information protection defined by the security of each individual resource from the set, which is intended to protect systems.

Today, the market of information technologies, represents the many software solutions aimed at evaluating the level of security of complex systems. The fundamental basis for determining an effective level of protection of acts of systemic and systematic approaches [4] to the solution of complex problems based on a modular partitioning system.

Information technology assessment is aimed at identifying the most vulnerable areas of complex distributed information and telecommunication systems, whose security may be compromised in the first place. Or disadvantages which the offender could use to harm security in General.

The initial stage of the analysis is the structuring of the system that is multifold (allocation of information objects, spatial and functional structuring). The latter says the restructuring of a complex system into functional subsystems.

After structuring system assesses the vulnerability, integrity, observability, privacy, security of each object separately with a detailed rationale for each criterion.

The next phase is model building event risks. Here, the main factor is to identify the level of risk each hazard that is identified at an early stage. The result of this analysis is a detailed description of risk events, their consequences, the threat level of possibility of realization (occurrence). Also important criterion is the degree of influence on the system and the probable loss of information.

After the statistical data are compared quantitatively with the available previously obtained (if any) is performed to determine the assessment of the level of probability of risk events, and in order to stress the dynamics of frequency of occurrence of risk events, and in order to stress the dynamics of frequency of occurrence of risk events according to detected direction.

Based on this, using a mathematical algorithm calculated the level of protection of complex distributed information and telecommunication systems, given the expected damage from each detected event risk as the mathematical expectation of the damage amount.

The described mechanism allows to evaluate the security of complex distributed information and telecommunication systems and based on this the user can define the

relevant measures to protect the last thing for today is a priority direction in the sphere of information technology.

REFFERENCES

1. KISELEV V., KOSTENKO A.: Cyberwar as a basis for hybrid operation. Military collector magazine. 11(2005)257, 3-6.
2. MATUSITZ J.: The Role of Intercultural Communication in Cyberterrorism. Journal of Human Behavior in the Social Environment 24(2014), 775-790.
3. OSTAPOV S. E.: Technology Zahist nformat: Navalny Oleshko, S.E. Ostapov, S. P. SEV, O. G. King. .:. National economic University 2013.
4. OVSYANNIKOV S. N.: A short course of probability theory and mathematical statistics: a textbook for students of the 2nd course of economic specialties, Ekon-inform, 2011.
5. GLUKHOV P. A., PONOMAREV A. P., YU. A., SHILENKOV M. V.: Approach to estimation and forecasting of the level of security of information and telecommunication systems, Proceedings of SPIIRAS 42(2015), 180-195.

Mykola SHEVCHUK[1], Maria MANDRONA[2]

Scientific supervisor: Volodymyr MAKSYMOVYCH[3]

BADANIE GENERATORA SEKWENCJI BITÓW PSEUDOLOSOWYCH OPARTYCH NA LFSR W RÓŻNYCH STOPNIACH MACIERZY FORMUJĄCEJ

Streszczenie: W artykule przedstawiono wyniki estymacji generatora LFSR (linear feedback shift register - liniowego sprzężenia zwrotnego) z różnymi stopniami wielomianów wejściowych, różnymi stanami początkowymi i różnymi stopniami macierzy wejściowych. Estymacje te przeprowadzono przy użyciu testów statystycznych NIST. Uzyskane wyniki umożliwiają optymalizację parametrów generatora przy danych parametrach sekwencji pseudolosowej.

Słowa kluczowe: generator sekwencji pseudolosowych, ochrona informacji, liczby pseudolosowe, charakterystyka statystyczna

INVESTIGATION OF THE PSEUDORANDOM BIT SEQUENCES GENERATOR BASED ON LFSR WITH DIFFERENT DEGREES OF THE FORMING MATRIX

Summary: The article presents the results of LFSR (linear feedback shift register) generator estimation with a different degrees of input polynomials, with different initial states and with a different degrees of input matrix, carried out with the use of NIST statistical tests. The received results allow to optimize the generator parameters at the given parameters of the output pseudorandom sequence.

Keywords: pseudorandom sequence generator, information protection, pseudorandom numbers, statistical characteristics.

1. Introduction

Pseudorandom number generator (PNG) and pseudorandom bit sequence generator (PSRBG) common in many areas of measurement technology systems, information security, in modeling different systems and processes. The requirements for their technical characteristics differ depending on the purpose of their application. Generating pseudorandom sequences and checking the randomness of

[1] Lviv Polytechnic National University, shevchuk.mykola93@gmail.com
[2] Lviv Polytechnic National University
[3] Lviv Polytechnic National University

generated sequence is the most important problems of modern cryptology. In modern cryptosystems, pseudorandom sequences generators are used to create key information and provide parameters for these systems.
The aim of this abstract is to use statistical tests of the National Institute of Standards and Technology (NIST) of the USA for testing pseudorandom bit sequence generators based on linear shift registers (LFSR) with different initial conditions, with different exponentiation of forming matrix and at different polynomials. Simplified block diagram of a generator based on LFSR is shown in Figure 1.

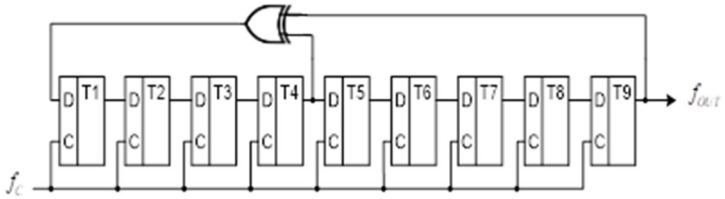


Figure 1. Simplified block diagram of the PSRBG based on LFSR

2. Pseudorandom bit sequences generators based on LFSR

The analysis of the literature, found that studies of generators built on LFSR performed many scientists in the results of which are shown in papers [1-9]. It is known that some pseudorandom bit sequences generators (PSRBG) based on linear feedback shift registers with linear loops - LFSR is not crypto-resistant [1 -2]. Despite this, they are often used as building blocks of more complicated devices such as the implementation of stream cipher (eg algorithms - A3, A5, A8, PIKE, SEAL, RC4) [2] and either a set of sequences for encryption keys. This type of generators is also used in communication systems [3] such as CDMA (Code Division Multiple Access), in the mobile communication systems, navigation systems, in the spread spectrum of communication systems, particularly in the Bluetooth technology [3-4], GSM, radar systems and some radio modems. Their main advantage is high speed and simplicity in hardware implementation.
Statistical characteristics of PSRBG based on LFSR considered in many papers [1, 2, 6-9]. However, there is not fully determined the minimum number of structural elements LFSR for devices based on them, when we will reach statistical security, and we can say that statistical characteristics of the output bit sequence are satisfactory;
The research if this paper is aimed for finding solution of these problem. We assume that the statistical characteristics PSRBG are satisfactory if the output bits sequence passes all NIST[12] tests and in complex compared generators also we should to consider: the recurrence of the builded output sequence (adaptability), complexity of construction and realization on programmable logic circuits (FPGA) and the maximum possible length of the encryption key, but in this paper, attention is focused on the research of statistical characteristics PSRBG with different conditions.
We should emphasize that the cryptostability of such generators, without cryptanalysis, is not guaranteed. As you know, statistical security is necessary, but it is not sufficient condition [10-12].
Options for constructing a PSRBG based on LFSR should be considered and based on the equation of its functioning.

$$Q(t+1) = T^{r} Q(t) \tag{1}$$

where, Q(t) and Q(t+1)- states of the register at moments of time t and t + 1 (before and after synchronization); T - square matrix of order n that has the form:

$$T_1 = \begin{vmatrix} a_1 & a_2 & \dots & a_{n-1} & a_n \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & & & \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix} \text{ or } T_2 = \begin{vmatrix} 0 & \dots & 0 & 0 & a_n \\ 1 & \dots & 0 & 0 & a_{n-1} \\ & & & & \\ 0 & \dots & 1 & 0 & a_2 \\ 0 & \dots & 0 & 1 & a_1 \end{vmatrix}; \quad (2)$$

n - degree of polynomial

$$F(x) = \sum_{i=0}^{n} a_i x^i, \quad a_n = a_0 = 1, \quad a_j \in \{0,1\}, \quad j = \overline{1,(n-1)}; \quad (3)$$

r - forming degree square matrix [1, 3, 4].
For polynomial F (x) = x ^{100} + x ^{37} +1, matrix T_1^1 (r=1), T_1^5 (r=5) and T_1^{10} (r=10) are:

$T_1 =$

	1	2	3	...	36	37	...	98	99	100
1	0	0	0	...	0	1	...	0	0	1
2	1	0	0	...	0	0	...	0	0	0
3	0	1	0	...	0	0	...	0	0	0
4	0	0	1	...	0	0	...	0	0	0
...	...	...	...	...	...	...	...	...	...	...
37	0	0	0	...	1	0	...	0	0	0
38	0	0	0	...	0	1	...	0	0	0
...	...	...	...	...	...	...	...	...	...	...
99	0	0	0	...	0	0	...	1	0	0
100	0	0	0	...	0	0	...	0	1	0

$T_1 =$

	1	2	3	...	33	34	35	36	37	...	94	95	96	97	98	99	100
1	0	0	0	...	1	0	0	0	0	...	0	0	1	0	0	0	0
2	0	0	0	...	0	1	0	0	0	...	0	0	0	1	0	0	0
3	0	0	0	...	0	0	1	0	0	...	0	0	0	0	1	0	0
4	0	0	0	...	0	0	0	1	0	...	0	0	0	0	0	1	0
5	0	0	0	...	0	0	0	0	1	...	0	0	0	0	0	0	1
6	1	0	0	...	0	0	0	0	0	...	0	0	0	0	0	0	0
7	0	1	0	...	0	0	0	0	0	...	0	0	0	0	0	0	0
8	0	0	1	...	0	0	0	0	0	...	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
28	0	0	0	...	1	0	0	0	0	...	0	0	0	0	0	0	0
29	0	0	0	...	0	1	0	0	0	...	0	0	0	0	0	0	0
30	0	0	0	...	0	0	1	0	0	...	0	0	0	0	0	0	0
31	0	0	0	...	0	0	0	1	0	...	0	0	0	0	0	0	0
32	0	0	0	...	0	0	0	0	1	...	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
99	0	0	0	...	0	0	0	0	0	...	1	0	0	0	0	0	0
100	0	0	0	...	0	0	0	0	0	...	0	1	0	0	0	0	0

$T_1 =$

	1	2	3	...	28	29	30	31	32	33	34	35	36	37	,,,	89	90	91	92	93	94	95	96	97	98	99	100
1	0	0	0	...	1	0	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	0	0	0	0	0	0
2	0	0	0	...	0	1	0	0	0	0	0	0	0	0	...	0	0	0	1	0	0	0	0	0	0	0	0
3	0	0	0	...	0	0	1	0	0	0	0	0	0	0	...	0	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	...	0	0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	1	0	0	0	0	0	0
5	0	0	0	...	0	0	0	0	1	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0	0	0	0
6	0	0	0	...	0	0	0	0	0	1	0	0	0	0	...	0	0	0	0	0	0	0	1	0	0	0	0
7	0	0	0	...	0	0	0	0	0	0	1	0	0	0	...	0	0	0	0	0	0	0	0	1	0	0	0
8	0	0	0	...	0	0	0	0	0	0	0	1	0	0	...	0	0	0	0	0	0	0	0	0	1	0	0
9	0	0	0	...	0	0	0	0	0	0	0	0	1	0	...	0	0	0	0	0	0	0	0	0	0	1	0
10	0	0	0	...	0	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0	0	0	0	0	0	0	1
11	1	0	0	...	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
12	0	1	0	...	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	1	...	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
38	0	0	0	...	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	...	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	...	0	0	1	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	...	0	0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	...	0	0	0	0	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	...	0	0	0	0	0	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	...	0	0	0	0	0	0	1	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	...	0	0	0	0	0	0	0	1	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	...	0	0	0	0	0	0	0	0	1	0	...	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	...	0	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
99	0	0	0	...	0	0	0	0	0	0	0	0	0	0	...	1	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	...	0	0	0	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	0	0	0

Schemes of generators based on LFSR are shown on Figures 2-4.

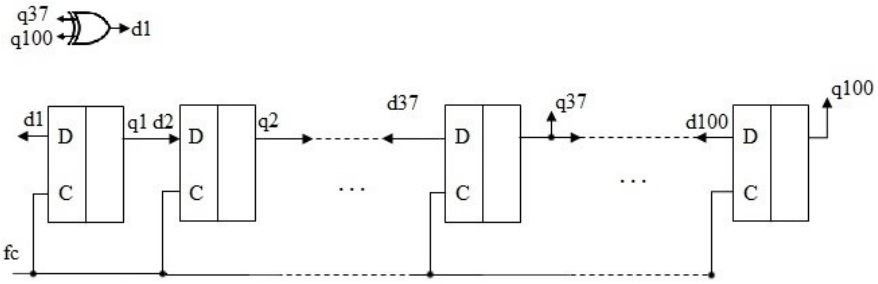


Figure 2. Generator pseudorandom bit sequence based on LFSR with a primitive polynomial F (x) = $x^{100} + x^{37} + 1$ matrix T_1, r = 1.

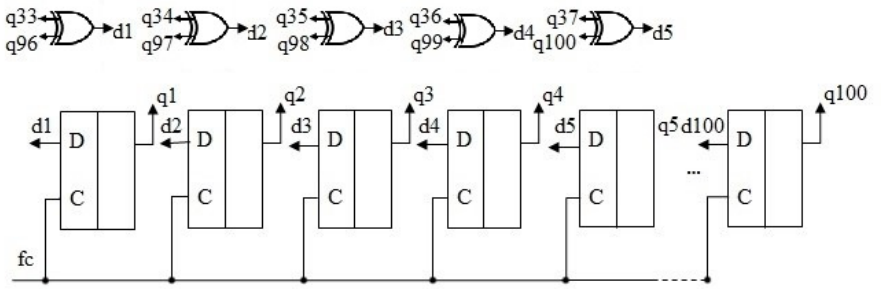


Figure 3. Generator pseudorandom bit sequence based on LFSR with a primitive polynomial F (x) = $x^{100} + x^{37} + 1$ matrix T_1 when r = 5.

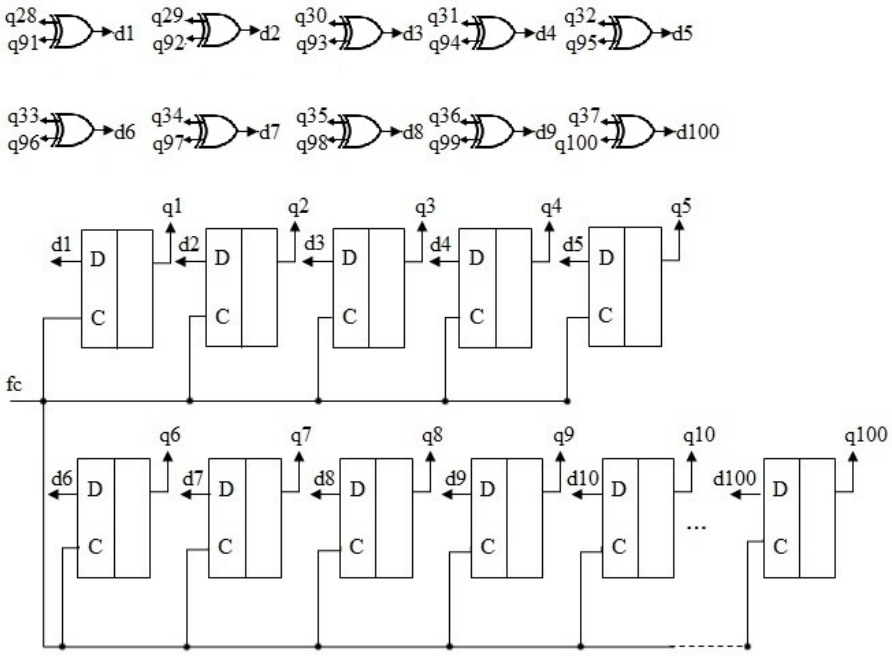


Figure 4. Generator pseudorandom bit sequence based on LFSR with a primitive polynomial F (x) = $x^{100} + x^{37} + 1$ matrix T_1 when r = 10.

2.1 Results of research characteristics of the pseudorandom bit sequence generator based on LFSR

With the help of simulation tests and NIST, we have been investigated statistical characteristics with different generator polynomials, with different powers of forming square matrix, and with different initial conditions. In the table 1 are the primitive polynomials which were investigated for PSRBG based on LFSR.

Table 1. Primitive polynomials for a generator based on LFSR

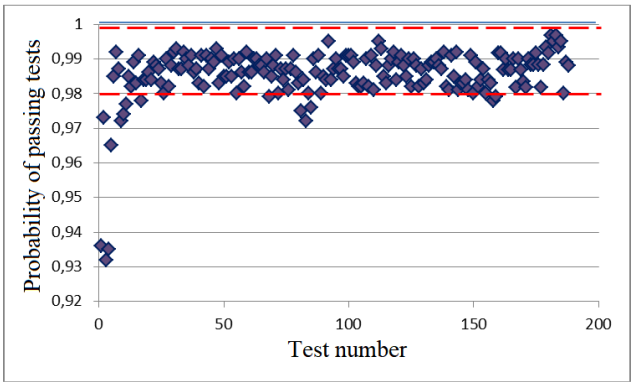
Symbolic designation	The value of a primitive polynomial
A	$F(x) = x^{100} + x^{37} + 1$
B	$F(x) = x^{150} + x^{53} + 1$
C	$F(x) = x^{201} + x^{14} + 1$
D	$F(x) = x + x^{270} + 1^{133}$
E	$F(x) = x^{322} + x^{67} + 1$
F	$F(x) = x^{378} + x^{43} + 1$

In the table 2 are the investigation results of the statistical characteristics of PSRBG based on LFSR at different values of primitive polynomials (Table 1) and degrees of the matrix and different initial conditions, namely the initial state 1 - the first bit "1" remaining bits "0"; state 2 - the last bit "0" is the rest of the bits "1"; state 3 - half bits "1" the other half "0".

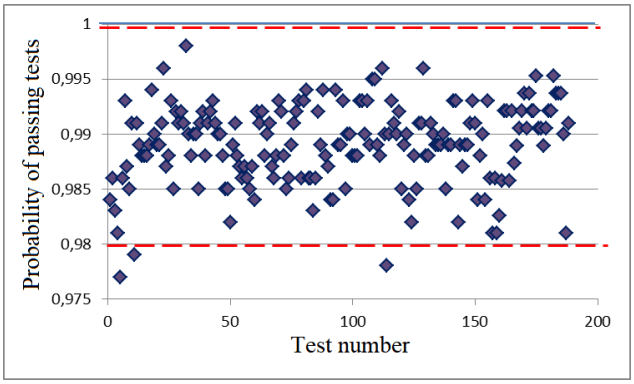
Table 2. Results of testing PSRBG

Conditional sign off an input polynomial	A			B			C			D			E			F		
The degree of matrix	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10	1	5	10
Initial state-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Initial state-2	-	-	-	-	-	-	-	-	-	-	+	+	-	+	+	-	+	+
Initial state-3	-	-	-	-	-	-	-	-	-	-	+	+	-	+	+	-	+	+

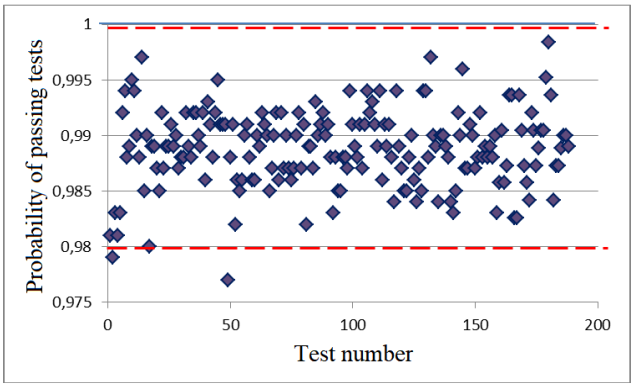
Below are statistical portraits of generator E (Figure. 5-8) with the initial states 1-3 and degrees of the matrix 1- a, 5 - b, 10 - c. Dashed lines indicate the limits of positive tests.



a



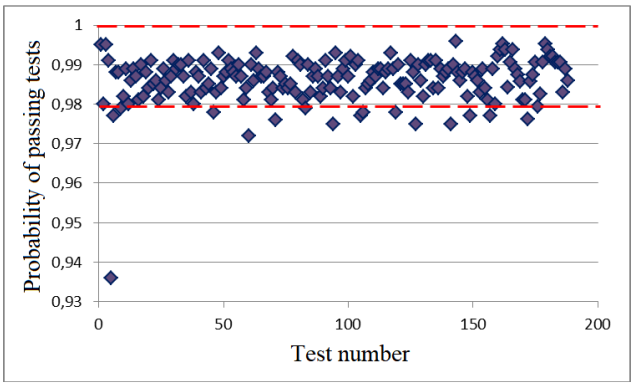
b



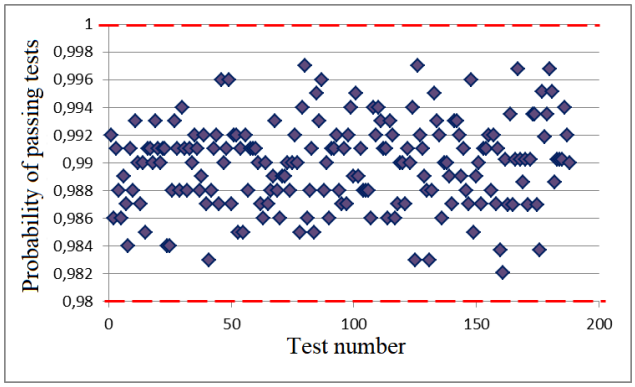
c

Figure. 5. Statistical portraits generator E, initial state-1

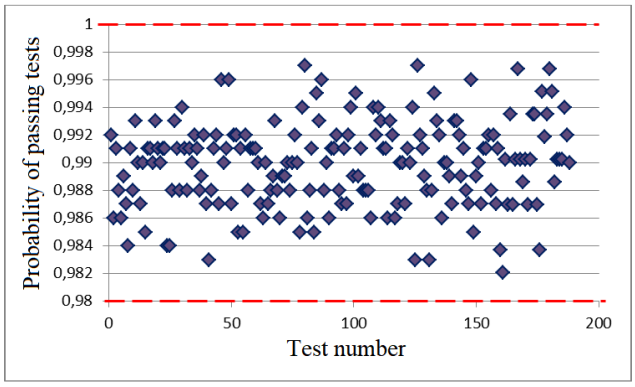
As can be seen from Figure 5, increasing the number of degree matrix leads to improved statistical characteristics PSRBG.



a



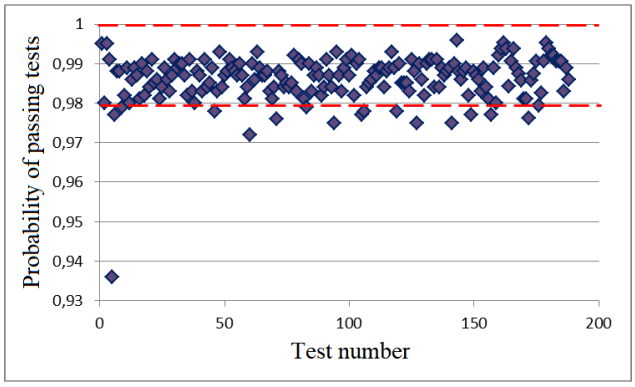
b



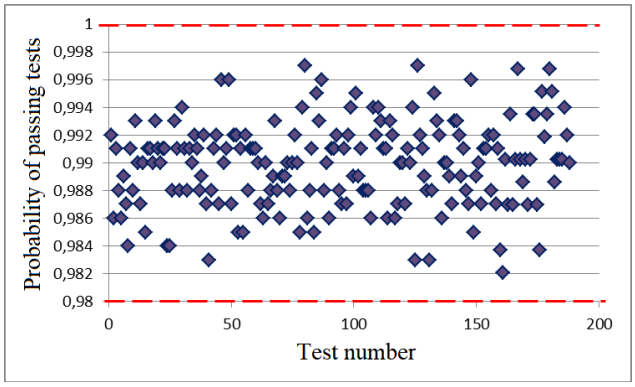
c

Figure 6. Statistical portraits generator E, initial state-2

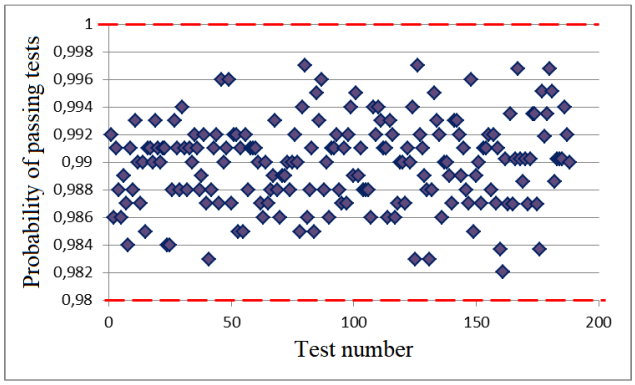
Comparing with the previous results (*Figure* 5), we observe a significant improvement in the statistical characteristics of the PSRBG. Variants of PSRBG (Figure 6b, c) have passed all the tests, so we can conclude that these generators are statistically safe.



a



b



c

Figure. 7. Statistical portraits generator E, initial state-3

2.2 Analysis of the research work

The results of testing the generator showed that statistical characteristics do not meet in the degree of randomness forming a square matrix r = 1 (table 2). Increase of the degree of matrix r is leading to a significant improvement of statistical characteristics. So, starting with generator D (270 bit) and higher degrees and forming matrix where r has degree bigger than 5, we can say that this PSRBG is statistically safe.

3. Conclusion & future work

The investigation PSRBG based on LFSR showed that even at large degrees of generators generator polynomials are not statistically safe, but when we change the degree of forming square matrix r, we can reach statistical safety. So investigated generators can be used as components of cryptographic systems.

REFERENCES

1. IVANOV M.A.: Cryptographic methods of information protection in computer systems and networks: a teaching manual, M.A. Ivanova IV Chugunkov - M.: NIUA MIFI, 2012. - 400 p.

2. SCHNEIER B.: Applied cryptography, protocols, and algorithms for language yshodnыe tekstы C, B. Schneier. - M.: Triumph, 2002. - 816 pp.
3. BABACAR ALASSANE NDAW, DJIBY SOW, MAMADOU SANGHARE: Construction of the Maximum Period Linear Feedback Shift Registers (LFSR) (Primitive Polynomials and Linear Recurring Relations) British Journal of Mathematics & Computer Science 11 (4): 1-24, 2015, Article no.BJMCS.19442, ISSN: 2231-0851 SCIENCEDOMAIN international.
4. MILOVANOVIC E., STOJCEV M., MILOVANOVIC I., NIKOLIC T.: Concurrent generation of pseudo random numbers with lfsr of fibonacci and galois type, Computing and Informatics, 34(2015), 941-958.
5. KITSOS P., SKLAVA N., ZERVAS N., KOUFOPAVLOU O.: A Reconfigurable Linear Feedback Shift Register (LFSR) for the Bluetooth System. The 8th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2001), 2(2001), 991-994.
6. NAS R.J.M., Van Berkel C.H.: High Throughput, Low Set-Up Time, Reconfigurable Linear Feedback Shift Registers. Proceeding of the 28th IEEE International Conference on Computer Design (ICCD), Amsterdam, October 2010, 31-37.
7. JHANSIRANI A., HARIKISHORE K., BASHA F.N., POORNIMA J., JYOTHIL M., SAHITHI M., SRINIVAS P.: Fault Tolerance in Bit Swapping LFSR Using FPGA Architecture. International Journal of Engineering Research and Applications, 2(2012)1, 1080-1087.
8. MAKSYMOVYCH V.: Research pseudorandom bit sequence generators based on LFSR, Maksymovychh VM, Shevchuk M., M. Mandrona, automation, measurement and control: Coll. Sci.-Tech. works. - Lviv: View of Lviv Polytechnic National University. 852(2016), 29-34.
9. MANDRON M.M.: Investigation of the Statistical Characteristics of the Modified Fibonacci Generators, MM Mandrona, VM Maksymovych, Journal of Automation and Information Sciences 10.1615, JAutomatInfScien.v46.i12.60 pages 48-53
10. MANDRONA M.M.: Comparative Analysis of Pseudorandom Bit Sequence Generators, MM Mandrona, VM Maksymovych, Journal of Automation and Information Sciences, DOI: 10.1615, JAutomatInfScien.v49.i3.90 pages 78-86
11. A Study of the Characteristics of the Fibonacci Modified Additive Generator with a Delay, [Maksymovych VM, Mandrona MM, Garasimchuk OI Kostiv Yu.M.], Journal of Automation and Information Sciences DOI: 10.1615 / JAutomatInfScien.v48.i11.70 pages 76-82
12. NIST SP 800- 22. A Statistic Test Suite for Random and Pseudorandom Number Generators for Cryptographic Application: [electronic resource], April 2000. Access: http: // csrc. nist. gov ./ publications / niatpubs // SP 800-22 rev 1 a. pdf .

Kazimierz SIKORA[1]

Opiekun naukowy: Stanisław ZAWIŚLAK[2]

IZOMORFIZM WYBRANYCH KLAS GRAFÓW

Streszczenie:W artykule omówiono program komputerowy do sprawdzania wybranych klas grafów o relatywnie małej liczbie wierzchołków. Wyniki przedstawiane są na ekranie. Zastosowano metodę pełnego przeglądu opartą o generowanie permutacji.

Słowa kluczowe: wizualizacja, metoda pełnego przeglądu, generowanie permutacji

ISOMORPHISM OF CHOSEN GRAPH CLASSES

Summary: In the paper, the computer program is described which performs checking of isomorphisms of chosen class of graphs. The results are visualized on the computer screen in special window. The full search method has been utilized which is based on permutation generation method.

Keywords: visualization, full search method, permutation generation

1. Introduction

Graphs theory is a branch of mathematics having versatile applications in other branches of knowledge. The field of graph theory was initiated in 1736 by Leonhard Euler via introduction of famous Königsberg bridges problem. It is under the constant development and it is taught at the Universities at the different directions of studies like e.g. mathematics, telecommunication and computer science. In Poland, several books are available on the market [7,8]. However, there are still problems which have not been solved. The problem which belongs to this area is detecting of isomorphism of arbitrary chosen graphs, which is discussed in the book fully dedicated to this problem solely [4].

The application of isomorphism checking can be applied in some practical problems e.g. virtual network modeling [5], in comparison of chemical compounds [3, 9] (fields of chemistry) and comparison of kinematical chains [14, 15] (field of

[1]University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, Student of Informatics on master level, email:
[2] Associate Professor, University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science,szawislak@ath.bielsko.pl;

mechanics). These chains are represented via graphs and they are obtained as a result of the automatic synthesis of mechanical systems. The review type papers, where comparisons of methods were done, are e.g. [1,3]. Several methods had been used to solve the considered problem:

a) neuronal networks [15],
b) quadratic forms [12],
c) discriminating invariants [2],
d) decision trees [10],
e) eigensystem [11].

In the present paper, the full search method was utilized. Obviously, it could not be utilized for graphs of e.g. n>100 vertices. However, the goal was to prepare the application which could help in didactics as well as which gives the results on the screens – so having visualization facility [18]. Moreover, it allows for listing of all possible solutions i.e. all permutation of vertices for which isomorphism holds.

2. Basic notions and definitions

The graph G(V,E) is a pair of sets: V – set of vertices (nonempty) and E – set of edges, moreover we usually denote |V| = n and |E| = m. The graph can be represented via different algebraic structures e.g. list of neighbors and incidence matrix as well as many others. In Fig. 1, some graphs are presented in columns: (a) paths or trees distinguished in the K_4 graph, (b) cycles distinguished in graph K_5, two self-complementary graphs distinguished in graph K_5.Graph K_5 is a clique spanning on 5 vertices, known as the Kuratowski's graph [16].

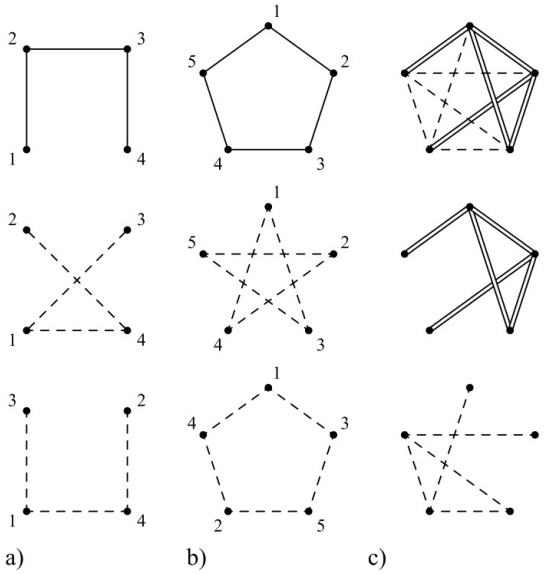


Figure 1. Sc-graphs for n = 4 (a) and n=5 (b)(c)

Self-complementary graphs are special class of graphs where so called complementary graph is isomorphic with the source one. Graph complementary to G

is G^* which is built on the same vertex set $V^* = V$ and the $E^* = \{\text{full set of edges}\} - E$. Full set of edges contains all possible graph edges i.e. $0.5 \cdot n \cdot (n - 1)$ in a graph having n vertices.

The graphs $G_1(V_1,E_1)$ and $G_2(V_2,E_2)$ are isomorphic if and only if from existence of bijection $V_1 \leftrightarrow V_2$ indicates that exists the bijection $E_1 \leftrightarrow E_2$ exists. In the language of normal people, isomorphism means that these two graphs have the same geometrical shape or form. Therefore, after some rearrangement these two graphs can be placed one on another showing that they are the same.

In table 1, we show the isomorphism of the first and the third graphs placed in first column i.e. Figure 1a. They are complementary graphs. Each of them has 3 edges as well as their edges together create a full set of edges in K_4. In this case we see that bijection of vertices is equivalent to bijection of edges - see Table 1.

The graphs are represented via lists of neighbors in the present paper.

Table 1. Isomorphism details for the graph shown in Fig. 1a

Bijection of vertices	Bijection of edges
1↔3 **2**↔1 **3**↔4 **4**↔2	**(1,2)** ↔ (3,1) **(2,3)** ↔ (1,4) **(3,4)** ↔ (4,2)
List of source graph	List of complementary graph
(**(1, 2)** **(2, 1, 3)** **(3, 2 ,4)** **(4, 3)**)	((3, 1) (1, 3, 4) (4, 1, 2) (2, 4))

Like we can see bijection of vertices: **1**↔3, **2**↔1, **3**↔4, **4**↔2, generate a bijection of edges **(1,2)** corresponds to the edge (3,1) etc. Obviously there two possible assignment – the second consists in reverse order of vertices in the second sequence.

Only the most relevant notions were discussed above, other needed definitions and theorems could be found in the cited references [4,7,8,13].

3. Algorithm

The algorithm consists in generation of all permutation of graph vertices. The full search approach consists in comparison of two structures representing graphs. The first structure remains as a source one. The second structure is changed via the consecutive permutation, identical structured and the adequate permutation are registered.

The algorithms given in the monograph of professor Lipski [13] for generations of permutations were utilized e.g. recursive or anti-lexicographic order.

4. Computer program

The computer program was prepared in Visual Studio 2015 environment, utilizing technology WPF as well as algorithmic languages C# and XAML. Full description of the program is given in thesis [17].

The program opens the window in the screen of constant dimension 440 x 830 pixels.

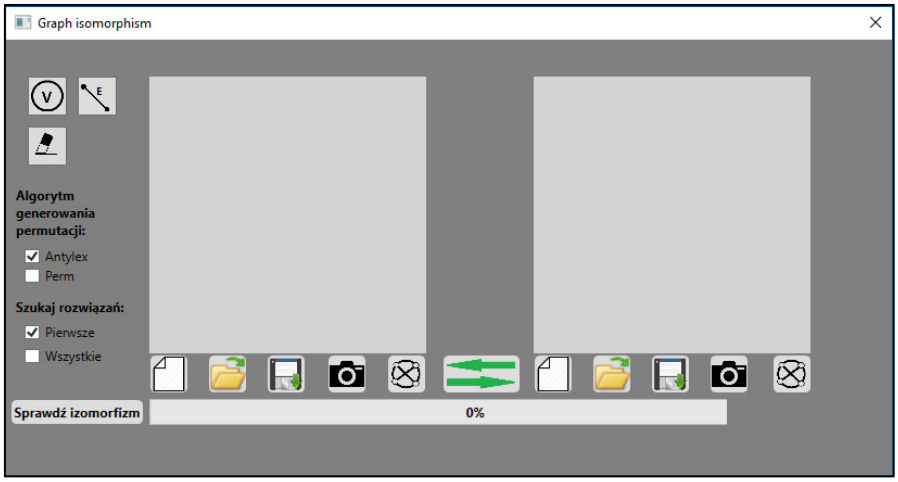


Figure 2. Form of the user interface, which appears on computer screens

The graph can be red (input) or drawn by a user by means of mouse-driven actions. The exemplary graphs will be given underneath. Very simple graph on n=4 is shown in Fig. 3.

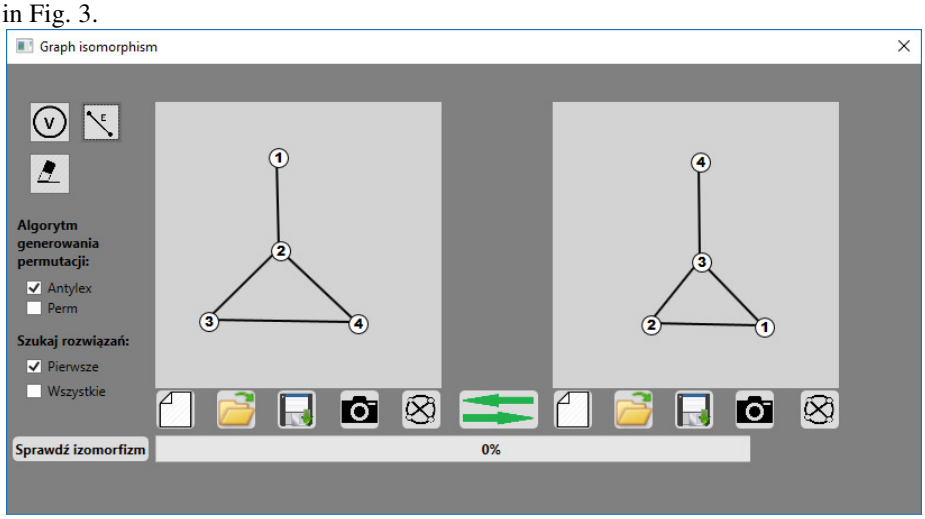


Figure 3. Exemplary simple graph shown on the computer screen

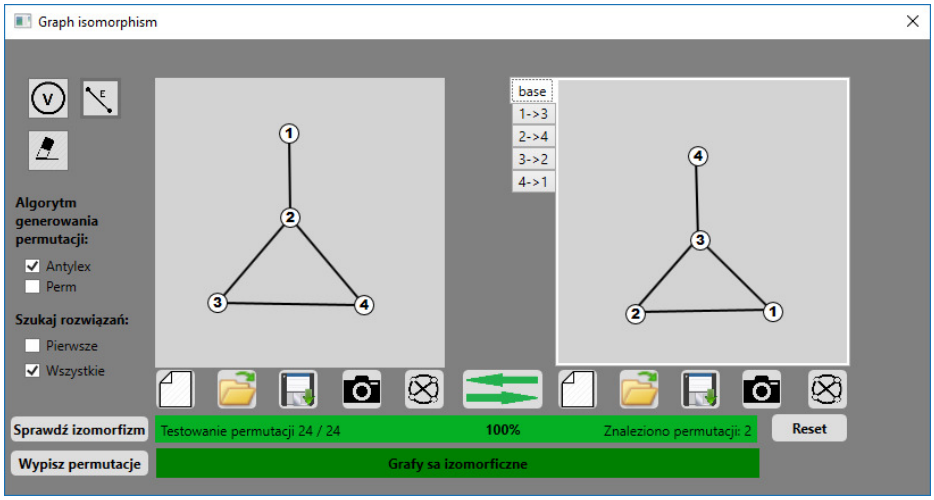


Figure 4. Exemplary graphs for detection of isomorphism

Like previously, one can check isomorphism via eye inspection. The vertices should be assigned: 1↔4 the highest placed ones, 2↔3. However, there two possibilities for assignment of vertices being ends of the horizontal edge. The programs checks:

$$P_4 = 4! = 1\ 2\ 3\ 4 = 24 \tag{1}$$

This information is given in the bottom bar of the programme window. Additionary, there is the information that two solutions. Therefore, same result was obtained via the discussed application (Fig. 5) like based on commonsense reasoning.

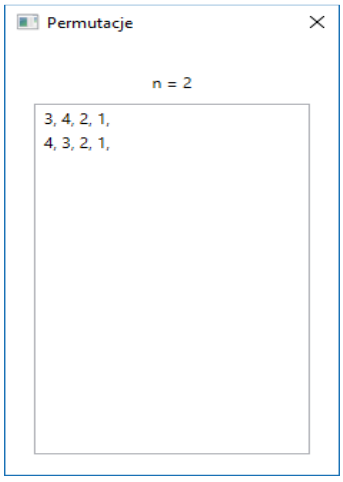


Figure 5. Result given out by the programme

The more complex graph of n=10 vertices is shown in Fig. 6.

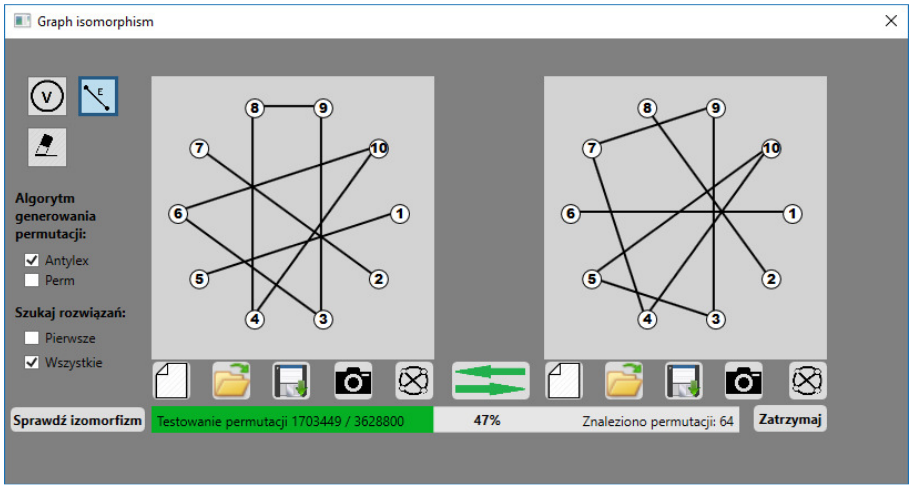


Figure 6. Exemplary graphs of n=10 vertices

Now, eye inspection is useless, layout of the graphs are so complex that we could not decide if the two graphs are isomorphic or not. The graph is disconnected. The programme checks:

$$P_{10} = 10\,! = 1\ 2\ 3\ \ldots\ 9\ 10 = 3628800 \tag{2}$$

permutations. We can see that 47% were analyzed. Till this moment, 64 solutions were found.

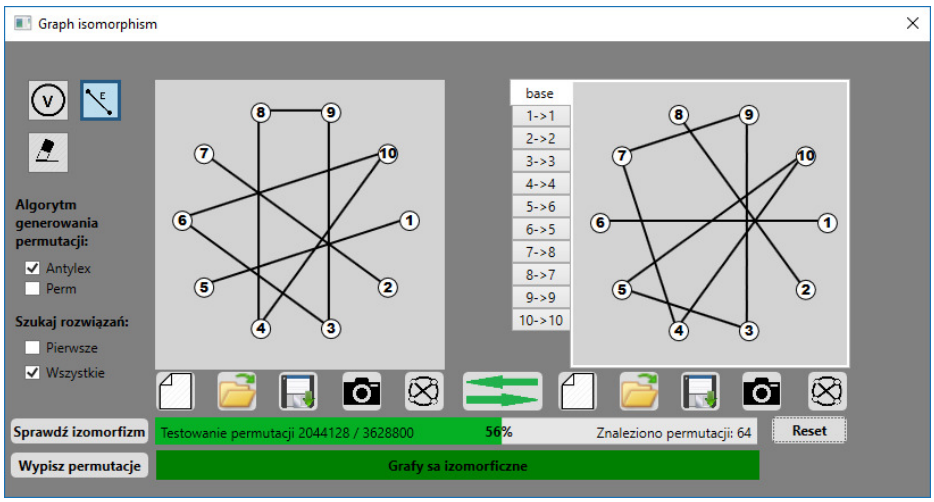


Figure 7. Result of checking isomorphism, positive conclusion and the exemplary assignment (bijection of vertices)

The programme is dedicated for the didactics of subject related to graph theory. The advantage is that the list of all permutations is obtained. Moreover, user can rearrange the layout of vertices on the screen which helps is explanation of the essence of the idea of isomorphism.

5. Conclusions

In the present paper, the problem of graphs isomorphism was discussed. The computer programme for solving this problem for small graphs was described. The program could be used in didactics of subjects related to discrete mathematics and graph theory. The programs confirms isomorphism as well as lists all the permutations which fulfill the isomorphism condition. It shows how many solutions are available.

REFERENCES

1. ARWIND V., TORAN J.: Isomorphism testing: perspectives and open problems, Bulletin of the European Association for Theoretical Computer Science, **86**(2005), 66-84.
2. DEHMER M. et al.: An efficient heuristic approach to detecting graph isomorphism based on combinations of highly discriminating invariants, Advanced in Computational Mathematics, **39**(2013), 311-325.
3. FORTIN S.: The graph isomorphism problem. University of Alberta, Canada, Technical Report (1996).
4. KOBLER J., SCHŐNINGU., TORAN J.: The graph isomorphism problem: its structural complexity. Springer Science & Business Media, Berlin 2012.
5. LISCHKA J.; HOLGER K.: A virtual network mapping algorithm based on subgraph isomorphism detection. In: Proceedings of the 1st ACM workshop on Virtualized infrastructure systems and architectures. ACM, (2009) 81-88.
6. ZEMLYACHENKO V.N. et al.: Graph isomorphism problem, Journal of Mathematical Science, **29/4**(2000), 1426-1481.
7. WOJCIECHOWSKI J., PIEŃKOSZ K.: Graphs and Networks (in Polish), PWN, Warszawa (2013).
8. WILSON R.J.: Introductory graph theory (in Polish), PWN, Warszawa (2017).
9. RAYMOND J. W., WILLETT P.: Maximum common subgraph isomorphism algorithms for the matching of chemical structures. Journal of computer-aided molecular design, **16.7**(2002) 521-533.
10. MESSMER B. T., BUNKE H.: A decision tree approach to graph and subgraph isomorphism detection. Pattern recognition, **32.12**(1999), 1979-1998.
11. HE, P. R.; ZHANG, W. J.; LI, Q. Some further development on the eigensystem approach for graph isomorphism detection. Journal of the Franklin Institute, **342.6**(2005) 657-673.
12. HE, P. R., et al. A new method for detection of graph isomorphism based on the quadratic form. Journal of Mechanical Design, **125.3** (2003) 640-642.
13. LIPSKI W.: Combinatorics for programmers(in Polish), WNT, Warszawa (2009).

14. MRUTHYUNJAYA T. S., BALASUBRAMANIAN H. R.: In quest of a reliable and efficient computational test for detection of isomorphism in kinematic chains. Mechanism and Machine Theory, **22.2** (1987) 131-139.
15. KONG, F. G.; LI, Q.; ZHANG, W. J. An artificial neural network approach to mechanism kinematic chain isomorphism identification. Mechanism and Machine Theory, **34.2** (1999) 271-283.
16. WOJNAROWSKI J., ZAWIŚLAK S.: Kazimierz Kuratowski—Biography and Genesis of the Theorem on Planar Graphs. In: Graph-Based Modelling in Engineering. Editors: S. Zawiślak and J. Rysiński, Springer International Publishing, Berlin (2017) 233-246.
17. SIKORA K.: Detecting of isomorphism of chosen classes of graphs (in Polish), Bachelor/Engineer Thesis, University of Bielsko-Biala, Bielsko-Biała (2016).
18. ZAWIŚLAK S., KOPEĆ J.: Dedicated computer programs for visualizing some graph theory problems as learning enhancement, APLIMAT, Bratislava (2017), 1715-1727.

Andrii STEFANIV[1], Taras DOLINSKII[2]

Supervisors: Ruslan KOZAK[3]

UŻYCIE ALGORYTMÓW UCZENIA MASZYNOWEGO APACHE SPARK MLLIB DO WYKRYWANIA WYŁUDZANIA (PHISHING-U) W DANYCH TEKSTOWYCH

Streszczenie: W tym artykule opisano użycie Apache Spark MLlib do wykrywania witryn phishingowych w danych tekstowych opartych na algorytmach uczenia maszynowego. Została użyta implementacja regresji logistycznej oraz algorytmów uczenia maszynowego na drzewach decyzyjnych z pomocą technologii Spark. Implementacja zostala porównana z podobnymi rozwiązaniami na Python-a, które korzystują z bibliotek uczenia maszynowego, aby pokazać, że technologia Apache Spark ma dobre wyniki rozwiązywania problemu wykrywania danych phishingowych.

Słowa kluczowe: MLlib, nauczanie maszynowe, regresja logistycznej, Spark, drzewo decyzyjne, algorytm

USING MACHINE LEARNING ALGORITHMS OF APACHE SPARK MLLIB FOR DETECTION OF PHISHING IN TEXT DATA

Summary: This paper describes using Apache Spark MLlib for detection of phishing web sites in text data based on machine learning algorithms. Used implementation of logistic regression and decision-tree machine learning algorithms on Spark technology and compared with similar solutions used Python machine learning libraries to show that Apache Spark technology has good performances to solve the problem of phishing data detection.

Keywords: MLlib, machine learning, logistic regression, spark, decision-tree, algorithm

1. Introduction

Today, computer technologies are rapidly developing and used in all areas of everyday life, it is often used by intruders, falsifying original sources of false information and attempting to steal user personal data, this term was called phishing. Personal data can be: numbers of cards, logins and password for internal banking systems, Internet shops, money transfer services. Theft of this information may lead to theft of users. In order to prevent, in the area of information technology security there are methods of preventing phishing. In this area machine learning algorithms are used quite effectively, the most popular of them are:

[1] Ternopil Ivan Pul'uj National Technical University, Department of Computer Science and Information Technologies, andrystefaniv@gmail.com
[2] Ternopil Ivan Pul'uj National Technical University, Department of Computer Science and Information Technologies, dtaras85@gmail.com
[3] Assoc.Prof, PhD, Ternopil Ivan Pul'uj National Technical University, Department of Cyber Security ruslan.o.kozak@gmail.com

- logistic regression;
- decision-tree.

These methods have many implementations in many programming languages. Within the framework of the research, using and comparing an implementation of algorithms is presented in Apache Spark MLlib [1].

2. Traditional methods of phishing detection

Traditional methods of detection fall into two categories, the network-level protection and the authentication protection. The first category of protection at a network level includes blacklist filters and whitelist filters which prevent phishing by blocking suspected IP addresses or domains from accessing the network. In addition, there are the Pattern Matching filters and the Rule-based filters which rely on manually entered and updated fixed rules for detection.

The second category, authentication protection, provides security on both user and domain levels. For a user-level protection, users will have to provide authentications before sending their messages such as verified email and password, while the authentication protection on a domain-level is created for emails servers [2].

3. Problem statement

Phishing is popular among attackers, since it is easier to trick someone into clicking a malicious link which seems legitimate than trying to break through a computer's defense systems. The malicious links within the body of the message are designed to make it appear that they go to the spoofed organization using that organization's logos and other legitimate contents. Overall the problems carried out in this research are as following: how to use Apache Spark MLlib for determine the best set of features to be used with phishing detection? Why Apache Spark MLlib is the best decision of this problem?

4. Overview Apache Spark MLlib

Apache Spark is a fast and general-purpose cluster computing system [3]. It provides high-level APIs in Java, Scala [4], and an optimized engine that supports general execution graphs. It also supports a rich set of higher-level tools including MLlib for machine learning [5]. MLlib is Spark's machine learning (ML) library. Its goal is to make practical machine learning scalable and easy. At a high level, it provides such tools as [6]:

- ML Algorithms: common learning algorithms such as classification, regression, clustering, and collaborative filtering
- Featurization: feature extraction, transformation, dimensionality reduction and selection
- Pipelines: tools for constructing, evaluating, and tuning ML Pipelines
- Persistence: saving and load algorithms, models, and Pipelines
- Utilities: linear algebra, statistics, data handling, etc.

5. Machine learning methods of phishing detection

These methods apply automated classifiers that rely on machine learning and data mining. The classifiers work beside the server and filter the received data into phishing or legitimate by examining different features of the phishing data [7].

A logistic regression is a widely-used method due to its easily-interpretable and practical results. This model is functional in predicting binary data (0/1 response) as it relies on statistical data and applies a generalized linear model. Despite of this method's simplicity, it has three shortcomings; first, it requires more statistical assumptions before being applied [8]. Second, it is more functional with variables that have linear relation than those with a complex relation. Last, the accurateness of the prediction rate is sensitive to the data completeness.
Decision Tree is a graphical model of classification that is comprised of nodes and arrows. The base node is called the Root from which the DT is initiated. Each node within the network contains an "If-then" rule, a class, and a feature, and leads to the next one using the arrows, referred to as edges. The decision tree ends with a leaf node called terminator. The tree can include one or more classifier stages and the internal nodes, which are bounded by the root and terminating nodes [9, 10].

6. Test data

For training of models and checking results, dataset of phishing website publicly available on the machine learning repository provided by UCI was used. The dataset was collected by analyzing a collection of 2456 websites among which some were used for phishing and others not. For each website included in the dataset, 30 attributes were given.

7. Experiments

In a solution of the problem of detecting phishing typical machine learning processes flow was used (Fig. 1). At first, featurization of all data from dataset was provided. The next step was model training with test dataset which contains 70% of input dataset, for maximum prediction result. The last step was sending last 30% of data for checking.
The accuracy value was get during an experiment and was a bit better than results of Python Sklearn implementation with default parameters which was got from opensource github repository, it means that technology Apache Spark with default parameters has a better algorithms realization and it can help provide more scalable and more faster systems for phishing data analysis. The implementation of Apache Spark MLlib partitions data by rows, allowing distributed training with millions or even billions of instances. Results of the experiment are presented in table 1.

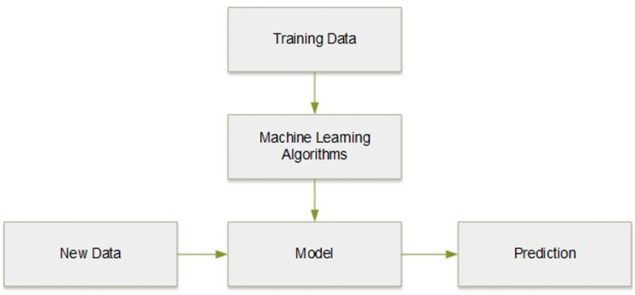


Figure 1. Base machine learning process flow

Table 1. Results of experiments

Method	Accuracy % (Apache Spark MLlib)	Accuracy % (Sklearn library)
Logistic regression	92.4	90.4
Decision Tree(DT)	92.3	90.5

8. Conclusions and future work

This paper describes using of linear regression and decision-tree algorithm with Spark technology for phishing detection. For testing of working system accuracy dataset of phishing website was used, publicly available on the machine learning repository provided by UCI. The implementation of this algorithm shows good results. In future experiments other algorithms such as CART and SVM, which can give better metrics results, will be used.

REFERENCES

1. SHANAHAN J.G., DAI L.: Large Scale Distributed Data Science using Apache Spark. KDD '15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM New York, NY, USA 2015, 2323-2324.
2. MOHAMMAD R., McCLUSKEY T.L.: Thabtah, Fadi An. Assessment of Features Related to Phishing Websites using an Automated Technique. In: International Conferece For Internet Technology And Secured Transactions. ICITST 2012. IEEE, London UK 2012, 492-497. ISBN 978-1-4673-5325-0
3. Big Data Processing with Apache Spark - Part 4: Spark Machine Learning. October, 2017, P. 10. Available at: https://www.infoq.com/articles/apache-spark-machine-learning.
4. SHORO A.G., SOOMRO T.R.: Big Data Analysis: Ap Spark Perspective.Global Journal of Computer Science and Technology: C Software & Data Engineering. 15(2015)1., Version 1.0., Global Journals Inc. (USA),7-14.
5. SHYAM R., BHARATHIGANESH H.B., SACHINKUMAR S., PRABAHARAN POORNACHANDRANB, SOMAN K.P.: Apache Spark a Big Data Analytics Platform for Smart Grid. Procedia Technology. 21(2015), 171-178.
6. Spark Machine Learning Library (MLlib). October, 2017, P. 8. Available at: https://spark.rstudio.com/mllib.html
7. SA'ID ABDULLAH AL-SAAIDAH. Detecting Phishing Emails Using Machine Learning Techniques. January, 2017, P. 59. Available at: https://meu.edu.jo/uploads/1/590422b4d5dd8_1.pdf
8. AKINYELU, A. A., ADEWUMI, A. O.: Classification of phishing email using random forest machine learning technique. Journal of Applied Mathematics. 2014, P. 6.
9. RAMANATHAN V., WECHSLER, H.: PhishGILLNET—phishing detection methodology using probabilistic latent semantic analysis, AdaBoost, and cotraining. EURASIP Journal on Information Security. 2012, P. 22. Available at: https://www.researchgate.net/ publication/ 257879512_PhishGILLNET-phishing_detection_methodology_using_probabilistic_latent_semantic_analysis_AdaBoost_and_co-training
10. Classification and regression - Spark 2.2.0 Documentation. October, 2017, P. 20. Available at: https://spark.apache.org/docs/2.2.0/ml-classification-regression.html

Vitalii SUSUKAILO[1]

Opiekun naukowy: Yuriy LAKH[2]

PROGNOZY STOSOWANIA SYSTEMÓW KONTROLI DOSTĘPU OPARTEJ NA ROLACH

Streszczenie: W pracy omówiono analizy ogólnych problemów dotyczących systemów kontroli dostępu. Zaproponowano system kontroli dostępu opartego na rolach z zastosowaniem technologii QR-code, ze zintegrowanym menedżerem haseł oraz określonymi polisami kontroli dostępu.

Słowa kluczowe: QR-kod, skaner, kontrola dostępu oparta na rolach (KDOR), menedżer haseł, QR kod dekoder

RBAC-Q FUTURE OF ROLE BASE ACCESS CONTROL SYSTEM

Summary: General problems of access control systems were analyzed. Role based access control system based on QR – code technology with integrated password manager and defined access control polices has been proposed.

Keywords: QR-code, scanner, role-based access control (RBAC), password manager, QR code decoder

1. Introduction

Nowadays more and more organizations should take care of information security. Tesla Crypt, WannaCry, NotPetya and other "popular" malware prompted the whole world to think about cybersecurity. Confidentiality, availability, integrity of information are three main principles, needed to be covered by organizations because in the 21st century appeared a new war area – cyberspace. Antiviruses, SIEM, firewalls became more and more popular. So different software need to be installed to protect information. Even for secure user authorization data any organization needs to have not only good RBAC systems, password managers but also perfect password management policy. There are three different processes needed to be involved and continuously improved to protect user from account compromising. In this article we will cover these three processes by one innovative concept of role-based access control system based on QR-codes. QR-code technology is a quick and easy way to share authorization data between users and information systems. Mostly we can see it in finance and media industries. From cybersecurity point of view all of us see it as an additional type of authorization on the popular web-sites. To define necessity of access control system based on QR-code technology let's get in touch with password managers, RBAC systems and access management policy in the next section.

[1] Lviv Polytechnic National University, Faculty of Computer Technologies, Automation and Metrology, Information Security department: mister.lembert@gmail.com
[2] Lviv Polytechnic National University, Faculty of Computer Technologies, Automation and Metrology, Information Security department: yurii.v.lakh@lpnu.ua

2. Problems with access control systems

The use of role based access control system means that managing user's privileges within a single system or application is widely accepted as the best practice. This kind of systems allows cybersecurity team to control users access in organization. With role-based access control organizations manage user's privileges for network, computer systems, web-services and physical access control. It's necessary to define: who will be network administrator and who – usual user, who has an access to the financial data and who will operate with organization infrastructure, who will have access to the executive and who to the accounting rooms. This is a good approach, which should be implemented everywhere, but it requires a lot of resources, not only financial but human also. Three or even four systems should be managed by two or three employees. For every system they should control that authorization data meet the requirements of company access management policies.

Physical access could be controlled by access card, biometric data control systems or code panel. Plastic cards nowadays can be easily duplicated or in the best case lost. Biometric access control systems are very expensive, so organization cannot install them into every door. And the code panel is not enough secure. Usually they have restored combination or access code could be guessed. That's why we cannot control access without security guard or the person who will control system state and check system logs.

Access in cyber area also may be determined in access control polices. Of course, organizations can automatize this process with rules in RBAC, for instance password should contain more than X symbols or it should have specific characters, but perhaps it will not work with every system and each user. To deny human factor we can use password managers like LastPass or KeePass. They allow user to set up passwords according to rules which are required by an organization and store all of them not in user's memory. But what user should do when he/she suddenly delete password database? Or forget password to the database? Or organization will not have good RBAC and user will have an access to the system, which he/she is not allowed to use more? So many problems, but let's try to solve them.

3. Information security management system

After a long time research we determined, what is needed to be covered by RBAC to be the most comfortable for user and useful for organization. Such system should:

- not provide user with access to the passwords;
- combine all types of access to the organization;
- have fast data transmission between server and endpoint;
- create difficult passwords according to the best practice rules;
- have password management system for external access;
- be monitored by one person;
- have multifactor authorization
- have notification system about access expiration;
- be not expensive.

In order to correspond that requests we create an innovative concept which will combine all the mentioned principles called "RBAC-Q"

4. RBAC-Q, what is needed?

RBAC-Q is based on the next hardware:

- QR – Code scanners or devices which can scan QR – Codes, which can transmit data via network;
- mobile phones with camera;
- high quality server (i.e. Dell PowerEdge);

For RBAC – Q you will need to have next software:

- application for IOS and Android operation systems (mobile agent);
- server-side application;
- application for the workstations (workstation agent).

Also, we need to have network connection between all of the mentioned before elements. And two databases connected to the server-side application.

5. General principles of RBAC-Q software architecture

Mobile agent for the following technology needs to have two main functions: QR-Code scanner and picture transmission option. Every agent needs to have its own identifier connected with server side application.

Table 1. Example of Server Side Database

User	Agent ID	Recovery password	Accesses Granted		
			Service	Login	Password
ilon.mask	12672188	*******	Gmail	il.ma@gmail.com	******
			Dropbox	il.ma@tesla.com	******
			Phycial access	Mobile agent ID	******

Server Side application needs to have QR - Code recognizing and converting technology, QR-Code generator, pseudo-random symbols generator, notification system, recovery system and database (Table 1.). QR - Code recognizing and converting technology is necessary to have because all data in database will be stored in text format. Pseudo-random symbols generator needs to be integrated into the system to generate passwords for users in different organization services. Notification system will inform responsible person about user's password expiration date.

Workstation agent needs to have connection with the server side to get user authorization QR-Code. Also it needs to have data recognition option to define area where should be inserted password for the requested service. In workstation agent should be also integrated ability to generate passwords on server's side, but only for additional resources.

6. RBAC-Q, how it works?

RBAC-Q as a physical access control system will work in the next way:

1. Once user generates QR-Code at his mobile phone. This process requires request to the server to get it.
2. Server sends QR –Code for the authorization based on agent ID.
3. User transmits QR-Code to the server via scanner.
4. Server verifies user's authorization data.
5. After successful verification access for user will be granted.

RBAC – Q as a role base access control system for web services, workstation and network will work in the next way:

1. User gets one-time QR code in his desktop application based on his login in the organization.
2. Scans the QR-code via mobile agent.
3. QR-Code will be automatically transmitted to the server.
4. After successful verification user will get access to the workstation, web-services and network.

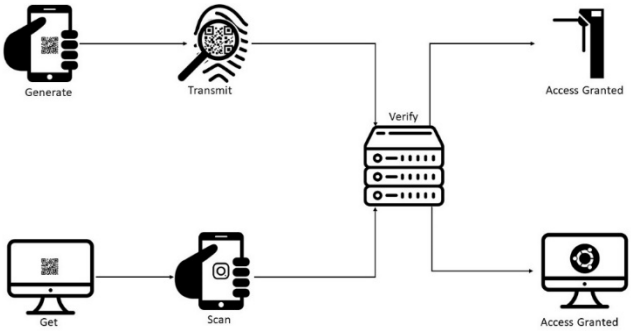


Figure 1. RBAC-Q work

7. Advantages of RBAC-Q system:

Based on information provided in the section 5 was defined that system combines all types of access to the organization. Also you can set up privileges for each user by defining which accesses he should have. RRBAC-Q can have integrated password management system which is not accessible for a user. Technology based on QR-code will have fast data transmission between server and endpoint. Notification system will inform responsible persons about password expiration data.

According to that information we can make a conclusion that RBAC-Q allows organization to have easy managed, high productive, multifunctional and secure access control system.

Acknowledgement

The research had been performed in the framework of International Project: Educating the Next generation experts in Cyber Security: the new EU-recognized Master's program (ENGENSEC)

Project number: 544455-TEMPUS-1-2013-1-SE-TEMPUS-JPCR

REFERENCES

1. Authentication Methods, Enisa, 2017, https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/authentication-methods
2. Password managers, SANS Institute, 2015, https://www.sans.org/
3. Elements of RBAC. Web-page. − https://www.ibm.com/support/knowledge center/ssw_aix_72/com.ibm.aix.security/rbac_elements_of.htm

Andrii SVERSTYUK[1]

Scientific supervisor: Vasyl MARTSENYUK[2]

METODA KONSTUOWANIA STEROWANIA OPTYMALNEGO DLA FAZY HYBRYDYZACJI POLIMERAZOWEJ REAKCJI ŁAŃCUCHOWEJ

Streszczenie: W artykule omówiono ogólną metodę wyznaczania sterowania optymalnego dla fazy hybrydyzacji reakcji łańcuchowej polimerazy. W rozważanym modelu korzystamy z równania Arrheniusa, które uwzględnia zależność prędkości reakcji od temperatury absolutnej. Do rozwiązania problemu istnienia rozwiązania optymalnego sterowania użyto Zasadę Maksimum Pontriagina. Przedstawiono konieczne warunki optymalności. Stosuje się metodę bezpośrednią do numerycznego wyznaczenia problemu sterowania optymalnego z zastosowaniem paczki klas programowych Java. Wynikowe sterowania optymalne przedstawiono w postaci wykresów dla fazy hybrydyzacji. Optymalizacja dotyczy minimalizacji tzw. odcinków starterowych..

Słowa kluczowe: bioinformatyka, polimerazowa reakcja łańcuchowa (PCR), faza odpalania, sterowanie opymalne, metoda bezpośrednia

ON DIRECT METHOD FOR THE CONSTRUCTING THE OPTIMAL CONTROLLER FOR ANNEALING STAGE OF POLYMERASE CHAIN REACTION

Summary: The paper uses a general methodology for solving optimal control problems for annealing stage of the Polymerase Chain Reaction. In the model investigated we use Arrhenius equation, which takes into account the dependence of reaction rate on the absolute temperature. Pontryagin maximum principle is used for the problem of existence of optimal control solution. Necessary conditions for optimality are presented. There is applied a direct method of numerical optimal control problem solving which is implemented as package of Java-classes. Results of optimal control problem solution for PCR annealing stage are presented graphically. Optimal control for PCR annealing stage is required to minimize consumption of the primer at this stage.

Keywords: bioinformatics, polymerase chain reaction (PCR), annealing stage, optimal control, direct method

[1] Państwowy Uniwersytet Medyczny w Tarnopolu, Katedra Informatyki Medycznej, specjalność: informatyka medyczna, email: sverstyuk@tdmu.edu.ua

[2] Prof., Dr hab., Akademia Techniczno-Humanistyczna w Bielsku-Białej, Wydział Budowy Maszyn i Informatyki, email: vmartsenyuk@ath.bielsko.pl

1. Introduction

Polymerase Chain Reaction (PCR) is a method of molecular biology, which is based on a significant increase in the concentration of small pieces of Deoxyribonucleic Acid (DNA) in biological material by amplification. PCR is widely used in biological and medical research for gene cloning, diagnostics of mutations, allocation of new gene sequencing, determination of genetically modified organisms [1, 2].
The reaction of PCR is based on numerous copying (selective amplification) of DNA studied with help of DNA polymerase enzyme. The resulting copies of DNA are identified by the method of electrophoresis.
PCR includes 20-35 cycles [1], each of which consists of three stages.
Double-stranded DNA matrix is heated to 94-96 °C (or 98 °C, especially if there are used thermostable polymerase) for 0.5-10 min until to DNA chains are divided. This stage is called denaturation. The hydrogen bonds between two chains are broken down. Sometimes they do pre-heating of the reaction mixture for 2-5 minutes for complete denaturation of the matrix and primers.
When the chains are broken, they reduce the temperature for primers can contact with single-chain matrix. This stage is called annealing. Annealing temperature depends on the primers and is usually chosen at 4-5 °C below than their melting point. Duration of stage is 0.5-2 min.
DNA polymerase replicates the matrix chain using primer as seed. This is so-called elongation stage. Elongation temperature depends on the polymerase. Taq polymerase and Pfu, which are most widely used, are the most active at 72 °C. Time of elongation depends both on the type of DNA polymerase, and the length of the fragment, which is amplified. The average rate of elongation is 1000 base pairs per 1 min. Upon completion of all cycles they often execute additional stage of final elongation to finish all single-stranded fragments. This stage lasts 10-15 minutes.
Effective PCR should require multistage cyclic regimens of temperature change. Each stage of the cycle (denaturation, annealing, elongation) should occur at certain temperatures and during the relevant period. Otherwise, the necessary changes of DNA can not happen. Fig. 1 shows an example of setting temperature conditions of relevant stages of PCR with help of software for Rotor-GeneTM 6000.
Thus the problem of mathematical modeling is to develop optimal temperature-like controller maximizing amount of one-chained DNA linked with primer and minimizing amount of primer inself.
In many works [3-7] they study different models of PCR, but it requires control problems for PCR models to be constructed and investigated.
So, the objective of this work is to investigate the control problem of PCR annealing stage considering temperature as controller.

2. Model Development

At annealing stage temperature of the mixture is reduced to 55 °C, primers join the single-stranded DNA target.

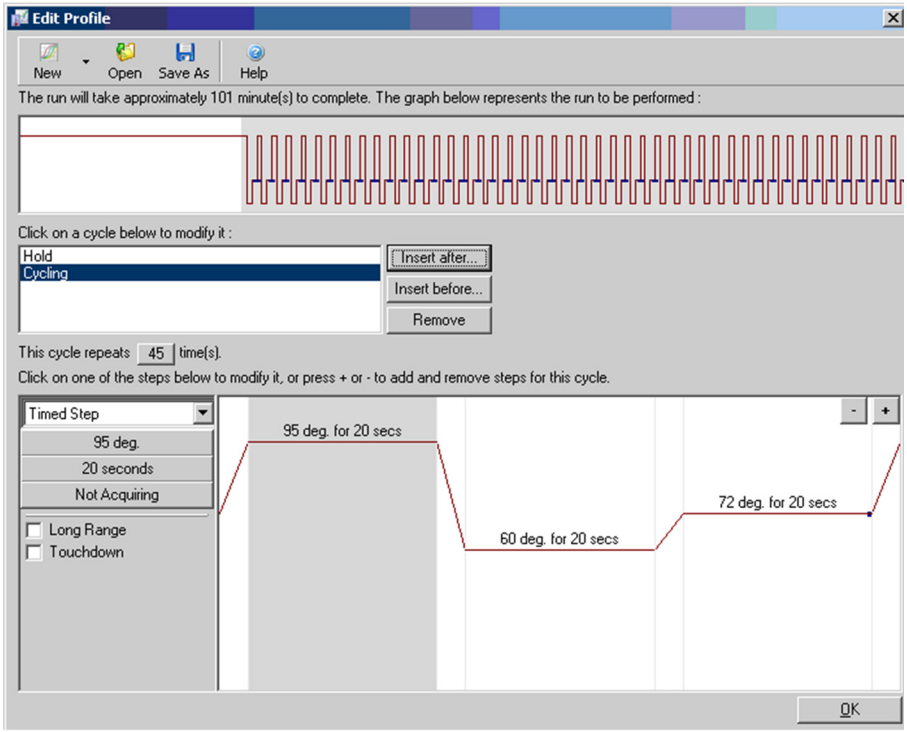


Figure 1. Setting the temperature regimes of PCR stages with help of software of Rotor-GeneTM 6000

A simplified chemical equation describing the process of accession of primer p to the single-stranded DNA s can be represented as

$$s + p \overset{k_1,\, k_{-1}}{\longleftrightarrow} s' \tag{1}$$

As a result of flowing annealing stage there is constructed single-stranded DNA, which is connected to the primer s'. In equation (1) k_1 and k_{-1} are the forward and reverse rate constants for annealing reaction.

Using equation (1) we construct a model of compartment of annealing stage in the following form (Fig. 2).

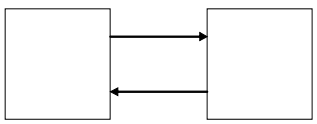


Figure 2. Compartment model of PCR annealing stage

As it is shown in Fig. 2 direct annealing reaction constant rate k_1 promotes the formation of single-stranded DNAs, which are connected with the primer s'. Constant k_{-1} presents reverse reaction of stage studied at which primers from formed earlier single-stranded DNAs which are connected with the primer s' disappear.

Temperature rregimen for PCR annealing stage is chosen so that $k_1 >> k_{-1}$.

The problem of optimal control for PCR annealing stage

A model of PCR annealing stage proposed in [7]

$$\frac{ds}{dt} = -k_1 sp + k_{-1} s \tag{2}$$

$$\frac{dp}{dt} = -k_1 sp + k_{-1} s' \tag{3}$$

$$\frac{ds'}{dt} = -k_1 sp + k_{-1} s' \tag{4}$$

In equations (2-4) s is a single-stranded DNA, p is a primer, s' is bound single-stranded DNA with primer, k_1, k_{-1} are forward and reverse reaction rate for annealing.

In annealing phase of PCR temperature acts as a controller [4, 5]. The dependence of the reaction rate k and the absolute temperature T is described by Arrhenius law [14]:

$$k = Ae^{-E_a / RT}, \tag{5}$$

where A characterizes the frequency of collisions of molecules, R is the universal gas constant, E_a is activation energy.

Based on the Arrhenius equation (5), the system of differential equations for the annealing stage (2) - (4) can be corrected as follows:

$$\frac{ds}{dt} = -k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s' \tag{6}$$

$$\frac{dp}{dt} = -k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s' \tag{7}$$

$$\frac{ds'}{dt} = k_1 e^{-\frac{r}{T}} sp - k_{-1} e^{-\frac{r}{T}} s' \tag{8}$$

with appropriate initial conditions:

$$s(t_1) = s_0,\ p(t_1) = p_0,\ s'(t_1) = s_0. \tag{9}$$

Here we denote $r = \frac{E_a}{R}$ as constant.

We assume that $T = T(t)$ is controller. Let $T(t) \in [T_e^{\min}, T_e^{\max}]$.

For annealing stage the objective is to get as many single-stranded DNA associated with primer s' as we can, while spending as little primer p as it is possible, i.e.

$$J(s, p, s') = \int_{t_1}^{t_2} (s'^2(t) - Wp^2(t))dt \to \inf_{T \in U} \tag{10}$$

Here $W>0$ is weight constant, U is a control set of piecewise continuous functions $T(t) \in [T_e^{\min}, T_e^{\max}]$.

Biologically significant area is

$$\Omega_1 = (s, p, s') \in R_+^3 \tag{11}$$

Thus, the purpose is to determine the optimal controller $T^* \in U$ satisfying:

$$J[T^*] = \inf_{T \in U} J[T] \tag{12}$$

That the optimal controller in problem (6) - (12) exists because function in cost criterion is a convex function and the trajectory of the system belongs to space L^∞ is obvious.

The necessary optimality conditions are obtained with help of Hamilton-Pontryagin function:

$$H = s'^2 - Wp^2 + \lambda_1(-k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s') + \lambda_2(-k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s') + \\ + \lambda_3(k_1 e^{-\frac{r}{T}} sp - k_{-1} e^{-\frac{r}{T}} s') \tag{13}$$

Denote:

$$\Phi(t) = \lambda_1(-k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s') + \lambda_2(-k_1 e^{-\frac{r}{T}} sp + k_{-1} e^{-\frac{r}{T}} s') + \\ + \lambda_3(k_1 e^{-\frac{r}{T}} sp - k_{-1} e^{-\frac{r}{T}} s' \tag{14}$$

Given (14) Hamilton-Pontryagin function is written as:

$$H = s'^2 - Wp^2 + e^{-\frac{r}{T}}[\Phi(t)]. \tag{16}$$

Hence we see that the maximum value of H will be reached at $T = T^*(t)$, where:

$$T^*(t) = \begin{cases} T_e^{\min}, \; if \; \Phi(t) > 0 \\ T_e^{\max}, \; if \; \Phi(t) < 0 \\ any \left[T_e^{\min}, T_e^{\max}\right], \; if \; \Phi(t) = 0 \end{cases} \tag{17}$$

Thus, the optimal trajectory $\left(s^*, \; p^*, \; s'^*\right)$ governed by T^* can be constructed as a result of the boundary value problem including combination of control system and adjoint system (18):

$$\frac{ds}{dt} = -k_1 e^{-\frac{r}{T^*}} sp + k_{-1} e^{-\frac{r}{T^*}} s'$$

$$\frac{dp}{dt} = -k_1 e^{-\frac{r}{T^*}} sp + k_{-1} e^{-\frac{r}{T^*}} s'$$

$$\frac{ds'}{dt} = k_1 e^{-\frac{r}{T^*}} sp - k_{-1} e^{-\frac{r}{T^*}} s'$$

$$\frac{d\lambda_1}{dt} = -\frac{\partial H}{\partial s} = k_1 e^{-\frac{r}{T^*}} p(\lambda_1 + \lambda_2 - \lambda_3)$$

$$\frac{d\lambda_2}{dt} = -\frac{\partial H}{\partial p} = 2Wp + k_1 e^{-\frac{r}{T^*}} s(\lambda_1 + \lambda_2 - \lambda_3)$$

$$\frac{d\lambda_3}{dt} = -\frac{\partial H}{\partial s'} = k_{-1} e^{-\frac{r}{T^*}} (\lambda_3 - \lambda_1 - \lambda_2) - 2s'$$

The boundary conditions:

$$s(t_1) = s_0, \; p(t_1) = p_0, \; s'(t_1) = s'_0;$$
$$\lambda_1(t_2) = 0, \; \lambda_2(t_2) = 0, \; \lambda_3(t_2) = 0.$$

It was proved that for sufficiently small values of $t_2 - t_1$ the solution of problem (18) is unique.

3. Outline of the solution

The numerical method offered implies reduction of infinite-type problem (6) - (8) to a finite optimization problem.

This is achieved by sampling time interval $t \in [t_1, t_2]$ using N nodes t_i such that $t_1 = t_1 < t_2 < \ldots < t_{N-1} = t_f$.

At any given time instant t_i the controller is unknown vector $\overline{T_i} \in R^m$. At each open interval $t \in (t_i, t_{i+1})$, $i = \overline{0, N-2}$ controller can be estimated by linear approximation:

$$T(t) = \overline{T}_i + \frac{t - t_i}{t_{i+1} - t_i}(\overline{T}_{i+1} - \overline{T}_i)$$

An array of controllers in the nodes t_i form a common vector:

$$\tilde{T} = \left[\tilde{T}_0^T, ..., \tilde{T}_{N-1}^T\right]$$

For a given initial approximation $\tilde{T}$ we can integrate system at $t \in [t_1, t_2]$ and get the trajectory $x(t, \tilde{T}, p)$. Thus, infinite dimensional problem (6) - (12) is approximated by finite-dimensional problems of nonlinear programming with respect to $\tilde{T}$, p:

$$J(\tilde{T}, p) = \int_{t_1}^{t_2} L(t, x(t, \tilde{T}, p), \tilde{T}, p)dt + \phi(x(t_f, \tilde{T}, p), p) \to \inf_{\tilde{T}, p}$$

under constraints:

$$\tilde{c} = [c(t_0)^T, ..., c(t_{N-1})^T, \psi^T]^T = 0$$

$$\tilde{d} = [d(t_0)^T, ..., d(t_{N-1})^T, \gamma^T]^T \le 0$$

4. Advantages and disadvantages

Methods of numerical solution of optimal control problems can be classified as direct and indirect [15, 16]. These methods are different approaches to find the solution of optimal control problem. Indirect methods try to solve the boundary problem of necessary optimality conditions. In contrast, direct methods do not require direct construction of the necessary conditions. Direct methods do not build adjoint system, control system and transversality conditions. For the purpose of comparison we investigated solutiob of optimal control problem using both approaches. The main disadvantage of using indirect methods is that even knowing a priori admissible state and control, we are not sure that the computed solution improves known one. Moreover, the indirect method requires initial approximate values for adjoint variables and numerical solution of the adjoint system in practice is poorly conditioned problem [17].

For the reasons given, we used the direct method proposed in [18] for PCR model, which allows us to find numerical solutions for control model (6) - (8).

5. Experimental Research

The direct method of numerical solution of optimal control problem is implemented in package Java-classes dyn.Opt [19]. We use this method in a separate process in try-block:

```
try {
     Process p = Runtime.getRuntime().exec ("java dyn.Opt");
    }
    catch (java.io.IOException ex) {
      System.err.println("Problems invoking class dyn.Opt:
"+ex);
    }
```

As an example, considered problem (6) - (8) can be described using the input text file. So the system state variables are defined through the commands:

```
state s p s_
variable control:
control T
constants:
real kone kminusone r  W
```

the number of time units:

```
nodes = 365
```

method of solving nonlinear programming:

```
method = dyn_sqp
```

method of integrating the system of differential equations:

```
ode = huen
```

file for initial data:

```
output_file = temperaturecontrol
the accuracy of the method:
epsilon = 1.0e-6
```

Control system (6-8) with the values of parameters are described in the block:

```
dynamic_equation:
kone = 0.205
kminusone = 0.01025
r = 0.02
ddt s = -kone*exp(-r/T)*s*p + kminusone*exp(-r/T)*s_
ddt p = -kone*exp(-r/T)*s*p + kminusone*exp(-r/T)*s_
ddt s_ = kone*exp(-r/T)*s*p - kminusone*exp(-r/T)*s_
```

Block of initial conditions (13):

```
initial_condition:
s = 0.002
p = 0.5
s_ = 0
```

Restrictions of inequality type:

```
inequality_constraint:
```

```
d = -T +330    # -T <= 330
d = T-367   # T <= 367
```

Restrictions of equality type:

```
terminal_condition:
psi = s_ - 0.002
psi = s - 0.00002
psi = p  - 0.498
```

Block of cost criterion:

```
cost_functional:
W = 0.2
initial_time = 0.0
final_time = 30
L =s_*s_ - W*p*p
```

The results of solving the problem of nonlinear programming (19) - (21) are shown in Fig. 3-6.

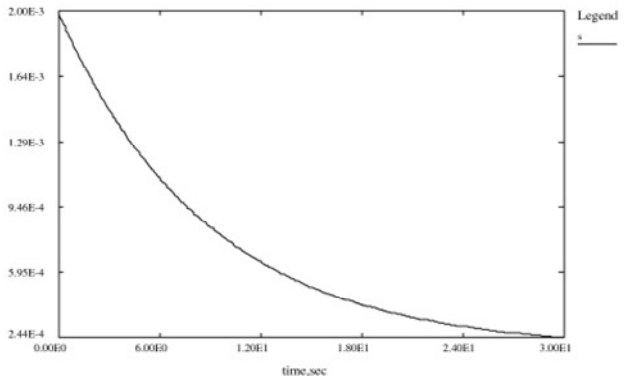


Figure 3. Numerical modeling of optimal control PCR annealing stage, change the number of single-stranded DNA

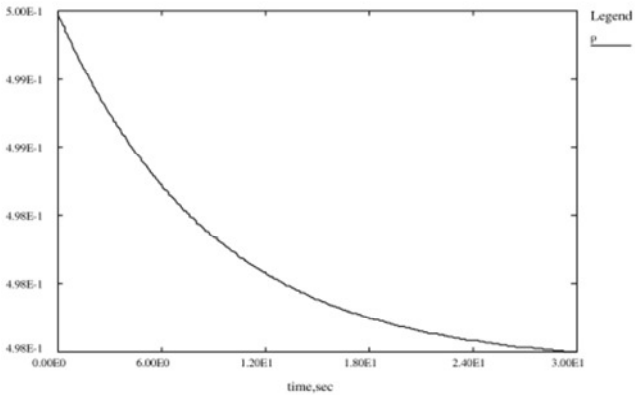


Figure 4. Numerical modeling of optimal control problem under PCR annealing: change of primer

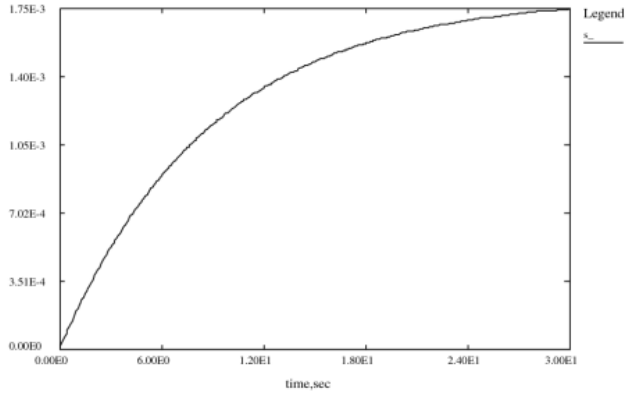


Figure 5. Numerical modeling of optimal control problem under annealing PCR, change the number of single-stranded DNA, which are connected with the primer

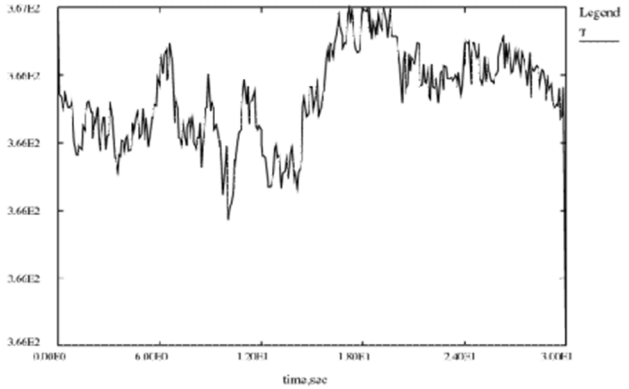


Figure 6. The optimal temperature

Analyzing the results of numerical solution of optimal control problem for PCR annealing stage we can monitor changes of the amount of single-chain DNA primer and changes of the amount of single-stranded DNA, which are connected with the primer for 30 seconds (Fig. 3-6).

With help of the results of numerical solution of optimal control problem for annealing stage of PCR using temperature as a controller (Fig. 6) we receive as many single-stranded DNA that are associated with primer s' as it is possible, while spending as little primer p as it is possible. This dependence helps us also to minimize the time required for implementation of annealing stage and can also be used to development of the new systems for PCR testing.

6. Outcomes

So, the problem of optimal control for annealing stage of PCR was considered. In contrary to previous well-known linear models we offer nonlinear system based on Arrhenius law. It allows us to investigate an influence of temperature applied on the flow of reaction

On the other hand, using such a model we can construct optimal control problem for annealing stage of PCR. Pontryagin maximum principle was applied to get necessary conditions for optimality and, in turn, in order to prove the existence and uniqueness of its solution.

These results were sort of the theoretical background for numerical calculation of optimal controller investigated. For the purpose of numerical solution of optimal control problem we developed and used special so called “direct” computational method.

Scheme for temperature applied that is constructed in such a way, can generally provide the minimal time for PCR. In further investigations we should pay attention on other parameters that can be considered as controllers on different stages of PCR.

Acknowledgment

The author would like to express his gratitude to the reviewer for the valuable comments

REFERENCES

1. WOLFE S.L.: An introduction to cell and molecular biology, University of California at Davis: ITP, 1995, 422-423.
2. SVERSYUK A.S., BIHUNYAK T.V., PEREVIZNYK B.O.: The survey of methods and models of polymerase-chaine reaction. Medical informatics and engineering, 3(2014), 97-100. [in Ukrainian].
3. AACH J., CHURCH G.M.: Mathematical models of diffusion-constrained polymerase chainreactions: basis of high-throughput nucleic acid assays and simple self-organizing systems. Journal of Theoretical Biology, 228(2004), 31-46.
4. PFAFFL M.W.: A new mathematical model for relative quantification in real-time RT–PCR. Oxford Journals Science & Mathematics Nucleic Acids Research, vol. 29, no. 900, 45-51.
5. XIANGCHUN X., SINTON D., DONGQING L. Thermal end effects on electroosmotic flow in capillary. Int. J. of Heat and Mass transfer, 47(2004)14-16, 3145-3157.
6. STONE E., GOLDES J., GARLICK M.: A multi-stage model for quantitative PCR. Mathematical biosciences and engineering, 2000, 1-17.

7. GARLICK M., POWELL J., EYRE D., ROBBINS T.: Mathematically modeling PCR: an asymptotic approximation with potential for optimization, Mathematical biosciences and engineering, 7(2010)2, 363-384.
8. LUKES D.L.: Differential Equations: Classical to Controlled. Academic Press, New York, 1982, vol. 162, 322 p.
9. PICCININI L.C., STAMPACCHIA G., VIDOSSICH G. Ordinary Differential Equations in Rn. Problems and Methods Ordinary. Berlin-Heidelberg-New York-Tokyo, Springer-Verlag, 1984, vol. XII, 385 p.
10. MACKI J., STRAUSS A.: Introduction to Optimal Control Theory. Springer-Verlag, New York, 1982, vol. XIV, 168 p.
11. FLEMING W.H., RISHEL R.W.: Deterministic and Stochastic Optimal Control. Springer Verlag, New York, 1975, vol. XIII, 222 p.
12. KAMIEN M.I., SCHWARTZ N.L.: Dynamic Optimization. North-Holland, Amsterdam, 1991, vol. 3, 272 p.
13. PONTRYAGIN L.S., BOLTYANSKII V.G., GAMKRELIDZE R.V., MISHCHENKO Ye.F.: Mathematical theory of optimal processes. Moscow, 1983, 393 pp. [in Russian]
14. KELLY K., KOSTIN M.: Non-Arrhenius rate constants involving diffusion and reaction. Journal of Chemical Physics; 1986, vol. 85, iss. 12, pp.7318-7335.
15. BETTS J.T.:. Practical Methods for Optimal Control Using Nonlinear Programming Society for Industrial and Applied Mathematics, 2001, 190 p.
16. O. von STRYK, BULIRSCH R.: Direct and indirect methods for trajectory optimization. Annals of Operations Research, 2(1992)37, iss. 1-4, 357-373.
17. BRYSON A.E. JR., HO Yu-Chi.:Applied optimal control. Hal-sted Press, New York, 1975, 481 p.
18. FABIEN B. C.: Some tools for the direct solution of optimal control problems. Advances in Engineering Software, (1998)29, 45-61.
19. MARTSENYUK V.P., SVERSTYUK A.S., GVOZDETSKA I.S. Optimal control problem of the elongation stage in the polymerase chain reaction. System Research and Information Technologies : the International Journal, 4(2015)4, 75-82 [in Ukrainian].

Viktoriia SYDORENKO[1], Tatiana ZHMURKO[2], Yuliia POLISHCHUK[3],

Opiekun naukowy: Sergiy GNATYUK[4]

MODELE DANYCH DO TWORZENIA INFRASTRUKTURY KRYTYCZNEJ ORAZ OKREŚLENIA ICH SPÓJNOŚCI

Streszczenie: W pracy omówiono problem możliwości naruszenia/zniszczenia państwowych systemów krytycznych oraz zasobów (zgromadzonych danych). Zatem, wiodące państwa świata rozwijają metody i środki służące do identyfikacji, systematyzacji oraz zapewnienia bezpieczeństwa obiektów/zasobów stanowiących infrastrukturę krytyczną. Utrata danych lub awaria działania tychże obiektów może spowodować istotne oraz nieodwracalne naruszenia bezpieczeństwa państwa. Jednakże, jak pokazano na podstawie analizy baz krajowych, w dzisiejszej Ukrainie - nie sporządzono wyczerpującej listy obiektów należących do krytycznej infrastruktury informacyjnej kraju. Ponadto, nie ma jasnego mechanizmu oraz zasad jak sformułować taką listę. A zatem, w niniejszym artykule zaproponowano uniwersalny model danych - mający służyć sformułowaniu infrastruktury informacji krytycznej mogącej znaleźć się na państwowej liście obiektów. Ponadto, na takiej podstawie opracowano model oraz listę krytycznych obiektów w dziedzinie lotnictwa cywilnego. W dalszych pracach planuje się opracować efektywne metody oraz narzędzia do identyfikacji oraz rankingu obiektów, a także listę – którą utworzono za pomocą zaproponowanego modelu danych.

Słowa kluczowe: infrastruktura krytyczna, krytyczny system informacji lotniczych, uniwersalny model danych, lotnictwo cywilne

DATA MODEL FOR FORMING CRITICAL INFRASTRUCTURE AND DETERMINING ITS CONNECTIVITY

Summary: state critical systems and resources can be damaged. By this means most of world leader states have attended to methods and means of identifying, systematization and security assurance for critical infrastructure objects. Loss or operational breakdown of these objects can cause significant or irreparably damage for national security of the state. However, as shown by the analysis of the domestic normative base, today in Ukraine an exhaustive list of objects of the critical information infrastructure of the state is not yet formed and there is no clear mechanism for the formation of this list. Given the above, the paper proposes a universal data

[1] National Aviation University, IT-security Academaic Depatment, v.sydorenko@ukr.net
[2]PhD, National Aviation University, IT-security Academaic Depatment, t.zhmurko@nau.edu.ua
[3] National Aviation University, IT-security Academaic Depatment, liya7954@gmail.com
[4] PhD, National Aviation University, IT-security Academaic Depatment, s.gnatyuk@nau.edu.ua

model for the formation of critical information infrastructure of the state objects list and, on the basis of the developed model, a list of critical objects in the field of civil aviation is formed. In further works is planned to develop efficient methods and tools for identifying and ranking objects, the list of which is formed using the proposed model of data.

Keywords: critical infrastructure, critical information infrastructure, critical aviation information systems, universal data model, civil aviation.

1. Introduction

Modern society entirely depends on information-communication systems and networks, the failure of which, can lead to chaos, significant financial losses and even mass deaths of people. However, the majority of humanity tends to take the most important services (in particular, their quality) as a matter, until something or somebody disturbs their work. To determine and generalize the most important and most vulnerable of state assets, relatively recently, the term critical infrastructure (CI) [1] was introduced into international law. Typically, this category relates to energy and transmission line, oil and gas line, seaports, high-speed and government communications channels, life-saving systems of megacities, high-tech enterprises and enterprises of the military-industrial complex, and also the central government authority. Recently, the issue of the objects and protection the safety of CI in general (at the state and international levels) became relevant.

2. Existing research analysis and problem statement

In the second half of the 1990s, the concepts of CI began to use in the relatively largely distributed large-scale information systems (data centers, united communications networks, etc.) [2]. Most developed countries made attempts to define the CI and develop a strategy for its protection independently. According to [3], the list of vital (critical) infrastructures is different for individual states and is determined according to their traditions, social and political beliefs, as well as the geographical and historical features of each state. An important component of the CI is its informational structure (i.e. critical information infrastructure, CII), whose concept of protection was first developed in the United States, and subsequently developed and adapted in pro-state states of the world [3-5]. According to [6], the CI for any state – is a large complex system of strategic scale, which is a significant number combination of different types elements, united by ties of different nature and has a common property (purpose, function), which different from the properties of individual elements of the whole population. Conducted in [1] the domestic normative base analysis shows that the sphere of protection the CII for our state is at the initial stage of formation. Although the existing domestic legislation specifies certain objects of the socio-economic sphere of Ukraine, extraordinary events that may lead to socially dangerous consequences, but they do not constitute a unified system [4].

According to [7] in Ukraine is ongoing the development of proposals for the formation the list of information-telecommunication systems of CI objects for the state. According to [7], the object of CI –is an enterprises and institutions (regardless of

ownership) of such industries as energy, chemical industry, transport, banks and finance, information technology and telecommunications (electronic communications), food, health care, communal-non-economy, which are strategically important for the functioning of the economy and the security of the state, society and population.

Among other branches of CI, civil aviation (CA) of the state needs special protection, where according to the guidance documents in this area (in particular [8]) it is necessary to identify and protect critical aviation information systems (CAIS). It is obvious, that unauthorized interference in transport system operation can lead to significant economic losses, human casualties and the destruction of nation-wide infrastructure. Operating conditions of CI rapidly and significantly change with the introduction of modern technologies for the processing, transmission and preservation of information, which increase the level of protection and simplification of formalities [9]. Consequently, the provision of CAIS protection is a general requirement for every state that is and wants to be part of the international aviation community. However, none of the ICAO or ECAC guiding documents (on the protection of international CA) includes a complete list of CAIS, which complicates the development of effective methods for protecting CAIS from various types of cyber threats. In the paper [10], was conducted the search and systematization of modern CAIS, divided into categories, analyzed their properties and basic characteristics. However, the unlimited number of objects and parameters of systems that are constantly changing, and the difficultly predicted behavior of objects with a large number of interconnections, are the main reasons for the difficulties of displaying objects of state authority (in particular, in the sphere of CA). According to this, the main purpose of work is to develop a data model for the establishment of CII object list for state.

3. Study part

According to [7], a full set (range) of CII systems categories in a particular industry was introduced:

$$\mathbf{S} = \{ \overset{n}{\underset{i=1}{}} \mathbf{S}_i \} = \{\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_n\}, \tag{1}$$

where $\mathbf{S}_i \subseteq \mathbf{S}$ $(i = \overline{1,n})$ – systems categories in some CI industry, n – total number of system categories.

Consider an example of the formation the list of CII objects for the CA industry (based on the CAIS system) in accordance to [10], $n=3$ taking into account (1), define the set of systems categories:

$$\mathbf{S}_{\mathbf{KAIS}} = \{ \overset{3}{\underset{\mathbf{i}=1}{}} \mathbf{S}_\mathbf{i} \} = \{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3\} =$$

$$= \{\mathbf{S}_{\mathbf{ISAO}}, \mathbf{S}_{\mathbf{BSPS}}, \mathbf{S}_{\mathbf{ISAA}}\} = \{\mathbf{ISAO}, \mathbf{BSPS}, \mathbf{ISAA}\},$$

where . $\mathbf{S}_1 = \mathbf{S}_{\mathbf{ISAO}} = \mathbf{ISAO}$. – set of aeronautical information system; $\mathbf{S}_2 = \mathbf{S}_{\mathbf{BSPS}} = \mathbf{BSPS}$ – set of on-board information systems for aircraft; $\mathbf{S}_3 = \mathbf{S}_{\mathbf{ISAA}} = \mathbf{ISAA}$ – a set of information systems of airlines and airports in accordance with [10].

Important, that each set can be represented in three forms: set with an index (I), for example $\mathbf{S}_1$; set with an index of the object name (IIO), for example $\mathbf{S}_{\mathbf{ISAO}}$, and set with of the object name (IO), for example $\mathbf{ISAO}$.
The set of categories $\mathbf{S}_i$ can be represented as the set of system:

$$\mathbf{S}_i = \{\overset{m_i}{\underset{j=1}{}} \mathbf{S}_{ij}\} = \{\mathbf{S}_{i1}, \mathbf{S}_{i2}, ..., \mathbf{S}_{im_i}\}, \tag{2}$$

where $\mathbf{S}_{ij} \subseteq \mathbf{S}_i$ $(i = \overline{1,n}$, $j = \overline{1,m_i})$ – system of i -th category, m_i – number of system i -th category.
Accordingly with (2), expression (1) can be presented such as:

$$\mathbf{S} = \{\overset{n}{\underset{i=1}{}} \mathbf{S}_i\} = \{\overset{n}{\underset{i=1}{}} \{\overset{m_i}{\underset{j=1}{}} \mathbf{S}_{ij}\}\} = \{\{\mathbf{S}_{11}, \mathbf{S}_{12}, ..., \mathbf{S}_{1m_1}\},$$
$$\{\mathbf{S}_{21}, \mathbf{S}_{22}, ..., \mathbf{S}_{2m_2}\}, ..., \{\mathbf{S}_{n1}, \mathbf{S}_{n2}, ..., \mathbf{S}_{nm_n}\}\}, (i = \overline{1,n}, j = \overline{1,m_i}).$$

Systematized data representation of set systems i -x category could be displayed using tab. 1, where I and IO are the type of representation the sets with an index or with the object name, respectively.

Table 1. Presentation the sets of i category systems

Sets of category (I) $\mathbf{S}_i$ $(i = \overline{1,n})$	Sets of category (IO) $\mathbf{S}_i$ $(i = \overline{1,n})$	Number of system category - i $(i = \overline{1,n})$	Number of system i category - j $(j = \overline{1,m_i})$	Sets of system (I) $\mathbf{S}_{ij}$ $(i = \overline{1,n}, j = \overline{1,}$	Sets of system (IO) $\mathbf{S}_{ij}$ $(i = \overline{1,n}, j = \overline{1,}$
$\mathbf{S}_i$	$\mathbf{S}_i$	i	m_i	$\mathbf{S}_{ij}$	$\mathbf{S}_{ij}$

For example, for the $\mathbf{S}_1$ category set, if $n = 1$, $m_1 = 5$ using the (2), represent the set of systems such as:

$$\mathbf{S}_1 = \mathbf{S}_{\text{ISAO}} = \mathbf{ISAO} = \{\overset{5}{\underset{j=1}{}} \mathbf{S}_{1j}\} = \{\mathbf{S}_{1.1}, \mathbf{S}_{1.2}, \mathbf{S}_{1.3}, \mathbf{S}_{1.4}, \mathbf{S}_{1.5}\} =$$
$$= \{\mathbf{S}_{\text{SAE}}, \mathbf{S}_{\text{RZZP}}, \mathbf{S}_{\text{SSP}}, \mathbf{S}_{\text{SOD}}, \mathbf{S}_{\text{SMZ}}\} =$$
$$= \{\mathbf{SAE}, \mathbf{RZZP}, \mathbf{SSP}, \mathbf{SOD}, \mathbf{SMZ}\},$$

where $\mathbf{S}_{1.1} = \mathbf{S}_{\text{SAE}} = \mathbf{SAE}$ – aviation telecommunication systems; $\mathbf{S}_{1.2} = \mathbf{S}_{\text{RZZP}} = \mathbf{RZZP}$ – radio navigational aids of flight operations; $\mathbf{S}_{1.3} = \mathbf{S}_{\text{SSP}} = \mathbf{SSP}$ – surveillance systems; $\mathbf{S}_{1.4} = \mathbf{S}_{\text{SOD}} = \mathbf{SOD}$ – data processing systems; $\mathbf{S}_{1.5} = \mathbf{S}_{\text{SMZ}} = \mathbf{SMZ}$ – meteorological support systems [10].
Similarly, for the set of categories $\mathbf{S}_2$ if $n = 2$, $m_2 = 7$ using (2), represent the set of systems such as:

$$\mathbf{S}_2 = \mathbf{S}_{\text{BSPS}} = \mathbf{BSPS} = \{\overset{7}{\underset{j=1}{}} \mathbf{S}_{1j}\} = \{\mathbf{S}_{2.1}, \mathbf{S}_{2.2}, \mathbf{S}_{2.3}, \mathbf{S}_{2.4}, \mathbf{S}_{2.5}, \mathbf{S}_{2.6}, \mathbf{S}_{2.7}\} =$$
$$= \{\mathbf{S}_{\text{SPS}}, \mathbf{S}_{\text{SZV}}, \mathbf{S}_{\text{NAVS}}, \mathbf{S}_{\text{SSPZ}}, \mathbf{S}_{\text{OSL}}, \mathbf{S}_{\text{SVI}}, \mathbf{S}_{\text{ABSK}}\} =$$
$$= \{\mathbf{SPS}, \mathbf{SZV}, \mathbf{NAVS}, \mathbf{SSPZ}, \mathbf{OSL}, \mathbf{SVI}, \mathbf{ABSK}\},$$

where $\mathbf{S}_{2.1} = \mathbf{S}_{SPS} = \mathbf{SPS}$ – air data system; $\mathbf{S}_{2.2} = \mathbf{S}_{SZV} = \mathbf{SZV}$ – communication systems; $\mathbf{S}_{2.3} = \mathbf{S}_{NAVS} = \mathbf{NAVS}$ – navigation systems; $\mathbf{S}_{2.4} = \mathbf{S}_{SSPZ} = \mathbf{SSPZ}$ – observing and collision avoidance systems; $\mathbf{S}_{2.5} = \mathbf{S}_{OSL} = \mathbf{OSL}$ – computer aircrafts systems; $\mathbf{S}_{2.6} = \mathbf{S}_{SVI} = \mathbf{SVI}$ – information display system; $\mathbf{S}_{2.7} = \mathbf{S}_{ABSK} = \mathbf{ABSK}$ – automatic on-board control systems [10].

Similarly, for the set of categories $\mathbf{S}_3$, if $n = 3,\ \ m_3 = 5$ using (2), represent the set of systems such as:

$$\mathbf{S}_3 = \mathbf{S}_{ISAA} = \mathbf{ISAA} = \{\overset{5}{\underset{j=1}{}} \mathbf{S}_{1j}\} = \{\mathbf{S}_{3.1}, \mathbf{S}_{3.2}, \mathbf{S}_{3.3}, \mathbf{S}_{3.4}, \mathbf{S}_{3.5}\} =$$

$$= \{\mathbf{S}_{CRS}, \mathbf{S}_{GDS}, \mathbf{S}_{IDS}, \mathbf{S}_{BSP}, \mathbf{S}_{DCS}\} = \{\mathbf{CRS}, \mathbf{GDS}, \mathbf{IDS}, \mathbf{BSP}, \mathbf{DCS}\},$$

where $\mathbf{S}_{3.1} = \mathbf{S}_{CRS} = \mathbf{CRS}$ – computer distribution system; $\mathbf{S}_{3.2} = \mathbf{S}_{GDS} = \mathbf{GDS}$ – global distribution system (booking); $\mathbf{S}_{3.3} = \mathbf{S}_{IDS} = \mathbf{IDS}$ – Internet Distribution Systems (IDS) or Alternative Distribution Systems (ADS); $\mathbf{S}_{3.4} = \mathbf{S}_{BSP} = \mathbf{BSP}$ – settlement system; $\mathbf{S}_{3.5} = \mathbf{S}_{DCS} = \mathbf{DCS}$ – dispatch control systems in accordance with [10].

Representation the sets of systems of i CAIS category, according to Table 1 is shown in the Table. 2

Table 2. Representation the sets of systems of i CAIS category

Sets of category (I) $\mathbf{S}_i\ (i = \overline{1,})$	Sets of category (IO) $\mathbf{S}_i\ (i = \overline{1,n})$	Number of system category - i $(i = \overline{1,n})$	Number of system category- j $(j = \overline{1,m_i})$	Sets of system (I) $\mathbf{S}_{ij}\ (i = \overline{1,n},\ j = \overline{1,m_i})$	Sets of system (IO) $\mathbf{S}_{ij}\ (i = \overline{1,n},\ j = \overline{1,m_i})$
$\mathbf{S}_1$	**ISAO**	*1*	$m_1 = 5$	$\mathbf{S}_{1.1}, \mathbf{S}_{1.2}, \mathbf{S}_{1.3}, \mathbf{S}_{1.4}, \mathbf{S}_{1.5}$	**SAE, RZZP, SSP, SOD, SMZ**
$\mathbf{S}_2$	**BSPS**	*2*	$m_2 = 7$	$\mathbf{S}_{2.1}, \mathbf{S}_{2.2}, \mathbf{S}_{2.3}, \mathbf{S}_{2.4}, \mathbf{S}_{2.5}, \mathbf{S}_{2.6}, \mathbf{S}_{2.7}$	**SPS, SZV, NAVS, SSPZ, OSL, SVI, ABSK**
$\mathbf{S}_3$	**ISAA**	*3*	$m_3 = 5$	$\mathbf{S}_{3.1}, \mathbf{S}_{3.2}, \mathbf{S}_{3.3}, \mathbf{S}_{3.4}, \mathbf{S}_{3.5}$	**CRS, GDS, IDS, BSP, DCS**

The sets of systems $\mathbf{S}_{ij}$ can be represented as the set of subsystem:

$$\mathbf{S}_{ij} = \{\overset{r_{ij}}{\underset{k=1}{}} S_{ijk}\} = \{S_{ij1}, S_{ij2}, ..., S_{ijr_{ij}}\}, \tag{4}$$

where $S_{ijk} \subseteq \mathbf{S}_{ij}\ \ (i = \overline{1,n},\ \ j = \overline{1,m_i},\ \ k = \overline{1,r_{ij}})$ – set of subsystem of system $\mathbf{S}_{ij}$, r_{ij} – subsystem number of ij system.

Due to (4), the expression (3) can be represented as follows:

$$\mathbf{S}=\{\overset{n}{\underset{i=1}{}}\mathbf{S}_i\}=\{\overset{n}{\underset{i=1}{}}\{\overset{m_i}{\underset{j=1}{}}\mathbf{S}_{ij}\}\}=\{\overset{n}{\underset{i=1}{}}\{\overset{m_i}{\underset{j=1}{}}\{\overset{r_{ij}}{\underset{k=1}{}}S_{ijk}\}\}\}=$$
$$=\{\{\{S_{111},S_{112},...,S_{11r_{11}}\},\{S_{121},S_{122},...,S_{12r_{12}}\},...,\{S_{1m_11},S_{1m_12},...,S_{1m_1r_{1m_1}}\}\},$$
$$\{\{S_{211},S_{212},...,S_{21r_{21}}\},\{S_{221},S_{222},...,S_{22r_{22}}\},...,\{S_{2m_21},S_{2m_22},...,S_{2m_2r_{1m_2}}\}\},...,$$
$$\{\{S_{n11},S_{n12},...,S_{n1r_{n1}}\},\{S_{n21},S_{n22},...,S_{n2r_{n2}}\},...,\{S_{nm_n1},S_{nm_n2},...,S_{nm_nr_{nm_n}}\}\}\}. \quad (5)$$

Systematized data representation of set subsystems $\ddot{y}$ system can be displayed using Table 1, the set of subsystem labeled as sets elements and are the lower level of system detail.

Table 3. Presentation the sets of subsystem $\ddot{y}$ systems

Sets of system (I) $\mathbf{S}_{ij}\ (i=\overline{1,n},\ j=\overline{1,m_i})$	Sets of system (IO) $\mathbf{S}_{ij}\ (i=\overline{1,n},\ j=\overline{1,m_i})$	Number of subsystem $\ddot{y}$ system - k $\left(k=\overline{1,r_{ij}}\right)$	Sets of subsystem (I) $S_{ijk}\ (i=\overline{1,n},\ j=\overline{1,m_i},\ k=$	Sets of subsystem (IO) $S_{ijk}\ (i=\overline{1,n},\ j=\overline{1,}$
$\mathbf{S}_{ij}$	$\mathbf{S}_{ij}$	r_{ij}	S_{ijk}	S_{ijk}

For example, for the set $\mathbf{S}_{1.1}$, if $n=1,\ m_1=1,\ r_{1.1}=5,$ using (4), represent the set of subsystems such as:

$$\mathbf{S}_{1.1}=\mathbf{S}_{\mathrm{SAE}}=\mathbf{SAE}=\{\overset{5}{\underset{k=1}{}}S_{1.1.k}\}=\{S_{1.1.1},S_{1.1.2},S_{1.1.3},S_{1.1.4},S_{1.1.5}\}=$$
$$=\{S_{\mathrm{SAPE}},S_{\mathrm{SANE}},S_{\mathrm{ZAR}},S_{\mathrm{SASZ}},S_{\mathrm{MTM}}\}=$$
$$=\{SAPE,SANE,ZAR,SASZ,MTM\},$$

where $S_{1.1.1}=S_{\mathrm{SAPE}}=SAPE$ – air aviation telecommunication systems; $S_{1.1.2}=S_{\mathrm{SANE}}=SANE$ – systems and networks of aviation ground telecommunication; $S_{1.1.3}=S_{\mathrm{ZAR}}=ZAR$ – tools of aviation broadcasting; $S_{1.1.4}=S_{\mathrm{SASZ}}=SASZ$ – systems of aviation satellite communication; $S_{1.1.5}=S_{\mathrm{MTM}}=MTM$ – magistral telecommunication networks in accordance with [10].

Similarly, for the set $\mathbf{S}_{1.2}$, if $n=1,\ m_1=2,\ r_{1.2}=4,\ \mathbf{S}_3$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{1.2}=\mathbf{S}_{\mathrm{RZZP}}=\mathbf{RZZP}=\{\overset{4}{\underset{k=1}{}}S_{1.2.k}\}=\{S_{1.2.1},S_{1.2.2},S_{1.2.3},S_{1.2.4}\}=$$
$$=\{S_{\mathrm{NDB}},S_{\mathrm{VOR}},S_{\mathrm{DME}},S_{\mathrm{ILS}}\}=\{NDB,VOR,DME,ILS\},$$

where $S_{1.2.1}=S_{\mathrm{NDB}}=NDB$ – Non-Directional Beacons (NDB); $S_{1.2.2}=S_{\mathrm{VOR}}=VOR$ – Very High Frequency Omni-Directional Range (VOR); $S_{1.2.3}=S_{\mathrm{DME}}=DME$ – Distance Measuring Equipment (DME); $S_{1.2.4}=S_{\mathrm{ILS}}=ILS$ – Instrument Landing Systems (ILS) in accordance with [10].

Similarly, for the set $\mathbf{S}_{1.3}$, if $n=1,\ m_1=3,\ r_{1.3}=9,$ using (4), represent the set of subsystems such as:

$$\mathbf{S}_{1.3} = \mathbf{S}_{\text{SSP}} = \mathbf{SSP} = \{\overset{9}{\underset{k=1}{}} S_{1.3.k}\} =$$
$$= \{S_{1.3.1}, S_{1.3.2}, S_{1.3.3}, S_{1.3.4}, S_{1.3.5}, S_{1.3.6}, S_{1.3.7}, S_{1.3.8}, S_{1.3.9}\} =$$
$$= \{S_{\text{PSR}}, S_{\text{SSR}}, S_{\text{MSSR}}, S_{\text{RADS}}, S_{\text{SMR}}, S_{\text{WRAD}}, S_{\text{MLAT}}, S_{\text{ADS}}, S_{\text{DF}}\} =$$
$$= \{PSR, SSR, MSSR, RADS, SMR, WRAD, MLAT, ADS, DF\},$$

where $S_{1.3.1} = S_{\text{PSR}} = PSR$ – Primary Surveillance Radars (PSR); $S_{1.3.2} = S_{\text{SSR}} = SSR$ – Secondary Surveillance Radars (SSR); $S_{1.3.3} = S_{\text{MSSR}} = MSSR$ – Monopulse Secondary Surveillance Radars (MSSR); $S_{1.3.4} = S_{\text{RADS}} = RADS$ – Radar Sites (PSR+ SSR); $S_{1.3.5} = S_{\text{SMR}} = SMR$ – Surface Movement Radars (SMR); $S_{1.3.6} = S_{\text{WRAD}} = WRAD$ – Weather Radars; $S_{1.3.7} = S_{\text{MLAT}} = MLAT$ – Multilateration Systems (MLAT); $S_{1.3.8} = S_{\text{ADS}} = ADS$ – Automatic Dependent Surveillance (ADS); $S_{1.3.9} = S_{\text{DF}} = DF$ – Direction Finders (DF) in accordance with [10].

Similarly, for the set $\mathbf{S}_{1.4}$, if $n = 1$, $m_1 = 4$, $r_{1.4} = 5$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{1.4} = \mathbf{S}_{\text{SOD}} = \mathbf{SOD} = \{\overset{5}{\underset{k=1}{}} S_{1.4.k}\} = \{S_{1.4.1}, S_{1.4.2}, S_{1.4.3}, S_{1.4.4}, S_{1.4.5}\} =$$
$$= \{S_{\text{ASYPR}}, S_{\text{SPPP}}, S_{\text{ESAN}}, S_{\text{SOPD}}, S_{\text{SOPA}}\} =$$
$$= \{ASYPR, SPPP, ESAN, SOPD, SOPA\},$$

where $S_{1.4.1} = S_{\text{ASYPR}} = ASYPR$ – automated air traffic control systems; $S_{1.4.2} = S_{\text{SPPP}} = SPPP$ – automated airspace planning systems; $S_{1.4.3} = = S_{\text{ESAN}} = ESAN$ – European Organization for the Safety of Air Navigation; $S_{1.4.4} = S_{\text{SOPD}} = SOPD$ flight data processing and transfer systems; $S_{1.4.5} = = S_{\text{SOPA}} = SOPA$ – systems for processing and transmitting aeronautical information in accordance with [10].

Similarly, for the set $\mathbf{S}_{1.5}$, if $n = 1$, $m_1 = 5$, $r_{1.5} = 3$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{1.5} = \mathbf{S}_{\text{SMZ}} = \mathbf{SMZ} = \{\overset{3}{\underset{k=1}{}} S_{1.5.k}\} = \{S_{1.5.1}, S_{1.5.2}, S_{1.5.3}\} =$$
$$= \{S_{\text{SCMAU}}, S_{\text{KRAMS}}, S_{\text{SADIS}}\} = \{SCMAU, KRAMS, SADIS\},$$

where $S_{1.5.1} = S_{\text{SCMAU}} = SCMAU$ – Centralized Meteorological System for UkSATSE; $S_{1.5.2} = S_{\text{KRAMS}} = KRAMS$ – comprehensive radio-aerodrome meteorological stations; $S_{1.5.3} = = S_{\text{SADIS}} = SADIS$ – Satellite Distribution System for Information Relating to Air Navigation (SADIS) in accordance with [10].

Similarly, for the set $\mathbf{S}_{2.1}$, if $n = 2$, $m_2 = 1$, $r_{2.1} = 4$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.1} = \mathbf{S}_{\text{SPS}} = \mathbf{SPS} = \{\overset{4}{\underset{k=1}{}} S_{2.1.k}\} = \{S_{2.1.1}, S_{2.1.2}, S_{2.1.3}, S_{2.1.4}\} =$$
$$= \{S_{DPPT}, S_{\text{DZP}}, S_{\text{TPT}}, S_{\text{POP}}\} = \{DPPT, DZP, TPT, POP\},$$

where $S_{2.1.1} = S_{DPPT} = DPPT$ – sensors-receivers of air pressure; $S_{2.1.2} = S_{\text{DZP}} = DZP$ – retardant flow sensors; $S_{2.1.3} = S_{\text{TPT}} = TPT$ – pressure transducer sensors; $S_{2.1.4} = S_{\text{POP}} = POP$ – hardware of processing and transformation the information into electrical signals in accordance with [11].

Similarly, for the set $\mathbf{S}_{2.2}$, if $n = 2$, m_2=2, r_{22}=3, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.2} = \mathbf{S}_{\text{SZV}} = \mathbf{SZV} = \{\overset{3}{\underset{k=1}{}} S_{2.2.k}\} = \{S_{2.2.1}, S_{2.2.2}, S_{2.2.3}\} =$$
$$= \{S_{BRS}, S_{\text{CPDLS}}, S_{\text{AKARS}}\} = \{BRS, CPDLS, AKARS\},$$

where $S_{2.2.1} = S_{BRS} = BRS$ – on-board radio stations; $S_{2.2.2} = = S_{\text{CPDLS}} = CPDLS$ – data transmission facility CPDLS; $S_{2.2.3} = S_{\text{AKARS}} = AKARS$ – data transmission facility ACARS in accordance with [10].

Similarly, for the set $\mathbf{S}_{2.3}$, if $n = 2$, m_2=3, r_{23}=8, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.3} = \mathbf{S}_{\text{NAVS}} = \mathbf{NAVS} = \{\overset{8}{\underset{k=1}{}} S_{2.3.k}\} =$$
$$= \{S_{2.3.1}, S_{2.3.2}, S_{2.3.3}, S_{2.3.4}, S_{2.3.5}, S_{2.3.6}, S_{2.3.7}, S_{2.3.8}\} =$$
$$= \{S_{\text{SNS}}, S_{\text{ISN}}, S_{\text{ARK}}, S_{\text{RV}}, S_{\text{BVOR}}, S_{\text{BD}}, S_{\text{BILS}}, S_{\text{DVKZ}}\} =$$
$$= \{SNS, INS, ARK, RV, BVOR, BD, BILS, DVKZ\},$$

where $S_{2.3.1} = S_{\text{SNS}} = SNS$ – satellite navigation systems; $S_{2.3.2} = S_{\text{ISN}} = INS$ – inertial navigation system; $S_{2.3.3} = S_{\text{ARK}} = ARK$ – automatic direction finder; $S_{2.3.4} = S_{\text{RV}} = RV$ – radioaltimeter; $S_{2.3.5} = S_{\text{BVOR}} = BVOR$ – on-board equipment of VOR system; $S_{2.3.6} = S_{\text{BD}} = BD$ – on-board rangefinders; $S_{2.3.7} = S_{\text{BILS}} = BILS$ – on-board equipment of ILS system; $S_{2.3.8} = S_{\text{DVKZ}} = DVKZ$ – Doppler velocity and drift angle gauge in accordance with [10].

Similarly, for the set $\mathbf{S}_{2.4}$, if $n = 2$, $m_2 = 4$, $r_{2.4} = 4$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.4} = \mathbf{S}_{\text{SSPZ}} = \mathbf{SSPZ} = \{\overset{4}{\underset{k=1}{}} S_{2.4.k}\} = \{S_{2.4.1}, S_{2.4.2}, S_{2.4.3}, S_{2.4.4}\} =$$
$$= \{S_{\text{TRA}}, S_{\text{TCAS}}, S_{\text{SRPZ}}, S_{\text{BMR}}\} = \{TRA, TCAS, SRPZ, BMR\},$$

where $S_{2.4.1} = S_{TRA} = TRA$ – transponders; $S_{2.4.2} = = S_{TCAS} = TCAS$ – on-board collision avoidance systems (TCAS); $S_{2.4.3} = S_{SRPZ} = SRPZ$ – the system of the early alternation of non-heaped lands with land; $S_{2.4.4} = S_{BMR} = BMR$ – n-board meteorological radar in accordance with [10].

Similarly, for the set $\mathbf{S}_{2.5}$, if $n = 2$, m_2=5, r_{25}=2, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.5} = \mathbf{S}_{\text{OSL}} = \mathbf{OSL} = \{\overset{2}{\underset{k=1}{}} S_{2.5.k}\} = \{S_{2.5.1}, S_{2.5.2}\} =$$
$$= \{S_{OBCH}, S_{MCDU}\} = \{OBCH, MCDU\},$$

where $S_{2.5.1} = S_{OBCH} = OBCH$ – calculators; $S_{2.5.2} = S_{MCDU} = MCDU$ – Multifunction Control and Dis play Unit — MCDU) in accordance with [12].

Similarly, for the set $\mathbf{S}_{2.6}$, if $n = 2$, $m_2 = 6$, $r_{26} = 5$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.6} = \mathbf{S}_{\text{SVI}} = \mathbf{SVI} = \{\overset{5}{\underset{k=1}{}} S_{2.6.k}\} = \{S_{2.6.1}, S_{2.6.2}, S_{2.6.3}, S_{2.6.4}, S_{2.6.5}\} =$$
$$= \{S_{DBSVS}, S_{DSVS}, S_{OSVS}, S_{\text{NSVSO}}, S_{ISVS}\} =$$
$$= \{DBSVS, DSVS, OSVS, NSVSO, ISVS\},$$

where $S_{2.6.1} = S_{DBSVS} = DBSVS$ – sensors and databases of the SVS (Synthetic Vision System) system; $S_{2.6.2} = S_{DSVS} = DSVS$ – SVS displays; $S_{2.6.3} = S_{OSVS} = OSVS$ – SVS calculators; $S_{2.6.4} = S_{\text{NSVSO}} = NSVSO$ – required SVS equipment; $S_{2.6.5} = S_{ISVS} = ISVS$ – other systems for displaying information in accordance with [13].

Similarly, for the set $\mathbf{S}_{2.7}$, if $n = 2$, $m_2 = 7$, $r_{27} = 4$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{2.7} = \mathbf{S}_{\text{ABSK}} = \mathbf{ABSK} = \{\overset{4}{\underset{k=1}{}} S_{2.7.k}\} = \{S_{2.7.1}, S_{2.7.2}, S_{2.7.3}, S_{2.7.4}\} =$$
$$= \{S_{APIL}, S_{SAU}, S_{\text{PILS}}, S_{\text{PNK}}\} = \{APIL, SAU, PILS, PNK\},$$

where $S_{2.7.1} = S_{APIL} = APIL$ – autopilots; $S_{2.7.2} = S_{SAU} = SAU$ – automatic control systems; $S_{2.7.3} = S_{\text{PILS}} = PILS$ – flight systems; $S_{2.7.4} = S_{\text{PNK}} = PNK$ – aerodrome navigation systems in accordance with [14].

Similarly, for the set $\mathbf{S}_{3.1}$, if $n = 3$, $m_3 = 1$, $r_{3.1} = 2$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{3.1} = \mathbf{S}_{\text{CRS}} = \mathbf{CRS} = \{\overset{2}{\underset{k=1}{}} S_{3.1.k}\} = \{S_{3.1.1}, S_{3.1.2}\} =$$
$$= \{S_{\text{DELTM}}, S_{\text{PANAM}}\} = \{DELTM, PANAM\},$$

where $S_{3.1.1} = S_{\text{DELTM}} = DELTM$ – unified Deltamatic system of «Delta» company; $S_{3.1.2} = S_{\text{PANAM}} = PANAM$ – unified Pana-Mac system of «Pan Am» company in accordance with [15].

Similarly, for the set $\mathbf{S}_{3.2}$, if $n = 3$, $m_3 = 2$, $r_{3.2} = 18$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{3.2} = \mathbf{S}_{\text{GDS}} = \mathbf{GDS} = \{\overset{18}{\underset{k=1}{}} S_{3.2.k}\} = \{S_{3.2.1}, S_{3.2.2}, S_{3.2.3}, S_{3.2.4}, S_{3.2.5}, S_{3.2.6}, S_{3.2.7}, S_{3.2.8}, S_{3.2.9}, S_{3.2.10}, S_{3.2.11}, S_{3.2.12}, S_{3.2.13}, S_{3.2.14}, S_{3.2.15}, S_{3.2.16}, S_{3.2.17}, S_{3.2.18}\} =$$
$$= \{S_{\text{AMDS}}, S_{\text{TGDS}}, S_{\text{SAB}}, S_{\text{TRES}}, S_{\text{APSS}}, S_{\text{ABCS}}, S_{\text{ACA}}, S_{\text{AXS}}, S_{\text{IBE}}, S_{\text{KUI}}, S_{\text{MER}}, S_{\text{NAV}}, S_{\text{PATH}}, S_{\text{RAD}}, S_{\text{AKF}}, S_{\text{TTI}}, S_{\text{WSMS}}, S_{\text{SIR}}\} =$$
$$= \{AMDS, TGDS, SAB, TRES, APSS, ABCS, ACA, AXS, IBE, KUI, MER, NAV, PATH, RAD, AKF, TTI, WSMS, SIR\},$$

where $S_{3.2.1} = S_{AMDS} = AMDS$ – Amadeus; $S_{3.2.2} = S_{TGDS} = TGDS$ – Travelport GDS; $S_{3.2.3} = S_{SAB} = SAB$ – Sabre; $S_{3.2.4} = S_{TRES} = TRES$ – TameliaRES; $S_{3.2.5} = S_{APSS} = APSS$ –

Avantik PSS; $S_{3.2.6} = S_{ABCS} = ABCS$ – Abacus; $S_{3.2.7} = S_{ACA} = ACA$ – AccelAero; $S_{3.2.8} = S_{AXS} = AXS$ – Axess; $S_{3.2.9} = = S_{IBE} = IBE$ – Internet Booking Engine; $S_{3.2.10} = S_{KUI} = KUI$ – KIU; $S_{3.2.11} = S_{MER} = MER$ – Mercator; $S_{3.2.12} = S_{NAV} = NAV$ – Navitaire; $S_{3.2.13} = S_{PATH} = PATH$ – Patheo; $S_{3.2.14} = S_{RAD} = RAD$ – Radixx; $S_{3.2.15} = S_{AKF} = AKF$ – Akeflite; $S_{3.2.16} = S_{TTI} = TTI$ – Travel Technology Interactive; $S_{3.2.17} = = S_{WSMS} = WSMS$ – WorldTicket Sell-More-Seats; $S_{3.2.18} = S_{SIR} = SIR$ – Syrena [10].

Similarly, for the set $\mathbf{S}_{3.3}$, if $n = 3$, $m_3 = 3$, $r_{3.3} = 8$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{3.3} = \mathbf{S}_{\text{IDS}} = \mathbf{IDS} = \{\overset{8}{\underset{k=1}{}} S_{3.3.k}\} =$$
$$= \{S_{3.3.1}, S_{3.3.2}, S_{3.3.3}, S_{3.3.4}, S_{3.3.5}, S_{3.3.6}, S_{3.3.7}, S_{3.3.8}\} =$$
$$= \{S_{\text{BKNG}}, S_{\text{OKT}}, S_{\text{EXP}}, S_{\text{ORB}}, S_{\text{HRS}}, S_{\text{TRAV}}, S_{\text{HOT}}, S_{\text{PRLN}}\} =$$
$$= \{BKNG, OKT, EXP, ORB, HRS, TRAV, HOT, PRLN\},$$

where $S_{3.3.1} = S_{\text{BKNG}} = BKNG$ – Booking.com; $S_{3.3.2} = = S_{\text{OKT}} = OKT$ – Oktogo; $S_{3.3.3} = S_{\text{EXP}} = EXP$ – Expedia.com; $S_{3.3.4} = S_{\text{ORB}} = ORB$ – Orbitz.com; $S_{3.3.5} = S_{\text{HRS}} = HRS$ – HRS.com; $S_{3.3.6} = S_{\text{TRAV}} = TRAV$ – Travelocity.com; $S_{3.3.7} = S_{\text{HOT}} = HOT$ – Hotels.com; $S_{3.3.8} = S_{\text{PRLN}} = PRLN$ – Priceline.com [10].

Similarly, for the set $\mathbf{S}_{3.4}$, if $n = 3$, $m_3 = 4$, $r_{3.4} = 8$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{3.4} = \mathbf{S}_{\text{BSP}} = \mathbf{BSP} = \{\overset{8}{\underset{k=1}{}} S_{3.4.k}\} =$$
$$= \{S_{3.4.1}, S_{3.4.2}, S_{3.4.3}, S_{3.4.4}, S_{3.4.5}, S_{3.4.6}, S_{3.4.7}, S_{3.4.8}\} =$$
$$\{S_{STD}, S_{\text{SAF}}, S_{\text{ODOC}}, S_{\text{ZVPR}}, S_{\text{PROCO}}, S_{\text{SABZ}}, S_{\text{PPKK}}, S_{\text{POV}}\} =$$
$$= \{STD, SAF, ODOC, ZVPR, PROCO, SABZ, PPKK, POV\},$$

where $S_{3.4.1} = S_{STD} = STD$ – standard transit documentation (STD); $S_{3.4.2} = S_{\text{SAF}} = SAF$ – standard administrative form (SAF); $S_{3.4.3} = S_{\text{ODOC}} = ODOC$ – document processing system; $S_{3.4.4} = S_{\text{ZVPR}} = ZVPR$ – system of reports on payments; $S_{3.4.5} = S_{\text{PROCO}} = PROCO$ – system of payment procedure; $S_{3.4.6} = S_{\text{SABZ}} = SABZ$ – c system of sanctions, administrative and bank charges; $S_{3.4.7} = S_{\text{PPKK}} = PPKK$ – credit card sales system; $S_{3.4.8} = S_{\text{POV}} = POV$ – system for returning tickets in accordance with [16].

Similarly, for the set $\mathbf{S}_{3.5}$, if $n = 3$, $m_3 = 5$, $r_{3.5} = 5$, using (4), represent the set of subsystems such as:

$$\mathbf{S}_{3.5} = \mathbf{S}_{\text{DCS}} = \mathbf{DCS} = \{\overset{5}{\underset{k=1}{}} S_{3.5.k}\} = \{S_{3.5.1}, S_{3.5.2}, S_{3.5.3}, S_{3.5.4}, S_{3.5.5}\} =$$
$$= \{S_{\text{SITA}}, S_{\text{TAIS}}, S_{\text{SAMDS}}, S_{\text{JKCS}}, S_{\text{HCS}}\} = \{SITA, TAIS, SAMDS, JKCS, HCS\},$$

where $S_{3.5.1} = S_{\text{SITA}} = SITA$ – SITA; $S_{3.5.2} = S_{\text{TAIS}} = TAIS$ – TAIS; $S_{3.5.3} = S_{\text{SAMDS}} = SAMDS$ – system from Amadeus; $S_{3.5.4} = = S_{\text{JKCS}} = JKCS$ – John Keells Computer Services; $S_{3.5.5} = S_{\text{HCS}} = HCS$ – Hitit Computer Services [10].

Representation of set subsystems ij system, according to Table 3, displayed in Table 4.

Table 4. Representation of set subsystems ij system CAIS

Sets of system (I) $\mathbf{S}_{ij}$ $(i=\overline{1,n},\ j=\overline{1,}$	Sets of system (IO) $\mathbf{S}_{ij}$ $(i=\overline{1,n},\ j=\overline{1,}$	Number of subsystem ij system k $\left(k=\overline{1,r_{ij}}\right.$	Sets of subsystem (I) S_{ijk} $(i=\overline{1,n},\ j=\overline{1,m_i},\ k=\overline{1,r_{ij}})$	Sets of subsystem (IO) S_{ijk} $(i=\overline{1,n},\ j=\overline{1,m_i},\ k=\overline{1,r_{ij}})$
$\mathbf{S}_{1.1}$	**SAE**	$r_{1.1}=5$	$S_{1.1.1}, S_{1.1.2}, S_{1.1.3}, S_{1.1.4}, S_{1.1.5}$	*SAPE, SANE, ZAR, SASZ, MTM*
$\mathbf{S}_{1.2}$	**RZZP**	$r_{1.2}=4$	$S_{1.2.1}, S_{1.2.2}, S_{1.2.3}, S_{1.2.4}$	*NDB, VOR, DME, ILS*
$\mathbf{S}_{1.3}$	**S S P**	$r_{1.3}=9$	$S_{1.3.1}, S_{1.3.2}, S_{1.3.3}, S_{1.3.4}, S_{1.3.5}, S_{1.3.6}, S_{1.3.7}, S_{1.3.8}, S_{1.3.9}$	*PSR, SSR, MSSR, RADS, SMR, WRAD, MLAT, ADS, DF*
$\mathbf{S}_{1.4}$	**S O D**	$r_{1.4}=5$	$S_{1.4.1}, S_{1.4.2}, S_{1.4.3}, S_{1.4.4}, S_{1.4.5}$	*ASYPR, SPPP, ESAN, SOPD, SC*
$\mathbf{S}_{1.5}$	**S M Z**	$r_{1.5}=3$	$S_{1.5.1}, S_{1.5.2}, S_{1.5.3}$	*SCMAU, KRAMS, SADIS*
$\mathbf{S}_{2.1}$	**S P S**	$r_{2.1}=4$	$S_{2.1.1}, S_{2.1.2}, S_{2.1.3}, S_{2.1.4}$	*DPPT, DZP, TPT, POP*
$\mathbf{S}_{2.2}$	**S Z V**	$r_{2.2}=3$	$S_{2.2.1}, S_{2.2.2}, S_{2.2.3}$	*BRS, CPDLS, AKARS*
$\mathbf{S}_{2.3}$	**N A V S**	$r_{2.3}=8$	$S_{2.3.1}, S_{2.3.2}, S_{2.3.3}, S_{2.3.4}, S_{2.3.5}, S_{2.3.6}, S_{2.3.7}, S_{2.3.8}$	*SNS, INS, ARK, RV, BVOR, BD, BILS, DVKZ*
$\mathbf{S}_{2.4}$	**SSPZ**	$r_{2.4}=4$	$S_{2.4.1}, S_{2.4.2}, S_{2.4.3}, S_{2.4.4}$	*TRA, TCAS, SRPZ, BMR*
$\mathbf{S}_{2.5}$	**O S L**	$r_{2.5}=2$	$S_{2.5.1}, S_{2.5.2}$	*OBCH, MCDU*
$\mathbf{S}_{2.6}$	**S V I**	$r_{2.6}=5$	$S_{2.6.1}, S_{2.6.2}, S_{2.6.3}, S_{2.6.4}, S_{2.6.5}$	*DBSVS, DSVS, OSVS, NSVSO, IS*
$\mathbf{S}_{2.7}$	**A B S K**	$r_{2.7}=4$	$S_{2.7.1}, S_{2.7.2}, S_{2.7.3}, S_{2.7.4}$	*APIL, SAU, PILS, PNK*
$\mathbf{S}_{3.1}$	**C R S**	$r_{3.1}=2$	$S_{3.1.1}, S_{3.1.2}$	*DELTM, PANAM*
$\mathbf{S}_{3.2}$	**G D S**	$r_{3.2}=18$	$S_{3.2.1}, S_{3.2.2}, S_{3.2.3}, S_{3.2.4}, S_{3.2.5}, S_{3.2.7}, S_{3.2.8}, S_{3.2.9}, S_{3.2.10}, S_{3.2.11}, S_{3.2.13}, S_{3.2.14}, S_{3.2.15}, S_{3.2.16}, S_{3.2.1}$	*AMDS, TGDS, SAB, TRES, APSS, ACA, AXS, IBE, KUI, MER, N, PATH, RAD, AKF, TTI, WSMS*
$\mathbf{S}_{3.3}$	**I D S**	$r_{3.3}=8$	$S_{3.3.1}, S_{3.3.2}, S_{3.3.3}, S_{3.3.4}, S_{3.3.5}, S_{3.3.6}, S_{3.3.7}, S_{3.3.8}$	*BKNG, OKT, EXP, ORB, HRS, TRAV, HOT, PRLN*

$\mathbf{S}_{3.4}$	**B S P**	$r_{3.4}=8$	$S_{3.4.1}, S_{3.4.2}, S_{3.4.3}, S_{3.4.4}, S_{3.4.5}, S_{3.4.6}, S_{3.4.7}, S_{3.4.8}$	*STD, SAF, ODOC, ZVPR, PROCO, SABZ, PPKK, POV*
$\mathbf{S}_{3.5}$	**D C S**	$r_{3.5}=5$	$S_{3.5.1}, S_{3.5.2}, S_{3.5.3}, S_{3.5.4}, S_{3.5.5}$	*SITA, TAIS, SAMDS, JKCS, HCS*

The set of subsystem of S_{ijk} system can be represented as a subset of subsystems:

$$\mathbf{S}_{ijk}=\{\overset{v_{ijk}}{\underset{p=1}{}} S_{ijkp}\}=\left\{S_{ijk1}, S_{ijk2}, ..., S_{ijkv_{ijk}}\right\}, \tag{6}$$

where $S_{ijkp}\subseteq\mathbf{S}_{ijk}$ $(i=\overline{1,n}$, $j=\overline{1,m_i}$, $k=\overline{1,r_{ij}}$, $p=\overline{1,v_{ijk}})$ – subset of subsystems $\mathbf{S}_{ijk}$, v_{ijk} – the number of subset of ijk -ї subsystem.

Taking into account (6), the expression (3) can be represented as follows:

$$\mathbf{S}=\{\overset{n}{\underset{i=1}{}}\mathbf{S}_i\}=\{\overset{n}{\underset{i=1}{}}\{\overset{m_i}{\underset{j=1}{}}\mathbf{S}_{ij}\}\}=\{\overset{n}{\underset{i=1}{}}\{\overset{m_i}{\underset{j=1}{}}\{\overset{r_{ij}}{\underset{k=1}{}}\mathbf{S}_{ijk}\}\}\}=\{\overset{n}{\underset{i=1}{}}\{\overset{m_i}{\underset{j=1}{}}\{\overset{r_{ij}}{\underset{k=1}{}}\{\overset{v_{ijk}}{\underset{p=1}{}}S_{ijkp}\}\}\}\}=$$
$$\{\{\{\{S_{1111}, S_{1112}, ..., S_{111v_{111}}\}, \{S_{1121}, S_{1122}, ..., S_{112v_{112}}\}, ..., \{S_{11r_{11}1}, S_{11r_{11}2}, ..., S_{11r_{11}v_{11r_{11}}}\}\}, \{\{S_{1211}, S_{1212}, ..., S_{121v_{121}}\}, \{S_{1221}, S_{1222}, ..., S_{122v_{122}}\}, ..., \{S_{12r_{12}1}, S_{12r_{12}2}, ..., S_{12r_{12}v_{12r_{12}2}}\}\}, ...,$$
$$\{\{S_{1m_11}, S_{1m_12}, ..., S_{1m_1v_{1m_11}}\}, \{S_{1m_121}, S_{1m_122}, ..., S_{1m_12v_{1m_12}}\}, ..., \{S_{1m_1r_{1m_1}1}, S_{1m_1r_{1m_1}2}, ..., S_{1m_1r_{1m_1}v_{1m_1r_{1m_1}}}\}\}\},$$
$$\{\{\{S_{2111}, S_{2112}, ..., S_{211v_{2111}}\}, \{S_{2121}, S_{2122}, ..., S_{212v_{2122}}\}, ..., \{S_{21r_{21}1}, S_{21r_{21}2}, ..., S_{21r_{21}v_{21r_{21}}}\}\}, \{\{S_{2211}, S_{2212}, ..., S_{221v_{2211}}\}, \{S_{2221}, S_{2222}, ..., S_{222v_{2222}}\}, ..., \{S_{22r_{22}1}, S_{22r_{22}2}, ..., S_{22r_{22}v_{22r_{22}2}}\}\}, ...,$$
$$\{\{S_{2m_21}, S_{2m_212}, ..., S_{2m_2v_{2m_21}}\}, \{S_{2m_221}, S_{2m_222}, ..., S_{2m_2v_{2m_22}}\}, ..., \{S_{2m_2r_{2m_2}1}, S_{2m_2r_{2m_2}2}, ..., S_{2m_2r_{2m_2}v_{2m_2r_{2m_2}}}\}\}\}, ...,$$
$$\{\{\{S_{n111}, S_{n112}, ..., S_{n1v_{n11}}\}, \{S_{n121}, S_{n122}, ..., S_{n12v_{n12}}\}, ..., \{S_{n1r_{n1}1}, S_{n1r_{n1}2}, ..., S_{n1r_{n1}v_{n1r_{n1}}}\}\}, \{\{S_{n211}, S_{n212}, ..., S_{n21v_{n211}}\}, \{S_{n221}, S_{n222}, ..., S_{n22v_{n222}}\}, ..., \{S_{n2r_{n21}}, S_{n2r_{n22}}, ..., S_{n2r_{n2}v_{n2r_{n22}}}\}\}, ...,$$
$$\{\{S_{nm_n11}, S_{nm_n12}, ..., S_{nm_nv_{nm_n1}}\}, \{S_{nm_n21}, S_{nm_n22}, ..., S_{nm_nv_{nm_n2}}\}, ..., \{S_{nm_nr_{nm_n}1}, S_{nm_nr_{nm_n}2}, ..., S_{nm_nr_{nm_n}v_{nm_nr_{nm_n}}}\}\}\}. \tag{7}$$

For example, for the set $\mathbf{S}_{1.1.1}$, if $n=1,\ m_1=1,\ r_{1.1}=1,\ v_{1.1.1}=3,$ using (6), represent the set of subsystems such as:

$$\mathbf{S}_{1.1.1}=\mathbf{S}_{\text{SAPE}}=\mathbf{SAPE}=\{\overset{3}{\underset{p=1}{}}S_{1.1.1.p}\}=\left\{S_{1.1.1.1}, S_{1.1.1.2}, S_{1.1.1.3}\right\}=$$
$$=\left\{S_{\text{NRPZ}}, S_{\text{CPDLC}}, S_{\text{ACARS}}\right\}=\left\{NRPZ, CPDLC, ACARS\right\},$$

where $S_{1.1.1.1}=S_{\text{NRPZ}}=NRPZ$ – ground-based radio communication «air-land»; $S_{1.1.1.2}=S_{\text{CPDLC}}=CPDLC$ – equipment for data transmission Controller-Pilot Data Link Communications (CPDLC); $S_{1.1.1.3}=S_{\text{ACARS}}=ACARS$ – equipment for data transmission Aircraft Communications Addressing and Reporting System (ACARS) in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{1.1.2}$, if $n=1,\ m_1=1,\ r_{1.1}=2,\ v_{1.1.1}=6,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{1.1.2}=\mathbf{S}_{\text{SANE}}=\mathbf{SANE}=\{\overset{6}{\underset{p=1}{}}S_{1.1.2.p}\}=$$
$$=\left\{S_{1.1.2.1}, S_{1.1.2.2}, S_{1.1.2.3}, S_{1.1.2.4}, S_{1.1.2.5}, S_{1.1.2.6}\right\}=$$
$$=\left\{S_{\text{ZPRZZ}}, S_{\text{SCGZ}}, S_{\text{ZRZ}}, S_{\text{AFTN}}, S_{\text{AMHS}}, S_{\text{MOD}}\right\}=$$
$$=\left\{\text{ZPRZZ}, SCGZ, ZRZ, AFTN, AMHS, MOD\right\},$$

where $S_{1.1.2.1}=S_{\text{ZPRZZ}}=\text{ZPRZZ}$ – hardware of wire (operative and telephone) and radio «earth-earth»; $S_{1.1.2.2}=S_{\text{SCGZ}}=SCGZ$ – voice commutation systems; $S_{1.1.2.3}=S_{\text{ZRZ}}=ZRZ$ – radio relay communication hardware; $S_{1.1.2.4}=S_{\text{AFTN}}=AFTN$ – Aeronautical Fixed

Telecommunication Network (AFTN); $S_{1.1.2.5} = = S_{\text{AMHS}} = AMHS$ –Air Traffic Service Message Handling System (AMHS); $S_{1.1.2.6} = S_{\text{MOD}} = MOD$ – data exchange network in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{1.1.3}$, if $n=1,\ m_1=1,\ r_{1.1}=3,\ v_{1.1.3}=2,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{1.1.3} = \mathbf{S}_{\text{ZAR}} = \mathbf{ZAR} = \{\overset{2}{\underset{p=1}{}} S_{1.1.3.p}\} = \{S_{1.1.3.1}, S_{1.1.3.2}\} =$$

$$= \{S_{\text{VOLM}}, S_{\text{ATIS}}\} = \{VOLM, ATIS\},$$

where $S_{1.1.3.1} = S_{\text{VOLM}} = VOLM$ – equipment of VHF radio broadcast VOLMET types; $S_{1.1.3.2} = = S_{\text{ATIS}} = ATIS$ – equipment of VHF radio broadcast ATIS types in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{1.4.1}$, if $n=1,\ m_1=4,\ r_{1.4}=1,\ v_{1.4.1}=7,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{1.4.1} = \mathbf{S}_{\text{ASYPR}} = \mathbf{ASYPR} = \{\overset{7}{\underset{p=1}{}} S_{1.4.1.p}\} =$$

$$= \{S_{1.4.1.1}, S_{1.4.1.2}, S_{1.4.1.3}, S_{1.4.1.4}, S_{1.4.1.5}, S_{1.4.1.6}, S_{1.4.1.7}\} =$$

$$= \{S_{ODSS}, S_{\text{OPD}}, S_{\text{MKS}}, S_{\text{ZVI}}, S_{\text{KGZ}}, S_{\text{PPR}}, S_{\text{ZBP}}\} =$$

$$= \{\text{ODSS}, OPD, MKS, ZVI, KGZ, PPR, ZBP\},$$

where $S_{1.4.1.1} = S_{ODSS} = \text{ODSS}$ – data processing of the surveillance system; $S_{1.4.1.2} = S_{\text{OPD}} = OPD$ – flight data processing; $S_{1.4.1.3} = S_{\text{MKS}} = MKS$ – monitoring and control of systems; $S_{1.4.1.4} = S_{\text{ZVI}} = ZVI$ – ecording and reproduction of information; $S_{1.4.1.5} = S_{\text{KGZ}} = KGZ$ – voice commutation; $S_{1.4.1.6} = S_{\text{PPR}} = PPR$ – support decision-making; $S_{1.4.1.7} = S_{\text{ZBP}} = ZBP$ ensuring flight safety in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{1.4.3}$, if $n=1,\ m_1=4,\ r_{1.4}=3,\ v_{1.4.3}=2,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{1.4.3} = \mathbf{S}_{\text{ESAN}} = \mathbf{ESAN} = \{\overset{2}{\underset{p=1}{}} S_{1.4.3.p}\} = \{S_{1.4.3.1}, S_{1.4.3.2}\} =$$

$$= \{S_{\text{ARTAS}}, S_{\text{SDDS}}\} = \{ARTAS, SDDS\},$$

where $S_{1.4.3.1} = S_{\text{ARTAS}} = ARTAS$ – ATM Surveillance Tracker And Server (ARTAS); $S_{1.4.3.2} = S_{\text{SDDS}} = SDDS$ – Surveillance

Data Distribution System (SDDS) in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{1.4.4}$, if $n=1,\ m_1=4,\ r_{1.4}=4,\ v_{1.4.4}=1,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{1.4.4} = \mathbf{S}_{\text{SOPD}} = \mathbf{SOPD} = \{\overset{1}{\underset{p=1}{}} S_{1.4.4.p}\} = \{S_{1.4.4.1}\} = \{S_{\text{IFPS}}\} = \{IFPS\},$$

where $S_{1.4.4.1} = S_{\text{IFPS}} = IFPS$ – EUROCONTROL Integrated Initial Flight Plan Processing System (IFPS) in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{2.1.1}$, if $n=2,\ m_2=1,\ r_{2.1}=1,\ v_{2.1.1}=3,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{2.1.1}=\mathbf{S}_{DPPT}=\mathbf{DPPT}=\{\overset{3}{\underset{p=1}{}}S_{2.1.1.p}\}=\{S_{2.1.1.1},S_{2.1.1.2},S_{2.1.1.3}\}=$$
$$=\{S_{\text{PST}},S_{\text{PDT}},S_{\text{KPPT}}\}=\{PST,PDT,KPPT\},$$

where $S_{2.1.1.1}=S_{\text{PST}}=PST$ – statistical pressure receivers; $S_{2.1.1.2}=S_{\text{PDT}}=PDT$ – dynamic pressure receivers; $S_{2.1.1.3}=S_{\text{KPPT}}=KPPT$ – combined pressure receivers in accordance with [11].

Similarly, for the set of subsystem $\mathbf{S}_{2.1.3}$, при $n=2,\ m_2=1,\ r_{2.1}=3,\ v_{2.1.3}=2,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{2.1.3}=\mathbf{S}_{\text{TPT}}=\mathbf{TPT}=\{\overset{2}{\underset{p=1}{}}S_{2.1.3.p}\}=\{S_{2.1.3.1},S_{2.1.3.2}\}=$$
$$=\{S_{\text{STATL}},S_{\text{DYNL}}\}=\{STATL,DYNL\},$$

where $S_{2.1.3.1}=S_{\text{STATL}}=STATL$ – pipelines of statistical lines; $S_{2.1.3.2}=S_{\text{DYNL}}=DYNL$ – pipelines of dynamic lines in accordance with [11].

Similarly, for the set of subsystem $\mathbf{S}_{2.6.1}$, if $n=2,\ m_2=6,\ r_{2.6}=1,\ v_{2.6.1}=5,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{2.6.1}=\mathbf{S}_{DBSVS}=\mathbf{DBSVS}=\{\overset{5}{\underset{p=1}{}}S_{2.6.1.p}\}=$$
$$=\{S_{2.6.1.1},S_{2.6.1.2},S_{2.6.1.3},S_{2.6.1.4},S_{2.6.1.5}\}=$$
$$=\{S_{BBSB},S_{RVM},S_{\text{BMRL}},S_{\text{BMDX}},S_{\text{BSIB}}\}=$$
$$=\{BBSB,RVM,BMRL,BMDX,BSIB\},$$

where $S_{2.6.1.1}=S_{BBSB}=BBSB$ – on-board synthetic vision database; $S_{2.6.1.2}=S_{RVM}=RVM$ – radio volumetric SVS; $S_{2.6.1.3}=S_{\text{BMRL}}=BMRL$ – board meteoradiolokator; $S_{2.6.1.4}=S_{\text{BMDX}}=BMDX$ – б on-board radar of millimeter wavelength range; $S_{2.6.1.5}=S_{\text{BSIB}}=BSIB$ – on-board system of infrared vision in accordance with [13].

Similarly, for the set of subsystem $\mathbf{S}_{2.6.2}$, if $n=2,\ m_2=6,\ r_{2.6}=2,\ v_{2.6.2}=5,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{2.6.2}=\mathbf{S}_{DSVS}=\mathbf{DSVS}=\{\overset{5}{\underset{p=1}{}}S_{2.6.2.p}\}=$$
$$=\{S_{2.6.2.1},S_{2.6.2.2},S_{2.6.2.3},S_{2.6.2.4},S_{2.6.2.5}\}=$$
$$=\{S_{PFD},S_{ND},S_{\text{HUD}},S_{\text{HMD}},S_{IDIS}\}=$$
$$=\{PFD,ND,HUD,HMD,IDIS\},$$

where $S_{2.6.2.1}=S_{PFD}=PFD$ – primary flight display (PFD); $S_{2.6.2.2}=S_{ND}=ND$ – navigation display (ND); $S_{2.6.2.3}=S_{\text{HUD}}=HUD$ – Head-Up Display – HUD; $S_{2.6.2.4}=S_{\text{HMD}}=HMD$ – Helmet-Mounted Display; $S_{2.6.2.5}=S_{IDIS}=IDIS$ – other displays in accordance with [13].

Similarly, for the set of subsystem $\mathbf{S}_{2.6.3}$, if $n=2,\ m_2=6,\ r_{2.6}=3,\ v_{2.6.3}=3,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{2.6.3}=\mathbf{S}_{OSVS}=\mathbf{OSVS}=\{\overset{3}{\underset{p=1}{}}S_{2.6.3.p}\}=\{S_{2.6.3.1},S_{2.6.3.2},S_{2.6.3.3}\}=$$

$$=\{S_{OSZ},S_{OGPP},S_{\text{OVP}}\}=\{OSZ,OGPP,OVP\},$$

where $S_{2.6.3.1}=S_{OSZ}=OSZ$ – image synthesis calculator; $S_{2.6.3.2}=S_{OGPP}=OGPP$ - grouping calculator, validation and approval of SVS system performance; $S_{2.6.3.3}=S_{\text{OVP}}=OVP$ – the display calculator in accordance with [13].

Similarly, for the set of subsystem $\mathbf{S}_{3.2.2}$, if $n=3,\ m_3=2,\ r_{3.2}=2,\ v_{3.2.2}=3,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{3.2.2}=\mathbf{S}_{TGDS}=\mathbf{TGDS}=\{\overset{3}{\underset{p=1}{}}S_{3.2.2.p}\}=\{S_{3.2.2.1},S_{3.2.2.2},S_{3.2.2.3}\}=$$

$$=\{S_{\text{APLL}},S_{\text{GALL}},S_{\text{WSPN}}\}=\{APLL,GALL,WSPN\},$$

where $S_{3.2.2.1}=S_{\text{APLL}}=APLL$ – Apollo; $S_{3.2.2.2}=S_{\text{GALL}}=GALL$ – Galileo; $S_{3.2.2.3}=S_{\text{WSPN}}=WSPN$ – Worldspan in accordance with [10].

Similarly, for the set of subsystem $\mathbf{S}_{3.5.2}$, if $n=3,\ m_3=5,\ r_{3.5}=2,$ using (6), represent the subset of subsystems such as:

$$\mathbf{S}_{3.5.2}=\mathbf{S}_{TAIS}=\mathbf{TAIS}=\{\overset{3}{\underset{p=1}{}}S_{3.5.2.p}\}=\{S_{3.5.2.1},S_{3.5.2.2},S_{3.5.2.3}\}=$$

$$=\{S_{TCRS},S_{TDCS},S_{\text{TTSH}}\}=\{TCRS,TDCS,TTSH\},$$

where $S_{3.5.2.1}=S_{TCRS}=TCRS$ – TAIS CRS (reservation and sale of transportation); $S_{3.5.2.2}=S_{TDCS}=TDCS$ – TAIS DCS (departure control); $S_{3.5.2.3}=S_{\text{TTSH}}=TTSH$ – TAIS Travel Shop (e-commerce) in accordance with [17].

Representation of set subsystems ijk system, according to Table. 3, displayed in Table 4.

Table 6. Representation of set subsystems ijk system CAIS

Sets of subsystem (I) $S_{ijk}\ (i=\overline{1,n},\ j=\overline{1,m_i},k=\overline{1,}$	Sets of subsystem (IO) $S_{ijk}\ (i=\overline{1,n},\ j=\overline{1,m_i},k=\overline{1,r_{ij}})$	Number of subset subsystem - $p\ \left(p=\overline{1,v_{ijk}}\right)$	Subsets of subsystem (I) $S_{ijkp}\ (i=\overline{1,n},\ j=\overline{1,m_i},\ k=\overline{1,r_{ij}},\ p=\overline{1,v_{ijk}})$	Subsets of subsystem (IO) $S_{ijkp}\ (i=\overline{1,n},\ j=\overline{1,m_i},\ k=\overline{1,r_{ij}},\ p=\overline{1,v_{ijk}})$
$\mathbf{S}_{1.1.1}$	**SAPE**	$v_{1.1.1}=3$	$S_{1.1.1.1},S_{1.1.1.2},S_{1.1.1.3}$	*NRPZ, CPDLC, ACARS*
$\mathbf{S}_{1.1.2}$	**SANE**	$v_{1.1.2}=6$	$S_{1.1.2.1},S_{1.1.2.2},S_{1.1.2.3},$ $S_{1.1.2.4},S_{1.1.2.5},S_{1.1.2.6}$	*ZPRZZ, SCGZ, ZRZ, AFTN, AMHS, MOD*
$\mathbf{S}_{1.1.3}$	**ZAR**	$v_{1.1.3}=2$	$S_{1.1.3.1},S_{1.1.3.2}$	*VOLM, ATIS*
$\mathbf{S}_{1.4.1}$	**ASYPR**	$v_{1.4.1}=7$	$S_{1.4.1.1},S_{1.4.1.2},S_{1.4.1.3},S_{1.4.1.4},$ $S_{1.4.1.5},S_{1.4.1.6},S_{1.4.1.7}$	ODSS, *OPD, MKS, ZVI, KGZ, PPR, ZBP*

$\mathbf{S}_{1.4.3}$	**ESAN**	$v_{1.4.3}=2$	$S_{1.4.3.1}, S_{1.4.3.2}$	*ARTAS, SDDS*
$\mathbf{S}_{1.4.4}$	**SOPD**	$v_{1.4.4}=1$	$S_{1.4.4.1}$	*IFPS*
$\mathbf{S}_{2.1.1}$	**DPPT**	$v_{2.1.1}=3$	$S_{2.1.1.1}, S_{2.1.1.2}, S_{2.1.1.3}$	*PST, PDT, KPPT*
$\mathbf{S}_{2.1.3}$	**TPT**	$v_{2.1.3}=2$	$S_{2.1.3.1}, S_{2.1.3.2}$	*STATL, DYNL*
$\mathbf{S}_{2.6.1}$	**DBSVS**	$v_{2.6.1}=5$	$S_{2.6.1.1}, S_{2.6.1.2}, S_{2.6.1.3}, S_{2.6.1.4}, S_{2.6.1.5}$	*BBSB, RVM, BMRL, BMDX, BSIB*
$\mathbf{S}_{2.6.2}$	**DSVS**	$v_{2.6.2}=5$	$S_{2.6.2.1}, S_{2.6.2.2}, S_{2.6.2.3}, S_{2.6.2.4}, S_{2.6.2.5}$	*PFD, ND, HUD, HMD, IDIS*
$\mathbf{S}_{2.6.3}$	**OSVS**	$v_{2.6.3}=3$	$S_{2.6.3.1}, S_{2.6.3.2}, S_{2.6.3.3}$	*OSZ, OGPP, OVP*
$\mathbf{S}_{3.2.2}$	**TGDS**	$v_{3.2.2}=3$	$S_{3.2.2.1}, S_{3.2.2.2}, S_{3.2.2.3}$	*APLL, GALL, WSPN*
$\mathbf{S}_{3.5.2}$	**TAIS**	$v_{3.5.2}=3$	$S_{3.5.2.1}, S_{3.5.2.2}, S_{3.5.2.3}$	*TCRS, TDCS, TTSH*

Depending on the possibility of details CII industry categories the S_{ijkp} subsystem can be presented in the form of subsets with an in-depth detail level. Therefore, it's necessary to present a complete set of categories of systems in the CII industry in general terms as follows:

$$\mathbf{S}=\{\overset{n_0}{\underset{i_1=1}{}}\{\overset{n_{i_1}}{\underset{i_2=1}{}}\{\ldots\{\overset{n_{i_1,i_2,\ldots,i_{l-1}}}{\underset{i_l=1}{}}\mathbf{S}_{i_1,i_2,\ldots,i_l}\}\}\}\}, \tag{8}$$

where $\mathbf{S}_{i_1,i_2,\ldots,i_l}\subseteq\mathbf{S}$ $(i_1=\overline{1,n_0}$, $i_2=\overline{1,n_{i_1}}$, $i_l=\overline{1,n_{i_1,i_2,\ldots,i_{l-1}}})$ – levels detalization of categories the $\mathbf{S}$system, l – number of levels detalization of categories the system. For example, in the paper for CA industry for CAIS system was determined the level of detalization $l=4$, which on accordance to (8), could be presented such as:

$$\mathbf{S}=\{\overset{n_0}{\underset{i_1=1}{}}\mathbf{S}_{i_1}\}=\{\overset{n_0}{\underset{i_1=1}{}}\{\overset{n_{i_1}}{\underset{i_2=1}{}}\mathbf{S}_{i_1,i_2}\}\}=$$
$$=\{\overset{n_0}{\underset{i_1=1}{}}\{\overset{n_{i_1}}{\underset{i_2=1}{}}\{\overset{n_{i_1,i_2}}{\underset{i_3=1}{}}\mathbf{S}_{i_1,i_2,i_3}\}\}\}=\{\overset{n_0}{\underset{i_1=1}{}}\{\overset{n_{i_1}}{\underset{i_2=1}{}}\{\overset{n_{i_1,i_2}}{\underset{i_3=1}{}}\{\overset{n_{i_1,i_2,i_3}}{\underset{i_4=1}{}}\mathbf{S}_{i_1,i_2,i_3,i_4}\}\}\}\}, \tag{9}$$

where $\mathbf{S}_{i_1,i_2,\ldots,i_l}\subseteq\mathbf{S}$ $(i_1=\overline{1,n_0}$, $i_2=\overline{1,n_{i_1}}$, $i_3=\overline{1,n_{i_1,i_2}}$, $i_4=\overline{1,n_{i_1,i_2,i_3}})$ – the level of detalization the $\mathbf{S}_{\text{KAIS}}$ category, moreover, on accordance with (1), (2), (4) and (6) $i_1=i, i_1=j, i_3=k, i_4=p$ and $n_0=n,\ n_{i_1}=m_i,\ n_{i_1,i_2}=r_{ij},\ n_{i_1,i_2,i_3}=v_{ijk}$.

To determine the coherence of obtained data, by way of universal model, the incident matrix Δ (10) was formed. This matrix shows influence ration $\lambda$for certan set of systems CAIS $\mathbf{Y}=\{\overset{m}{\underset{i=1}{}}Y_i\}=\{Y_1,Y_2,\ldots,Y_m\}$, where $Y_i\subseteq\mathbf{Y}$ $(i=\overline{1,m})$, where m – total number of systems, and the set of threats for object of CI $\mathbf{X}=\{\overset{n}{\underset{j=1}{}}X_i\}=\{X_1,X_2,\ldots,X_n\}$, where $X_i\subseteq\mathbf{X}$ $(j=\overline{1,n})$, where n – total numbers of threats. Incident matrix determined the

ratio of $\Delta=\left(\lambda_{ij}\right)$, that determined the possibility of a certain threat X_j affect a particular system CAIS Y_i (where $\lambda_{ij}=1$, if $\left(Y_i,X_j\right)\in 1$, and $\lambda_{ij}=0$, if $\left(Y_i,X_j\right)\notin 1$).

	X_1	X_2	…	X_n
Y_1	λ_{11}	λ_{21}	…	λ_{n1}
Y_2	λ_{12}	λ_{22}	…	λ_{n2}
…	…	…	…	…
Y_m	λ_{1m}	λ_{2m}	…	λ_{nm}

(10)

After that, the sets of vertices of complexes, that characterized the list of possible threats for particular system $K_Y(X;\lambda)$, and the list of systems that may be affected by a certain threat $K_X\left(Y;\lambda^{-1}\right)$ could be realized. But if it is necessary to consider the complex as a whole, it is expedient to use the notion of communication chain, which shows the fact - two simplexes may not have a common facet, but can be linked by a sequence of intermediate simplexes. The simplex complex – is a mathematical generalization of the planar graph concept, which reflects the multidimensional nature of the binary relation of the system. Due to the fact that simplex complex is a set of simplexes connected by common faces, then for the characteristic of the connection could be taken the value of the face which common to the two simplexes. So, if the sets **Y** and **X** include m and n elements, properly, in this case matrix Δ is the matrix with size $(m\times n)$, which consist of zero and units. Product $\Delta\Delta^T$ – is the number, that in place $(i,\ j)$ it's scalar product of rows i and j in matrix Δ. It's equal to unit, which are on the same places in the rows i and j in matrix Δ and correspond to value $(q+1)$, where q - the dimension of common facet of simplexes σ_p і σ_r, presented by rows i and j. In this case, for determine the q- common facet in each pair Y- simplexes in $K_Y(X;\lambda)$ it necessary: create the matrix $\Delta\Delta^T$ by size $(m\times m)$; estimate $\Delta\Delta^T-\Omega$, where $\Omega=\left(\omega_{ij}\right)$, and $\omega_{ij}=1$ for $i,j=\overline{1,m}$. Integers on matrix diagonal is dimensions of simplex Y, and Q - the analysis performed by checking other combinations of columns and rows. The analysis for $K_X\left(Y;\lambda^{-1}\right)$ performed by creating the matrix $\Delta^T\Delta-\Omega'$, where Ω' – the matrix with size $(n\times n)$, that consist of units. The integers on the diagonal of the matrix are the dimensions of simplexes X also, and Q - the analysis performed by checking other combinations of columns and rows.

For example, created list of identified critical objects (namely, the set S_{ij})was used for analyzing the impact of possible threats. Created, on the basis (10), the incident matrix of relation $\Delta_{KAIS\_THREATS}$ at $i=\overline{1,17},\ j=\overline{1,19}$, characterize the possibility that certain threat X_j affected on certain CAIS KAIC Y_i, where: Y_1 – aviation telecommunication systems; Y_2 – radio navigational aids of flight operations; Y_3 – surveillance systems; Y_4 – data processing systems; Y_5 – meteorological support

systems; Y_6 – air data system; Y_7 – communication systems; Y_8 – navigation systems; Y_9 – observing and collision avoidance systems; Y_{10} – computer aircrafts systems; Y_{11} – information display system; Y_{12} – automatic on-board control systems; Y_{13} – computer distribution system; Y_{14} – global distribution system (booking); Y_{15} – Internet Distribution Systems (IDS); Y_{16} – settlement system; Y_{17} – dispatch control systems, X_1 – aviation disasters; X_2 – nuclear accidents; X_3 – accidents in power supply systems; X_4 – release of hazardous substance; X_5 – system failures; X_6 – accidents and emergency occurrences due to negligence, organizational mistakes; X_7 – accidents at objects of high-threat; X_8 – meteorological or extreme weather conditions; X_9 – hydrological threats; X_{10} – seismic threats; X_{11} – geological threats; X_{12} – heliophysical threats; X_{13} – fires (forest, steppe, peat); X_{14} – pidemics and pandemics, epizootics, epiphytoties; X_{15} – terrorist acts; X_{16} – actions of criminals and saboteurs; X_{17} – military operations during the war; X_{18} – cybercrafts on; Information Telecommunication System X_{19} – threats of functioning of state authorities, military, law enforcement agencies and intelligence agency

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}
Y_1	1	1	1	0	1	1	0	1	1	1	0	1	0	0	1	1	1	1	0
Y_2	1	0	1	0	1	1	0	1	1	1	0	0	0	0	1	1	1	1	0
Y_3	1	0	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	0
Y_4	1	0	0	0	1	1	0	1	1	1	1	1	0	0	1	1	0	1	0
Y_5	1	1	0	1	1	0	0	1	1	1	1	1	1	0	0	1	0	0	0
Y_6	1	0	1	0	1	0	0	1	1	1	1	1	1	0	1	1	0	0	0
Y_7	1	0	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1
Y_8	1	0	1	0	1	1	0	1	1	1	1	1	0	0	1	1	1	1	0
Y_9	1	0	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0
Y_{10}	1	0	0	0	1	1	0	1	0	1	0	0	0	0	0	1	0	0	0
Y_{11}	1	1	1	0	1	1	0	1	1	1	0	0	0	0	1	0	1	1	0
Y_{12}	1	1	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Y_{13}	1	0	1	0	1	1	0	1	0	0	0	1	0	1	1	1	1	1	1
Y_{14}	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	1	0	1	1
Y_{15}	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	1	0	1	0
Y_{16}	1	0	1	0	1	1	0	0	0	0	0	0	0	1	1	1	0	1	0
Y_{17}	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1

Figure 1. Identity matrix $\Delta_{KAIS\_THREATS}$

a)

$(Y_1) < X_1, X_2, X_5, X_3, X_6, X_8, X_9, X_{10}, X_{12}, X_{13}, X_{16}, X_{17}, X_{18} > \sigma_{12}$
$(Y_2) < X_1, X_5, X_3, X_6, X_8, X_9, X_{10}, X_{15}, X_{16}, X_{17}, X_{18} > \sigma_{10}$
$(Y_3) < X_1, X_3, X_6, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{15}, X_{16}, X_{18} > \sigma_{11}$
$(Y_4) < X_1, X_3, X_6, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{16}, X_{18} > \sigma_{10}$
$(Y_5) < X_1, X_2, X_4, X_3, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{16} > \sigma_{10}$
$(Y_6) < X_1, X_3, X_5, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{15}, X_{16} > \sigma_{10}$
$(Y_7) < X_1, X_3, X_4, X_5, X_6, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{15}, X_{16}, X_{17}, X_{18}, X_{19} > \sigma_{14}$
$(Y_8) < X_1, X_3, X_5, X_6, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{15}, X_{16}, X_{17}, X_{18} > \sigma_{12}$
$(Y_9) < X_1, X_3, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{12}, X_{13}, X_{15}, X_{16}, X_{17}, X_{18} > \sigma_{13}$
$(Y_{10}) < X_1, X_5, X_6, X_8, X_{10}, X_{16} > \sigma_5$
$(Y_{11}) < X_1, X_3, X_4, X_5, X_6, X_8, X_9, X_{10}, X_{15}, X_{17}, X_{18} > \sigma_{10}$
$(Y_{12}) < X_1, X_2, X_3, X_5, X_7, X_8, X_9 > \sigma_6$
$(Y_{13}) < X_1, X_3, X_5, X_6, X_8, X_{12}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{19} > \sigma_{11}$
$(Y_{14}) < X_1, X_3, X_5, X_6, X_8, X_{14}, X_{15}, X_{16}, X_{18}, X_{19} > \sigma_9$
$(Y_{15}) < X_1, X_3, X_5, X_6, X_8, X_{14}, X_{15}, X_{16}, X_{18} > \sigma_8$
$(Y_{16}) < X_1, X_3, X_5, X_6, X_{14}, X_{15}, X_{16}, X_{18} > \sigma_7$
$(Y_{17}) < X_1, X_2, X_3, X_4, X_5, X_6, X_8, X_9, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{19} > \sigma_{15}$

b)

$(X_1) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{16}$
$(X_2) < Y_1, Y_5, Y_{11}, Y_{12}, Y_{17} > \sigma_4$
$(X_3) < Y_1, Y_2, Y_6, Y_7, Y_8, Y_9, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{12}$
$(X_4) < Y_5, Y_7, Y_{17} > \sigma_2$
$(X_5) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{16}$
$(X_6) < Y_1, Y_2, Y_3, Y_4, Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{13}$
$(X_7) < Y_9, Y_{12} > \sigma_1$
$(X_8) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{17} > \sigma_{15}$
$(X_9) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{11}, Y_{12}, Y_{17} > \sigma_{11}$
$(X_{10}) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{11} > \sigma_{10}$
$(X_{11}) < Y_3, Y_4, Y_5, Y_6, Y_7, Y_8 > \sigma_5$
$(X_{12}) < Y_1, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{13}, Y_{17} > \sigma_9$
$(X_{13}) < Y_3, Y_5, Y_6, Y_9, Y_{17} > \sigma_4$
$(X_{14}) < Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_4$
$(X_{15}) < Y_1, Y_2, Y_3, Y_4, Y_6, Y_7, Y_8, Y_9, Y_{11}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{13}$
$(X_{16}) < Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{14}$
$(X_{17}) < Y_1, Y_2, Y_7, Y_8, Y_9, Y_{11}, Y_{13}, Y_{17} > \sigma_7$
$(X_{18}) < Y_1, Y_2, Y_3, Y_4, Y_7, Y_8, Y_9, Y_{11}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17} > \sigma_{12}$
$(X_{19}) < Y_7, Y_{13}, Y_{14}, Y_{17} > \sigma_3$.

Figure 2. Set of complexes vertices: a) $K_Y(X;\lambda)$*; b)* $K_X(Y;\lambda^{-1})$

a)

Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	Y_{16}	Y_{17}	
12	10	9	9	7	8	11	11	11	5	10	5	9	7	7	6	11	Y_1
	10	8	8	5	7	10	10	10	5	9	4	8	7	7	6	9	Y_2
		11	10	8	9	10	10	10	5	7	3	7	6	6	5	9	Y_3
			10	7	8	10	10	9	5	7	3	7	6	6	5	8	Y_4
				10	8	8	7	7	4	5	4	4	3	3	2	8	Y_5
					10	9	9	9	4	6	4	6	5	5	4	8	Y_6
						14	12	11	5	9	4	10	8	7	6	12	Y_7
							12	11	5	9	4	9	7	7	6	10	Y_8
								13	5	9	5	9	7	7	6	11	Y_9
									5	4	2	4	4	4	3	4	Y_{10}
										10	5	7	6	6	5	9	Y_{11}
											6	3	3	3	2	5	Y_{12}
												11	9	8	7	11	Y_{13}
													9	8	7	9	Y_{14}
														8	7	8	Y_{15}
															7	7	Y_{16}
																15	Y_{17}

b)

при $q=15$, $\{Y_{17}\}$ $Q_{15}=1$;
при $q=14$, $\{Y_7\}$ $Q_{14}=1$;
при $q=13$, $\{Y_9\}$ $Q_{13}=1$;
при $q=12$, $\{Y_1\}\{Y_7,Y_8\}$ $Q_{12}=2$;
при $q=11$, $\{Y_3\}\{Y_{13},Y_{17}\}$ $Q_{11}=2$;
при $q=10$, $\{Y_1,Y_2,Y_7,Y_8,Y_9\}\{Y_3,Y_4\}\{Y_5\}\{Y_6\}\{Y_{11}\}$ $Q_{10}=5$;
при $q=9$, $\{Y_{13},Y_{14},Y_{17}\}$ $Q_9=1$;
при $q=8$, $\{Y_{13},Y_{14},Y_{15},Y_{17}\}$ $Q_8=1$;
при $q=7$, $\{Y_{13},Y_{14},Y_{15},Y_{16},Y_{17}\}$ $Q_7=1$;
при $q=6$, $\{Y_{12}\}$ $Q_6=1$;
при $q=5$, $\{Y_1,Y_2,Y_3,Y_4,Y_7,Y_8,Y_9,Y_{10}\}$ $Q_5=1$.
$Q_Y=\{1,1,1,2,2,5,1,1,1,1,1\}$
$\phi_{KAIS}=1,39$

Figure 3. q-th simplex value of complex $K_Y(X;\lambda)$ - CAIS system

a)

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	
16	4	12	2	16	13	1	15	11	10	5	9	4	4	13	14	7	12	3	X_1
	4	3	1	4	2	0	4	4	2	0	2	1	0	2	2	2	2	0	X_2
		12	1	12	10	1	11	8	6	2	6	2	4	11	10	7	10	3	X_3
			2	2	1	-1	2	2	1	1	2	1	0	1	2	1	1	1	X_4
				16	13	1	15	11	10	5	9	4	4	13	14	7	12	3	X_5
					13	0	12	8	8	3	7	2	4	12	12	7	12	3	X_6
						1	1	1	0	-1	0	0	-1	0	0	0	0	-1	X_7
							15	11	10	5	9	4	3	12	13	7	11	3	X_8
								11	9	5	8	4	0	9	9	6	8	1	X_9
									10	5	7	3	-1	8	9	5	7	0	X_{10}
										5	5	2	-1	4	5	1	3	0	X_{11}
											9	4	1	8	9	5	7	2	X_{12}
												4	0	3	4	1	2	0	X_{13}
													4	4	4	1	4	2	X_{14}
														13	12	7	12	3	X_{15}
															14	6	11	3	X_{16}
																7	7	2	X_{17}
																	12	3	X_{18}
																		3	X_{19}

b)

при $q=16$, $\{X_1,X_5\}$ $Q_{16}=1$;
при $q=15$, $\{X_1,X_5,X_8\}$ $Q_{15}=1$;
при $q=14$, $\{X_1,X_5,X_{16}\}$ $Q_{14}=1$;
при $q=13$, $\{X_1,X_5,X_6\}\{X_{12}\}$ $Q_{13}=2$;
при $q=12$, $\{X_1,X_3,X_5\}\{X_6,X_{12},X_{18}\}$ $Q_{12}=2$;
при $q=11$, $\{X_1,X_5,X_8,X_9\}$ $Q_{11}=1$;
при $q=10$, $\{X_1,X_5,X_8,X_{10}\}$ $Q_{10}=1$;
при $q=9$, $\{X_1,X_5,X_8,X_{12},X_{16}\}$ $Q_9=1$;
при $q=7$, $\{X_1,X_3,X_5,X_6,X_8,X_{15},X_{17},X_{18}\}$ $Q_7=1$;
при $q=5$, $\{X_1,X_5,X_8,X_9,X_{10},X_{11},X_{12},X_{16}\}$ $Q_5=1$;
при $q=4$, $\{X_1,X_2,X_3,X_8,X_9\}\{X_{12},X_{13},X_{16}\}\{X_3,X_6,X_{14},X_{15},X_{18}\}$ $Q_4=3$;
при $q=3$, $\{X_1,X_3,X_5,X_6,X_8,X_{15},X_{16},X_{18},X_{19}\}$ $Q_3=1$;
при $q=2$, $\{X_1,X_4,X_5,X_8,X_9,X_{12},X_{16}\}$ $Q_2=1$;
при $q=1$, $\{X_1,X_3,X_5,X_7,X_8,X_9\}$ $Q_1=1$;
$Q_X=\{1,1,1,2,2,1,1,1,1,3,1,1,1\}$
$\phi_{THREATS}=1,13$

Figure 4. q- th simplex value of complex $K_X(Y;\lambda^{-1})$ - the threats for CI object of state

According to research, the relation q- the sets connections of threats has a higher connectivity in comparison with similar relations q- the sets connections of CAIS systems, which indicates that the realization of one threat can trigger a cascade effect on other related threats and lead to heavy, and sometimes, devastating consequences for a particular CAIS system. In addition, the structural vector of the ratio for CAIS systems and the structural vector of the ratio for the threats of CI of the State (whereby could obtain and compare the measure (numerical value) of the complexity $\phi$of the complexes of these relations.) was shown The calculated measure of complexes complexity ϕ_{KAIS} and $\phi_{THREATS}$ attests about the big CAIS system “complexity". Note that this definition of complexity deals only the statistical complexity of selected complexes.

4. Conclusion

In the paper developed universal data model, which, due to the multi-level detail of critical aviation information systems, the hierarchical representation of sets, which describing systems and their components, as well as the introduction of the criticality incidence matrix of critical infrastructure, its simplex complexes and Q-analysis, allows to formalize the process of forming a list of critical information infrastructure objects of the state and determine its connectivity (the relation of q-bonds between sets of cyber threats and critical aviation information systems). For example the list of CII objects for the CAIS industry was formed at the level of detail $l=4$ (were highlighted 3 set of categories, 17 sets of systems, 97 sets of subsystems, 45 subsets of subsystems). The resulting list of identified critical objects can be used for further evaluation of vulnerabilities, analysis of possible threats and development of appropriate methods and means of protecting the objects from cyber threats.

REFERENCES

1. ГНАТЮК С., РЯБИЙ М., ЛЯДОВСЬКА В.: Визначення критичної інформаційної інфраструктури та її захист: аналіз підходів, Зв'язок, №4, С. 3-7, 2014.
2. ГНАТЮК С., ЛЯДОВСЬКА В.: Критерії визначення елементів критичної інфраструктури держави, Матеріали XXIII всеукр. наук.-практ. конф. «Інноваційний потенціал світової науки — XXI сторіччя». Запоріжжя: Вид-во ПГА, С. 55-57, 2013.
3. WENGER A., MAUER V., CAVELTY M.: Internatio-nal critical information infrastructure protection hand-book 2008-2009, Center for Security Studies, ETH Zurich, 2009.
4. ДОВГАНЬ О.: Критична інфраструктура як об’єкт захисту від кібернетичних атак, Матеріали наук.-практ. конф. «Інформаційна безпека: виклики і загрози сучасності». К: НА СБ України, С. 17-20, 2013.
5. БІРЮКОВ Д., КОНДРАТОВ С.,Захист критичної інфраструктури: проблеми та перспективи впровадження в Україні К: НІСД, 96 с, 2012.

6. C. Keating, R. Rogers, R. Unal, D. Dryer, R. Safford, W. Peterson, G. Rabadi, «System of Systems Engineerings», Engineering Management Journal, Vol. 15, № 3, p. 36-45, 2003.
7. Постанова про затвердження Порядку формування переліку інфо рмаційно-телекомунікаційних систем об'єктів критичної інфраструктури держави. Постанова Кабінету Міністрів України від 23.08.2016, №563. [Електронний ресурс]. Режим доступу: http://zak on3.rada.gov.ua/laws/show/563-2016-%D0%BF.
8. Doc 8973 ICAO «Керівництво з авіаційної безпеки» (Restricted), Вид. 9, 818 с, 2014.
9. КОРЧЕНКО О., БУРЯЧОК В., ГНАТЮК С.: Кібернетична безпека держави: характерні ознаки та проблемні аспекти, Безпека інформації, т. 19, №1, С. 40-45, 2013.
10. ГНАТЮК С., ВАСИЛЬЄВ Д.: Сучасні критичні авіаційні інформаційні системи, Безпека інформації, Т. 2, №1, С. 51-57, 2016.
11. ХАРЧЕНКО В., ОСТРОУМОВ І.: Авіоніка: навчальний посібник. К. : НАУ, 272 с., 2013.
12. Класифікація систем керування літаком. [Електронний ресурс]. Режим доступу: http://stud ope dia.su/10_106505_l--klasifikatsiya-sistem-keruvannya-litakom--c--.html. [Дата доступу: липень 2017].
13. Виникнення та еволюція комп'ютерних систем бронювання. [Електронний ресурс]. Режим доступу: http://pidruchniki.com/15801117/turizm/viniknenny a_evolyutsiya_kompyuternih_sistem_ bronyuvannya.
14. Посібник з BSP для Агентів Місцеві Процедури. [Електронний ресурс]. Режим доступу: http://w ww.iata.org/Sites/FMC/Files/ukraine-local-procedures-bsp-ukranian.pdf. [Дата доступу: липень 2017].
15. TAIS Airline Solution — система обслуживания авиапассажиров. [Электронный ресурс]. Режим доступа: http://tais.ru/solution/icarus/.
16. ЩЕРБАК Л., ГНАТЮК С., СИДОРЕНКО В., ШАХОВАЛ О.: Метод визначення рівня важливості об'єктів критичної інформаційної інфраструктури в галузі цивільної авіації, Безпека інформації, т. 23, №1, С. 27- 38, 2017.
17. СИДОРЕНКО В., ГНАТЮК С., ЮДІН О.: Експериментальне дослідження методу визначення рівня важливості об'єктів критичної інформаційної інфраструктури в галузі цивільної авіації, Захист інформації, т. 19, №2, С. 155-172, 2017.

Grygoriy TRIL [1], Hrystyna DANYLEVYCH [2]

Scientific Supervisor: Olexander BELEY[3]

'INTELIGENTNE' ANALIZY W ZARZĄDZANIU PROCESOWYM PRZEDSIĘBIORSTWEM

Streszczenie: W artykule rozważa się cechy aplikacji służącej do ekstrakcji danych (data mining) w procesie zarządzania podmiotami społeczno-gospodarczymi. W artykule zestawiono wszystkie czynniki oraz kryteria, które powinny być brane pod uwagę przy ocenie w zarządzaniu procesowym przedsiębiorstwem. Zaproponowano zastosowanie modelowania synergetycznego w oparciu o nieliniowe modele dynamiczne. Takie rozwiązania mają związek z własnościami procesów chaotycznych, które w adekwatny sposób odzwierciedlają obecne systemy socjalno-ekonomiczne. Możliwości formalnego opisu procesów losowych (stochastycznych) oraz określenia ich parametrów pozwoli biznesowi (zarządom) na uzyskiwanie stałego efektu synergetycznego

Słowa kluczowe: zarządzanie procesem, analiza inteligentna, efekt synergii, model dynamiczny, systemy chaotyczne

THE INTELLIGENT ANALYSIS IN PROCESS MANAGEMENT OF ENTERPRISE

Summary: The article considers the features of application of data mining in the process management of social-economic agents. The account of all the factors and criteria to take for evaluation of process management, proposed to use synergetic modeling based on nonlinear dynamic models. This is due to the properties of chaotic processes inherent in current socio-economic system. The possibility of formalization of random processes and their parameters will allow businesses to get permanent the synergetic effect.

Keywords: process management, intelligent analysis, synergetic effect, dynamic model, chaotic systems.

1. Formulation of the problem

[1] M.Sc., Lviv University of Trade and Economics, manager «McDonald's Corporation», postgraduate, bokalori3@gmail.com

[2] State university "The University of Banking", Lviv's educational institute, department of the "economic & information technology", student, kiber2@ukr.net

[3] Economic Dr, State university "The University of Banking", Lviv's educational institute, department of the "economic & information technology", associate professor, tiger_oles@i.ua

Effective management of the enterprise depends on the effective performance of each business-process. The effectiveness of process control parameters displayed for its implementation. Thus, the system performance execution processes of commercial enterprise based on information about: product quality, stability and reproducibility parameters of a particular product; quality process, its effectiveness and specific resource consumption, stability and reproducibility of process parameters; customer satisfaction, opportunity and feasibility of the client's needs, that provided.
The large number of multiparameter and multidimensional data for each process in the enterprise necessitates the creation and maintenance of multidimensional data stores. For processing and data mining need to implement multi-dimensional database with the possibility of intellectual processing.
We can confidently say that managing social-economic systems, we are dealing with random processes and their interaction is cyclical [1]. For control of the cyclic chaotic processes in enterprises we propose to use a complex of nonlinear modeling method. This set of methods can be associated with synergistic modeling, because it will determine the boundaries of these processes and the path of sustainable development. This is in turn suggests the existence of economic benefit as a result of a synergistic design and synergetic whole [2].

2. Analysis of recent research and publications

In a world of information and communication economy changes the content of the management in enterprises. All of the more important is used system and process management. The system allows managing all aspects of social-economic agents as structural units are connected. All elements economies have clearly defined final goals and initial objectives.
According to current trends in the management of [3], reminiscent of an engineering project, any enterprise defined processes. To perform these processes use resources and input information, and the result is a production of the original information.

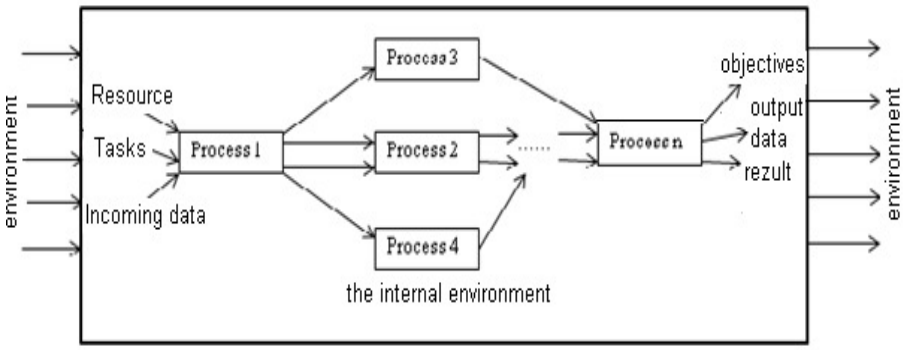


Figure 1. The system of process management of enterprises in information and communications economy

The goals and objectives of the company should be structured in a tree hierarchy of objectives under the control of reaching. Thus the main goal can be achieved if each tree its objectives clearly defined quantitative measure distance, and according to this

indicator, defined administrative center unambiguous responsibility. It follows that the solution of the problem of organizational structure of the enterprise process is to find bijective mapping topology formed for enterprise management objectives to tree topology management responsibility centers of the company.

From the standpoint of process control of system approach is to ensure that the head of the process continuously or at intervals prescribed course of the process controls and coordinates decisions on process parameters, which are formed within the regulated criteria. According to the scheme of the process control algorithm (Figure 2) process manager is plan the allocation of resources to achieve these goals with maximum efficiency. The implementation process checked the executing of information, which comes from places control points. Director of operational management process is the process of changing the planned allocation of resources, changing plans, timing and requirements according to the results of the ongoing changes. Operation manager process depends on the external and internal environment. Therefore, his decision is subject to strong influence of uncertainty and requires significant of operational formalization, what is possible through the detailed process approach.

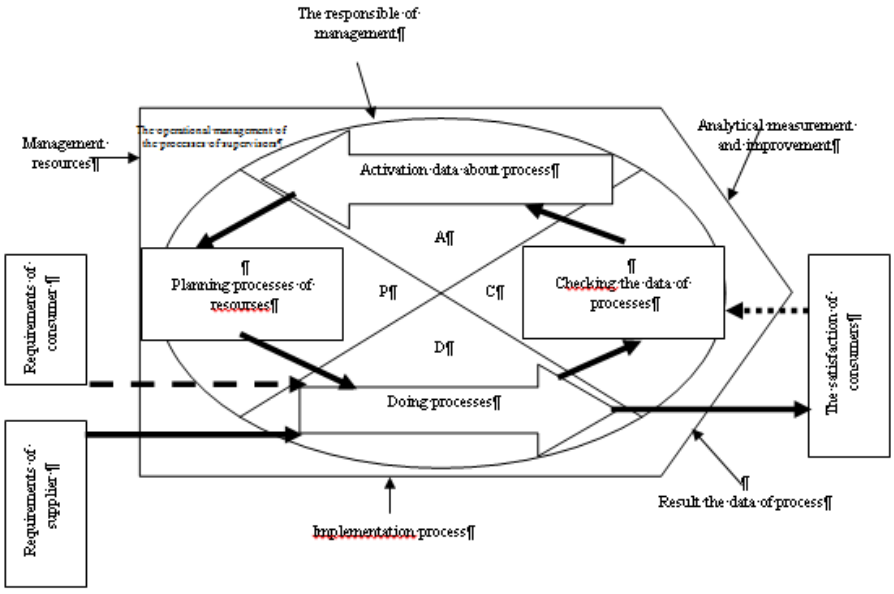


Figure 2. The data and resources of streams in the system of process management: A - Activation process; C - Checking processes; D - Doing processes; P - Planning processes of resources

Figure 2 is the algorithm of information systems of process management, which decomposes into components: the strategic management of the entire enterprise; the top management; the operational management of the processes of supervisors.

The approach is seen as an algorithm of actions leader in process management system to achieve the goal. Elements cycle management process comply with ISO 9000: 2015 (Figure 2).

As used in the literature [4] the term "limits of the process" establishes the area of responsibility and competence of the process. All values that cross these boundaries (inputs, outputs, resources) must have its data.
The system manages sales processes enterprise manager, which plans its activities; the responsibilities, authorities and relationships between processes. It is an analysis of the input parameters controls the process and set of commercial processes, deciding on the results of the analysis of their progress.
If the process of enterprise is proposed that a tree is the responsibility centers pyramid of process management – a system of interconnected cycle management PDCA (Plan-do-check-act), owners of managerial and technological business-processes. Then the quality of the implementation of this bijective mapping on a "structure follows strategy" ultimately will be determined by the ability and desire to create a model management company of process-oriented system management. It is proposed that model should allow management solution to maximize the economic efficiency of business on established on the balanced (*E)* topology tree set goals (*L)* key indicators (*V)* chosen strategy of business enterprise through the provision (*S)* for control indicators (*R*) pyramid of process management of acceptable rules handling (*W)* is based on the integration of management (*Abr)* and technological (*Tbp)* processes for each business enterprise management matrix ($b \in R$) [4]:

$$M = \max_{E \in S} \{E(S, V, L), P[R_{(b)}, W_{(n)}, Abr, Tbp]\} \quad (1)$$

We believe that the pyramid of process management in tree centers of responsibility for achieving the objectives tree will be relevant because it includes only those competencies that are necessary to achieve the objectives of the strategic objectives tree company.

3. Problem definition

For analyze the performance of each process and calculate the efficiency of the enterprise as a whole, taking into account all factors and parameters optimization model we see the need to use information technology and multidimensional data warehouse. This data warehouse will be streamlined multi-dimensional arrays, which are also often called OLAP-cubes. The technology OLAP (online analytical processing) allows for rapid removal of relevant information from large data sets and formation of appropriate reports.
From the above it is clear that the majority of real processes described by many parameters, properties, and other attributes. In particular, the description of the sales process may need information about your products or groups of supplier and buyer, the city where held sales and prices, quantity of goods sold and total. In addition, the tracking process in time to enter an attribute such as dates. If you collect all this information in the table, it will be difficult for visual analysis and interpretation. Moreover, it may be excessive if the same product sold in the same day in different cities, you have to repeat several times the same essence "city-good" indicating the various amounts and quantities. After all it will be difficult to extract from this table useful information to analyze the current state of sales and finding ways to optimize

the trade process. These issues occur for one simple reason: in two-dimensional tables are stored multidimensional data.

To avoid problems with multidimensional and parametric data decomposition we will make information a few simple tables, bind them some set of relations and transform into a relational model that uses classic database. This will be possible thanks to the multidimensional model of data representation that is implemented using multidimensional cubes.

Multidimensional cube we see as a coordinate system, the axes of which are to measure certain parameters of the model. For axes are delayed value measurement – the date, name of goods, title companies, buyers name or other individuals. In such OLAP-system measurements of each set of values will match the table in which you can place a number of indicators related to specific parameters. Thus, between the business-process objects and their numerical characteristics will be set unambiguous communication. The principle of multidimensional cube data mining commercial enterprise we have shown in Fig. 3. In box the first cell will house data related to the sale of cement PAT «Budivelnyk» at November 3, the second cell – for sale PAT «Patriot» at November 6, and the third cell – for sale boards PAT «Specbud» at November 4.

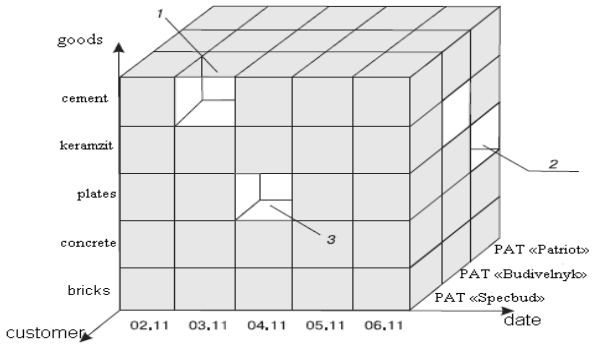


Figure 3. The structure of the multidimensional cube for Data Mining of trade enterprise

The multidimensional view of the measurement (Date, Goods and Buyer) shown in Fig. 4. Data in this case may be price, quantity, and amount. Then the selected segment will contain information about the number plates on how much and at what price the company has acquired PAT «Budivelnyk» at November 3.

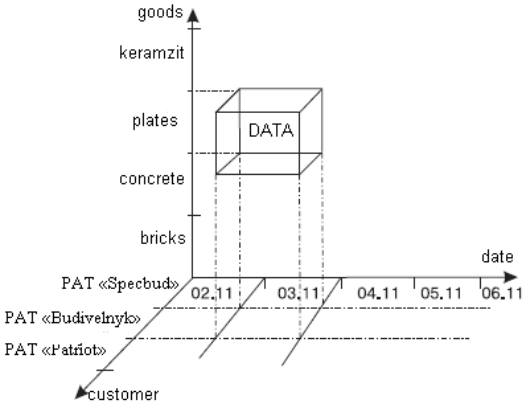


Figure 4. Indicators and data in a multidimensional cube for Data Mining of enterprises

Data in this case may be price, quantity, and amount. Then the selected segment will contain data on the number of boards on how much and at what price the company has acquired PAT «Budivelnyk» at November 3. Thus, the information in multidimensional data of warehouse is logically coherent. This is not just a set of records and numbers, which in the case of the relational model you want to receive from different tables, and the entire structure of the "who, what and how much was sold at any given time". The benefits of data mining based on multidimensional data warehouse we have identified the following features:

- Presentation of data in multidimensional cubes clearer than a set of standardized tables of the relational model structure which represents only a database administrator;
- The possibility of constructing analytical queries to the system widely used multidimensional data warehouse;
- In some cases, the use of multivariate models can significantly reduce the duration of the search data stores, providing performance analytic queries in real time. This is due to the fact that pre-aggregated data is calculated and stored in multidimensional cubes with detailed, so spend time calculating aggregates with a request not need.

Using multidimensional data model is accompanied by certain difficulties. So, for its implementation require more memory. This is due to the fact that the implementation of physical multidimensionality uses a large amount of technical information, because the amount of data that can be supported multidimensional model, typically less than several tens of gigabytes. In addition, the multi-dimensional structure harder to modify. If necessary, you can embed another dimension to perform physical restructuring of the entire multidimensional cube. Based on this we can conclude that the use of storage systems, which are based on multi-dimensional data representation, appropriate only in cases where the amount of data used is relatively small, and the multidimensional model has a stable set of measurements.

For the enterprise, as well as for any socio-economic system, is characterized by unstable dynamic processes. Very often, these processes can be compared to random processes. Such processes are becoming increasingly characteristic of today's businesses in the context of globalization, and the level of uncertainty in the economic balance observed only increases. In addition to all socio-economic systems are characterized by a certain cyclical (pattern) execution processes. In this regard, we

propose to examine the functioning of the enterprise as a chaotic system that contains an infinite number of unstable periodic cycles.
To address the question of stabilization processes and parameters of such a system, in our opinion, there should be used attractors. Their use would solve the problem of stabilization of any modern enterprise system by solving nonlinear equations that allow the search parameters of a chaotic system. In addition, the use of attractors will achieve a synergistic effect by stabilizing the socio-economic system and thus to anticipate the conditions and parameters achieve economic benefits.
Based on the Lorenz equations and offered him a strange attractor we tried to build own attractor for our proposed trade enterprise [5].

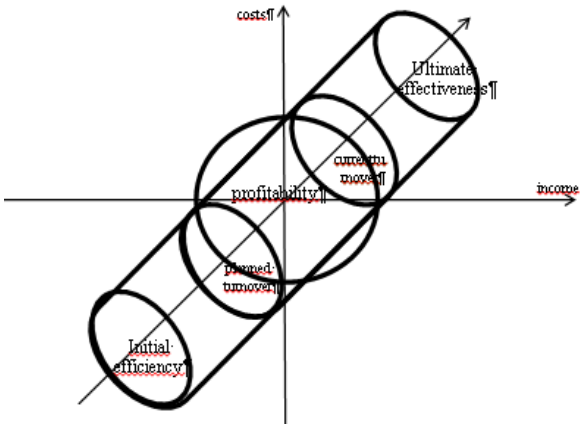


Figure 5. Indicators and data in a multidimensional cube for Data Mining of enterprises

To search parameters we propose to build a model for synergistic search parameters using OGY-method. It comes down to stabilize unstable periodic solutions through chaotic system of ordinary differential equations by applying discrete control actions in the form of feedback at the intersection of Poincare within a point of equilibrium which corresponds to desired cycle. View feedback linearization mapping is defined in the fixed point of the Poincare [6].
Note that the method used Pirahas` feedback with delay, and the delay time is close to the desired period of unstable periodic solution. So, consider smoothed autonomous system of nonlinear processes:

$$X = F(x,\mu), x \in X \subset R^n, \mu \in L \subset R^k, F \in C^\infty, \quad (2)$$

Set in the phase space (*X)* smooth vector fields (*F*), dependent on the coordinates of system parameters (*μ)*, which lie in the *(L)* space (R^k).
Let unstable limit cycle is the desired solution of the family system (1), for the same parameter ($\mu=\mu^*$) is regular or singular attractor.
Construct Poincare section *S*, which passes through the loop $x_0 = x^*(t,\mu^*)$ transversely $x^*(t,\mu^*)$ to it (perpendicular to the tangent at the point x_0 cycle $x^*(t,\mu^*)$). Consider the Poincare controlled display $x \to P(x,\mu)$, which $P(x,\mu)$ is the first point of returning to the plane trajectory *S* system (2) beginning at a point *x* on the

vector parameter value μ, which is in this case the vector control parameters. By applying a sequence of guided maps, obtain discrete dynamical system:

$$x_{n+1} = P(x_n, \mu_\eta), \tag{3}$$

Where $x_n = x(t_n)$, t_n - at time *n*-th intersection plane *S*, and μ_η - important vector control parameters in the interval between t_n and t_{n+1}.
Replace now display (3) close to it linearized at the point (x_0, μ^*) of reflection, that is:

$$y_{(n+1)} = Ay_n + Bu_n, A = \frac{\partial P(x_0, \mu^*)}{\partial x}, B = \frac{\partial P(x_0, \mu^*)}{\partial \mu}, \tag{4}$$

Where $y_n = x_n - x_0$, $u_n = \mu_n - \mu^*$ (fig. 6).

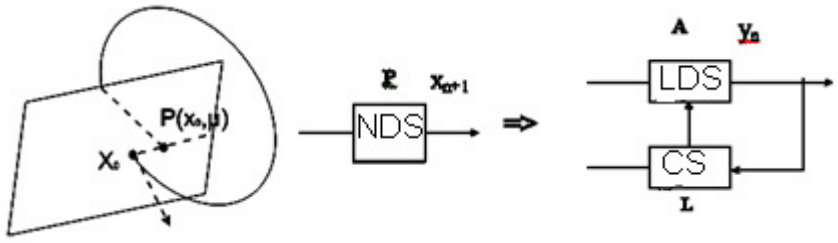


Figure 6. Illustration the linearization and display of Poincare

For a linear system (4) select stabilizing (y_n) management as a linear feedback as: $u_n = -Ly_n$. Then, (3) we find that:

$$y_{n+1} = (A - BL)\, y_n \tag{5}$$

Thus, the fixed point (x_0) display Poincare, and therefore sought an unstable cycle system (2) will be stabilized if determine the matrix (*L*) so, the matrix (*A-BL*) had its own meaning smaller units, that is, the condition of stability of linear discrete dynamical system (5).
Advantage of the OGY-method is that the stabilization of fixed point display and Poincare limit cycle system of differential equations achieved it insignificant manager's actions in discrete time. However, the major drawback is that the fixed point of the Poincare mapping is unsustainable. Therefore, the applicability of the method should not only know exactly matrix (*A*), but also its own values and own vectors corresponding stable and unstable variety that point. On the every iteration of trajectory should correct in the direction of sustainable diversity fixed point.
The big problem is also the choice of starting point. In OGY-method implicitly assumes that the system (2) has a chaotic attractor. This means that it is the closure of all periodic path contained therein; which implies that any path for any initial condition must eventually falls into a small interval desired cycle. But this is not so. Many singular attractors of this property have the persistence of the Feigenbaum` attractor, which coexists with an infinite number of unstable cycles and lies at a finite distance from each [7].

Based on the principles of smooth nonlinear control system processes the chaotic system of our company, our attractor, which depends on vector control parameters *u*, will have the following form:

$$\begin{gathered} X = F(x, \mu, u), \\ x \in X \subset R^n, \mu \in L \subset R^n, u \in L \subset R^n, F \in C^\infty \end{gathered} \tag{6}$$

Suppose you need to stabilize unstable limit cycle period (*T*), which is the solution of system (6) at (*u=0)* and (*μ=μ*)*. Thus we consider values (*u=0* and *μ=μ**) system (6), which has a singular attractor. Then the task of stabilizing the cycle ($x^*(t, \mu^*)$) can be solved in some cases the choice of a simple law of feedback delay, which has the form:

$$u(t) = K(x(t) - (x - T)) , \tag{7}$$

Where *K* – the matrix of transfer coefficient (the matrix of gain).
If the initial condition is set that the point (x_0) is sufficiently small neighborhood in orbit cycle, the solution (*x(t)*) of system:

$$x(t) = F(x(t)), \mu^*, K(x(t) - x(t - T))) , \tag{8}$$

Feedback (6), when ($\mu = \mu^*$) can converge to the desired unstable cycle ($x^*(t, \mu^*)$).
Analytical study of asymptotic properties of solutions of a closed system (8) is quite a challenge. Because until recently were known only numerical and experimental results that came out of properties and areas of applicability of the Pirahas` method.
The problem of finding sufficient conditions that guarantee the applicability of the method is still unresolved. In addition, the major drawback control law (6) is its sensitivity to timing delay (*T*). That is, if desired cycle time (*T*) previously known (namely, this situation is typical of the chaotic system of differential equations), then obtain the necessary convergence is possible only in exceptional cases, correctly calculating the value of the period of heuristic methods.

4. Statement of the main material

The analysis of the definition of synergistic effect allows us to assert that the most promising areas of improvement practices determine the effectiveness of the management in conditions of high uncertainty solve the problem is try to integrate the approaches outlined in order to use the full value that has each of them separately.
As one possible solution to this problem the authors propose the development of empirical correlation formulas or models which reflect the dependence functioning performance management system of the enterprise on the state of its business.
A common hypothesis in this case is the claim that value performance is determined by the level of development of the enterprise as an object of management. State control system thus can be solved in the form of parametric characteristics, which are multidimensional cubes. Accordingly, the basis for the formation of the final list of performance management system can serve as a link between overcrowding parametric change characteristics and changes in performance. However, the final list

of indicators to include only those that are most sensitive to any changes of the enterprise.
The positive side of this approach, in our view, it is a complete orientation to practice. Identifying a list of performance management systems now based on the establishment of correlation relationships between specific parameters of the system and its specific characteristics, the second - includes elements of chance and subjectivity in the formation of a set of indicators.
In analysis studied the relationship of 20 parametric performance management systems and studied 27 indicators of their effectiveness.
The main stage of preparing a list of performance management systems is a step of regression analysis. This method is known, allows formalizing the procedure for selecting the most significant indicators by calculating partial regression coefficients and coefficients of determination and determines their significance.
For parametric specification of management system were merged into three groups. Each group analyzed the correlation matrix and partial coefficients of determination of the factors included in the regression equation for each performance indicator.
Carrying step regression analysis allowed dividing all the performance indicators in terms of their information content into three groups. Analysis of the data shows that the introduction of OLAP-technology to the most informative performance hit or misses many of the indicators is allocated based on an analysis by the method of analysis groupings, and correlation analysis.
Table 1 shows data characterizing density connection between changes in these structural factors. As the number of managed objects and scope of their activities reflected the relationship of performance indicators in management systems.

Table 1. The matrix of coefficient pair correlation for individual performance indicators of management

Performance indicators	**The amount of facilities management**	**Volume facilities of management activities**
Productivity	-0,51	0,49
The level of expenditure rotation	0,57	-0,20
Assets of current assets	-0,8	0,09
Assets of fixed assets	-0,3	0,32

Synergetic effect of the introduction of process management and information technology can be viewed in different ways. In our opinion, there are three main areas to measure the synergistic effect of improving the management of socio-economic systems:
- With a combined synergistic effect derived from the activities on the basis of modern information technology;
- The share of cumulative effect, which is caused by process management based on synergistic simulation of random processes;

- For immediate synergistic effect of administrative work, this is also a messy process.
In our opinion, completely reject the first two methodological approaches to evaluating the effectiveness of management is not entirely true, because in certain situations they are eligible for use as basic conceptual terms. Regarding the third direction of research management efficiency, we then asked to review its basic position from the point of forming integral index of efficiency of management.
We are of the opinion that this line of research management efficiency involves two approaches: information and organizing. At the heart of the first of them is the idea of measuring the direct effect of control on the quantity and quality of information produced by the machine control particular company. Information of course, is a direct product of administrative work, but the practical use of such a proposal is extremely difficult, because there is no objective basis for comparison of different types of information. In addition, the quantity and quality of information provided to even reduce to a common denominator, not give answers to questions about the effectiveness of information potential and its impact on economic outcomes of facility management.
Organizational approach to management effectiveness is based on the premise that the effect of control is the level of the production process. However, the main methodological shortcoming of this approach is that high levels of the production process creates only prerequisite for its effective functioning, but does not guarantee the same efficiency.
Definitely final performance of the control system must match, that carry a single methodological framework. The criterion of economic efficiency of enterprise management systems must meet the following requirements: to display the final results of economic activity targeted managed object and the degree of their achievement; record spending regulator to achieve a managed object goals.
Based on the existing requirements under the criteria of economic efficiency management should understand the results management system that achieves a managed object goals at minimum cost management.
Generalized model of economic efficiency of management system that satisfies the above listed requirements can be formulated as follows:

$$EE = \sqrt{(\frac{R_b}{R_p} \cdot \frac{P_b}{P_p}) \times (\frac{\frac{NP_b}{CM_b}}{\frac{NP_p}{CM_p}})}, \quad (9)$$

Where: R_b, R_p - return (ratio of net profit to turnover) in the planning and the base period; P_b, P_p - productivity (the ratio of turnover by number of employees) in the planning and the base period; NP_b, NP_p - net profit in the planning and the base period; CM_b, CM_p - cost of maintaining the control system in the planning and the base period.
This increase in economic efficiency of enterprise management system can occur while performing one of the following four conditions: the effectiveness and efficiency of economic management while increasing; the effectiveness of economic activity increases and profitability management is stable; the effectiveness of economic activity is stable, and cost management increases; increasing efficiency of

economic activities to a greater degree than the reduced level of economic management. Considered a generalized indicator of the effectiveness of enterprise management systems sets the value depending on the efficiency of business operations (collectively functional business processes) the level of enterprise cost management entity. Its practical meaning is expressed in the fact that it is aimed at achieving not only high-efficiency operation of its management components. In addition, it can be used to assess the dynamics of economic efficiency of a particular commercial enterprise, and for comparative assessment of different options for improving control systems.

5. Conclusions

We can see that no matter how complete and comprehensive system is not a performance, it is not able to answer questions about the outcome of performance management system (regardless of the filling system integrators). This is due, firstly, various orientation of individual indicators. Second, due to the fact that any attempt to use a set of indicators to measure or assess the development of the functional process leads to a clash with another unsolved problem prioritization of individual indicators. Really existing practical difficulties unambiguous assessment of management efficiency through a system of indicators put economics before the need to search index received such an assessment. Such studies are carried out widely enough and now to build and implement effective of systems management.

REFERENCES

1. MIKHAILOV A.S., LOSKUTOV A.YU.: Foundations of Synergetics II. Chaos and Noise, 2nd revised and enlarged edition, Springer Series in Synergetics. Springer, Berlin-Heidelberg 1996.
2. HAKEN H.: Advanced Synergetic: Instability Hierarchies of Self-Organizing Systems and Devices. Springer-Verlag, NewYork 1993.
3. BARNETT W.A., HINICH M.: Has chaos been discovered with economic data? Nonlinear Dynamics and Evolutionary Economics. Oxford University Press, Oxford, 1993, 254–265.
4. KAPLAN R.S., NORTON D.P.: The Balanced Scorecard – Measures then drive Performance. The Harvard Business Review, Vol. 70(1992)1, 71-79.
5. LORENZ E.N.: Deterministic non-periodic flow. Journal of Atmospheric Science, (1963)20, 130–141.
6. OXLEY L.: Economics on the Edge of Chaos: How does economics deal with complexity and the implications for systems management. University of Canterbury, New Zealand, 2004.
7. ZHANG W.: Stability Versus Instability in Urban Pattern Formation. Occasional Paper Series on Socio-Spatial Dynamics, (1990)1, 41-56.

Ekaterina TRYFONOVA [1]

Scientific Supervisor: Alla A. KOBOZEVA [2]

WYRYWANIE NARUSZENIA INTEGRALNOŚCI OBRAZÓW CYFROWYCH Z ZASTOSOWANIEM SZUMU PERLINA

Streszczenie: W artykule zaprezentowano metodę wykrywania zafałszowania obrazu cyfrowego. Metodę tę opracowano na podstawie ogólnego podejścia stosując kilka różnych procedur m.in. analizę macierzową, teorię perturbacji. Metoda pozwala na rozwiązanie problem wykrycia naruszenia integralności sygnału cyfrowego. Przez zafałszowanie obrazu cyfrowego – w niniejszej pracy – rozumie się naruszanie integralności obrazu cyfrowego, które jest oparte na modelowaniu wizualnym naturalnych zjawisk takich jak: chmury oraz niebo, jako rezultat generowania szumu Perlina.

Słowa kluczowe: obraz cyfrowy, fałszerstwo, szum Perlina, rozkład na wartości osobliwe

DETECTION INTEGRITY VIOLATIONS OF DIGITAL IMAGE BY PERLIN NOISE

Summary: The paper presents a method for detecting falsification of a digital image, developed on the basis of a general approach based on the use of matrix analysis, perturbation theory, and provides an opportunity to solve the problem of detecting a violation of the integrity of a digital signal. Under the falsification of a digital image in a work is understood a violation of the integrity of a digital image, which is based on the modeling of a visually realistic natural phenomenon, clouds and sky, as a result of Perlin noise generation.

Keywords: digital image, forgery, Perlin noise, singular decomposition.

1. Formulation of the problem

Thanks to a significant increase in the computing power of computers in recent times and the active development of computational methods of computer graphics, the tools of various graphic editors are constantly enriching, the task of realizing the realistic falsification of the digital image has become extremely simple. So, the task of proving

[1] Odessa National Polytechnic University, Department of Informatics and Information Security Management, Master of Science, katikkatik@gmail.com

[2] Prof. D.Sc., Odessa National Polytechnic University, Department of Informatics and Information Security Management, Head of Department, alla_kobozeva@ukr.net

the authenticity, revealing the violation of the integrity of the digital image, the solution of which in many areas of human life, such as: forensic examination, medical diagnostics, military intelligence, electronic document management and others, is crucial, it becomes more difficult year after year. The purpose of this work is to increase the effectiveness of the method of detecting falsification of a digital image.

2. Analysis of recent research and publications

At present, in order to solve the task, active work is being done to develop methods for detecting the falsification of digital images [1-4].
The greatest attention is paid to the study of falsification, which is realized with the help of an affine transformation of transfer, as the most simple and widespread method of falsification [1,2]. The next most common method of falsification is violation of digital image integrity as a result of composition of two affine transformations: transfer and stretching [3] or rotation of some rectangular image area [4].
One of the most complex methods of falsification for research and detection is the violation of digital image integrity, based on the modeling of visually realistic natural phenomena that are an integral part of almost any natural scene: objects such as clouds and sky.
To date, a variety of cloud modeling methods have been proposed and implemented [5]. One of the most common methods for modeling visually realistic clouds, implemented in common and commonly available graphical editors, is a method based on the implementation of Perlin noise.

3. Problem definition

The paper carries out research of digital images saved in the format without loss, the following types. The first is digital images containing the sky and clouds produced by modern digital cameras. The second is digital images obtained as a result of noise generation of the Pearl to simulate a visually realistic natural phenomenon, clouds and sky. The third is the digital images simulated when replacing the part of the main image with the artificial region obtained as a result of the generation of the Pearl noise.
To achieve the goal, the following tasks must be solved:
1) determine the characteristic signs of the parameters of the image, the presence or absence of which allows to establish violations of the integrity of the digital image for the presented method of falsification;
2) based on certain features of the digital image parameters, develop a method for increasing the effectiveness of the method of detection of falsification.

4. Statement of the main material

In [1], on the basis of perturbation theory and matrix analysis, a general approach has been developed that provides an opportunity for solving the problem of detecting the violation of the digital signal integrity.

According to the bases, the mathematical parameters that contain information about the state, and their perturbation - information about the change of the state of the digital signal is proposed. Various methods of perturbation (in particular, various ways of breaking the integrity) of a digital signal result in various characteristic perturbations of mathematical parameters signaling the corresponding effect.

Let $F = (f_{y,x})$ $R \times C$ the matrix, $R \geq C$ with elements $f_{y,x}$, $y = \overline{1, R}$, $x = \overline{1, C}$. Singular value decomposition [6]:

$$F = U \Sigma V^T . \tag{1}$$

Columns $u_1, ..., u_C$ of the matrix U are called the left singular vectors of the matrix F . The columns $v_1, ..., v_C$ of the matrix V are called the right singular vectors of the matrix F . Values $\sigma_1, ..., \sigma_C$ - singular values [6].

The singular value decomposition is not unique in the general case. According to [7], the vector is called lexicographically positive if its first non-zero component is positive. The singular value decomposition is called normal if the columns of the matrix are lexicographically positive and the diagonal elements of the matrix are in a non-increasing order. In accordance with Theorem [7], the matrix has a unique normal singular value decomposition if its singular numbers are pairwise distinct and nonzero. Thus, singular values and singular vectors obtained as a result of normal singular value decomposition determine the matrix unambiguously.

Consequently, a set of singular vectors and singular values obtained as a result of the normal singular value decomposition of the digital image matrix carry all information about the state of a digital signal [1].

To solve the problem of detecting the violation of the digital image integrity, based on the visually realistic natural phenomenon simulation, clouds and sky, as a result of the generation of Perlin noise, consider the following forms for representing the parameters of the digital image matrix: a discrete function of singular values block of an image matrix; discrete function of the ratio of singular values block of the image matrix; discrete speed function for changing the ratio of singular values block of the image matrix.

The discrete function of singular values for a block $F_{k,p} = (f_{y,x})$, $x = \overline{(p-1)n+1, pn}$, $y = \overline{(k-1)n+1, kn}$, $n \times n$, of an image matrix $F = (F_{k,p})$, $k = \overline{1, \left[\frac{R}{16}\right]}$, $p = \overline{1, \left[\frac{C}{16}\right]}$, $R \times C$, defined as follows:

$$\Omega_i(F_{k,p}) = \sigma_i, \quad i = \overline{1, n} . \tag{2}$$

In accordance with the geometric interpretation of the matrix singular value decomposition, the singular values represent the lengths of the ellipsoid semi-axes. The ratio of the smallest to the largest length of the ellipsoid semi-axis, that is, the ratio of the smallest to the largest singular value determines the condition number of the matrix, which is related to the eccentricity. Thus, the definite condition matrix number represents the degree of splicing of the ellipsoid only with respect to one of the smallest semi-axes.

The concept of the condition number may be expanded, in the research of the degree deviation each semi-axis from the smallest for the block of the digital image matrix. Consequently, the discrete function of the ratio of singular values for a block $F_{k,p} = (f_{y,x})$, $y = \overline{(k-1)n+1, kn}$, $x = \overline{(p-1)n+1, pn}$, $n \times n$, of an image matrix $F = (F_{k,p})$, $k = \overline{1, \left[\frac{R}{16}\right]}$, $p = \overline{1, \left[\frac{C}{16}\right]}$, $R \times C$, is defined as follows:

$$\overline{\Psi}_i(F_{k,p}) = \frac{\sigma_i}{\sigma_n}, \quad i = \overline{1, n}. \tag{3}$$

For further research, the discrete function of the ratio of singular values for a block in a logarithmic scale is used:

$$\Psi_i(F_{k,p}) = \log_{10}\left(\frac{\sigma_i}{\sigma_n}\right), \quad i = \overline{1, n}. \tag{4}$$

To study the rate of change in the ratio of block image matrix singular values, the derived function (4) is determined. Since the function is discrete, for the numerical differentiation to be performed, a divided first-order difference is applied.
The discrete rate function of change in the ratio of singular values for a block $F_{k,p} = (f_{y,x})$, $y = \overline{(k-1)n+1, kn}$, $x = \overline{(p-1)n+1, pn}$, $n \times n$, of an image matrix $F = (F_{k,p})$, $k = \overline{1, \left[\frac{R}{16}\right]}$, $p = \overline{1, \left[\frac{C}{16}\right]}$, $R \times C$, is defined as follows:

$$\Phi_i(F_{k,p}) = \log_{10}\left(\frac{\sigma_{i+1}}{\sigma_i}\right), \quad i = \overline{1, n-1}. \tag{5}$$

Thus, the main steps of the method for detection integrity violations of digital image by Perlin noise are as follows:

a) to construct of a digital image matrix $F = (F_{k,p})$, $k = \overline{1, \left[\frac{R}{16}\right]}$, $p = \overline{1, \left[\frac{C}{16}\right]}$,
b) $R \times C$;
c) to divide the resulting matrix on blocks with size 16×16, $F_{k,p} = (f_{y,x})$, $y = \overline{16 \cdot (k-1)+1, 16 \cdot k}$, $x = \overline{16 \cdot (p-1)+1, 16 \cdot p}$;
d) to build a discrete rate function of change in the ratio of singular values for each block of the image matrix $\Phi_i(F_{k,p}), \quad i = \overline{1, 15}$;
e) to allocate the block of image matrix for which the condition is fulfilled:

$$\overline{=}{F} = \left\{ F_{k,p} \Big| \max_{i=1}^{15} \left(\left| \Phi_i(F_{k,p}) \right| \right) > 10, k = \overline{1, \left[\frac{R}{16}\right]}, p = \overline{1, \left[\frac{C}{16}\right]} \right\}, \tag{6}$$

represent blocks that are the result of integrity violations of digital image by Perlin noise.

On the obtained results of the forms for digital image parameters representation, which uniquely determine the digital signal, - singular values, consider the method of detecting falsification, which is based on the simulation of a visually realistic natural phenomenon, clouds and sky, as a result of the generation of Perlin noise.
Consider the digital images of the sky and clouds that are obtained by modern digital cameras. Let the part of the main image (OZ) be replaced by the artificial region, which was obtained as a result of the generation of Perlin noise (ZO) (no further image processing is performed for greater clarity of the conclusions below).
Thus a digital image constructed, the example of which is based on the image (Fig. 1a), demonstrates the integrity of the digital image, presented in (Fig. 1b) is stored without loss.
The graphs of a discrete rate function of change in the ratio of singular values for block of the image matrix for the OZ and of the image containing ZO are clearly demonstrated the area of integrity violation (Fig. 1d, e).

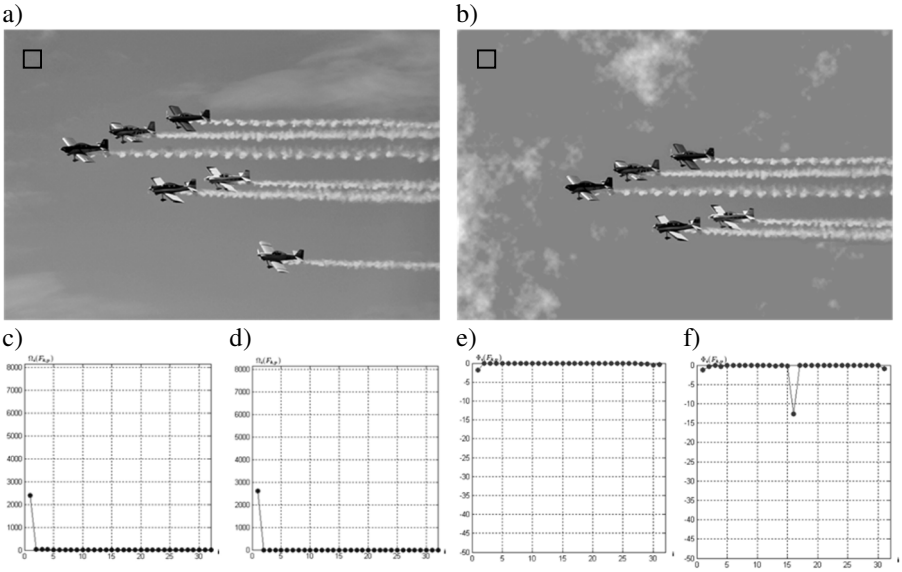


Figure 1 – Example of integrity violations of digital image by Perlin noise a) – OZ F^I *; b) – OZ with ZO* F^P *; c) –* $\Omega(F^I_{2,2})$*; d) –* $\Omega(F^P_{2,2})$*; e) –* $\Phi(F^I_{2,2})$*; f) –* $\Phi(F^P_{2,2})$

5. Conclusions

As a result of this work, in order to increase the effectiveness of the method of detecting the falsification of a digital image, based on the violation of the integrity of a digital image, based on the simulation of a visually realistic natural phenomenon, clouds and sky, as a result of Pearl noise generation, it is proposed to use a new form of representation of the parameters of a digital image.
Based on the discrete function of the rate of change in the ratio of singular values, the main steps of the method of detecting the falsification of a digital image of the type under consideration are presented.

REFERENCES

1. КОБОЗЕВА А.А.: Матричный анализ – основа общего подхода обнаружения фальсификации цифрового сигнала, А.А. Кобозева, О.В. Рыбальский, Е.А. Трифонова, Вісник Східноукраїнського національного університету ім. В.Даля. №8(126). Частина1.2008.62–72.
2. ТРИФОНОВА К.О.: Підвищення ефективності методу виявлення несанкціонованого втручання «copypaste» в цифрове зображення, К.О. Трифонова, Вісник Національного університету, Львівська політехніка, Автоматика, вимірювання та керування. Львів. 741(2012), 227–230.
3. ТРИФОНОВА Е.А.: Метод идентификации и локализации масштабирования в цифровом изображении, Е.А. Трифонова , Информатика и математические методы в моделировании. Одесса. 1(2013), 22–34.
4. ТРИФОНОВА К.О.: Адаптація інваріантних до афінних перетворень моментів цифрового зображення для виявлення фальсифікації, В.В. Мурова, К.О. Трифонова, Тези доповідей V міжнародної науково-практичної конференції, Інформаційні управляючі системи та технології Одеса. – 20–22 вересня 2016р. С. 40–41.
5. HARRIS M.: Simulation of cloud dynamics on graphics hardware. M. Harris, W. Baxter, T. Scheuermann, A. Lastra, Proceedings of Graphics Hardware 2003, 92–101.
6. ДЕММЕЛЬ ДЖ.: Вычислительная линейная алгебра. Теория и приложения, Дж. Деммель. 2001.
7. BERGMAN C.: Unitary Embedding for Data Hiding with the SVD / C. Bergman, J. Davidson, Security, Steganography, and Watermarking of Multimedia Contents VII, SPIE 5681(2005), 619–630.

Joanna WALUS[1], Paweł RUDYK[2]

Opiekun naukowy: Stanisław ZAWIŚLAK[3]

EWOLUCYJNE UJĘCIE 2-KRYTERIALNEGO PROBLEMU MINIMALNEGO DRZEWA NAPINAJĄCEGO GRAFU

Streszczenie: W pracy omówiono algorytm oraz program komputerowy do wyznaczania minimalnego drzewa napinającego grafu w ujęciu dwukryterialnym. Zastosowano algorytm ewolucyjny uproszczony. Wyniki działania przedstawiono także graficznie. Wyznaczono zbiór rozwiązań kompromisowych tzw. front Pareto. Jedną z opcji problemu jest wprowadzenie ograniczenia na maksymalny stopień wierzchołka drzewa.

Słowa kluczowe: optymalizacja dyskretna, problem z ograniczeniami, algorytm ewolucyjny, drzewo napinające

EVOLUTIONARY APPROACH TO BI-CRITERIA PROBLEM OF MINIMAL SPANNING TREE IN A PARTICULAR GRAPH

Summary: In the paper, the problem of spanning tree is considered. The algorithm and the computer program are described for the multicriteria version of the task. Simplified evolutionary algorithm was utilized. Results of the calculations are presented in graphical form, as well. The so called Pareto front of solutions was distinguished. The constrained version of the problem was also considered – taking into account the restriction on the vertices degrees of the considered graph.

Keywords: discrete optimization, constrained problem, evolutionary algorithm, spanning tree

1. Introduction

Graph theory is a branch of mathematics which was originated in 1736 by Leonhard Euler. Nowadays, the algorithmic approach to graph theory is extremely important due to development of networks. Weighted graph can be interpreted as a net, therefore graph-based methods are widely used for network analysis. One of the popular

[1] University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, jwalus1313@gmail.com
[2] University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, vagrant326@gmail.com
[3] Assoc. Professor [Dr hab. inż.], University of Bielsko-Biala, Faculty of Mechanical Engineering and Computer Science, szawislak@ath.bielsko.pl

network topologies is just tree-like layout of nodes and edges. Plain problem of spanning tree has been successfully solved. Three algorithms are known: elaborated by Kruskal, Prim and Boruvka, respectively. But in many cases more than criteria should be taken into account. In such a case, the standard algorithms are not adequate, moreover in case of introducing some constraints to the problem, they also could not be simply modified.
Therefore, in such cases the evolutionary approach is relatively easy to implement as well as it could be effective. The multi-criteria problems can be solved via AI-based algorithms. The obtained solution are not straight simple, on contrary the solution set have to be additionally analyzed to choose final one.

2. Problem formulation

In the paper, the spanning tree problem is considered. In graph theory, tree is a sub-graph of a particular graph which does not contain cycles as well as it is built of *n-1* edges, where *n* denotes number of tree vertices. Graph G(V,E) consists of two sets: V – set of vertices (non-empty) and E – set of edges. If E is empty than graph consists of separate vertices, only. Graph is named connected if for each pair of its vertices, the path exists – which connects these chosen vertices. We consider connected graphs. The spanning tree of a particular graph is a tree induced on all graph vertices, so in this case numbers of vertices of graph and tree are equal.

Weighted graph is a graph in which weights are assigned to the edges. So, we consider weight functions f_1, f_2 assigning the natural numbers in the range [1,99]. Weight of a particular tree is a sum of weights of all its branches.

We consider a weighted clique K_n. Then, according to Cayley's formula, number of all spanning trees is equal to: $n^{(n-2)}$. It means, that for some n (e.g. $n > 30$) the full search approach to tree enumeration is purposeless. The goal of our consideration is distinguish the multi-criteria minimal spanning trees. We propose to use the evolutionary approach to solve this problem.

The interpretation of the problem would be building of the tree-like net where weights assigned to branches mean e.g. length and cost. Cost is not a direct cost of the wire or optical fiber needed for building a network but also costs of works which depends on drilling holes, ground digging etc.

3. Algorithm

The simplified evolutionary algorithm was utilized. The advantages of the evolutionary approach are: consideration of population of working points so called chromosomes, flexibility, possibility of adaptation to the particular considered problem, special tailored evolutionary operations. In turn, the drawbacks consist in lack of precise knowledge about the reaching the final solution, need of numerical experiments, duplication of elements inside a population. The evolutionary algorithm repeats the sequence of actions until the termination condition is fulfilled. After random generation of initial population, actions are as follows: mutation, evaluation and succession. In case of two considered criterial, the applied succession consists in division of population into two equal subpopulations, performance of tournaments in

both parts according to two different target functions and final reshuffling of the population.

Special introductory generation of population is needed. A chromosome means in our case a tree. So, the procedure of tree generation was prepared. The evolutionary operation which was utilized is mutation. In the proposed version mutation of tree means removal of a randomly chosen branch. The result of this action is division of tree into two separate subtrees so the graph is disconnected. Next step consist in repair step i.e. adding of new randomly generated branch which connects two components of connectivity into one spanning tree.

4. Program description

Computer program was written in C# language.

4.1. Application specification

Our application may be used to solve two-criterial minimal spanning tree problem using evolutionary algorithms. It allows to set all important parameters of this algorithm. Our application also allows user to see results in two modes – automatic or "step-by-step" control. User has constant view on trees generated in each generation. They can analyze the results thanks to charts presenting all specimens and Pareto-fronts. After simulation, user is able to save results to file.

4.2. How to use application

Application must get a set of input data. This set consist of two text files which contain value of F1 and F2 edge weights. Data can be loaded from local disk using "Load F1" and "Load F2" buttons, so the exemplary data were generated and saved. Application won't allow any further steps if user doesn't load data. Satisfying this condition will unlock some disabled options. Figure 1 shows an example of F1 weights values – collected in the given text file.

After loading F1 and F2 weights, user can change some parameters, which affect simulation:

- Max. vertex degree – allows to create constraint on number of edges connected to one vertex - therefore the problem is with the constraint,
- Population size – number of specimens in one generation,
- Generations number – this parameter affects quality of result. More generations mean more attempts to find best solution,
- Mutation probability – mutation algorithm randomly change one edge of the tree. In first step, algorithm selects random edge and deletes it. Then it marks two subgraphs created by deleting bridge between them. Lastly, algorithm choose two random vertices, which degree doesn't exceed constraint made by the user and connects them with new edge,
- Generation interval – allows to set time delay between generations which allows for monitoring of the running of the programme,

- Succession type – our application allows user to choose one of two types of succession:

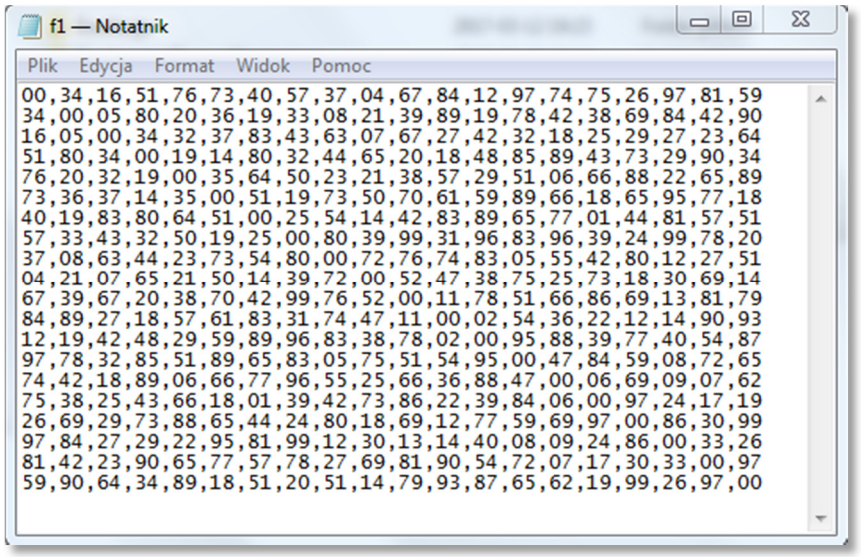


Figure 1. Input data example

Tournament succession divides population in few equal groups. Number of groups depends on number of criteria considered by algorithm. All members of groups are then compared with other random members by criterion important in this group. Specimen with better result succeeds to next generation. Because of random opponent selection, this option allows succession of many copies of one specimen. Custom Succession initially sorts population by sum of normalized criteria ratings and divides it in ten equal groups. With descending sum of ratings in each group, there is bigger probability of elimination from population. In next step, application randomly eliminates half of the population, using probability from previous step. Free slots are then occupied by copies of specimens which survived.

Execution type – application can be controlled in step-by-step mode, which allows user to see every generation individually or in automatic mode which shows all generations with some time delay.

Start button runs the simulation. Application menu is shown on Figure 2. Application gives preview of three the most important specimens of every population – best, average and worst specimen as shown on Figure 3. Each of these charts is made of vertices and edges connecting them written on the circle. Index of vertex is written under every point. Color of line representing an edge reflects sum of its weights. Green colors mean good, low values, while red ones mean high values. Clicking on any of these charts opens new window with zoomed chart. Title of this window, shown on Figure 4, is the Prűfer code of that tree.

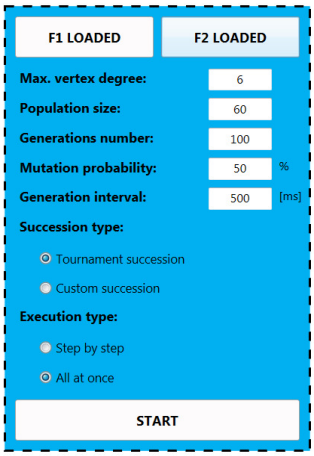


Figure2. Application menu

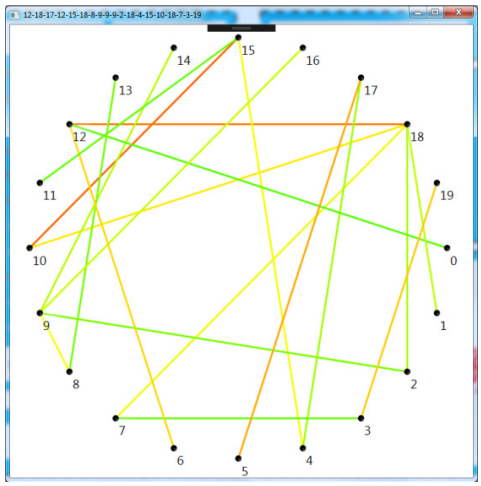


Figure 3. Zoomed specimen

Generation chart shows results of the algorithm. Each point on the chart represents one specimen. Location of these points mean value of F1 and F2 functions. Best specimens can be found in lower left corner. Color of the point allows to roughly estimate specimen generation. Red colors mean early generations while green colors mean final generations. Thanks to this chart, user can determine the rate of improvement of the results as well as the overall trend of subsequent generations.

Pareto front graph shows best specimens from all generations. There can be many such individuals because it is a multi-criteria problem and specimen with best rating in one criterion doesn't have to be the best one in other criteria. Color of point in this graph allows to estimate specimen generation, just like in generation chart. Red colors mean early generations while green colors mean final generations. Clicking on any point in this graph opens a new window with tree represented by this point. Population and Pareto Front diagrams are shown on Figure 5. In original version the displays are shown in colors therefore the user can observe the present and former generations simultaneously on the screen. The same, what more important, can be said about the Pareto front. Black-white printing makes unable to show this advantages.

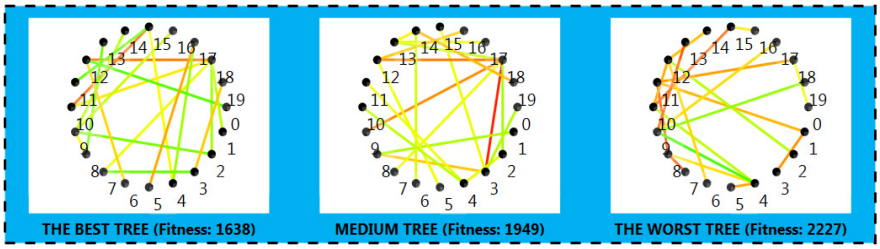


Figure 4. Preview of most important specimens

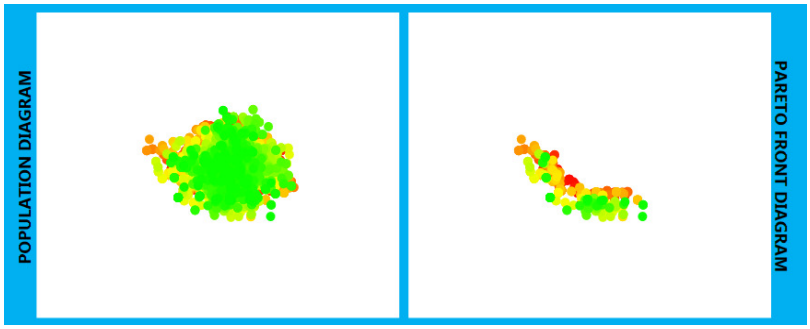


Figure 5. Population and Pareto Front diagrams

User can use application in step-by-step mode. It allows user to see every generation individually using "Step forward" and "Step backward" buttons. These buttons can be found in navigation bar, shown on Figure 6, along with index of currently displayed generation.

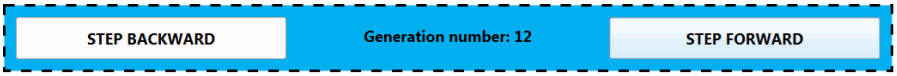


Figure 6. Navigation bar in step-by-step mode

Automatic mode may be used to save time. It runs the simulation and generates needed results without any attention.

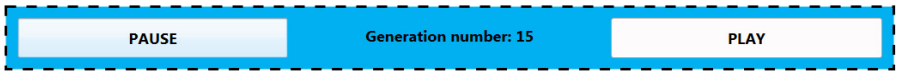


Figure 7. Navigation bar in automatic mode

Results will be displayed with some time delay, which can be configured by user, who can also stop the simulation at any time using Stop button. After that, simulation may be continued or restarted. Application can be controlled using navigation bar shown on Figure 7.

After the simulation user can save results to file. Example of such file is shown on Figure 8. The codes of trees are listed. The values of total weights of the adequate trees are given in two columns.

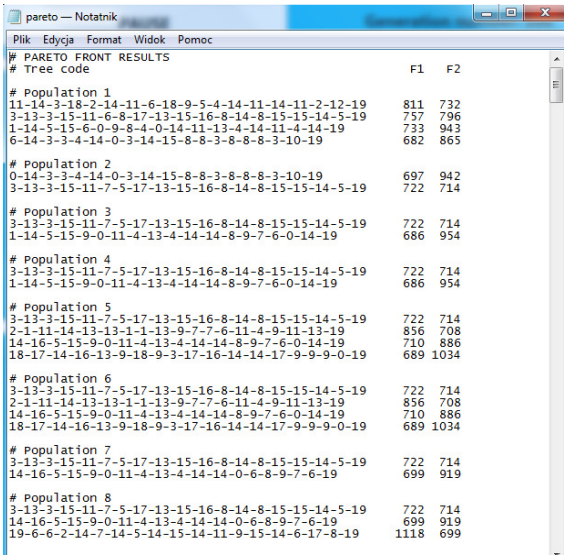


Figure 8. Saved results

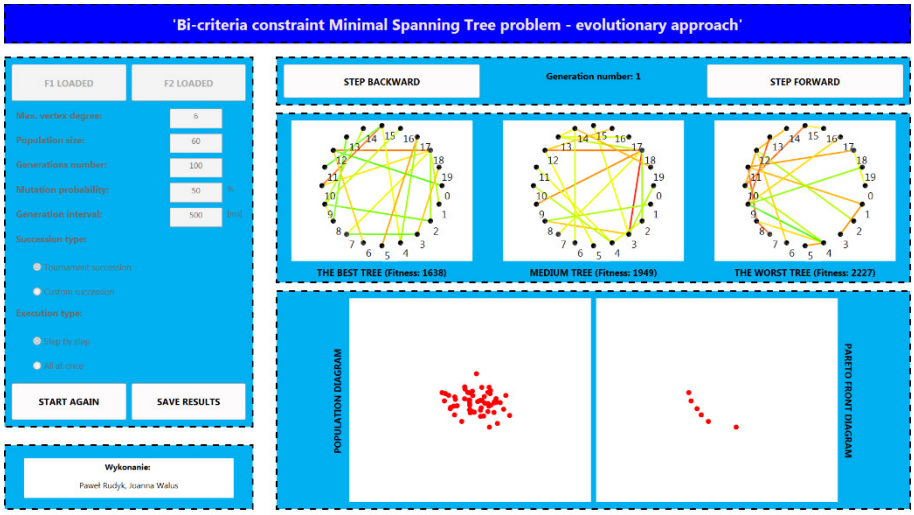


Figure 9. General view of application window

The general view of the program is given in Figure 9. All the above described panels are displaced in the screen. The user can choose options and observe dynamical changes of results in two sub-windows in the lower right corner of the screen. In this particular case the initial Pareto front for the start population and the population itself are shown.

5. Summary

The discussed application, thanks to evolutionary algorithms, allows solving two-criterial minimal spanning tree problem. The Pareto front being a set of compromise solution was distinguished in every case. Two options of succession can be chosen by a user. Some parameters of the algorithm can be set by a user which could be a useful tool for observation of behavior of the algorithm and layout of obtained solutions.

REFERENCES

1. ARABAS J.: Wykłady z algorytmów ewolucyjnych, WNT, Warszawa (2001).
2. CHRISTOFIDES N.: Graph Theory. An Algorithmic Approach, Academic Press Inc., New York (1975).
3. CHARTRAND G., OELLERMAN O.R.: Applied and Algorithmic Graph Theory, McGraw-Hill, New York (1993).
4. CRAVEIRIHA J., et al.: A bi-criteria minimum spanning tree routing model for MPLS/overlay networks. Telecommunication Systems, 52.1 (2013), 203-215.
5. EHRGOTT M., GANDIBLEUX X.: A survey and annotated bibliography of multiobjective combinatorial optimization. Or Spectrum, 22.4(2000), 425-460.
6. FERNÁNDEZ F. R.; HINOJOSA M. A.; PUERTO J.: Multi-criteria minimum cost spanning tree games. European Journal of Operational Research, 158.2(2004), 399-408.
7. OSYCZKA A.: Evolutionary algorithms for single and multicriteria design optimization, Physica-Verlag a Springer-Verlag Company, Heidelberg (2002).
8. PAGACZ A., RAIDL G., ZAWISLAK S.: Evolutionary Approach to Constrained Minimum Spanning Tree Problem - commercial software based application. IX National Conference Evolutionary Computation and Global Optimization 2006, Ed.: Jarosław Arabas; Prace Naukowe. Elektronika, Politechnika Warszawska, z.156, Oficyna Wydawnicza Politechniki Warszawskiej, (2006).
9. RAIDL, G. R.; JULSTROM, Bryant A. Edge sets: an effective evolutionary coding of spanning trees. IEEE Transactions on evolutionary computation, 7.3(2003), 225-239.
10. SILVA C. G.; CLI'MACO J. C.: A note on the computation of ordered supported non-dominated solutions in the bi-criteria minimum spanning tree problems. Journal of Telecommunications and Information Technology, (2007), 11-15.
11. WOJCIECHOWSKI J., PIEŃKOSZ K.: Grafy i sieci, PWN, Warszawa, 2013.
12. WILSON R.J.: Introductory graph theory, (Polish translation: „Wprowadzenie do teorii grafów"), PWN, Warszawa (2017).

Olga WESELSKA[1], Oleksandr SZMATOK[2]

Opiekun naukowy: Oleksandr JUDIN [3]

NOWOCZESNE METODY WYKRYWANIA UKRYTYCH INFORMACJI W OBRAZACH STATYCZNYCH

Streszczenie: W pracy przeprowadzono symulację ataków steganograficznych na steganosystem w celu identyfikacji ukrytych informacji w obrazie statycznym. Ukryte informacje są osadzone w przestrzeni obrazu statycznego metodą najmniej znaczącego bitu (LSB). Metody badania opierają się na teorii ochrony steganograficznej, metodach osadzania danych w statycznych obrazach, metodach statycznej analizy procesów losowych, przemian afinicznych oraz falkowych.

Słowa kluczowe: steganografia, steganokontener, najmniej znaczący bit, przestrzenny obwód obrazu, korelacja, przekształcenia afiniczne, ataki steganograficzne, pasywne ataki, transformacja falkowa.

MODERN METHODS FOR DETECTING INFORMATION IN STATIC IMAGES

Summary: In this paper, simulations of steganographic attacks on steganosystems were performed to identify information in a static image. Hidden information is embedded in the spatial space of the static image by the least significant bit (LSB) method. The test methods are based on the theory of steganographic protection, methods of embedding data in static images, static methods of analysis of random processes, affine and wavelet transformations.

Keywords: steganography, steganocontainer, least significant bit, spatial image area, correlation, affine transformations, steganographic attacks, passive attacks, continuous wavelet transform.

1. Wstęp

Steganografia jest skutecznym sposobem ochrony informacji. Jednocześnie metody stegonograficzne mogą być wykorzystywane przez intruzów do niewidocznego przekazywania informacji przez kanały komunikacyjne. Takimi intruzami mogą być

[1] National Aviation University: Institute of Computer Information Technologies, email olga_veselskaya@ukr.net
[2] Candidate of Sciences (Technical) National Aviation University, Institute of Computer Information Technologies, email sh_al_st@mail.ru
[3] Professor, Doctor of Science (Technical), National Aviation University, Institute of Computer Information Technologies, email kszi@ukr.net

terroryści, szpiedzy, cyberprzestępcy, konkurenci, którzy prowadzą szpiegostwo w przedsiębiorstwach i używają metod steganograficznych do przekazywania poufnych informacji lub rozprzestrzeniania ukrytego wirusa.
Zagrożenie utraty poufnych informacji przy użyciu metod steganograficznych jest wystarczająco duże, ponieważ nikt nawet nie zauważy faktu jej wykradzenie. Dlatego konieczne jest opracowanie skutecznego sposobu ujawniania ukrytych informacji, ciągłe przeszukiwanie wszystkich informacji transmitowanych przez kanały komunikacyjne i niszczenie ukrytych informacji w przypadku ich wykrycia.

2. Znaczenie

Celem pracy jest opracowanie metody wykrywania faktu obecności ukrytych informacji za pomocą ataków steganograficznych, takich jak przekształcenia afiniczne i falkowe, w celu przesyłania statycznych obrazów w kanale komunikacyjnym. Obraz statyczny jest pojemnikiem steganograficznym.
Nowością naukowych badań jest konieczność analizowania obrazów przesyłanych do lub z przedsiębiorstwa lub Internetu, w celu ujawnienia obecności ukrytych informacji w tym obrazie, czyli faktu uprowadzenia informacji poufnych lub rozprzestrzeniania złośliwego oprogramowania. Osiągnięcie tego celu oznacza rozwiązanie następujących zadań:

- wprowadzenie do pojemnika ukrytych informacji za pomocą metody steganograficznej;
- korelacja i analiza spektralna pustego i napełnionego pojemnika;
- symulacja przemian afinicznych i falowych.

3. Symulacja pasywnych ataków na stegosystemy.

Przed przystąpieniem do modelowania ataków pasywnych na stegosystemy wprowadzamy informacje do pojemnika, używając metod steganograficznych.
W tym eksperymencie jako kontener użyto statycznych obrazów w formacie BMP o rozmiarach 640x480 pikseli. Informacja – tekst formatu ASCII w różnych rozmiarach: 37,4 Kb (100% pojemności pojemnika) 28,2 Kb (75% pojemności pojemnika) 18,5 Kb (50% pojemności pojemnika), 8,9 Kb (25% pojemności pojemnika) wbudowano w kontener przy użyciu metody najmniej znaczącego bitu (NZB).
Obraz złożony za pośrednictwem funkcji dyskretnej, którą określa kolor wektora dla każdego piksela obrazu, gdzie wartość koloru określa wektor trzyczęściowy w przestrzeni kolorów. Najczęstszym sposobem przenoszenia kolorów jest model RGB.
Sposób NZB - najczęstszy wśród steganograficznych metod przestrzennych [1] ma niską odporność do ataków steganograficznych typu pasywnego i aktywnego. Jego główną wadą jest duża wrażliwość na najmniejsze zniekształcenia pojemnika.
Atak pasywny na stegano kontener jest metodę przechwytywania pojemnika, który jest przekazywany przez lokalną i globalna sieć, prowadząc steganoanalizę w celu ujawnienia faktu obecności ukrytych informacji.
Przeprowadzimy analizę korelacji pustego i napełnionego pojemnika, w celu określenia obecności ukrytych informacji w pojemniku. Wystarczy, aby utrzymać

współczynnik korelacji wypełnionego i pustego pojemnika, i określić fakt obecności ukrytych informacji w pojemniku. Przeprowadzimy ocenę korelacji pustego kontenera z kontenerem, do którego wkłada się informację różnej objętości (25 - 100%) w tabieli 1.

Tabela 1. Korelacja wypełnionego pojemnika

Metoda / objętość	25% (8,9 Кб)	50% (18,5 Кб)	75% (28,2 Кб)	100% (37,4 Кб)
NZB–8 bit	0.99991941	0.99991943	0.99991941	0.99991946
NZB–7 bit	0.99990035	0.99987989	0.99985878	0.99983898
NZB–6 bit	0.99982276	0.99971975	0.99961449	0.99951592
NZB–5 bit	0.99951787	0.99908387	0.99864401	0.99823402
NZB–4 bit	0.99835477	0.99662219	0.99484726	0.99314205
NZB–3 bit	0.99383535	0.98710237	0.98032035	0.97441410
NZB–2 bit	0.97583826	0.94994154	0.92475879	0.90364336
NZB–1 bit	0.90594949	0.82671980	0.75838305	0.70636937

Na podstawie analizy uzyskanych wyników można zauważyć, że przy wprowadzaniu nawet niewielkiej ilości informacji do pojemnika, jego korelacja jest znacznie zmniejszona, co niewątpliwie wskazuje na zmianę jego początkowych cech statycznych.

Na podstawie wyników eksperymentu możemy stwierdzić, że ustalając pewien próg dla każdej metody, łatwo jest określić fakt obecności ukrytych informacji w pojemniku.Ale ten schemat działa tylko wtedy, gdy mamy w oryginale oryginalny obraz analizowanego obrazu, co jest niemożliwe biorąc pod uwagę dużą liczbę obrazów w całym internecie.Dlatego też nie można zastosować ogólnego algorytmu określania obecności ukrytej wiadomości. W rezultacie przechodzimy do statycznej i spektralnej steganoanalizy. [2]

Metody statyczne są oparte na koncepcji "naturalnego" pojemnika. Istotą tych metod jest oszacowanie prawdopodobieństwa istnienia osadzania steganograficznego z nieznanym układem steganosystemu na podstawie kryterium oceny bliskości kontenera objętego dochodzeniem do "naturalnego".

W celu analizy obrazu formatu BMP dla obecności w nich ukrytych danych, obliczymy oczekiwanie matematyczne (MO), dyspersje, średnie odchylenie (SO) po rządkach otrzymanego masywu dla dalszego analizy

W rezultacie uzyskują się funkcje, które reprezentują statyczną charakterystykę całego masywu kontenera.. Korelujemy uzyskane funkcje w celu określenia charakterystyk najbardziej wpływających na proces wprowadzania informacji do pojemnika (Tabela 2.).

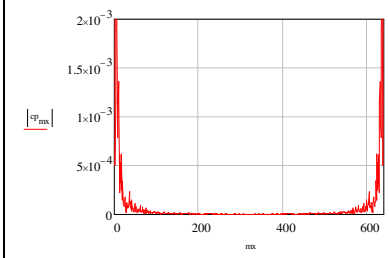
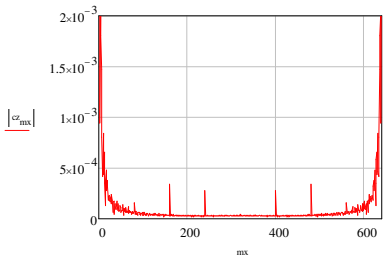
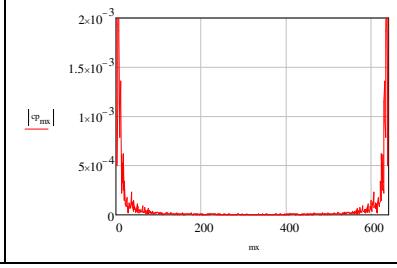
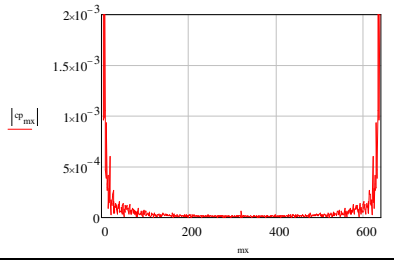
Na podstawie wyników uzyskanych w tabeli 3 stwierdzamy, że po informacji w pojemniku, najbardziej uszkodzona jest oczekiwanie matematyczne, a zatem w przyszłości przeprowadzimy analizę widma dokładnie tej charakterystyki statycznej.

Korzystając z szybkiej transformaty Fouriera, skonstruujemy spektrum otrzymanej funkcji wartości oczekiwanej zgodnie z warunkami tabeli 3.

Tabela 2. Korelacja funkcji

Współczynnik korelacji funkcji MO po rządkach pustych i wypełnionych pojemników.	Współczynnik korelacji funkcji dyspersji po rządkach pustych i wypełnionych. pojemników.	Współczynnik korelacji funkcji SO po rządkach pustych i wypełnionych. pojemników
$\mathrm{corr}(m_{ii}, mz_{ii}) =$	$\mathrm{corr}(m_{ii}, mz_{ii}) =$	$\mathrm{corr}(m_{ii}, mz_{ii}) =$
0.99979393	0.99997726	0.99997788
0.99991527	0.99999641	0.99998687
0.99969952	0.99996948	0.99996949
0.99990726	0.99998598	0.99998971
0.9998306	0.99998989	0.99999045
0.99978187	0.99999078	0.99999336
0.99468547	0.99997235	0.99995342
0.99955782	0.99997066	0.9999575

Tabela 3. Widmo pustego i wypełnionego pojemnika

№ bit	Widmo pustego pojemnika.	Widmo wypełnionego pojemnika
1		
8		

Na podstawie wyników z tabeli 3. widać, że na widmie napełnionego pojemnika pojawiają się wybuchy w pewnych częstotliwościach.

Zauważmy, że ukryte informacje były takie same dla wszystkich, więc rozpryski są na tej samej częstotliwości.

Rozpryski pojawiają się tylko wtedy, gdy informacje są wprowadzane do starszych fragmentów steganokontenera. Gdy wprowadzimy wiadomość informacyjną do niższych bitów, rozpryski są trudniejsze do określenia dla oka.

Po otrzymaniu wyników, w sprawie dostępności ukrytych informacji, możemy przystąpić do aktywnych ataków skierowanych przeciwko systemom ukrytego przekazu informacji..

4. Atak na system transmisji ukrytych informacji

Celem aktywnego ataku jest usunięcie ukrytych informacji lub naruszenie synchronizacji między procesem wprowadzania ukrytych informacji i procesu ich wydobywania. W wyniku takiego ataku ukryte informacje pozostają niezmienione, ale sam właściciel danych traci zdolność wyodrębniania go z kontenera. Najbardziej niebezpiecznymi aktywnymi atakami są ataki geometryczne. Są to ataki w postaci: przesunięcia każdego piksela obrazu wzdłuż osi współrzędnych przestrzennych przez niezdefiniowaną liczbę pikseli; obrót każdego piksela obrazu do nieokreślonego kąta; zmiany skali pikseli obrazu na nieokreśloną wartość [3]
Ataki geometryczne można modelować za pomocą przekształcenia afiniczne przekształcenia afinicznego (z ang. affinity – pokrewieństwo) z nieznanym dla detektora parametrem.
W sumie istnieje sześć przekształceń afinicznych: skalowanie, zmiana proporcji (rozciąganie / kompresja), obroty, przesuwanie i obcięcia[4]. Ataki te skutkują utratą synchronizacji pomiędzy wprowadzeniem procedury ukrytych informacji w obrazie i jego procedury wydobycia.
Również w pracy zbadano wpływ ataków na wbudowaną wiadomość przy użyciu prostej manipulacji obrazami, formatowania i kompresji obrazów z wbudowanym cyfrowym znakiem wodnym (CWZ)[5]. W oparciu o przekształceniach afinicznych można modelować geometryczne najgroźniejszych ataków na cyfrowy znak wodny naruszać efektywności detektora CWZ
Wyniki ataków geometrycznych prawidłowego funkcjonowania CWZ detektora przedstawiono w tabeli 4..

Tabel 4. Wyniki ataków afinicznych na poprawność działania detektora CWZ

Typ atak	NZB							
	1 bit	2 bit	3 bit	4 bit	5 bit	6 bit	7 bit	8 bit
Przekazywanie kontenerów	–	–	–	–	–	–	–	–
Powiększ	+	+	+	+	+	+	+	+
Rozciąganie / kompresja	+	+	+	+	+	+	+	+
Kąt pochylenia	–	–	–	–	–	–	–	–
Obróć się	–	–	–	–	–	–	–	–
Równoległy transfer	+	+	+	+	+	+	+	+
Odbicie	–	–	–	–	–	–	–	–
Obcięcie	+	+	+	+	+	+	+	+
Transformacja falkowa	+	+	+	+	+	+	+	+
Reformat bmp-png	–	–	–	–	–	–	–	–

Reformat bmp-tiff	–	–	–	–	–	–	–	–
Reformat bmp-jpeg (rgb)	+	+	+	+	+	+	+	+
Formatowanie bmp-jpeg 2000	–	–	–	–	–	–	–	–
Reformat bmp-gif	+	+	+	+	+	+	+	+

Uwaga: "+" to udany atak, "-" to nieudany atak.

Po analizie obrazów, można stwierdzić, że ataki przeciwko stegodetektora metodą NZB od 1 do 8 bitów, niezależnie od wielkości ukrytej informacji w oparciu o równoległe przeniesieni, obcięcia i transformacji falkowej całkowicie doprowadziły do awarii detektora. Należy również zauważyć, że przy użyciu ataky afinicznego, takiego jak skalowanie, rozciąganie / ściskanie, prowadzi do częściowego uszkodzenia dekodera. Podczas konwersji obrazów z jednego formatu na inny, pozytywny wynik uzyskuje się przez konwersję na format gif i jpeg.

5. Wnioski

W wyniku badań zwrócono szczególną uwagę na ataki steganograficzne i metody steganoanalizy, możliwość ich wykorzystania w celu rozwiązania problemu bezpieczeństwa informacji w systemach komputerowych.

Przeprowadzono krzyżową korelację pojemników, analizę statyczną i spektralną. Wyniki korelacji pojemników wykazały, że informacje, które wprowadzono do pustego pojemnika metodą steganografii, zmieniły charakterystykę statyczną. Ale aby wykryć fakt obecności informacji w pojemniku jednej korelacji nie wystarczy. ponieważ w celu analizy potrzebujemy oryginalnego obrazu, co jest niemożliwe w rzeczywistych warunkach. Analiza statycznych cech napełnionego pojemnika wykazała, że największa zmiana wartości oczekiwanej linii pojemnika ma miejsce podczas wprowadzania informacji do pojemnika Na podstawie uzyskanych wyników można wizualnie ujawnić fakt obecności informacji w statycznym obrazie Przy użyciu szybkiej transformaty Fouriera skonstruowano widmo otrzymanej funkcji matematycznego oczekiwania Zbadane zostało działanie ataków na wbudowaną wiadomość za pomocą prostych manipulacji obrazem, z wbudowanymi informacjami.

LITERATURA

1. KONAKHOVICH G. F., PUZIRENKO A. YU: Steganografia komputerowa. Teoria i praktyka. MK-Press, 2006 -288 str.
2. GRIBUNIN VG, OKOW IN, TURINTSEV IV Digital Steganografia M .: Solon-Press, 2009. - 265 str.
3. VATOLIN D., RATUSHNIAK A., SMIRNOV M., YUKIN V. Metody kompresji danych. Archiwizuj urządzenie, kompresję obrazu i wideo. - Moskwa: DIALOG-MEPHY, 2002. - 384 str.
4. YAGLYM IM, ASHKINUZ VG Pomysły i metody geometrii afinicznej i projekcyjnej. Część 1. Geometria afiniczna.1996.
5. WESELSKA O., SZMATOK O. Użycie algorytmu transformacji falkowej w stegoanalizie. VI Międzyuczelnianej Konferencji Studentów oraz Doktorantów, 2 grudnia 2016 r.: - Bielska-Biała: ATH,2016 r. – S. 415 – 420

Maryna YESINA[1], Olga AKOLZINA[2]

Supervisor: Olena KACHKO[3]

PROPOSALS OF THE EXPERT ESTIMATIONS TECHNIQUE USAGE FOR THE COMPARING AND ESTIMATION NTRU-LIKE CRYPTOGRAPHIC SYSTEMS

Summary: The possibility of encryption type cryptographic transformation properties comparative analysis methods usage is considered in this paper. The methods of comparative analysis – analytic hierarchy process and variations of weight indices method are studied, analyzed and applied. Conclusions and recommendations on the use of the cryptographic primitives estimation methods as well as recommendations for the possibility of applying encryption algorithms are made.

Keywords: cryptographic primitives analysis, cryptographic primitives, comparison analysis methods, NTRU, NTRUPrime.

TECHNIKI ESTYMACYJNE DLA PORÓWNANIA ORAZ ESTYMACJI SYSTEMÓW KRYPTOGRAFICZNYCH

Streszczenie: W artykule omówiono możliwości zastosowania zaawansowanych narzędzi kryptograficznych: transformację własności kryptograficznych wybranymi metodami szyfrowania. Omówiono także metodę analizy porównawczej – tj. metodę hierarchicznego procesu analitycznego oraz zmienności wskaźników wagowych. Zaproponowaną metodę przeanalizowano oraz zastosowano. Sformułowano wnioski oraz zalecenia odnośnie użycia prymitywów kryptograficznych w metodach estymacji, a także rekomendacje co do możliwości zastosowania algorytmów szyfrowania.

Słowa kluczowe: analizy prymitywów kryptograficznych, prymitywy kryptograficzne, metody analizy porównawczej, asymetryczny kryptosystem z kluczem publicznym i prywatnym

[1] V. N. Karazin Kharkiv National University, department of information systems and technologies security, PhD in Technical Sciences, rinayes20@gmail.com

[2] V. N. Karazin Kharkiv National University, department of information systems and technologies security, student, aos3@i.ua

[3] JSC «Institute of Information Technologies», Kharkiv, head of the programming department, Kharkiv National University of Radio Electronics, PhD in Technical Sciences, professor, iit@iit.kharkov.ua

1. Introduction

In 2016 there were the series of important events, that have significantly affected to the intensive development of post-quantum cryptography. To them should be referred the statement on the Internet – Alfred J. Menezes and Neal Koblitz articles [8], organization and conduction by NSA and NIST USA VII international conference on post quantum cryptography, which took place in February 2016 in Japan [13, 15]. An extremely important event was the publication in the USA report «Report on Post – Quatum Cryptography. NISTIR 8105 (DRAFT)» [10], in which fully confirmed the possibility of asymmetric cryptographic primitives successful quantum cryptanalysis and the main problems and opportunities, and stages of their decision are identified.

NIST USA announced a competition to develop the standards of post-quantum asymmetric cryptographic primitives [13], understanding the need to find new electronic signature and asymmetric encryption type cryptographic transformation, which will be relevant and can be applied in post-quantum period. The specified one due to two factors. First, there is significant progress in the development of quantum computers, including experimental demonstration of physical qubits realization are carried out, which can be scaled up to larger systems.

Second, likely transition to post-quantum cryptography will not be easy, because it is unlikely to be a simple replacement of the current asymmetric cryptographic primitives standards. Significant efforts will be needed to develop, standardize and implement a new post-quantum cryptosystems. Therefore, should be a significant transition stage, when as current and post-quantum cryptographic primitives are used.

The European Union has also started the preparation of a new post-quantum standards. A new direction "Quantum-Safe Cryptography" are formed by European Organization for Standardization ETSI in the cluster "Security" [4, 11, 16]. According to the results of these studies are predicted the groups standards for post-quantum period adoption. ETSI has published a group report "Quantum-Safe Cryptography. Quantum-Secure infrastructure" [4], in which fixed bases of perspective infrastructure, provided mechanisms, described primitives types, that will be used. Separately requirements are nominated and estimation criteria are formed for future candidates.

ES and encryption type cryptographic transformations, and their application algorithms are allocated among the set of asymmetric cryptographic primitives. The specified one is explained by their wide application in a significant number of applications and potential large losses in case of discrediting ES and encryption type cryptographic transformations, that are used at present [5, 7, 9, 2, 3, 25].

Our experience obtained during the conducting research on projects AES and NESSIE [1, 12], and in national standards for hash function DSTU 7564:2014 developing and adopting, and block symmetric encryption algorithm DSTU 7624:2014 [7, 9] etc., allows to conclude, that the extremely important problem is substantiation of the estimation criteria system choice and comparison of each cryptographic primitives with other, and development and application the scientifically based techniques of them analysis and comparison in accordance with the nominated requirements. These methods and developed on their basis technique or techniques should take into account all requirements, that are nominated for asymmetric cryptographic primitives and allow to help make the decision about winners based on use the unconditional and conditional criteria system, as partial and integral.

The objective of these proposals are the substantiation, development and experimental confirmation of methodical bases application possibilities of system unconditional and conditional criteria selection and application, and methods and technique of comparative analysis and making the decision on asymmetric post-quantum encryption type cryptographic primitives.

2. Problem formulation

After detailed analysis, it was determined that the first time techniques of estimation and comparative analysis of cryptographic primitives type block symmetric cipher (BSC), streaming symmetric cipher (SSC), electronic signature (ES) and cryptographic protocol were proposed in [22, 23], and detailed in [24]. They are based on the use of unconditional and conditional partial and integral criteria system, and indicators, that allow to assess the degree of nominated to the candidate requirements fulfillment. In our opinion the main task of these techniques are the formalization of decision-making processes regarding fulfillment of nominated to them requirements, taking into account the strengths and weaknesses of cryptographic primitives, that are candidates for the post-quantum standard, reduce the influence of subjective factors in decision-making, including unauthorized influence of outside organizations, etc. For example, following techniques can be applied to estimate and compare the ES and encryption type cryptographic primitives, which are the candidates for the post-quantum standard in our case.

At the formal level such estimation and comparison techniques ES and encryption type cryptographic primitives can be summarized. But, since to these cryptoprimitives are nominated different requirements, then for each of the primitives they may be supplemented or simplified and display the entire spectrum of nominated requirements. Also, these techniques can ensure transparency of decision-making, experts independent, and help to substantiate making appropriate decisions and confidence in them.

Further in research technique we'll mean a fixed set of methods, methods of practice, tested and studied for the expedient implementation of specified work, that leads to a predetermined outcome [24].

In research in the broad sense we'll mean the search of new knowledge or a systematic investigation in order to establish the facts. In a narrow sense this is the scientific method (process) of study anything.

3. The achievements state of estimation methods and techniques and comparative analysis of cryptographic primitives development and application

From described above follows the necessity and actuality of solving the problem, a great extent automation and significantly reduce decision-making subjectivity relatively the benefits of the cryptographic primitives of certain set, for example, encryption type cryptographic primitives. The solution of tasks certain components of this problem is contained in [22, 24].

Later in the criterion will understand the sign on which basis is carried out the assessment, anything determination or classification [24], that is, in fact, we will understand the measure of estimation. The previous researches and [24] allow to substantiate the conclusion, that the estimation and comparison of cryptographic primitives should implement using two sets of criteria: unconditional and conditional (the estimation can be carried out in two stages) [21, 24].
On the first stage it is checked the conformity cryptographic primitives to unconditional criteria requirements – partial and integral, and in the second, using conditional criteria – partial conditional criteria and integral conditional criterion. It is possible to compare different encryption type cryptographic primitives by using partial conditional criteria and integral conditional criterion.

3.1. Cryptographic primitives estimation by unconditional criteria

Analysis of the application state, development and estimation experience of the different type cryptographic transformations properties on the prospect of the post-quantum period, the achieved results in the practical solution of cryptanalysis tasks and various attacks implementing, allow as basic to choose the following unconditional evaluation criteria (concerning the estimation of encryption type cryptographic transformations) [14-16] – Table 1.

Table 1. Unconditional estimation criteria of post-quantum encryption type cryptographic transformation

№	Unconditional criteria	Denotation
1	Reliability, simplicity and transparency of mathematical base (mathematical transformations) used in the implementation of post quantum encryption type cryptographic transformation.	$W_{\delta 1}$
2	Practical security of encryption type cryptographic transformation in the mechanism "semantically secure encryption" implementation against known attacks using a quantum computer and cryptanalyst access to the 2^{64} selected ciphertexts for security model IND–CCF2.	$W_{\delta 2}$
3	The validity of real security (stability) of encryption type cryptographic transformation against all known and potential cryptanalytic attacks of post quantum period based on the use of common parameters and keys with the necessary size and properties (128-bit keys and more classical stability (security)), including statistical security.	$W_{\delta 3}$
4	Theoretical security of encryption type cryptographic transformation in post quantum period against existing force, analytical and special attacks for existing threats models (at least for the model IND-CCF2 for encryption).	$W_{\delta 4}$
5	The possibility of replacing existing standardized cryptographic primitives to the post quantum ones and application in the existing cryptographic systems and protocols in certain conditions and restrictions.	$W_{\delta 5}$

6	Computational efficiency – complexity of direct I_{dir} and reverse I_{rev} encryption type cryptographic transformation, and generating asymmetric key pairs I_{key} is not above polynomial, providing the necessary complexity (performance) values I_{dir}, I_{rev}, I_{key} in practical use in applications with their hardware and software, and program implementation.	$W_{\delta 6}$
7	The performance of limitations for minimum and maximum lengths of private and public key, sizes and unprofitability of ciphertext, the absence of weak private keys for post quantum period security models.	$W_{\delta 7}$

Since the presented partial criteria are unconditional, then the selection criterion is a logical variable yes/no (1/0), so unconditional criterion can be written as [21]:

$$(W_{\delta 1}, W_{\delta 2}, W_{\delta 3}, W_{\delta 4}, W_{\delta 5}, W_{\delta 6}, W_{\delta 7}) \in (1, 0) \tag{1}$$

Given the described above partial unconditional criteria $W_{\delta 1}$–$W_{\delta 7}$ and condition (1), cryptographic transformation accordance function can be presented as:

$$f_{af} = \mathrm{W}_{\delta 1} \wedge \mathrm{W}_{\delta 2} \wedge \mathrm{W}_{\delta 3} \wedge \mathrm{W}_{\delta 4} \wedge \mathrm{W}_{\delta 5} \wedge \mathrm{W}_{\delta 6} \wedge \mathrm{W}_{\delta 7} = \mathrm{W}_{\delta}. \tag{2}$$

Therefore, the quality of cryptographic transformation can be estimated using unconditional integral criterion – cryptographic transformation accordance function to requirements $f_{af} \in (0,1)$ and on $f_{af} = 1$ cryptographic transformation, that estimated, complies with the requirements.

Introduced thereby integral criterion allows to establish, whether the considered ES type cryptographic transformation complies with considered discussed requirements. If it is yes, it can be reasonably recommended for use.

Provided a positive estimation of cryptographic transformation by integral unconditional criterion, further comparison and estimation can be made based on the conditional criteria and integral conditional criterion [24].

3.2. Cryptographic primitives estimation by conditional criteria

Researches have shown that qualitative and quantitative comparison of any type cryptographic transformations can be carried out using generalized conditional preference criterion [21, 24]or integral conditional criterion.

As the main partial conditional criteria for encryption type cryptographic transformations estimation it is suggested to use the following (Table 2).

Table 2. Conditional estimation criteria of post-quantum encryption type cryptographic transformation

№	Conditional criteria	Denotation
1	Additional security features: perfect forward secrecy; resistance to side-channel attack; resistance to multi-key attacks; resistance to failures.	W_{y1}
2	Stability requirements 1) classic security 128-bit / 64-bit quantum protection (stability reserve AES-128); 2) classic security 128-bit / 80-bit quantum protection (stability reserve SHA-256/SHA3-256); 3) classic security 192-bit / 96-bit quantum protection (stability reserve AES-192); 4) classic security 192-bit / 128-bit quantum protection (stability reserve SHA-384/SHA3-384); 5) classic security 256-bit / 128-bit quantum protection (stability reserve SHA2-512, SHA3-512).	W_{y2}
3	Encryption errors. The low percentage of encryption errors.	W_{y3}
4	The possibility of multiple encryption.	W_{y4}
5	Flexibility and simplicity: 1) additional scheme options (optimization, asynchronous or implicitly authenticated key exchange, etc.); 2) cross-platform; 3) the possibility of parallelization; 4) structure intelligibility.	W_{y5}
6	Correctness verification. Checking the correctness of basic and optimized implementations.	W_{y6}
7	Effectiveness verification: calculation of time needed for key generation, encryption and decryption (testing is carried out on optimized versions).	W_{y7}
8	Test conditions The main platforms: 1) NIST PQC Reference Platform; 2) Intel x64; 3) Windows or Linux, the GCC compiler; 4) Additional testing of other conditions (8-bit processors, digital signal processors, dedicated CMOS etc.)	W_{y8}
9	Possibility and conditions of free distribution post quantum encryption type cryptographic transformation.	W_{y9}
10	Confidence level to the post quantum encryption type cryptographic transformation at different levels of use.	W_{y10}
11	Perspective and justification the use of post quantum encryption type cryptographic transformation.	W_{y11}
12	Easiness in use: the possibility of imbedding encryption type cryptographic transformations to the majority applied protocols, such as TLS or IKE etc. and persistence to wrong use.	W_{y12}

It is important to choose the method of clotting the partial conditional criteria to integral conditional criterion in their application. The conducted analysis and practical researches have shown [8, 10, 14, 17-21] that as a method of clotting the partial conditional criteria can choose the analytic hierarchy process based on pairwise comparisons and method of determining weight indices.
When using the analytic hierarchy process based on pairwise comparisons, obtained statements expressed in integers taking into account nine-point scale, defined in [21, 22].

4. The analytic hierarchy process based on pairwise comparisons and features of its use for the cryptographic primitives estimation

Analytic hierarchy process (AHP) – the systematic approach to the complex problems of making decision mathematical tool. AHP does not prescribe to the decision making person (DMP) any "right" decision, and allows him to interactively find this option (alternative), which the best agrees with its understanding of the problem essence and requirements to its solution [18-20].
For use the analytic hierarchy process must choose a conditional criteria system. With such indicators set, using the conditional criteria can calculate the integral conditional criterion value, and, consequently, make the comparison by integral conditional criterion.
The elements pairwise comparison method [21, 24] can be described as follows. The set of paired comparisons matrices is constructed. Paired comparisons are carried out in terms of the dominance of one element over another. Obtained statements are expressed in integers, considering the nine-point scale [21, 22].
In pairwise comparison the expert compares investigated objects of their importance pairwise, establishing the most important object in each pair. All possible objects pairs expert represents in a record of each combination (object 1 – object 2, object 2 – object 3, etc.) or in the matrix form [6, 21].
More detail, this estimation method and the procedure for its implementation is described in [21, 26].

5. Method and procedure of estimation and comparative analysis cryptographic primitives based on weight indices

In the case, when get information about parameters comparable systems importance using informal methods is not possible, necessary to use formalized methods. These include methods based on the determination of weighting coefficients [6, 14, 17-19, 21].
Let us consider the general problem formulation for cryptoprimitives estimation technique based on method of determining weight indices. Let there are [2]: k systems, which is necessary to estimate; m indicators, according to which systems are estimated; n experts, that carry out the evaluation.

We define some partial indicators, at which cryptographic primitives can be evaluated (concerning the estimation of encryption type cryptographic transformations) [13, 15]:

x_1 – additional security features: perfect forward secrecy; resistance to side-channel attack; resistance to multi-key attacks; resistance to failures;

x_2 – stability requirements: classic security 128-bit / 64-bit quantum protection (stability reserve AES-128); classic security 128-bit / 80-bit quantum protection (stability reserve SHA-256/SHA3-256); classic security 192-bit / 96-bit quantum protection (stability reserve AES-192); classic security 192-bit / 128-bit quantum protection (stability reserve SHA-384/SHA3-384); classic security 256-bit / 128-bit quantum protection (stability reserve SHA2-512, SHA3-512);

x_3 – encryption errors, the low percentage of encryption errors;

x_4 – the possibility of multiple encryption;

x_5 – flexibility and simplicity: additional scheme options (optimization, asynchronous or implicitly authenticated key exchange, etc.); cross-platform; the possibility of parallelization; structure intelligibility;

x_6 – correctness verification, checking the correctness of basic and optimized implementations;

x_7 – effectiveness verification: calculation of time needed for key generation, encryption and decryption (testing is carried out on optimized versions);

x_8 – test conditions, the main platforms: NIST PQC Reference Platform; Intel x64; Windows or Linux, the GCC compiler; Additional testing of other conditions (8-bit processors, digital signal processors, dedicated CMOS etc.);

x_9 – possibility and conditions of free distribution post quantum encryption type cryptographic transformation;

x_{10} – confidence level to the post quantum encryption type cryptographic transformation at different levels of use;

x_{11} – perspective and justification the use of post quantum encryption type cryptographic transformation;

x_{12} – easiness in use: the possibility of imbedding encryption type cryptographic transformations to the majority applied protocols, such as TLS or IKE etc. and persistence to wrong use.

For the evaluation, we will use the following methods of determining weight indices [14, 20, 21]: using the Fishburn scale; based on the ranking method; based on the points attribution method; based on the numerical method.

Detailed estimation process using these methods is presented in [6, 21].

6. Description and analysis of modern encryption algorithms general parameters

6.1 Analysis of NTRU encryption algorithm

NTRU – the first public key cryptosystem not based on factorization or discrete logarithmic problems. NTRU is based on the shortest vector problem in a lattice.

Operations are based on objects in a truncated polynomial ring $R=\mathbf{Z}[x]/(x^n-1)$, polynomial degree at most $n-1$.
NTRU parameters are as follows: n – the polynomials in the ring R have degree $n-1$ (non-secret); q – the large modulus to which each coefficient is reduced (non-secret); p – the small modulus to which each coefficient is reduced (non-secret); f – a polynomial that is the private key; g – a polynomial that is used to generate the public key h from f (secret but discarded after initial use); h – the public key, also a polynomial; r – the random "blinding" polynomial (secret but discarded after initial use); d – coefficient.
The encryption of message m is carried out according to the formula $c = rh + m$.
Decryption is performed as follows: using a private polynomial f it is calculated polynomial $a = f \cdot e (\mathrm{mod}\, q)$. Then the polynomial $b = a \bmod p$ is calculated. Another personal polynomial f_p is used to compute $c = f_p \cdot b (\mathrm{mod}\, p)$, where c is an outgoing message m. More details about the NTRU algorithm is described in [2].

6.2 Analysis of NTRUPrime encryption algorithm

The NTRUPrime cryptosystem is proposed as one of the alternative variants of the asymmetric NTRU method in order to get rid of the weaknesses inherent in NTRU, which are associated with undesirable structural properties of the ring $\mathbf{Z}_q[x]/(x^n-1)$: in many cases, a ring of this type has a subrings and a factor-rings of a high order. Unlike NTRU, NTRUPrime uses a ring $\mathbf{Z}_q[x]/(x^n-x-1)$, which provided that the proper selection of numbers q and n, is a field that does not contain its own subfields. In addition, the Galois group of polynomial $x^n - x - 1$ over the field Q is a symmetric group S_n, which excludes the possibility of attacking a certain type on the cryptosystem.
In NTRUPrime, the public key is calculated by the formula $h = g / 3f$ that it matters to create an effective secret key transfer protocol. However, to construct an asymmetric encryption system, it is desirable to use the traditional formula $h = 3g / f$
The decryption of messages in the cryptosystem NTRUPrime occurs correctly on condition $q > 48t$. Details about the NTRUPrime algorithm is described in [3].

6.3 Analysis of NTRUPrime(IIT) encryption algorithm

The given asymmetric encryption scheme is a modification of the NTRU scheme, and differs from the latter only in two aspects:
1. Instead of the ring $\mathbf{Z}_q[x]/(x^n-1)$ used in NTRU, a field $\mathbf{Z}_q[x]/(x^n-x-1)$ is used, as in the NTRUPrime cryptosystem [3]. According to [3], this prevents cryptosystem attacks of some kind and precludes the use of (at least potentially) weaknesses of the standard NTRU cryptosystem that are associated with the existence of non-trivial subrings or truncated rings of ring $\mathbf{Z}_q[x]/(x^n-1)$.

2. In the proposed scheme, polynomials F and r are arbitrary t-*small*, that is, they have $2t$ non-zero coefficients equal to ± 1, whereas in [2] each of these polynomials has exactly t nonzero coefficients equal to 1 and -1 correspondingly. A similar remark is also valid for a polynomial g, which is an arbitrary small polynomial in a modified cryptosystem and has the same number of non-zero coefficients, which are equal 1 and -1 in NTRU. This difference is not significant, however, it provides the opportunity to expand the amount of key space in comparison with NTRU without losing the effectiveness of algorithms implementation for key generation and encryption-decryption of messages.

In this algorithm, the secret key is any pair of polynomials (f, g), where $f = (1 + 3F) \bmod q$, $F, g \in R / 3$, $\| F \|_1 = 2t$, and the corresponding public key is a polynomial $h = 3g / f \in R / q$.

Encryption of the message m is carried out according to the formula $c = m + rh$, where r – the random equal probability t-small polynomial, h – public key, and the addition and multiplication are carried out in the field R / q.

To retrieve a message m by message c using a secret key (f, g), you must calculate $m' = (cf \,(\bmod q)) \bmod 3$ and put $m' = (m' f^*) \bmod 3$. That is, only polynomials f and f^* are used to decrypt messages, and where f^* there is an inverse to an element $f \bmod 3$ in the ring $R / 3$.

In the NTRUPrime(IIT) using the appropriate estimates as specified in the description of the algorithm, can (it is allowed) to significantly weaken the condition for decryption of messages in comparison with NTRUPrime, namely, to replace it with a condition $q > 32t$. This, in turn, allows you to reduce the value q compared to NTRUPrime, while maintaining the decryption correct.

More details about the NTRUPrime(IIT) algorithm is described in [25].

7. Encryption type cryptographic transformations by unconditional criteria estimation

According to the results of Table 3, we can see that all encryption type cryptographic transformations are positively estimated by an integral unconditional criterion.

Table 3. Results of comparative analysis regarding unconditional criteria

Estimation criterion / Cryptoprimitive	$W_{\delta 1}$	$W_{\delta 2}$	$W_{\delta 3}$	$W_{\delta 4}$	$W_{\delta 5}$	$W_{\delta 6}$	$W_{\delta 7}$	W_{δ}
NTRU	1	1	1	1	1	1	1	1
NTRUPrime	1	1	1	1	1	1	1	1
NTRUPrime(IIT)	1	1	1	1	1	1	1	1

All defined cryptographic primitives are received a positive estimation by an integral unconditional criterion, so further comparison and estimation can be carried out based on conditional criteria and integral conditional criterion.

8. Encryption type cryptographic transformations by conditional criteria estimation

Comparing encryption type algorithms relatively conditional criteria, construct for this goal tree (fig. 1).

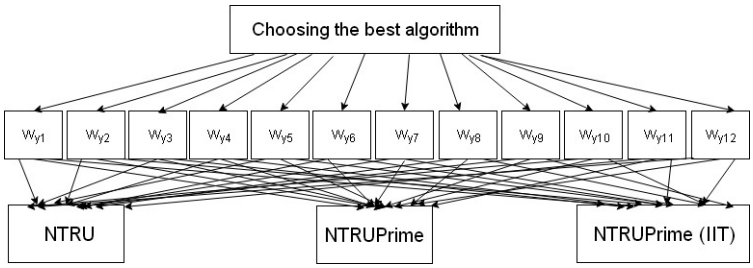


Figure 1. Goal tree

After this do the estimation of each criterion. For this construct the pairwise comparisons matrix relative to the compared cryptographic primitives for each criterion [21, 22].

To calculate the resulting priorities vector multiply the level 1 priority vector and the level 1 acquired values matrix (fig. 2).

$$v := (0.1623 \quad 0.2551 \quad 0.1371 \quad 0.0201 \quad 0.0563 \quad 0.1079 \quad 0.0978 \quad 0.0172 \quad 0.0358 \quad 0.0370 \quad 0.0185 \quad 0.0548)$$

$$M := \begin{pmatrix} 0.0594 & 0.2302 & 0.7686 \\ 0.0333 & 0.0333 & 0.0333 \\ 0.1046 & 0.2583 & 0.6370 \\ 0.4285 & 0.4285 & 0.1429 \\ 0.4285 & 0.4285 & 0.1429 \\ 0.0333 & 0.0333 & 0.0333 \\ 0.7009 & 0.2021 & 0.0971 \\ 0.4285 & 0.4285 & 0.1429 \\ 0.1046 & 0.6370 & 0.2583 \\ 0.1046 & 0.6370 & 0.2583 \\ 0.1046 & 0.6370 & 0.2583 \\ 0.0856 & 0.2969 & 0.6174 \end{pmatrix}$$

$$vv1 := v \cdot M = (0.159 \quad 0.219 \quad 0.304)$$

Figure 2. The priorities resulting vector calculation

Let us consider the obtained numerical results. The investigated encryption type cryptographic primitives can arrange the places, that they occupied on the results of comparison (1 – the best, 3 – the worst): NTRUPrime(IIT) – 0,304; NTRUPrime – 0,219; NTRU – 0.159.

Further we consider the practical application of the method of determining weight indices variations [14, 17-20]. According to each of specified method of determining weight indices variants [14, 17-21], we conduct the encryption type cryptographic primitives estimation [2, 3, 25]. Similarly to the results of the analytic hierarchy process based on pairwise comparisons, we obtain results for methods of determining weight indices. We obtain the following results after the estimation (Table 4).

Table 4. Results of cryptographic primitives estimation

Methods for the weight indices determining			
using the Fishburn scale	based on the ranking method	based on the points attribution method	based on the numerical method
NTRUPrime(IIT) – 0,304 NTRUPrime – 0,219 NTRU – 0,159	NTRUPrime – 0,096 NTRUPrime(IIT) – 0,096 NTRU – 0,074	NTRUPrime(IIT) – 0,095 NTRUPrime – 0,094 NTRU – 0,071	NTRUPrime – 0,09 NTRUPrime(IIT) – 0,09 NTRU – 0,077

9. The analysis of encryption algorithms researches results

According to the determined methods of expert estimation, the results shown in the previous sections were obtained.

One can assume, that the results of the encryption algorithms estimation, by different methods have been obtained almost identical – almost the same encryption algorithms arrangement from the best to the worst. Numeric scatter of weight indices values for one algorithm is almost negligible, only numeric values for all encryption algorithms in the analytic hierarchy process based on pairwise comparisons differ from weight indices values for these encryption algorithms according to other estimation methods. That is substantiated by more strong influence of the subjective experts opinion to the estimation result in defined method.

Figure 3 graphically shows the results of encryption algorithms estimation by different estimation methods.

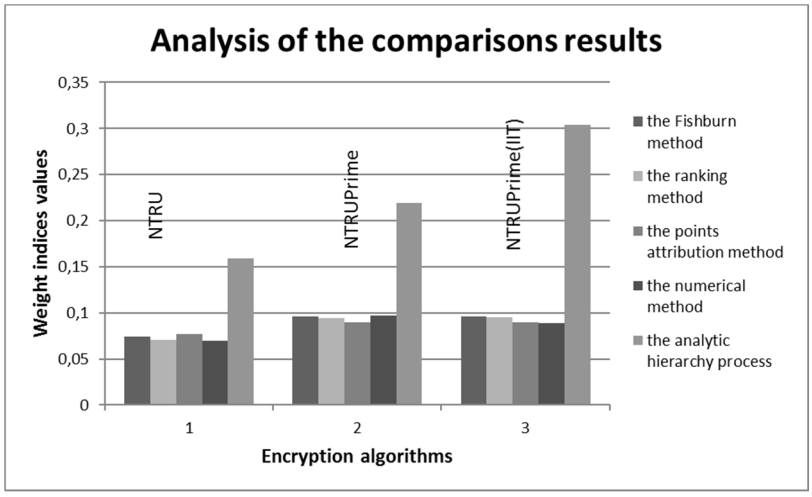


Figure 3. The results of the encryption algorithms estimation by different methods

After that, a summary table containing the average estimation results of the corresponding cryptographic primitives is constructed (Table 5).

Table 5. The averaged results of cryptographic primitives estimation

Cryptographic primitive	Averaged estimate
NTRU	0,0902
NTRUPrime	0,1192
NTRUPrime(IIT)	0,1348

And we give a graphic representation of the obtained averaged results in the form of a diagram (Fig. 4).

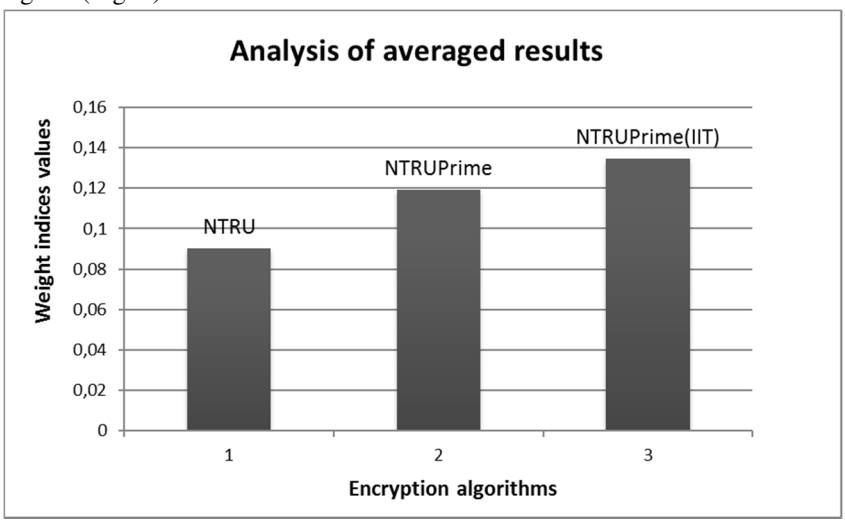


Figure 4. Averaged results of estimating encryption algorithms by different methods

10. Conclusions

1) In connection with the specific requirements for cryptographic transformations, including for encryption algorithms, the main criteria should be divided into two classes: conditional and unconditional. As the main criterion for integral evaluation can be and is recommended to use the integral unconditional criterion, that is derived by partial unconditional criteria.
2) The research results allow to conclude, that in terms of estimation objective the best use the variations of weight indices determining method, because the experts subjectivity has the a significant impact to the result in the analytic hierarchy process based on pairwise comparisons.
3) The comparative analysis results of encryption algorithms HIII NTRU, NTRUPrime and NTRUPrime(IIT) allowed to make the following conclusions and recommendations: the NTRU algorithm is on the last place for all estimation methods, NTRUPrime and NTRUPrime(IIT) algorithms divide among themselves the first places. Despite on the fact that the NTRU algorithm is faster than the other two, it is vulnerable to attacks of a certain kind, unlike NTRUPrime and NTRUPrime(IIT). This is due, first of all, to the mathematical apparatus used in these algorithms. Therefore, in practice it is recommended to use NTRUPrime and NTRUPrime(IIT) algorithms, since they are more secure and only slightly slower than NTRU.

4) To obtain more precise estimation results and for obtaining approximately same estimation results by different estimation methods, it is necessary to perform the estimation procedure several times and carefully approach to the choice of experts that will conduct the estimation.

REFERENCES

1. Mode of access Electron resource – AES: the Advanced Encryption Standard: *https://competitions.cr.yp.to/aes.html*, 10.01.2017.
2. American National Standard for Financial Services – Lattice-Based Polynomial Public Key Establishment Algorithm for the Financial Services Industry – ANSI X9.98–2010, 2010. – 284 p.
3. Mode of access Electron resource – Daniel J. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, Christine van Vredendaal: NTRU Prime: *https://ntruprime.cr.yp.to/ntruprime-20160511.pdf*, 14.09.2017.
4. Mode of access Electron resource – ETSI GR QSC 001 V.1.1.1 (2016-07). Quntum-Safe Cryptography (QSC); Quantum-safe algorithmic framework: *https://portal.etsi.org/webapp/workProgram/Report_WorkItem.asp?wki_id=46690*, 13.03.2017.
5. Information technology – Security techniques – Digital signatures with appendix – Part 3: Discrete logarithm based mechanisms: ISO/IEC 14888-3 (Edition 3) : 2016. – 130 p.
6. Mode of access Electron resource – Gorbenko I., Yesina M., Ponomar V.: Proposals of comparative analysis and decision making during the competition regarding the certain benefits of asymmetric post quantum cryptographic primitives. COMPUTER SCIENCE AND CYBERSECURITY. V. N. Karazin Kharkiv National University, 2017, Issue 1(5), 53–70: *http://periodicals.karazin.ua/cscs/article/view/8307.*
7. Mode of access Electron resource – Kalyna: *http://www.slideshare.net/oliynykov/kalyna*, 10.01.2017.
8. Mode of access Electron resource – Koblitz Neal, Menezes Alfred J.: A riddle wrapped in an enigma: *https://eprint.iacr.org/2015/1018.pdf*, 20.09.2016.
9. Mode of access Electron resource – Kupyna: *https://ru.wikipedia.org/wiki/Kupyna*, 10.01.2017.
10. Mode of access Electron resource – Lily Chen, Stephen Jordan, Yi-Kai-Liu, Dustin Moody, Rene Peralta, Ray Perlner, Daniel Smith-Tone: Report on Post-Quatum Cryptography. NISTIR 8105 (DRAFT): *http://csrc.nist.gov/publications/drafts/nistir-8105/nistir_8105_draft.pdf*, 13.03.2017.
11. Mode of access Electron resource – Mosca M.: “Setting the Scene for the ETSI Quantum-safe Cryptography Workshop”. E-proceedings of “1st Quantum-Safe-Crypto Workshop”, SophiaAntipolis, Sep 26-27, 2013:

http://docbox.etsi.org/Workshop/2013/201309_CRYPTO/eproceedings_Crypto_2013.pdf, 13.03.2017.

12. Mode of access Electron resource – NESSIE: New European Schemes for Signatures, Integrity, and Encryption: *https://competitions.cr.yp.to/nessie.html*, 20.09.2016.
13. Mode of access Electron resource – Post-quantum crypto project: *http://csrc.nist.gov/groups/ST/post-quantum-crypto/index.html*, 20.04.2017.
14. Mode of access Electron resource – Procedures for Determining the Weights of Selection Factors in the Weighted-Matrix Delivery Decision: *http://www.tcrponline.org/PDFDocuments/tcrp_rpt_131AppF.pdf*, 23.05.2016.
15. Mode of access Electron resource – Proposed Submission Requirements and Evaluation Criteria for the Post-Quantum Cryptography Standardization Process: *http://csrc.nist.gov/groups/ST/post-quantum-crypto/documents/call-for-proposals-draft-aug-2016.pdf*, 20.04.2017.
16. Mode of access Electron resource – Quantum Safe Cryptography and Security. An introduction, benefits, enablers and challenges. ETSI White Paper No. 8, 2015: *http://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf*, 13.03.2017.
17. Mode of access Electron resource – Roszkowska Ewa: Rank ordering criteria weighting methods – a comparative overview. Optimum. Studia ekonomiczne, **5**(2013)65, 14–33: *http://repozytorium.uwb.edu.pl/jspui/bitstream/11320/2189/1/02_Ewa%20ROSZKOWSKA.pdf*, 23.05.2016.
18. Mode of access Electron resource – Saaty Thomas L.: Decision making with the analytic hierarchy process. Int. J. Services Sciences, **1**(2008)1, 83–98: *http://www.colorado.edu/geography/leyk/geog_5113/readings/saaty_2008.pdf*, 23.05.2016.
19. SAATY T. L.: The Analytic Hierarchy Process. New York: McGraw Hill, 1980.
20. Mode of access Electron resource – The methods of expert estimations: *http://booksforstudy.com/19650323/ekonomika/metodi_ekspertnih_otsinok.htm*, 23.05.2016.
21. YESINA M., GORBENKO Y. (supervisor): Methods of cryptographic primitives comparative analysis, Inżynier XXI wieku ("Engineer of XXI Century" – the VI Inter University Conference of Students, PhD Students and Young Scientists: University of Bielsko-Biala, Poland, December 02, 2016). – Bielsko-Biała: Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej w Bielsku-Białej, 2016, 451–462. – ISBN 978-83-65182-51-7. – Chapter in monograph.
22. ANDREYCHIKOV A. V., ANDREICHIKOVA O. N.: Analysis, synthesis, solutions planning in the economy. Finance and statistics, Moscow 2002. (in Russian).

23. GORBENKO I. D., KUZNETSOV O. O., POTII O. V., GORBENKO YU. I., HANZIA R S., PONOMAR V. A.: Post-quantum cryptography and mechanisms for its implementation. Radiotechnics, 186 (2016), 32–52. (in Ukrainian).
24. GORBENKO YU. I.: Methods of constructing and analyzing cryptographic systems : monograph, Kharkov 2015. (in Ukrainian).
25. KACHKO O. G., MAKUTONINA L. V, AKOLZINA O. S.: The optimization of NTRU-like algorithm for asymmetric encryption with “inconvenient parameters”. Mathematical and computer modeling. Series: Engineering, 15 (2017), 79–85. (in Ukrainian).

Genadiy ZHYROV[1], Olena RUDNITSKA[2]

Scientific Supervisor: Yurii KHLAPONIN[3],

THE MODEL OF FUZZY NEURAL PRODUCTION NETWORK IN THE INFORMATION SECURITY SYSTEMS

Summary: The article dwells upon the issue of information security assessment. There has been elaborated the method of measurement allowing to obtain security information from each of the dataways based on the identified threat model. There has been presented a model of fuzzy neural production network for the system of information security assessment.

Keywords: neural network, information security system, threat model, information security level.

MODEL ROZMYTY PRODUKCYJNEJ SIECI NEURONOWEJ W SYSTEMIE BEZPIECZEŃSTWA INFORMACJI

Streszczenie: Artykuł dotyczy kwestii oceny bezpieczeństwa informacji. Opracowano metodę pomiarową pozwalającą na uzyskanie informacji o bezpieczeństwie z każdego sposobu wydobycia/pozyskania informacji na podstawie zidentyfikowanego modelu zagrożenia. Przedstawiono model rozmytej produkcyjnej sieci neuronowej dla systemu oceny bezpieczeństwa informacji.

Słowa kluczowe: sieć neuronowa, system bezpieczeństwa informacji, model zagrożenia, poziom bezpieczeństwa informacji.

1. Introduction

The information security assessment is currently determined by the system of quantitative and qualitative measurements that gives ground for solving the

[1] Candidate of Technical Sciences, Senior Research Fellow, leading researcher research center Military Institute of Taras Shevchenko National University of Kyiv, genna-g@ukr.net
[2]M.Sc., Senior Lecturer, PhD student, Kyiv National University of Construction and Architecture, Department of Cyber Security and Computer Engineering, olena.rudnitska@gmail.com
[3] Doctor Technical Sciences, Senior Research Fellow, Head of Department cyber security and Computer Engineering, Kyiv National University of Construction and Architecture, y.khlaponin@knuba.edu.ua

information security problem on the basis of regulations and requirements existing in the state [1,2,3].
The task of the overall information security assessment, based on the methodology of both quantitative and qualitative performance measures, is becoming more and more relevant provided that the information technology develops fast, the number of information security threats, with their degree of uncertainty and realization, increases rapidly, and the systems of information security together with their specific orientation become more complex.
The methodology of a qualitative information security assessment is based on the results of data measurement and expert analyses, the quality of which greatly depends on the experts' qualification and training that cause a low level of information security. The estimates of expert analyses can be fuzzy or vaguely defined to be described through mathematical functions. Apart from that, this information can be affine; the estimate of parameters will be carried out using different scales. However, it is often possible to describe these systems through the heuristic approach where fuzzy rule-based structures and various functions are used.
Based on the theory of the rule-based fuzzy models (networks), the article suggests a connectionist model of information security assessment system.

2. Research results

The information about the system's parameters, inputs, outputs, and the state can be unreliable, fuzzy or ill-defined. Rule-based fuzzy models belong to the most common type o fuzzy models used for description, analysis, and modeling of complex ill-defined systems and processes. The algorithms of fuzzy logic by Mamdani, Larsen, Tsukamoto, and Takagi-Sugeno have currently become the most widespread [4].
By "rule-based fuzzy model" we shall denote a conformal set of separate fuzzy rules **"if A, then B"** (where A and B are antecedent and consequent of the corresponding rule represented fuzzy statements), aimed at determining the degree to which statements of the fuzzy rules are true, based on the antecedents with the known degree of truth from the corresponding rules.
In order to formulate simple fuzzy statements in antecedents and consequences of the fuzzy rules it is needed to enter a membership function of the corresponding fuzzy sets.
For information security assessment there can be used rule-based fuzzy models and fuzzy consequent algorithms created on their basis. A series of pre-defined membership functions which covers the base set of input and output variables can be set to the information security expert in advance, or it will be needed for him to create membership functions on his own.
Information can be represented in the form of an added sound (linguistic information), processed information (information that is being passed around in the information telecommunication system (ITS)), or it can be stored on multiple data storage devices (paper, magnetized medium, and other physical storage medium).
For internet security assessment the expert needs to take into consideration all the possible technical channels of information leakage. The state of each channel will correspond to the particular state of internet security resulting from certain types of threats.

In spite of the numerous advantages of the rule-based fuzzy models, they have certain drawbacks:

- input set of the fuzzy rules is formulated by an expert and may appear to be incomplete or inconsistent;
- choice of membership function's type and parameters in fuzzy rule statements is subjective;
- no chance for automatic knowledge acquisition.

Formally, rule-based fuzzy models can be presented as fuzzy networks that are structurally identical to multilayer neural networks, in which each separate level of the fuzzy consequent in the rule-based fuzzy model is implemented by the elements of each layer.

In order to overcome these drawbacks it has been proposed to create adaptive rule-based fuzzy models and also implement different components of these models using the neural network technology.

The neural network structure (NNS) of information security assessment, shown in the Figure 1., includes *m*-neural ensembles (layers) which are determined by the state of quantity of information security related to certain types of threats. The security corresponds to the neural layer; the number of classes is determined by the parameters which are measured and compared with requirements in order to define the state of information security for each of the technical channels of the information leakage based on the tested threat models.

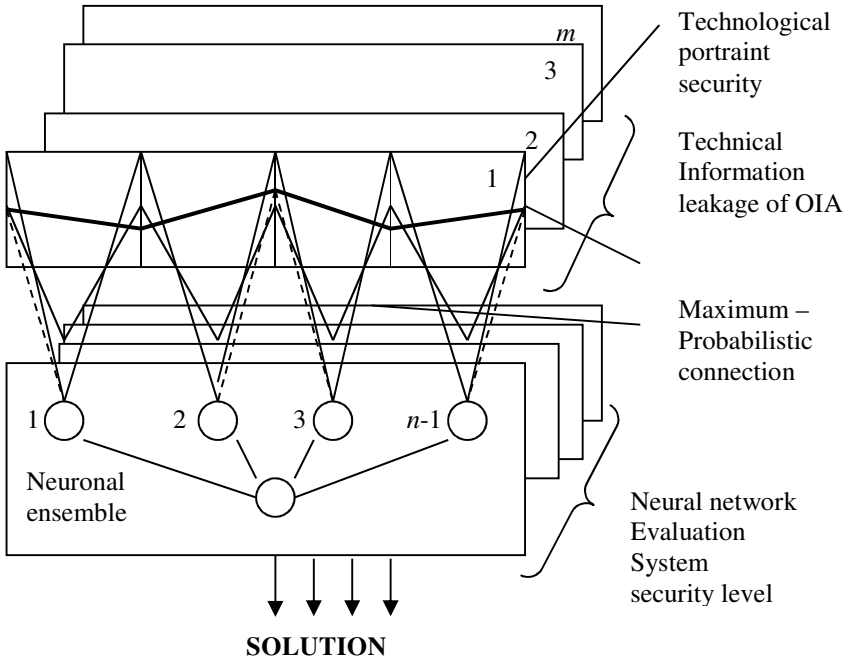


Figure 1. Neural network structure

According to the results of the examination and initial instrumental (calculated) security evaluation for each of a possible technical channel there are binary matrices

of $n \times M_l$, where M_l is a total quantity of threats that characterizes the *l*-technical channel.

As a result, there is a set of matrices TK^*_{RC}; TK^*_{AEC}; TK^*_{AC}; TK^*_{VAC}; TK^*_{CSEMRI}; TK^*_{PC}; TK^*_{OC} by which a formal description of an information activity object (IAO) is being determined.

The validity check is to be carried out, and the final list of the possible technical channels of information leakage in the object is to be determined.

According to the above mentioned results, there is a multitude $\{m_l\}_{\text{TCIL}}$, $l = \overline{1, L}$, де $L \le 7$.

Security states that correspond to the technical channels of information leakage, determined in IAO at a certain point, can be represented as dynamical systems. They are called events and are presented as processed security profiles. Thus, this term presupposes that a determined processed state or configuration of states will be taken over by the dynamic distributed network.

The evolutionary changes occurring in between events are not taken into consideration; it is believed that the dynamics of the system is being developed discretely from event to event. This system is called a Discrete Event System [5].

Environment for connectionist system assessment of the information security level can be represented as a set of discrete-event systems with connected discrete technological security. Getting the required number of instrumental measurements and special studies for each of the technical information leakage channel obtained in information activity object, you need to develop a procedure for processing measurements to receive automatically information about the state of technological protection of each channel according to the approved model threats.

The results of instrumental measurements and special studies for each of the technical information leakage channel are perceived by sensor matrix as a set of observations: $X = (X_1, X_2, ..., X_i, ..., X_m)$, $i = \overline{1, n}$.

In a separate sensor channel there is a reduction of sample space X, which results in a sequence of discrete variables U_k, $k = \overline{0, n-1}$ which take the values $Z_1, Z_2, ..., Z_\gamma$.

It is necessary to synthesize the structure of the neural classifier that realizes a crucial function $\gamma(U)$ in the reduced sample space *U* [6].

The sequence of discrete variables U_k, $k = \overline{0, n-1}$, which take values z_a, $a = \overline{1, r}$, can be approximated by vectors $\Xi, \Phi(0)_\mu$ and $\Xi, \Phi(k)_\mu$.

The structure of the simplest neural system of assessing the level of security - a set of *M*+1 neural network ensembles of the first layer. The ensemble consists of *n*-neurons, the level of excitement of which is defined as

$$Y_\mu(k) = \sum_{a=1}^{n} \Xi_a(k) \Phi_a(k)_\mu$$

where $\Xi, \Phi(0)_\mu$ and $\Xi, \Phi(k)_\mu$. - vectors, which can be approximated by a sequence of discrete variables U_k, $k = \overline{0, n-1}$, taking values z_a, $a = \overline{1, r}$ [5].

Each neuron performs encoding process which is determined by the so-called labeled lines method in which a certain value of the process is provided in conformity with

identified (labeled) lines $Z_1, Z_2, ..., Z_a, ..., Z_k$ and therefore a certain value setting process meets the most excited synoptic connection $\Xi_a(k) = 1$.

Unlike typical neuron with equivalent synoptic connections, the neuron that carries out coding using method of labeled lines synoptic connections have precedence. Synoptic entrance with a large number corresponds with more important process parameters.

Another difference is that in the *k* moment only one synoptic communication is excited and thereby the task of administration and management of threshold Θ is greatly simplified using the weight function *w*. Actually

$$Y_\mu(k) = \sum_{a=1}^{n} \Xi_a(k) \Phi_a(k)_\mu - \Theta_a(k)_\mu$$

For each set of technical channel from multitude $\{m_l\}_{\text{TCIL}}$ should determine its importance. [8] Despite the difference in the number of threats according to [5], which define every possible technical channel, this number does not determine the gravity of the technical channel. Therefore, to determine the importance of factors it will be rational to use the ratio between the total number of threats to a particular channel and the number of threats which received value "1" on the results of expert surveys.

Then the formula for determining the importance coefficient of each technical coefficients of the channel list has the form:

$$\lambda_l = \frac{M_l^1}{M_l}, \quad l = \overline{1, L},$$

where M_l^1 - number of threats of *l* technical channel, which received "1" based on the results of the expert survey; M_l - the total number of threats that characterize *l* - technical channel.

Thus, a set of values of the importance of every possible technical information leakage is formed - $\{m_l\}_{\text{TCIL}}$, $l = \overline{1, L}$. Synoptic entry of more number corresponds to the setting of technical channel with higher ratings.

It is necessary to carry out technical channels ranging according to the importance and form the final list of possible technical information leakage to the information activity object:

$\{m_l\}_{\text{TCIL}}$, $l = \overline{1, L}$, де $L \leq 8$.

3. Conclusions

1. When solving the problems related to the assessment of the security information it is necessary to assess the degree of conformity to received signals (technological protection portraits - results of instrumental measurements and special studies for each of the technical information leakage) reference, that is, to determine a decision criterion.

2. A quantitative measure of conformity to choose differently according to the nature of the research.

3. A procedure for handling such measurements gives opportunity to receive security information about each channel automatically according to the approved model threats.
4. The model of fuzzy neural network system of assessment of the production of information security, which is different from the known by the fact that the neural network structure is focused on solutions of specific task - the creation and the certification of the information activity object or a creation of comprehensive information security system in ITS. The demand of problem orientation of neural network (NN) leads to the realization of the principle of adequacy of its structure and environment.
5. The initial structuring of neural network should be conducted by formal synthesis methods, which help to determine optimal structure, including the number of neural layers and neural ensembles, the number of neural elements in each layer, the presence of deterministic relations between them and the weight numbers.

REFERENCES

1. Law of Ukraine "On protection of information in telecommunication systems".
2. BORISOV V.V., KRUGLOV V.V., FEDULOV A.S.: Nechetkye model and Networks, 2012, 284.
3. KRIVUTSA V.G., BERKMAN L.N., TOLIUPA S.V. Infocommunication new generation of information network, 2012, 288.
4. KOROLOV A.P., TOLIUPA S.V., THORZHEVSKYJ I.V. The need to build neural systems of technical diagnostics electronic equipment. Scientific works KVIUZ, 3(2000), 51-60.
5. LENKOV S.V., PEREGUDOV D.A., HOROSHKO V.A. Methods of information and protection means, 2010, 464.
6. KHLAPONIN Y.I. Model of assessing the level of information security based on neural network. Modern information technology security and defense, 1(2014), 96-100.

Ruslana ZIUBINA[1], Yuliia BOIKO[2]

Opiekun naukowy: Olexandr YUDIN[3]

METODY IDENTYFIKACJI I UWIERZYTELNIANIA SYGNAŁÓW AUDIO

Streszczenie: Stwierdzono, że identyfikacja przez rozpoznanie głosu sprzyja podwyższeniu pewności użycia systemów obrony krytycznej informacji. Przeprowadzono ocenę efektywności opracowanych metod: np. efektywnej szerokości spektrum oraz największej informacyjnej wagi głównego tonu - w zadaniach identyfikacji sygnałów audio.

Słowa kluczowe: częstość głównego tonu; metody identyfikacji spikera; rozpoznanie sygnałów

METHODS OF IDENTIFICATION AND AUTHENTICATION OF AUDIO SIGNALS

Summary: Determined that identification by voice contributes to the reliability of the use of protection systems critical information. An evaluation of the effectiveness of the developed methods effective spectral width and the largest scales the information of the fundamental tone in the tasks of identification of audio signals.

Keywords: pitch frequency; methods of speaker identification; signal recognition

1. Introduction

The process of active informatization of society, increasing amounts of information transmission, conducting many operations in cyberspace have significantly increased requirements to the Information and Communication Systems of protection of information resources. In terms of efficient transfer of critical information and effective protection is a topical issue identification and confirmation of identity on the basis of modern technologies and methods of biometrics in other fields of activity of the state and society. Biometrics holds a special place in the process of ensuring reliable access to physical and information resources in the ICS of general or special purpose. Biometrics technologies are successfully used in the different spheres of the human activity, such as bank operations, registration of citizenship, law enforcement agencies and others [1].

[1] National Aviation University: Department of Computerized Systems of Information Protection, email kszi@ukr.net
[2] Candidate of Sciences (Technical) National Aviation University, Department of Information Technology Security, email julia_boyko2010@ukr.net
[3] Professor, Doctor of Science (Technical), National Aviation University, Institute of Computer Information Technologies, email kszi@ukr.net

2. The main part

The most common biometrical characteristics are the fingerprints, the structural features of the human face, the hand geometry, the iris, the signature and the voice. The choice of the biometric characteristic depends on the system in which it would be used, from the quantity of people who will be identified and from the level of security of the object or system.[2]

The usage of the modern systems of identification a person by voice becomes very popular due to the low cost of the reader. For identification a person by fingerprints in the device should be mounted a special reader, what significantly increases the cost of the entire device. The scan of retina and face geometry also requires additional reader devices. At the same time the microphone is in every mobile device, even in the cheapest one, what makes the identification a person by the voice quite relevant direction of research.

In the process of personal identification the system determines four questions:

- what a person know (password);
- what a person has (key, passport);
- who a person is (biometric characteristics);
- what a person is doing (talk, write).

Unlike to other systems of biometric identification precisely the voice biometrics is answering to all these questions, that in its turn increases the reliability of the security system's work.

The basis of the work of systems of biometric identification is the mathematical statistics, while mathematical statistics is based on two main features – the mistake of first kind or "false alarm" and the second kind or "skipping signal". In biometrics there are also theirs identical concepts, namely FAR (False Acceptable Rate) and FRR (False Rejection Rate). The first concept describes the false coincidence of biometrics of two people, and the second one – the probability of rejection of access to a person who has a permit. Thus, the system is deemed the better, the lower the meaning of FRR is in case of the same meanings of FAR.

Voice biometrics work in two modes – active and passive. In the active mode (it depends on the text) the system offers to repeat random, pre-recorded phrases, and in the passive mode (independent from the text) the system identifies a person by his free speech.

Systems of biometric determination of a human voice solve two main tasks: checking the sample of the voice among pre-recorded samples in database (identification of a personality); confirming a personality by the special characteristics of the voice (verification of a personality). As the main purpose of the work is the identification of the voice, which in turn is more complicated task compared to verification – it is necessary to define the basic methods of decision-making in a process of identification of the voice and to determine among them the most optimal for the task.

The main purpose of decisions making is to choose one that will lead to the most favorable consequences. Such rules of making decisions are called optimal.

The results of decision making are evaluated according to the degree of their compliance to the intended purpose. This compliance is determined using quantitative measure that define the win or loss from the made decision. This measure is called a the loss function (penalty) or the win function (objective function).[3]

One of the main tasks of identification of the voice is the definition of the feature space, which will help to held analysis and identification. Methods of the effective width of the spectrum and the largest scales of the fundamental tone [4] in the tasks of identification and authentication of audio signals lie in the amplitude-frequency domain and uses the following informative characteristics:

- the intensity, amplitude;
- energy;
- the frequency of the fundamental tone;
- formant frequency [3].

Given that the methods identify the speaker's voice in conditions of high noise, the influence of which is making significant changes on most of the characteristics, the basic characteristic is the frequency of the fundamental tone, and values of mathematical expectation of each component of the frequency spectrum of the signal. As a result of the analysis of existing methods of determining the fundamental frequency, for methods of the effective width of the spectrum and the largest scales of the fundamental tone in the tasks of identification and authentication of audio signals selected the most convenient method for the determination of the fundamental frequency, namely the method of spectral harmonic analysis. The basis of this method is the idea of spectral analysis and determining maximum values of amplitudes. The frequency of the fundamental tone can be determined by finding the frequency harmonic components with the maximum amplitude and compute the greatest common divisor of these frequencies for the harmonic components. This divisor is determined by the recording frequency in a histogram for each harmonic. The harmonic with the greatest amplitude represents the greatest common divisor frequency, and, therefore, is the frequency of the fundamental tone. The advantages of this method include: ease of implementation, resistance to additive and multiplicative noise, the ability to control certain characteristics. [5]

Using the methods of voice identification gave the following results:

1) Probability of correct identification of the female voice $S_{ex1}(t)$ and $S_{ex2}(t)$ in the aggregate $_{DB}$ is 100% in the case of 5.22 dB for Method of the effective width of the spectrum and 5,26 dB for Metod of the largest scales of the fundamental tone, nd only after that starts to decrease (Fig. 1.).

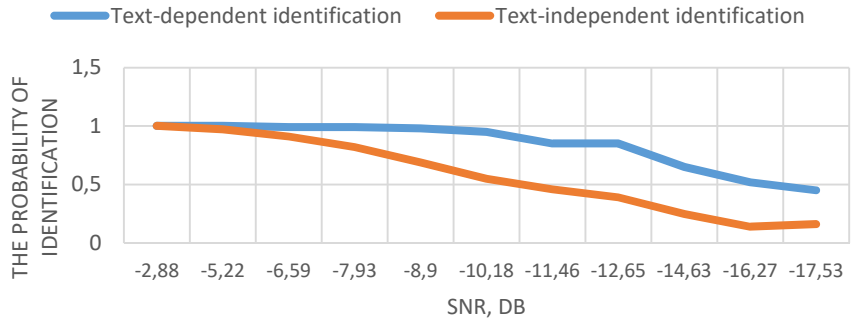


Figure 1. The dependence of the probability of correct identification of female voices in the plural DB depending on the level of SNR for Method of the effective width of the spectrum

2) The frequency of the fundamental tone and overtones are used as the main information components in the process of identifying a person by voice and give you the opportunity to neglect the rest of the spectrum. Using of the Metod of the largest scales of the fundamental tone gives the opportunity to identify the voice when the ratio noise/signal is 5.21 dB, which reduces the efficiency independent from the text identify only 2.1%, and does not affect the probability of text-dependent identification (Fig.2.)

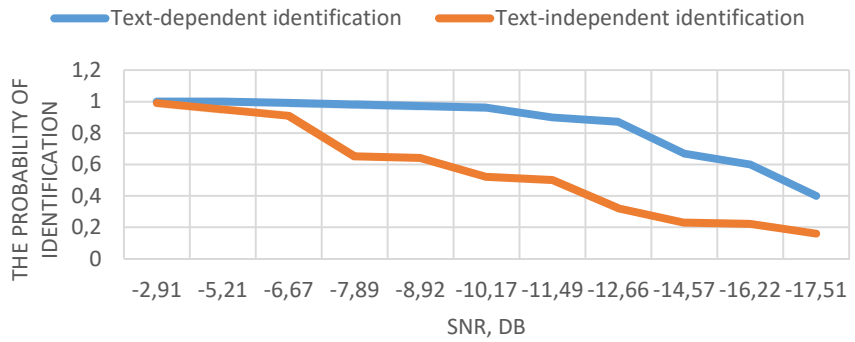


Figure 2. The dependence of the correct identification probability of female voices in the plural DB depending on the level of SNR for Metod of the largest scales of the fundamental tone

3) Samples of male and female voices are in different frequency ranges, so it is advisable to divide them into two subgroups *A* and *B*. The result of the experiment showed that the accurate identification of voices is possible when the SNR is of 5.29 dB for the text-dependent case and to 4.27 dB-independent text for the Method of the effective width of the spectrum. A certain dynamics can be traced to the Metod of the largest scales of the fundamental tone indices, respectively, is 5.7 dB and 5.24 dB. So, based on the foregoing, we can conclude that the split the database into two sets with different groups of samples will give the opportunity to improve the system performance on average by 27% for various values of SNR.

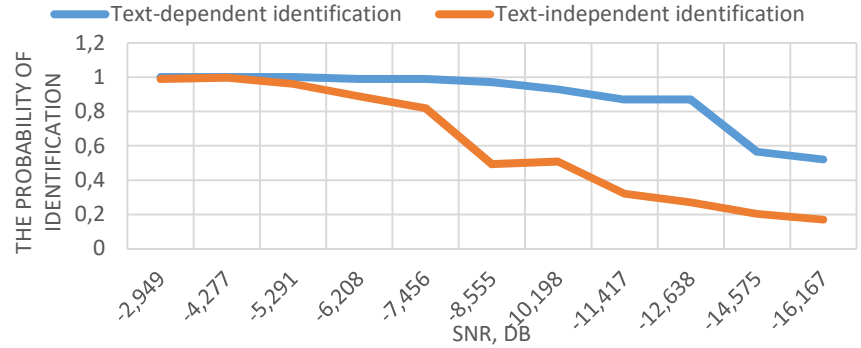


Figure 3. The dependence of the correct identification probability of female voices in the plural B depending on the level of SNR for Method of the effective width of the spectrum

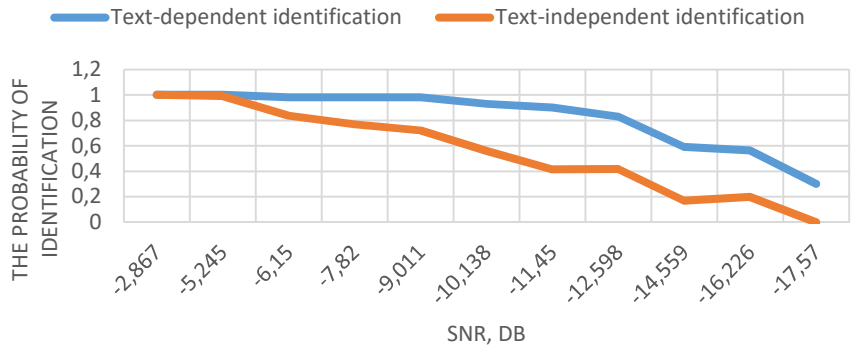


Figure 4. The dependence of the correct identification probability of female voices in the plural Б depending on the level of SNR for Metod of the largest scales of the fundamental tone

4) Checking the operation of the system for plural *B* confirmed the authenticity of the identification and the need for the division of the plural *DB* for two subsets. Identification of female voices among the male lower possible even when the noise exceeds the signal by more than 10 dB for both text-dependent and text-independent methods.

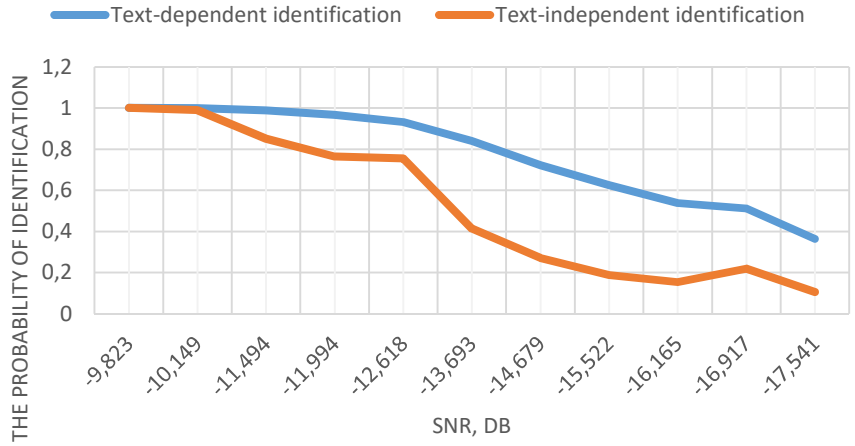


Figure 5. The dependence of the correct identification probability of female voices in the plural A depending on the level of SNR for Method of the effective width of the spectrum

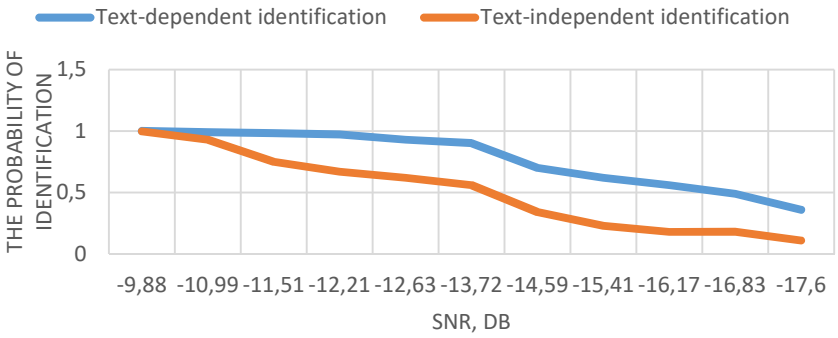


Figure 6. The dependence of the correct identification probability of female voices in the plural A depending on the level of SNR for Method of the largest scales of the fundamental tone

3. Conclusions

Such indicators identification of the speech signal demonstrate the versatility of the developed Methods of the largest scales of the fundamental tone and the effective width of the spectrum, and allow you to choose between text-dependent and text-independent identification depending on the chosen threshold of decision.

REFERENCES

1. Юдин А. К. и др. Основные понятия и математические аспекты в задачах канального кодирования: многоальтернативные правила, Кибернетика и системный анализ. – 2016. – №. 52,№ 6. – С. 53-59.
2. Юдин А. К. и др. Метод определения информативных составляющих на основании построения последовательного правила принятия решения, Кибернетика и системный анализ. – 2016. – №. 52,№ 2. – С. 167-173.
3. Крашенинников В. Р. Основы теории обработки изображений, учеб. пособие. – 2003.
4. Юдін О. К., Зюбіна Р. В. Оцінка ефективності методів ефективної ширини спектру та найбільшої інформаційної ваги основного тону в задачах ідентифікації та автентифікації аудіо сигналів, Наукоємні технології, 35(2017)3.
5. Юдін О. К., Зюбіна Р. В. Класифікація методів ідентифікації частоти основного тону Наукоємні технології, 33(2017)1. doi.org/10.18372/2310-5461.33.11553

INDEKS NAZWISK

INDEX OF NAMES