Informatics Control Measurement
in Economy and Environment Protection

4/2018

Informatyka Automatyka Pomiary

w gospodarce i ochronie środowiska

www.e-IAPGOS.pl

ISSN 2083-0157

Kwartalnik Naukowo-Technicznу

fot. Andriy Samila

Narodowy Uniwersytet „Politechnika Lwowska”
(Lwów, Ukraina)

4/2018

październik – grudzień

Wydanie pod redakcją naukową
prof. dr hab. inż. Waldemara Wójcika

INFORMATYKA AUTOMATYKA POMIARY

W GOSPODARCE i OCHRONIE ŚRODOWISKA

Informatics Control Measurement in Economy and Environment Protection

p-ISSN 2083-0157, e-ISSN 2391-6761, www.e-iapgos.pl

INTERNATIONAL PROGRAMME COMMITTEE – RADA PROGRAMOWO-NAUKOWA

EDITORIAL OFFICE – REDAKCJA

Redakcja czasopisma
Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska
Instytut Elektroniki i Technik Informacyjnych
Politechnika Lubelska
ul. Nadbystrzycka 38A
20-618 Lublin
tel. +48 81 53 84 309
fax: +48 81 53 84 312
www.e-iapgos.pl
iapgos@pollub.pl

PUBLISHER – WYDAWCA

Politechnika Lubelska
ul. Nadbystrzycka 38D
20-618 Lublin
tel. +48 81 53 84 100
www.pollub.pl

PRINTING HOUSE – DRUKARNIA

Agencja Reklamowa TOP
87-800 Włocławek
ul. Toruńska 148
tel. +48 54 423 20 40
nakład: 100 egzemplarzy

OTHER INFORMATION – INNE INFORMACJE

Czasopismo jest indeksowane w bazach:

BazTech:	baztech.icm.edu.pl
IC Journals Master List:	www.journals.indexcopernicus.com
Google Scholar	scholar.google.pl
POL-index	pbn.nauka.gov.pl

Czasopismo znajduje się w wykazie czasopism naukowych opublikowanym w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 9 grudnia 2016 r. – część B, pozycja 685 – z liczbą punktów przyznawanych za publikację równą 7.

Zasady publikowania artykułów, przygotowania tekstów, zasady etyczne, procedura recenzowania, wykazy recenzentów oraz pełne teksty artykułów dostępne są na stronie internetowej czasopisma:

www.e-iapgos.pl

W celu zwiększenia oddziaływania czasopisma w środowisku naukowym redakcja zaleca:

- w artykułach publikowanych w IAPGOS cytować artykuły z renomowanych czasopism międzynarodowych (szczególnie indeksowanych w bazach Web of Science oraz Scopus) używając oficjalnych skrótów nazw czasopism,
- w artykułach publikowanych w innych czasopismach (zwłaszcza indeksowanych w bazach Web of Science oraz Scopus) cytować prace publikowane w IAPGOS – zwłaszcza posługując się numerami DOI, np.:
  Kluszczyński K. *Modelowanie – umiejętność czy sztuka?* Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOS, 1/2016, 4–15, DOI: 10.5604/20830157.1193833.

CONTENTS – SPIS TREŚCI

DOI: 10.5604/01.3001.0012.7980

THE SOURCES OF RADIATION IN THE SHORT-WAVE RANGE ON THE BASIS OF II-VI HETEROLAYERS

Mikhail Slyotov[1], Alexey Slyotov[2]
[1]Yuriy Fedkovych Chernivtsi National University, Physical, Technical and Computer Science Institute, Department of Optics, Publishing and Printing
[2]Yuriy Fedkovych Chernivtsi National University, Physical, Technical and Computer Science Institute, Department of Electronics and Power Engineering

Abstract. *The possibility of obtaining zinc selenide and zinc sulfide layers of hexagonal modification by isovalent substitution method is shown. They are characterized by intensive luminescence which is formed by the dominant annihilation of bound excitons for α-ZnSe and recombination on donor-acceptor pairs for α-ZnS. The resulting radiation covers the violet wavelength range. Quantum radiation efficiency reaches η = 10–12% for α-ZnSe and η = 5–8% for α-ZnS. The radiation is characterized by high temperature stability and repeatability of characteristics and parameters.*

Keywords: cadmium sulfide, zinc selenide, zinc sulfide, hexagonal modification, reflection spectrum, photoluminescence

ŹRÓDŁA PROMIENIOWANIA W ZAKRESIE KRÓTKOFALOWYM NA PODSTAWIE HETEROWARSTW GRUP II-VI

Streszczenie. *Pokazano, że możliwe jest uzyskanie heterogennych warstw selenku i siarczku cynku o modyfikacji heksagonalnej za pomocą metody izowalentnego podstawienia. Charakteryzują się one intensywną luminescencją, która powstaje w wyniku anihilacji związanych ekscytonów dla dominującego pasma α-ZnSe i rekombinacji na parach donor-akceptor w przypadku α-ZnS. Otrzymane promieniowanie pokrywa fioletowy zakres optyczny. Sprawność kwantowa promieniowania wynosi η = 10–12% dla α-ZnSe i η = 5–8% dla α-ZnS. Promieniowanie charakteryzuje się wysoką stabilnością temperaturową oraz powtarzalnością charakterystyk i parametrów.*

Słowa kluczowe*: siarczek kadmu, selenek cynku, siarczek cynku, modyfikacja heksagonalna, mieszanie widma, fotoluminescencja*

Introduction

An important direction of optoelectronics is creation of various types of systems for the formation, transmission and processing of information. The use of optical communication in them can greatly improve the speed, significantly increase the throughput, the density of recording information and its noise immunity, and, most importantly, the ability to visualize the transmitted information. Therefore, it is important to obtain materials and create devices on their basis that extend the optical region to the short-wave range. Such requirements are met by wide-gap II-VI compounds the direct band nature of which assures high efficiency of generation and recombination processes in various types of light-emitting diodes and photodetectors [10]. Therefore, the study of the possibility of obtaining *α-ZnSe* and *α-ZnS* layers with a time-stable structure and the investigation of their optical properties and photoluminescence appears to be relevant. The heterostructures based on such layers can play an important role in the creation of various types of optoelectronic devices with the stable and repetitive properties.

1. Measurement procedure

The optical properties and photoluminescence (PL) of the *α-ZnSe* and *α-ZnS* heterolayers were studied. For this purpose a universal optical setup was used, Fig. 1. It allowed integrated measurements to be carried out under the identical conditions of optical reflection spectra and luminescence, using the classical procedure and the *λ*-modulation method [8]. The latter method significantly increased the sensitivity to spectral distribution features. In the case of optical reflection studies, this provided an opportunity to determine the type and parameters of the band structure of the resulting material. PL measurements made it possible to establish both important properties, characteristics and parameters of generation and recombination processes, and the opportunities of practical application of the obtained *α-ZnSe* and *α-ZnS* heterolayers. When constructing PL plots, the hardware function of the setup was taken into account.

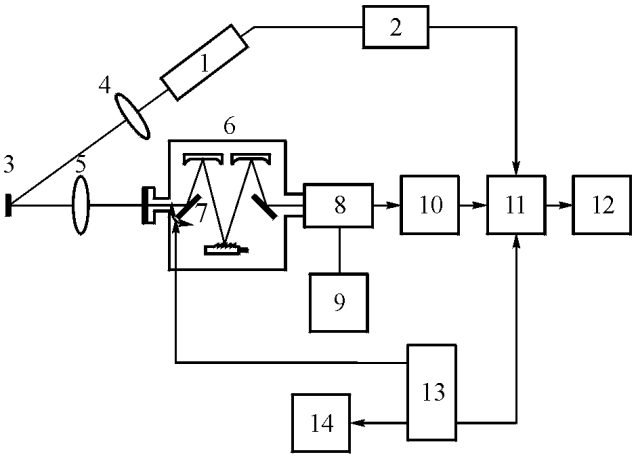


Fig. 1. Schematic diagram of experimental universal optical setup: (1) source of radiation (laser or lamp) with its power source (2); (3) sample; (4, 5) lenses; (6) monochromator MDR-23; (7) oscillating mirror (modulator); (8) photodetector with its power source (9); (10) selective amplifier; (11) synchronous detector; (12) plotting device; (13) sound generator; (14) millivoltmeter

2. Formatting of heterolayers components

The *α-ZnSe* and *α-ZnS* heterolayers of hexagonal modification were obtained on *α-CdS* crystals by isovalent substitution method [7]. Isothermal annealing took place consecutively in two steps in saturated vapours of *Zn* and *Se* elements. The diffusion character of the processes of substitution by isovalent elements causes growth of the layer deep into the substrate, which assures the necessary conditions for it to repeat the crystalline structure of the substrate.

The reaction of formation of heterolayers is described by the following equations

$$\alpha\text{-}CdS_S + Zn_G \rightarrow \alpha\text{-}ZnS_S + Cd_G\,, \quad (1)$$

$$\alpha\text{-}ZnS_S + Se_G \rightarrow \alpha\text{-}ZnSe_S + S_G\,, \quad (2)$$

where indices "*S*" and "*G*" correspond to the solid and gaseous states of reagents.

Also, isovalent substitution method provided formation by reaction (1) on the surface of *α-CdS* single crystals of hexagonal *α-ZnS* heterolayers. Note that at the present time the properties of the resulting *α-ZnSe* and *α-ZnS* heterolayers remain poorly understood.

3. Results and discussions

The *α-ZnSe* and *α-ZnS* heterolayers, which were obtained by isovalent substitution method, have hexagonal lattice structure. Their formation in the case *α-ZnSe* is confirmed by the generation on the surface of *α-CdS* of a film characterized by zinc selenide yellow-red colour. But in the case of *α-ZnS* heterolayers, obtained films were characterized by the absence of colour. X-ray studies showed a pattern of diffraction maxima typical of the hexagonal lattice. Formation of a hexagonal modification of the crystalline structure of the obtained heterolayers is also confirmed by the studies of differential optical reflection spectra, Fig. 2. The structure of energy bands at point Γ of the Brillouin zones is exhibited which is typical of the hexagonal lattice. The main maximum corresponds to the optical transitions between the bands of allowed energies E_g = 2.89 eV for *α-ZnSe* and E_g = 3.86 eV for *α-ZnS* [3]. For the first time, the values of the valence band splitting into subbands with the participation of the crystalline field Δ_{CR} = 0.07 eV and the spin-orbital interaction Δ_{SO} = 0.37 eV were determined. In the case of *α-ZnS* these values are as follows: Δ_{CR} = 0.055 eV and Δ_{SO} = 0.092 eV.

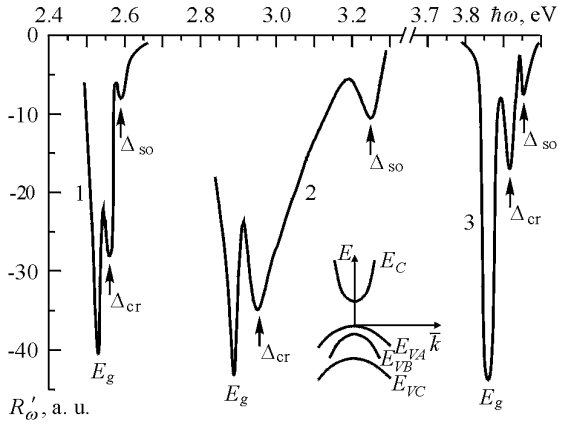


Fig. 2. Differential optical reflection spectra of basic α-CdS substrates (1) and annealed α-ZnSe (2) and α-ZnS (3) heterolayers. The inset shows the energy band structure of hexagonal structure semiconductors

The *α-ZnSe* heterolayers are characterized by intensive photoluminescence in a short-wave range with the photon energies $\hbar\omega$ = 2.72–3.3 eV at 300 K. Determination of its quantum efficiency by the known method gives the values of η = 10–12% [5]. For the widely used *β-ZnSe* crystals its value does not exceed 0.1–0.4%.

The radiation of the obtained *α-ZnSe* heterolayers is characterized by maxima $\hbar\omega_m$ = 2.82 eV and $\hbar\omega_m$ = 2.92 Ev in the violet range which correspond to two photon energy regions, Fig. 3 curve 1. Their spectral ranges are divided by the value of E_g = 2.89 eV. In the range $\hbar\omega \geq E_g$ the luminescence has the following features: a) independence of the position of maximum $\hbar\omega_m$ of photoexcitation level L; b) quadratic dependence of intensity I on L, that is, $I \sim L^2$, Fig. 3 curve 4. Spectral shape is well approximated by the analytic expression which describes interband transitions of free charge carriers [9]

$$N = (\hbar\omega)^2 \sqrt{\hbar\omega - E_g}\, \exp\left(-\frac{\hbar\omega - E_g}{kT}\right), \qquad (3)$$

where k is the Boltzmann constant, T is temperature. When substituting the value E_g = 2.89 eV, the band A is obtained, Fig. 3. Its mismatch with the experimental curve indicates the presence of the second component. It is determined by the optical transitions involving a subband that was split by crystalline field Δ_{CR} = 0.07 eV. The substitution of the value $E_g + \Delta_{CR}$ into (3) yields the calculated band B the shape of which is in good agreement with the experiment, Fig. 3.

This testifies to the active role of generation and recombination processes in the region $\hbar\omega > E_g$, with the participation of the main allowed energy bands and the valence subband, that is split into Δ_{CR} due to the action of the crystalline field. The presence of these components is confirmed by the experiment in the study by experimental λ-modulation. A typical differential curve of the first derivative N'_ω with two singularities is observed, the position of which is consistent with the analytical calculations of bands A and B, Fig. 3 curve 2.

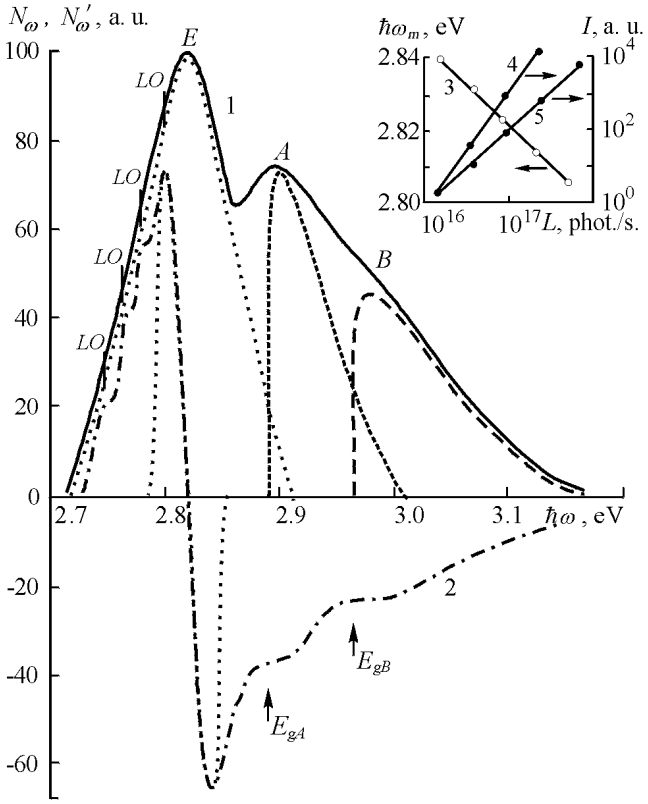


Fig. 3. The normal (1) and λ-modulated (2) photoluminescence spectrum of the α-ZnSe heterolayers. In the inset – maximum position (3) and intensity (4, 5) versus photoexcitation level L: 4 – $I \sim L^2$, 5 – $I \sim L^{1.5}$, T = 300 K

The singularity E on the curve 2 at its intersection with the axis of energies $\hbar\omega$ = 2.82 eV corresponds to the maximum of the dominant band E in the region $\hbar\omega < E_g$. The position of the maximum $\hbar\omega_m$ is shifted to the region of lower energies with increasing L (Fig. 3, curve 3), and the intensity $I \sim L^{1.5}$ (Fig. 3, curve 5). These properties of band E are characteristic of radiation due to the annihilation of excitons [6]. Its observation at 300 K is caused by the processes of isovalent impurities formation in the formation of *α-ZnSe* heterolayers by the isovalent substitution method. These are the residual (not completely substituted) atoms of the base substance.

As a result of short-range potential of isovalent impurities, the exciton is localized thereupon, that is, bound excitons are formed. So, the resulting heterolayers are characterized by high quantum radiation efficiency. It is also defined by "purification effect" of the obtained material *α-ZnSe* [8]. At the same time, excitons localized on isovalent impurities are temperature-resistant. Therefore, in experiment one could observe radiation even at T = 580 K, though under such conditions the PL of *β-ZnSe* crystals is completely quenched. PL intensity of the resulting *α-ZnSe* heterolayers obtained by isovalent substitution decreased insignificantly. It should be noted that after cooling the nature of the spectra and the intensity were completely restored.

The *α-ZnS* heterolayers obtained by isovalent method are characterized by intensive photoluminescence at 300 K. Its quantum efficiency is η = 7–8%. In so doing, the radiation spectrum covers a wide range of photon energies from 2.5 to 3.30 eV. The maximum falls on $\hbar\omega_m$ = 2.90 eV, Fig. 4.

The specific feature of radiation is the shift of the maximum towards higher photon energies with increase in photoexcitation level from 10^{16} phot/sec to 10^{18} phot/sec. This testifies to recombination on donor-acceptor pairs (DAP) [9]. According to studies of the second derivative of the PL spectra, the main components of the radiation spectrum were determined, Fig. 4, curve 2. Their presence is confirmed by calculations according to the well-known Alentsev-Fock method [1]. In Fig. 4 they are shown by corresponding dashed lines.

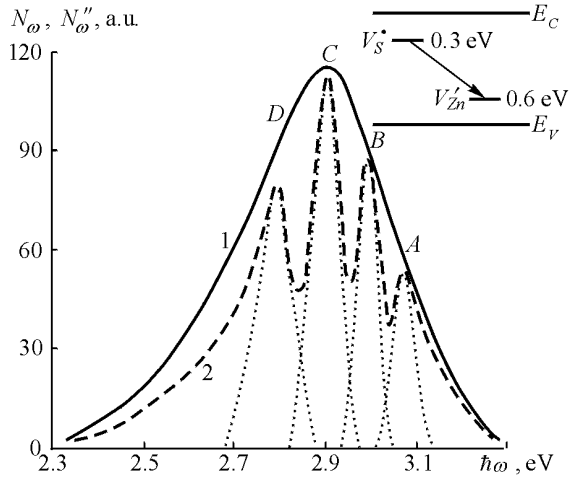


Fig. 4. The normal (1) and λ-modulated at a frequency of 2Ω (2) PL spectra of α-ZnS heterolayers. On the inset – the energy structure of simple centres, T = 300 K

The investigations of the differential spectra of optical reflection made it possible to reveal the specific features at the energies of 3.86 eV, 3.56 eV, 3.26 eV and 2.90 eV, Fig. 5.

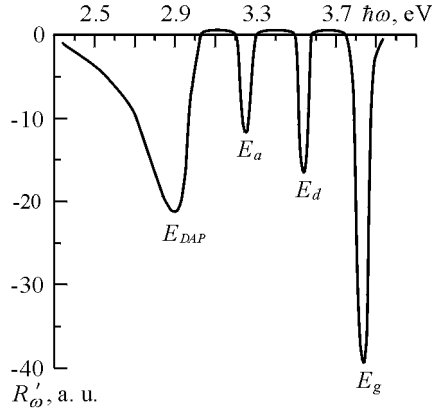


Fig. 5. λ-modulated spectra of optical reflection of α-ZnS heterolayers, T = 300 K

They correspond to the bandgap width of *α-ZnS* and the depth of the donor E_d and acceptor E_a states. The theoretical analysis of the spectral distribution of luminescence with the participation of the DAP (V'_{Zn} $V_S^\bullet$) was carried out according to the known equation [6, 10]

$$\hbar\omega = E_g - E_d - E_a + E_k = E_g - E_d - E_a + \frac{e^2}{4\pi\varepsilon_0\varepsilon R_i}, \quad (4)$$

where E_a and E_d are the depths of corresponding acceptor and donor levels; ε_0 is dielectric constant, ε is the dielectric permittivity of the material (semiconductor), R_i is the distance between the DAP partners.

The revealed components of the *α-ZnS* luminescence spectra are located in a non-equidistant fashion with the maxima at photon energies 3.072 eV (*A*), 2.989 eV (*B*), 2.904 eV (*C*), 2.795 eV (*D*). They were used to calculate the corresponding distances between the DAP partners with the participation of which the radiation transitions take place. They are 7 Å, 8.4 Å, 10.5 Å, 15.6 Å. The high-energy edge of the PL spectrum corresponds to the distance of R_i = 4.1 Å.

The resulting *α-ZnSe* and *α-ZnS* heterolayers are characterized by the anisotropy of optical properties. It is due to the hexagonal structure of crystalline lattice. This determines polarization of the optical transmission and reflection, as well as radiation of heterolayers in the short-wave range.

Important for practical application of the resulting *α-ZnS* and *α-ZnSe* heterolayers is a weak temperature dependence of their radiation (high-temperature properties). Integrated research has shown that maximum intensity of their luminescence is reduced not more than by a factor of 2 on heating from room temperatures to *T* = 480–500 K. Moreover, the character of the spectral distribution of radiation is almost unchanged when heated in the specified temperature range. Experimentally, as an example, typical dependences of their intensity and spectrum for *α-ZnSe* can be found in Fig. 5. A similar type of the dependence of the above characteristics on temperature is also valid in the case of *α-ZnS*.

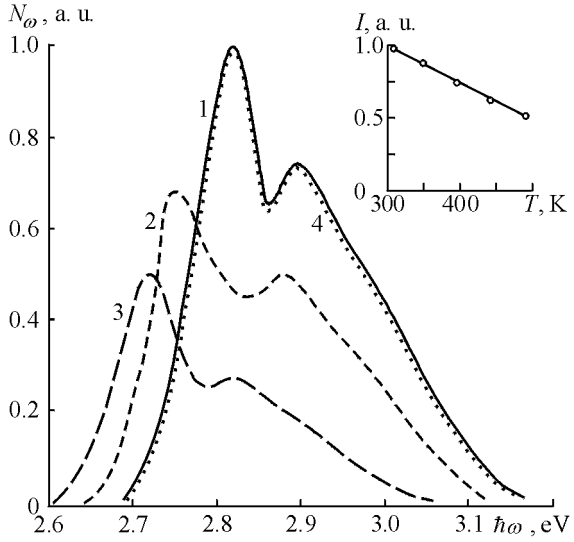


Fig. 6. PL spectra of α-ZnSe at T = 300 K (1), 490 K (2), 490 K (3) and after cooling to T = 300 K (4). On the inset – the dependence of radiation integral intensity versus temperature.

Note that the wavelength in the maximum of the given *α-ZnSe* radiation spectra when heated to 490 K changes its position from λ_m = 0.439 μm to λ_m = 0,455 μm. Such shear value is in good agreement with the experimentally determined temperature coefficient of the change in the band gap of zinc selenide with a hexagonal modification of the crystal lattice, namely $dE/dT = 5.3\cdot10^{-4}$ eV/K. In the case of *α-ZnS* heterolayers the position of maximum changes from λ_m = 0.427 μm to λ_m = 0.440 μm, which corresponds to the well-known value of its shear ratio $dE/dT = 3.8\cdot10^{-4}$ eV/K.

Note that the resulting luminescence spectra of the investigated heterolayers testify to permanence of radiation range 0.37–0.46 μm with a change in higher than room temperature range. It is noteworthy that the temperature mode used for the investigations revealed high temperature stability of the characteristics and parameters of materials obtained.

It was also found that the variation of temperatures in the range of their high values (up to 500 K) leaves unchanged the mechanism that determines the processes of formation of radiation of the grown heterolayers. For *α-ZnS*, the established properties of their luminescence at *T* = 300–500 K indicate the above discussed recombination on donor-acceptor pairs. The corresponding radiation is also formed by the components *A*, *B*, *C* and *D*, which

are determined by certain distances between the partners of the DAP. Note that their respective contribution to the total spectrum does not undergo fundamental (strong) changes. In the case of *α-ZnSe*, radiation is formed by the components that are due to interband recombination of free charge carriers and annihilation of bound excitons. The latter determines the nature of the dominant band with a maximum of 2.82 eV (0.439 μm) at 300 K. The analysis of its properties at $T = 390$ K also revealed the dependence of luminescence intensity I on the excitation level L, which is described by $I \sim L^{1.5}$. The specific shear of maximum $\hbar\omega_m$ with increase in L (see Fig. 3) additionally confirms the exciton nature of the dominant band. It is determined by the annihilation of bound excitons. The possibility of such localization and its temperature stability are due to the properties characteristic of isovalent impurities [2]. This is indicative of the fact that the proposed and employed isovalent substitution method allows not only growing heterolayers of unstable hexagonal crystallographic modification, but also their doping with residual isovalent impurities.

It has been experimentally established that the character of the spectra and the change in radiation intensity are repeated at multiple measurements, Fig. 5. Investigations conducted in the course of two years did not reveal the features distinct from those considered during operation of heterolayers at high temperatures, which indicates not only the stability, but also the repeatability of characteristics. Such high temperature and temporal stability and repeatability of properties, characteristics of parameters can be used not only in the information and communications systems of various types, but also in the manufacture of various sensors, in particular for temperature recording.

4. Conclusions

The use of isovalent substitution method allows obtaining zinc sulfide and zinc selenide heterolayers of hexagonal modification. They are characterized by intensive luminescence with quantum efficiency $\eta = 6$–8% and $\eta = 10$–12%, respectively. The radiation covers the violet optical range $\Delta\lambda = 0.37$–0.46 μm. It is formed by recombination on donor–acceptor pairs for *α-ZnS* and dominant annihilation of bound excitons in isovalent impurities and interband recombination of free charge carriers in the case of *α-ZnSe*. The resulting radiation is characterized by high temperature stability and repeatability of characteristics and parameters, which can be used in various information and communications systems and sensors.

References

[1] Berezovskii M. M., Makhnii V. P., Slyotov M. M.: Influence of an isovalent Cd impurity on photoluminescence of ZnSe crystals. Journal of Applied Spectroscopy 62(6)/1995, 1075–1079.

[2] Fistul V. I.: Doping impurity atoms in semiconductors (state and condition). Fiz.-Mat. Lit., Moscow 2004.

[3] Gorkavenko T. V., Zubkova S. M., Makara V. A., Rusina L. N.: Temperature dependence of the band structure of ZnS, ZnSe, ZnTe, and CdTe wurtzite-type semiconductor compounds. Semiconductors 41(8)/2007, 886–896.

[4] Gribkovskii V. P.: The theory of light absorption and emission in semiconductors. Nauka i Teknika, Minsk 1975.

[5] Khomyak V. V., Slyotov M. M., Shtepliuk I. I., Lashkarev G. V., Slyotov O. M., Marianchuk P. D., Kosolovskiy V. V.: Annealing effect on the near-band edge emission of ZnO. Journal of Physics and Chemistry of Solids 74/2013, 291–297.

[6] Koh Era, Langer D. W.: Luminescence of ZnSe near the band edge under strong laser light excitation. J. Luminescence 1-2/1970, 514–527.

[7] Makhniy V. P., Baranjuk V. Ye., Demich M. V., Melnik V. V., Malimon I. V., Slyotov M. M., Sobistchanskiy B. M., Stets E. V.: Isovalent substitution – a perspective methods of producing heterojunction optoelectronical devices. Proc. SPIE 4425, 2000, 272–276.

[8] Makhniy V. P., Slyotov M. M., Stets E. V., Tkachenko I. V., Gorley V. V., Horley P. P.: Application of modulation spectroscopy for determination of recombination center parameters. Thin Solid Films 450/2004, 222–225.

[9] Peter Y. Yu, Cardona M.: Fundamentals of semiconductors. Physics and materials properties. (4th Ed). Springer-Verlag, Berlin Heidelberg 2010.

[10] Tamargo M.C.: II-VI Semiconductor Materials and their Application. Taylor & Francis, New York 2002.

Prof. Mikhail Slyotov
e-mail: m.slyotov@chnu.edu.ua

Doctor of Science (Physics and Mathematics), professor of Department of Optics, Publishing and Printing (Chernivtsi National University). Area of interests: optical process in wideband II-VI and III-V semiconductors, heterostructures and optical sensors, the influence of hetero- and isovalent impurities on physica; and technical properties on diffusion and epitaxial layers. Author of over 150 scientific papers, including 12 patents.

ORCID ID: 0000-0003-0037-3934

Ph.D. Alexey Slyotov
e-mail: o.slyotov@chnu.edu.ua

Ph. Doctor in Physics and Mathematics, assistant professor of the Department of Electronics and Power Engineering (Chernivtsi National University). Area of interests: investigation of physical properties in wideband II-VI semiconductors, heterostructures and optical sensors, the influence of isovalent impurities on physical and technical properties on diffusion layers. Author of over 50 scientific papers, including 6 patents.

ORCID ID: 0000-0002-2135-9544

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.7982

ANALYSIS OF THE SEPARATION METHODS OF OPTICAL SPECTRA FOR INDIVIDUAL COMPONENTS

Victor Makhniy, Oksana Kinzerska, Illia Senko
Yuriy Fedkovych Chernivtsi National University, PTCSI

Abstract. *A comparative analysis of various methods for the decomposition of broad bands into individual components has been carried out. It is shown that the most universal are the methods of modulation spectroscopy and direct differentiation of conventional spectral characteristics, which, unlike the widely used Alentsev-Fock method, are applicable to spectra of any type (luminescence, transmission, absorption, etc.). The features and capabilities of the direct differentiation method are used to identify the structure in the emission spectra and transmission of ZnSe <Al> single-crystal substrates.*

Keywords: optical spectra, separation methods, luminescence, transmission

ANALIZA METOD SEPARACJI WIDM OPTYCZNYCH DLA INDYWIDUALNYCH SKŁADNIKÓW

Streszczenie. *Przeprowadzono analizę porównawczą różnych metod dekompozycji szerokich pasm na poszczególne składniki. Pokazano, że najbardziej uniwersalne są metody spektroskopii modulacyjnej i bezpośrednie różnicowanie konwencjonalnych charakterystyk spektralnych, które w przeciwieństwie do powszechnie stosowanej metody Alentseva-Focka mają zastosowanie do widm dowolnego rodzaju (luminescencja, transmisja, absorpcja itp.). Cechy i możliwości bezpośredniej metody różnicowania są stosowane do identyfikacji struktury w widmach emisyjnych i transmisyjnych monokrystalicznych składników ZnSe <Al>.*

Słowa kluczowe: widma optyczne, metody separacji, luminescencja, transmisja

Introduction

Measurement and study of optical spectral characteristics are widely used in the emission of the energy structure of the allowed bands and local centers of semiconductors, since they largely determine their physical properties and areas of practical use [3]. At the same time, the composition of local centers of real semiconductor materials is rather complicated due to the fact that they are formed by various types of point defects – intrinsic, impurity, and their associates [7]. This in turn leads to the formation of wide structureless bands in the optical spectra, which causes certain difficulties in interpreting of the experimental results. Note that the interaction of free or bound charge carriers also causes the expansion of narrow bands, the structure of which is difficult to reveal even at cryogenic temperatures.

In this regard, the separation of integral spectra into individual components is an urgent task. Today there are many ways to solve it, however, not every way leads to the desired results when used in real conditions.

Among the best-known ones, we should highlight the Alentsev-Fock method [2], which is based on the deformation of complex luminescence bands when the excitation conditions change and it does not require information on the shape and position of the maxima of their components. Meanwhile, this method is applicable only to the analysis of the emission spectra, and only in case of their noticeable reaction to the level of excitation. In addition, a serious problem is to take into account the overlap of individual bands, especially when there are many of them and the selection of individual horizontal sections becomes almost impossible. These deficiencies are devoid of modulation spectroscopy methods based on transformations of spectra caused by small periodic changes in the physical parameters of a sample or a light wave [4]. The effectiveness of the use of various structural methods requires the fulfillment of conditions, among which we can single out the following [8]. First, usually, for standardization, special samples are required, the creation of which in some cases presents certain difficulties. Secondly, mechanical devices that modulate the parameters of the light beam (amplitude, wavelength, phase or degree of polarization) must ensure a high temporal stability of the modulation frequency. This requirement becomes even more stringent when using higher harmonics of the fundamental modulation frequencies. This is due to the fact that the application of the latest strong effects, as well as another - the deterioration of the signal-to-noise ratio after a sharp decrease in the level of the useful signal.

Many proposed physical problems that can be reliably measured at sufficiently high signal levels. In particular, by this method the author [1] with high accuracy determined the width of the forbidden zone E_g of a series of indirect-gap semiconductors –Si and SiGe solid solutions. Meanwhile, despite the fact that in this work the differential method was applied only to the spectra of the edge luminescence, it can be very promising for analyzing the spectral characteristics of other types (transmission, absorption, photoconductivity, etc.), and more wide range of energies.

In this paper, we study the possibility of applying of the above mentioned methods to identify the structure of complex bands in the optical spectra of ZnSe<Al> and ZnSe <Al>:Gd semiconductor substrates.

1. Samples and research methods

The choice of objects for research is due to several reasons. First, zinc selenide for many decades has continued to remain one of the most promising materials for functional electronics. Secondly, Al is widely used as an alloying element for ZnSe, since, unlike many other donor impurities, it does not exhibit amphotericity and does not form deep centers. Third, the dependence of the form of a wide low-energy luminescence band of ZnSe <Al> crystals on the level of excitation allows the use of the Alentsev-Fock method for its separation. Fourth, studies show that this band is also present in the spectra of pure ZnSe crystals annealed in a zinc melt containing 0.01–5 mol%. Al [11] At the same time, its shape depends on the aluminum content, annealing time, cooling modes and the "biography" of the original crystal, which indicates the redistribution of the roles of the recombination centers included in an ensemble of ZnSe<Al> crystals, whose composition apparently remains unchanged.

Finally, a feature of the ZnSe<Al>:Gd samples is that, while maintaining a wide low-energy band, a narrower band of edge radiation also appears in them [9]. This allows a comparative analysis of the results obtained using the same separation methods for spectra of different nature and complexity.

Spectral studies were carried out on substrates cut from a bulk crystal grown from the melt stoichiometric composition by the Bridgman method with an additive of ~0.1 mol % Al. The doping technology of zinc selenide substrates with rare-earth impurities, including Gd, is described in detail [10]. Optical transmission T_ω and luminescence N_ω spectra were measured at room temperature on a universal installation containing

an MDG-23 monochromator, a PEM-79 photodetector, and radiation sources – a halogen lamp and an N_2-laser. Their differentiation was carried out on a computer using a special program.

2. Results and discussion

First, we will carry out a general analysis of the usual photoluminescence spectrum of ZnSe <Al> crystals, which is shown in Fig. 1 at 300 K.

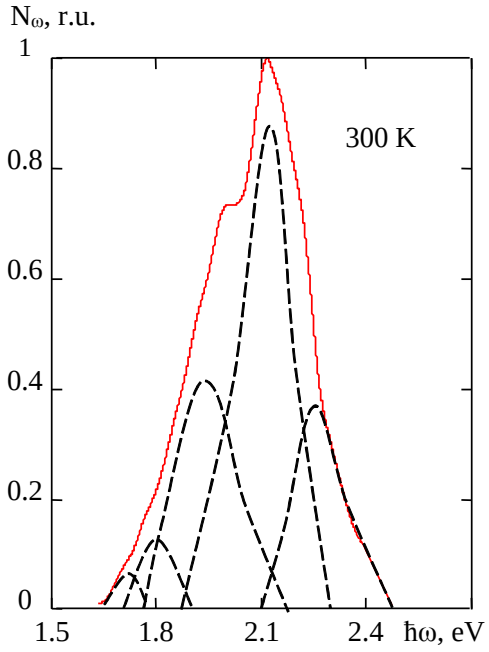


Fig. 1. The ordinary photoluminescence spectrum of ZnSe <Al> crystals is a solid line, the dotted line is the individual bands calculated by the Alentsev-Fock method

Its abnormally large half width ($\Delta\hbar\omega_{1/2} \simeq 0.45$ eV) as well as the presence of a clear maximum and a bend in the energy range of 2.0 and 2.1 eV, respectively, indicate that this band is not elementary. In this regard, there is a need for its decomposition into components, which was carried out according to the Alentsev-Fock method [2]. As a result of this procedure, it was possible to distinguish 5 bands with maxima near 1.70, 1.80, 1.95, 2.10 and 2.25 eV, which are shown in Fig. 1 (dotted line). Meanwhile, a detailed analysis showed that the half-width of all selected bands is about the same and at 0.20–0.22 eV, which significantly exceeds $kT \simeq 0.026$ eV.

So great values $\Delta\hbar\omega_{1/2}$ of these components indicate a strong electron-phonon interaction of the centers that are involved in their formation. In this case, the shape of the bands can be described using the ratio obtained in [5, 6]. So great values of these components indicate a strong electron-phonon interaction of the centers that are involved in their formation. In this case, the shape of the bands can be described using the ratio obtained in [5, 6]

$$I(y) = I_0\theta^{1/4} e^{\frac{-E_0}{kT(1-\delta)}} y\psi[\frac{3}{4},\frac{1}{2},\frac{(1-y)^2}{\theta}] e^{-\frac{(1-y-2\delta)^2}{\theta}} \quad (1)$$

where I_0 – constant; $y = \hbar\omega / E_0$; E_0 – optical energy of ionization or neutralization of the center of the glow; $\theta = (\frac{a^*\hbar\Omega}{E_0})^2 cth(\frac{\hbar\Omega}{2kT})$; $\delta = \theta E_0 / 4kT$; a^* – electron-phonon interaction constant; $\hbar\Omega$ – oscillation energy localized at the center; $\psi(a,c,x)$ – degenerate hypergeometric function. Due to the complexity of the expression (1), in practice, its simplified variants are used much more often, they are discussed below.

Formula (1) shows that the shape of the luminescence band is determined mainly by the exponent $e^{-\frac{(1-y-2\delta)^2}{\theta}}$ and is almost gaussian.

The condition is used to determine the position of the maximum emission band

$$\left.\frac{dI(y)}{dy}\right|_{y=y_m} = 0$$

Since the exponent in (1) determines the dependence $I(y)$, for a successful solution of the problem, it suffices to take the derivative only of the exponential part of the equation, that is, the exponent $\left(e^{-\frac{(1-y-2\delta)^2}{\theta}}\right)' = 0$ At the same time we get $y_m \simeq (1-2\delta)$ or taking into account the ratio for δ and θ will have

$$h\upsilon_m = E_0\left[1 - \frac{E_0}{2kT}\left(\frac{a^*\hbar\Omega}{E_0}\right) cth\left(\frac{\hbar\Omega}{2kT}\right)\right]. \quad (2)$$

at temperature, when $\hbar\Omega \ll 2kT$, will get

$$\hbar\omega_m \simeq E_0 - a^{*2}\hbar\Omega. \quad (3)$$

From (3) it is clear that the derivatives $a^{*2}\cdot\hbar\Omega$ are equal to energy which is actually Stoks displacement Δ.

The width of the emission band at the level $1/e$ will be determined by the condition $\ln(I_m / I_{1/e}) = 1$. Given only the exponential factor in (1), we obtain

$$(y_{1/e} - y_m)^2 = \theta. \quad (4)$$

On the other side,

$$y_{1/e} - y_m = \frac{E_{1/e}}{2E_0}. \quad (5)$$

Equating the right sides of (5) and (6), we get $E_{1/e} = 2E_0\theta^{1/2}$, and given the ratio for θ

$$E_{1/e} = 2a^*\hbar\Omega\left[cth\left(\frac{\hbar\Omega}{2kT}\right)\right]^{1/2}. \quad (6)$$

at temperature, when $\hbar\Omega \ll 2kT$,

$$E_{1/e} = 2a^*(2kT\hbar\Omega)^{1/2}$$

It should be noted that (2) and (6) are approximate, since in the corresponding calculations we took into account only the exponential multiplier of the dependence $I(y)$.

The positions of its maximum

$$\hbar\omega_m \simeq E_0(1-2\delta) \quad (7)$$

Using expert value $E_{1/e} - \hbar\omega_m$ from formulas (2) and (6) it is easy to determine the parameters θ and E_0, which also allow to calculate the Frank-Condon displacement

$$\Delta = \frac{a^{*2}\hbar\Omega}{2} \quad (8)$$

and the activation energy of the glow centers

$$E_a = E_g - \hbar\omega_m - \Delta. \quad (9)$$

Table 1. Parameters of the centers responsible for the "elementary" luminescence bands of ZnSe <Al> crystals

ħω, eV	E_0, eV	Δ, eV	E_a, eV
2.25	2.54	0.18	0.32
2.10	2.46	0.18	0.42
1.95	2.30	0.18	0.56
1.80	2.16	0.18	0.70
1.70	2.0	0.18	0.86

Values E_0, Δ and E_a, characterizing centers responsible for the formation of bands with maximum $\hbar\omega_m$ at 300 K, are shown in Table 1. Let's pay attention to the same values of the Franck-Condon offset for all the above bands, which is a consequence of the proximity of their widths at the level $1/e$.

Thus, the application of the Alentsev-Fock method together with the Kopylov-Pikhtin model made it possible not only to isolate individual components in a wide luminescence band of ZnSe <Al> crystals, but also to determine the main parameters of recombination centers. On the other hand, the "elementary" bands calculated by the Alentsev-Fock method are structureless symmetric gaussian curves (dotted line in Fig. 1), which do not always correspond to reality. In most cases, their shape is usually asymmetric due to the interaction of recombination centers with phonons. This should lead to the appearance of a certain structure on the "wings" of individual bands. The latter is simply inevitable in our case, since the Kopylov-Pikhtin model used to calculate the model is based on the assumption of a strong electron-phonon interaction.

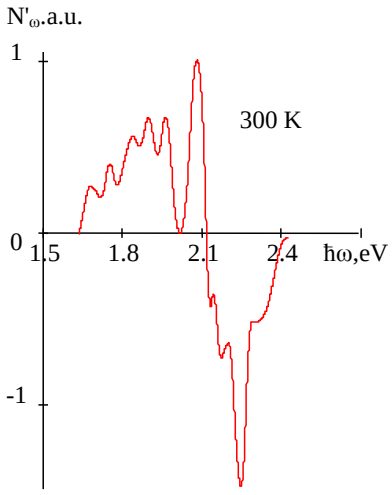


Fig. 2. The first derivative of the usual photoluminescence spectrum of ZnSe <Al> crystals

One of the ways to solve this problem is to use the method of direct differentiation mentioned earlier, the result of which is applied to the usual luminescence spectrum (Fig. 1), it is shown in Fig. 2. Despite the rather complex shape of the curve, its analysis allows us to identify some features of the "elementary" bands. So, in particular, the intersection point of the curve with abscissa axis corresponds to the maximum of the normal spectrum N'_ω, and minima - the energy position of the maxima of the corresponding "elementary" bands, Table 1. Equidistant peaks on the curve N'_ω located to the left of the minima near 1.95 and 2.25 eV are caused by scattering by LO-phonons, since the distance between them is ~ 30 meV and is consistent with the energy of the longitudinal optical phonon of ZnSe [3, 11]. The interaction of these centers with LO-phonons is also confirmed by the presence of sections with an equidistant structure on the curve (Fig. 4) obtained by differentiating the usual optical transmission spectrum, Fig. 3. The result is in favor of the wider possibilities of differential methods (which include modulation [4, 8]) as compared with the Alentsev-Fock method, which is not applicable to optical transmission spectra at all.

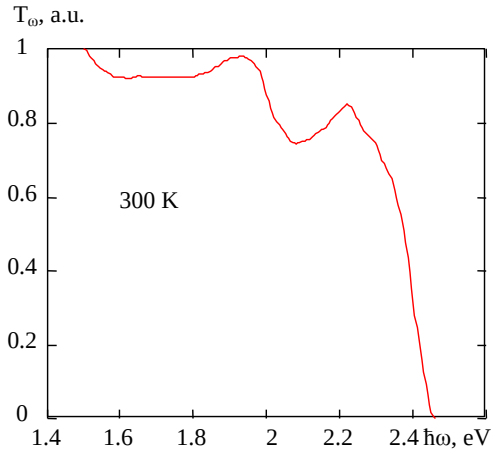


Fig. 3. The usual transmission spectrum of single-crystal ZnSe <Al> substrates with a thickness of 1 mm

We also pay attention to the presence in the spectra T'_ω and another intense band with a peak near 2.4 eV, which is absent on the curve N'_ω, Fig. 2. Meanwhile, the clarification of the nature of the centers responsible for its appearance requires additional research beyond the scope of this work.

The results obtained above confirm the legitimacy and perspectivity of applying the method of direct differentiation of broad optical spectra for separation into components. At the same time, the efficiency of its use decreases with the expansion of the studied spectral range, due to the substantial complication of the energy structure of the obtained curves and the overlap of individual bands. The transition from the first to higher derivatives, although it increases the separation capacity, further aggravates the situation due to the appearance of additional structural elements.

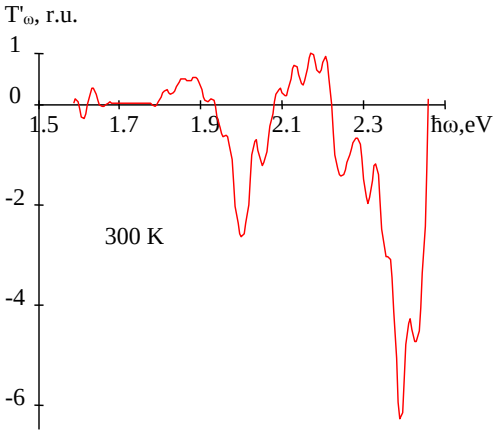


Fig. 4. The first derivative of the usual transmission spectrum of the ZnSe <Al> substrates shown in Fig. 3

Considering the above mentioned, it is reasonable to assume that the direct differentiation method is more appropriate to use for the analysis of relatively narrow bands, as which we chose the edge luminescence B band of the substrates ZnSe<Al>:Gd, Fig. 5.

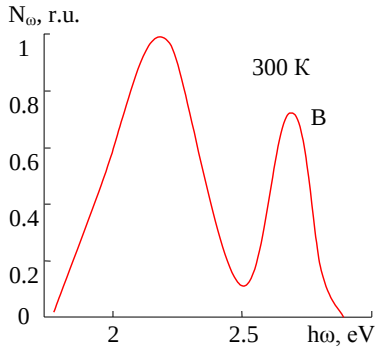


Fig. 5. The usual photoluminescence spectrum of ZnSe <Al>: Gd crystals

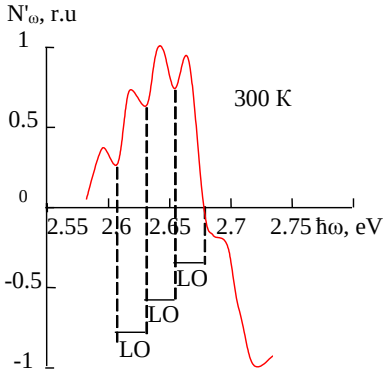


Fig. 6. The first derivative of the edge band of the usual photoluminescence spectrum of ZnSe <Al>: Gd crystals

As a result of its differentiation, a complex curve is obtained (Fig. 6), on which a number of equidistant minima are observed with an energy distance corresponding to the energy of the LO-phonon ZnSe. Intersection of N'_ω curve with the abscissa axis at 2.68 eV, corresponds to the maximum emission of a free exciton, and the bend at 2.7 eV corresponds to the band gap of ZnSe at 300 K [3, 11]. We note that the edge luminescence spectra of the ZnSe<Al>:Gd samples measured in [9] by the modulation method, unlike their direct differentiation (Fig. 6), do not exhibit phonon replications. This demonstrates the advantage of using the differential method for detecting and isolating the fine structure of optical spectra.

References

[1] Emelianov A. M.: Differential method for analyzing the luminescence spectra of semiconductors. Fizika i tehnika polyprovodnikov 44(9)/2010, 1170–1175.
[2] Fok M. V.: Decomposition of complex spectra into individual bands using the Alentsev method. Trydu FIAN SSUR. 59/1972, 3–24.
[3] Gavrilenko V. I., Grehov A. M., Korbytyak D. V., Litovchenko V. G.: Optical properties of semiconductors. Directory naykova dymka, Kiev 1987.
[4] Kardona M.: Modulation spectroscopy. Mir, Moscov 1972.
[5] Kopulov A. A., Pihtin A. I.: The effect of temperature on the optical absorption spectra of deep centers in semiconductors. Fizika tverdogo tela 16/1974, 1837–1842.
[6] Kopulov A. A., Pihtin A. I.: The shape of the absorption and luminescence spectra at deep centers in semiconductors (oxygen in gallium phosphide). Fizika i tehnika polyprovodnikov 8/1974, 2390–2403.
[7] Makhniy V. P.: Physics and chemistry of point defects in semiconductors. CHNU, Chernivtsi 2014.
[8] Makhniy V. P.: Principles and methods of modulation spectroscopy. Ruta, Chernivtsi 2001.
[9] Makniy V. P., Vakhnyak N. D., Kinzerska O. V., Senko I. M.: Luminescence of crystals ZnSe<Al>:Gd. Semiconductor Physics, Quantum Electronics and Optoelectronics 21(1)/ 2018, 80–82.
[10] Patent 107292 UA, MPK C30B/2500 (20.16.01) C30B 31/00 Method of doping ZnSe<Te> crystals with rare earth elements. Makhniy V. P., Kinzerska O. V., Senko I. M, Yuriy Fedkovych Chernivtsi National University, 2015, 12805, stated 24.12.2015; posted 25.05.2016, Bul. No. 10, 2016.
[11] Vaksman Yu.F.: Luminescence of single crystals of zinc selenide and radiating structures based on them. Diss. of doc. phys. and math. science 1993.

Prof. Viktor Makhniy
e-mail: vpmakhniy@gmail.com

Viktor P. Makhniy, born in 1946, defended his Doctoral Dissertation in Physics and Mathematics in 1992 and became full professor in 1994. Professor at the Department of Optics, Publishing and Printing of Yuri Fedkovych Chernivtsi National University, Ukraine. Authored over 450 publications, 40 patents, 9 textbooks and wrote 2 monographies. The area of his scientific interests includes technology, physics and applications of wide-band semiconductor compounds and devices based on them.

ORCID ID: 0000-0002-7073-6860

Ph.D. Oksana Kinzerska
e-mail: oksanakinzerska@gmail.com

Oksana V. Kinzerska, born in 1980, defended her PhD dissertation in Physics and Mathematics in 2012. Assistant Professor at the Department of Physics of Semiconductors and Nanostructures of Yuri Fedkovych Chernivtsi National University, Ukraine. Authored 44 publications, 3 patents, 2 methodical recommendations. The area of her scientific interests includes technology, physics and the application of wide-band semiconductor compounds of semi-magnetic and rare-earth elements.

ORCID ID: 0000-0003-1722-1216

M.Sc. Illia Senko
e-mail: senko.illia14@gmail.com

Illia M. Senko, born in 1993, PhD student at the Department of Semiconductors and Nanostructures of Yuri Fedkovych Chernivtsi National University, Ukraine. Authored 11 publication and 1 patent. The area of his scientific interests is the influence of rare-earth elements on physical properties of zinc selenide.

ORCID ID: 0000-0002-4665-6879

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8013

FEATURES OF THE MANAGEMENT OF DATA ENCRYPTION KEYS IN THE CLOUD STORAGE MS SQL AZURE

Olexander Beley
Lviv Polytechnic National University, Institute of Computer Science and Information Technologies / Department of Computer-Aided Design

Abstract. *The main principles of data security and access organization in the Microsoft Azure cloud storage are considered. A role of hierarchy and access keys are presented. We describe the setup and the use of their keys (BYOK) for transparent data encryption (TDE) using Azure Key Vault keyring.*

Keywords: communication equipment, data communication, cyberspace, data transfer

CECHY ZARZĄDZANIA KLUCZAMI SZYFROWANIA DANYCH PRZECHOWYWANYCH W CHMURZE MS SQL AZURE

Streszczenie. *Uwzględniono główne zasady bezpieczeństwa danych i organizacji dostępu w chmurze Microsoft Azure. Przedstawiono zagadnienia hierarchii ról i kluczy dostępu. Zostały opisane dostosowywanie i używanie własnych kluczy (BYOK) do przezroczystego szyfrowania danych (TDE) przy użyciu magazynu kluczy platformy Azure.*

Słowa kluczowe: sprzęt komunikacyjny, komunikacja danych, cyberprzestrzeń, transfer danych

Introduction

With the spread of the Internet, the technology of data processing has undergone significant changes. Previously, a computer without any software was just a pile of scrap metal. With the appearance of cloud technologies, even a simple mobile phone with the access to the Internet can help solve complex tasks. Cloud technology is a browser-based mailbox interface with the ability to create and edit online office documents, the solution to complex mathematical calculations, for which the power of one personal computer is not enough.

Currently the cloud computing technology is one big concept that includes many different concepts: software, infrastructure, platform, data, and workplace. The most important feature of cloud technologies is to meet the needs of users who need remote data processing.

Of course, for an average computer user, cloud-based technology is not something you can do nothing without. However, cloud computing is essential for business. Its main advantage is the ability to save on expensive software. After all, you do not have to install expensive office packages and specialized data processing software on each computer. In addition, cloud computing can allow all employees of the enterprise to use, in general, only one operating system, with the access to their workplaces through much cheaper terminals.

However, the concept of cloud technology is a subject to considerable criticism. The main drawback is the security, because not everyone considers it to be reliable to store personal data on a remote server. However, cloud computing has significant prospects, since Microsoft, Apple and Google almost simultaneously began to implement cloud-based technologies in their designs and are not going to abandon them before too long.

Flexibility, scalability and cost-effectiveness contribute to cloud-computing enterprises. This is the answer to the constantly changing economic, financial and technical conditions in which modern enterprises have to work. Constant changes require new ways of thinking, working and doing things. In this new reality, the development of a hybrid cloud model is based. Successful businesses, from small businesses to multinational corporations, recognize the importance of an information system that provides secure data access and effective administration. Cloud-based systems need to be quickly rebuilt to provide cost-effective efficiency with a positive return on investment. This combination of requirements is best served by the various IT services offered in the hybrid cloud.

The cloud storage solution for modern applications that provides stability, availability and scalability to meet customer needs is the MS Azure repository. The Azure repository provides many different security features, such as: the storage account can be protected by RBAC and Azure Active Directory; the data transmitted between the application and Azure can be protected by encrypting the client, HTTPS or SMB; automatic data encryption can be configured when writing to the Azure repository using the "Encrypting the repository" function; for OS drives and data disks used in virtual machines, the Azure disk encryption can be configured; delegated access to data objects in the Azure repository service can be provided with signed URLs; also you can use the analytics to track the authentication method used when accessing the repository.

The security features of Microsoft Azure repository include storage keys, data encryption during data transfer, inaccessible data encryption, and repository analytics. We can protect your storage account with role-based access control. Restricting access according to security principles (the principle of providing access only in those cases and to the extent that knowledge of such information will be necessary, as well as the principle of minimum rights) is extremely important for organizations that need to apply security policies for access to data. These rights are granted by assigning an appropriate role to RBAC groups and applications for a specific area. We can assign users the rights, for example, the rights to a member of the storage account, with the built-in RBAC roles.

The public signature provides delegated access to resources in your storage profile. This means that the client can get a limited right to work with objects in your storage profile for a certain period of time and with a certain set of permissions. We may grant these limited rights without notifying access keys to your account. To access the repository resources with SAS, client just has to pass SAS to the appropriate constructor or method, which contains all the information needed to access the authenticated storage repository in its query parameters.

The Azure data repository for data protection includes the use of: transport layer encryption for data transfer to or from the Azure storage service; SMB 3.0 encryption for Azure file resources; client-side encryption, which encrypts the data before being transferred to the repository and decrypts it after the data is received from the repository service.

1. Basic principles of security management in the Microsoft Azure Data Warehouse

The main recommendations for protecting and encrypting data in Azure include: Multi-Factor Authentication; role-based access control (RBAC); encryption of Azure virtual machines; use of hardware security models; safe management of workstations; SQL data encryption; data protection during transmission; application of file-level encryption.

Verifying the authenticity of the user is the first step in providing access to and management of data in Microsoft Azure. To do this, the Multi-Factor Authentication (MFA) Azure method

is used with additional tools other than the user name and password. This method helps to protect access to data and applications without compromising the user's login process. By incorporating Azure MFA for our users, we add a second level of security for signing in and transactions. A transaction may refer to a document hosted on a file server or in SharePoint Online. Azure MFA also minimizes the risk that third parties will be able to access data using compromised credentials. Enterprises that do not use this additional level of protection are more vulnerable to attacks through stealing credentials, which may lead to compromising data.

The main key store solution for Azure services is the Azure Key Vault, which provides general key management capabilities. Keys are stored by users and services. Azure Key Vault supports the creation of user keys or the import of custom keys using the scripts used by our encryption keys.

When implementing encryption of inactive data, several encryption keys are used. Asymmetric encryption can be used to provide the credentials and authentication necessary for key management and access to them. Symmetric encryption is more efficient for mass encryption and decryption, which ensures much more reliable encryption and better performance. Restricting the use of the encryption key reduces the risk of its damage, as well as the cost of re-encrypting if it is necessary to replace the key.

The AES256 (DEK) symmetric key is used to encrypt a partition or data block, which may include many sections and many data encryption keys. Encrypting each data block with another key creates additional complexity for executing attacks on encrypted data. When creating a new DEK key, re-encrypting this key requires only data in its associated block.

The Asymmetric Encryption Key (KEK) is used to encrypt data keys, the use of which allows you to encrypt the data encryption keys directly and manage them. Since KEK keys are needed to decrypt the DEK keys, it can actually be considered as the single point to control the DEK keys. Data encryption keys encrypted using key encryption keys are stored separately, and only the entity that has access to key encryption keys may receive any encryption key encrypted with the KEK key.

The client's encryption model is the encryption process that we perform with the service program in Azure or an application that works in the user's data processing center. When using this encryption model, we will receive encrypted data in the form of a large binary code without the possibility of decrypting it and also without the access to encryption keys. In this model, the keys are managed by the appropriate application and this process is opaque for the Azure service.

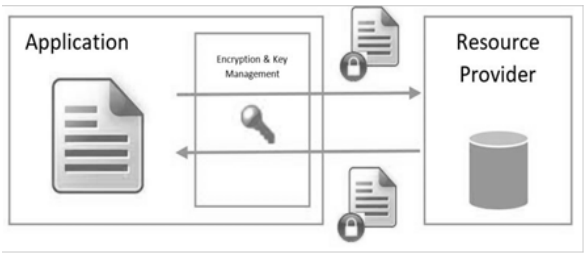


Fig. 1 Model encryption client

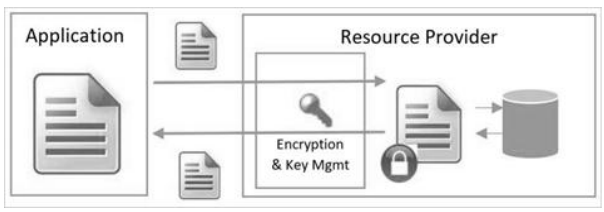


Fig. 2. Models for managing encryption keys on the server side

Server-side encryption models are Azure encryption. In this model, we perform encryption and decryption operations. In the Azure repository, you can retrieve data using plain text operations, and then perform internal encryption and decryption. We may use Microsoft encryption keys or our own.

If you want to encrypt inactive data and manage encryption keys, you can use server-side encryption with the help of user-managed keys in the repository. Some services can store root key encryption keys in Azure Key Vault and store the encryption key of encrypted data in the internal repository along with the data. In this scenario, we can transfer your own keys (BYOK) to the key store or create keys to use them to encrypt the desired resources. In doing so, we perform encryption and decryption operations using the custom key as the root key for all encryption operations.

When Azure disks are encrypted, the Azure key store is used. Thanks to this, we can manage the keys and secrets of disk encryption as a part of the key store subscription. Encryption is performed for all inactive data on virtual drives in the Azure storage service. The Azure key store should be used for key audit and policy use. There are many risks associated with the lack of suitable secret key protection tools used to encrypt data. If intruders have the access to secret keys, they will be able to decrypt data and gain access to confidential information.

To encrypt data stored on disk using a secret key, secure key creation, storage, access control and encryption keys are required. Although some points may differ, the implementation of inactive data encryption can be described using the concepts illustrated in the following scheme (see Fig. 1).

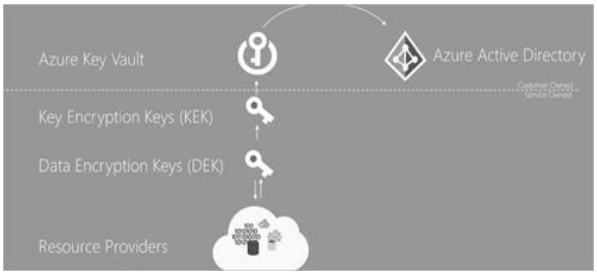


Fig. 3. Azure key storage

Since most attacks are aimed at the user, one of the main targets of the attackers is the end point. If the attacker damages the end point, they will be able to use user credentials to access the organization's data. Most end-to-end attacks occur because users are administrators on their local workstations.

All Azure storage services (BLOB storage, queue storage, spreadsheet and Azure file service) support server-side inaccessible encryption, and some services support customer-managed client-side encryption keys. The Azure BLOB storage and the Azure file service also support the 2048-bit RSA-controlled RSA keys in Azure Key Vault. When using client-side encryption, data is encrypted and transmitted as an encrypted large binary object.

Currently, the SQL Azure database supports inaccessible data encryption in client-side and on Microsoft managed services. Server-side encryption support is currently provided using SQL functions that implement transparent data encryption. Inactive data encryption can be enabled at the database level and at the server level. The SQL Azure database supports the client-managed 2048-bit RSA key in Azure Key Vault.

Azure role-based access control (RBAC) is used to assign users, groups, and access rights within a particular area. A role can be a subscription, a group of resources or a separate resource. We can assign users rights using the built-in RBAC roles in Azure: the role of "Member of the storage account", the role of "Member of the classical storage account", the role of "Participant of the virtual machine".

2. Transparent encryption of data using its own key

Using Your Own Keys (BYOK) for Transparent Data Encryption (TDE) allows you to encrypt the database encryption key (DEK) using an asymmetric key called protector TDE. With TDE keys, we can manage and store data in the Azure Key Vault cloud system for managing external keys. The TDE encryption key stored on the bootstrap database page is encrypted and decrypted with the TDE fuse. The protector TDE is stored in the key store of Azure Key Vault. If the server's access to the key store is cancelled, it will lose the ability to decrypt and read the encrypted database. Protector TDE is configured at the logical server level with the inheritance of all databases that are associated with this server.

TDE with BYOK support allows us to provide: higher transparency and precision with the ability to independently control the TDE fuse; centralized protector's TDE control (with other keys and permissions for all other Azure services) due to their location in Key Vault; division of responsibilities for managing keys and data in an organization; increasing customer confidence, as the Key Vault principle does not allow Microsoft employees to see or receive encryption keys; key change support.

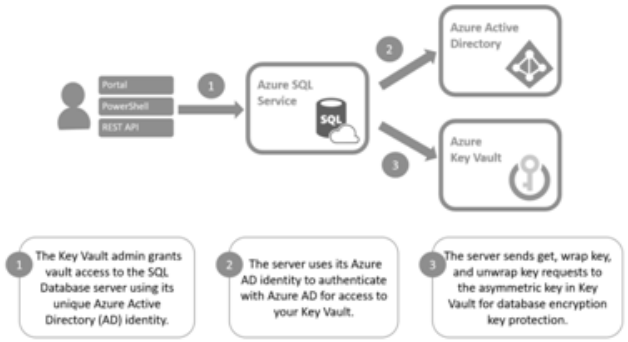


Fig. 4. The work of TDE with support of BYOK

When we first set up TDE to use a key logger with Key Vault, the server sends Key Vault encryption keys from all databases that support TDE to create a key packet request. Key Vault returns the encryption key for the encrypted database and this key is stored in the user's database. Saved in the Azure Key Vault protector, TDE never leaves the Azure Key Vault. A logical server can only send key operation requests to the TDE protector key material within Key Vault, and never accesses or caches the TDE protector. The Key Vault Administrator has the right to revoke Key Vault permissions of the server at any point, in which case all connections to the server are cut off.

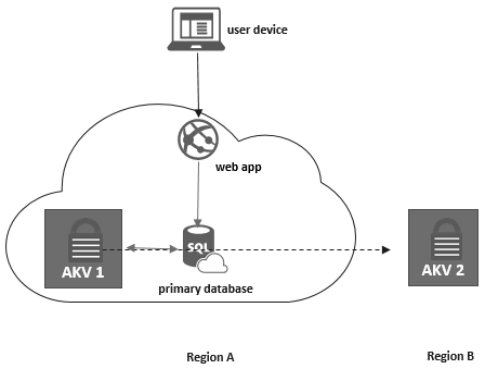


Fig. 5. Setting up a geographic emergency recovery for Azure Key Vault

The Azure Key Vault can be configured in a variety of ways: for a stand-alone database or logical server without geo-replication; for the database or logical server, configured denial handling groups or geo-replicated. At the same time, for each geo-replicated copy, a local Azure Key Vault is required within the refusal group to properly process geographically distributed bounces. In the first case, the high level of accessibility for a database and logical server without geo-replication can be configured by creating two different keystrokes for the server in two different regions, which will store the same key material. To do this, we create a protector TDE in the primary repository located in the same region with the logical server and a clone that is a key to the key store in another Azure region. Now the server will be able to use a secondary repository if primary problems arise during the operation of the database.

For geo-replicated SQL Azure databases, we create the appropriate Azure Key Vault configuration: one source database with a repository and one repository database for the repository in the same region; there must be at least one and no more than four recipient databases; secondary replica databases recipient (threading) are not supported.

For a new deployment of a SQL Server with a geographic disaster recovery, we will need to do the following: create two logical SQL servers in the same regions that previously created the repositories; select the TDE area for the logical server, and then for each SQL logical server select AKV in the same region and the key that can now be used as the protector of TDE. Each server will use a local copy of the protector's TDE. When we perform this operation on the portal, we will receive an AppID for the logical SQL server that allows you to assign a logical permission from the SQL server to access the key repository. After the active geo-replicate action is performed, the recipient database will be created.

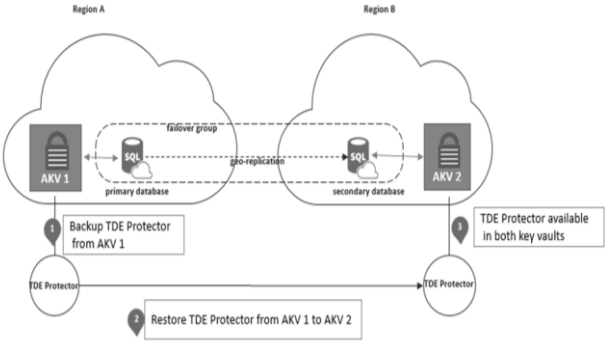


Fig. 6. Deploying SQL-Server with geographical disaster recovery using Azure Key Vault

In a script for SQL Geographic Disaster Recovery Database, you must create and maintain two Azure Key Vault repositories with identical content in the same regions that will be used for georeplication of the SQL database before turning on TDE with the keys stored in Azure Key Vault and the managed client. "Identical Content" here means that both key repositories must contain copies of one protector TDE so that both servers have access to the protector TDE used for all databases. We synchronize both key holders, i.e., we place identical copies of protector TDE in them after changing the keys; maintain the old versions of the keys that were used for log files or backups; preserve the same key properties for all subsequent TDE fuses; maintain the same permissions for accessing SQL servers in the repositories.

To restore a backup copy encrypted with a protector TDE from Key Vault, you need to make sure that the key is in the original repository with the same name. When the protector TDE changes for the database, the old database backups are not updated to the latest version of the protector TDE. We recommend you to keep all older versions of protector TDE in Key Vault to restore database backups. Saved backups of logs remain encrypted using the original TDE encoder, even if the TDE protector has changed and the database already uses the new TDE fuse. Both the key will be required to restore the database. If the log file uses a TDE fuse stored in the Azure Key Vault, this key will be required during the recovery, even if the database is transferred to the TDE fired service.

Supporting the creation of own keys allows us to independently manage the keys for TDE and establish who and when can access them. The cloud-based external key management system Azure Key Vault has become the first key management service in which transparent data encryption is integrated with the support for the creation of its own keys. Support for creating your own keys allows you to protect the key encryption database asymmetric key stored in Key Vault. This asymmetric key never leaves Key Vault. If the server has permission to access the key repository, the server sends requests for basic key operations in the corresponding Key Vault. The asymmetric key is configured at the server level and inherited by all databases on that server.

We can manage such key management tasks as changing a key and setting permissions for key repositories. You can also delete the keys and enable auditing for all encryption keys. Key Vault provides centralized key management and uses strict tracking with hardware security modules. Key Vault supports division of key management tasks and data to ensure compliance with regulatory requirements.

For operations within Azure, the database does not need to be decrypted. The TDE parameters are transmitted transparently from the source database to the recipient database. This applies to all the following operations: heavens; recovery at a certain point in time through the self-service interface; restore of a remote database; active geo-replication; creation of a copy of the database.

To configure TDE on the Azure portal, we connect on behalf of the owner, member, or SQL Azure security administrator. Transparent data encryption is configurable at the database level. To include TDE in the database, we enter the Azure portal with an administrator account or a member of Azure. The settings for transparent data encryption are displayed in the database user information section. By default, it is managed by a transparent data encryption service. For a server that has a database, the TDE certificate is created automatically.

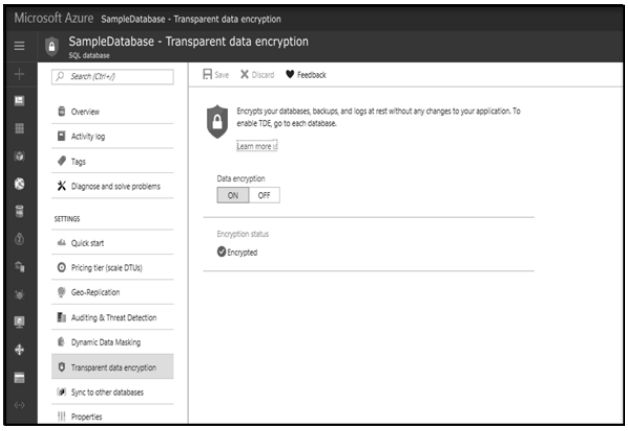


Fig. 7. Set up the protector TDE on the Azure portal

The primary key of transparent data encryption (also called transparent data protection protector) is set at the server level. To use TDE to support our own keys creation and to protect databases using the key stored in Key Vault, we use TDE parameters for our server.

Thanks to BYOK support, users can independently manage their keys and perform operations such as changing keys, setting keys per key retention, removing keys, turning on audits and reporting through all protectors' TDE through Azure Key Vault. Key Vault supports centralized key management, utilizes carefully crafted hardware security modules (HSM), and allows you to share key management and data management responsibilities to ensure compliance with legal requirements.

Conclusion

A complete solution for encrypting inactive data assumes that data is never stored in unencrypted form. During use, when the server loads data into memory, data can be stored locally in a variety of ways, including the Windows paging file, a crash dump, and logging that the application can perform. To ensure that this data is encrypted during storage, IaaS applications can use Azure Disk Encryption on the Windows Azure IaaS virtual machine and virtual disk.

In IaaS applications, we often have to encrypt Azure disks and inactive data encryption settings provided by any Azure service used. In some cases, such as with non-standard encryption requirements or using non-Azure storage, the IaaS application developer may need to implement inactive data encryption on its own. IaaS solution developers can provide better integration with Azure management and meet user expectations by using specific components of Azure. We use the Azure Key Vault service to provide secure key storage and to provide our users with consistent key management options for most of the services of the Azure platform. In addition, our solutions must use the credentials of Azure-managed services to provide service accounts with access to encryption keys.

Protecting user data stored in Azure services is of particular importance to Microsoft. All services hosted in Azure are committed to providing inactive data encryption options. Basic services, such as Azure storage, Azure SQL Database and key analytics services already provide inactive data encryption options. Some of these services support user-managed keys and client-side encryption, as well as service-managed keys and encryption. Microsoft Azure services are constantly improving the availability of inactive data encryption, and new parameters will appear in the preliminary and public versions in the near future.

References

[1] Amies Alex H. S., Qiang Guo Tong, Guo Ning Liu.: Developing and Hosting Applications on the Cloud. IBM Press, 2012.
[2] Armbrust M., et al.: A view of cloud computing. Commun. ACM 53/2010, 50–58.
[3] Armbrust M., et al.: Above the Clouds: A Berkeley View of Cloud Computing, 2009.
[4] Buyya R., et al.: Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Gener. Comput. Syst. 25/2009, 599–616.
[5] Casanova H.: Simgrid: a toolkit for the simulation of application scheduling. Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on, 2001, 430–437.
[6] Dean J., Ghemawat S.: MapReduce: Simplified Data Processing on Large Clusters. ACM OSDI, Dec. 2004.
[7] Hamdaqa Mohammad L. T., Ladan T.: A Reference Model for Developing Cloud Applications. Presented at the CLOSER, 2011.
[8] Krauter K., et al.: A taxonomy and survey of grid resource management systems for distributed computing. Software: Practice and Experience 32/2002, 135–164.
[9] Qaisar E. J.: Introduction to cloud computing for developers: Key concepts, the players and their offerings. Information Technology Professional Conference (TCF Pro IT), IEEE TCF, 2012, 1–6.
[10] Quiroz A., et al.: Towards autonomic workload provisioning for enterprise Grids and clouds. 10th IEEE/ACM International Conference on Grid Computing, 2009, 50–57.
[11] Rajkumar Buyya R. N. C.: Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities. Presented at the 7th High Performance Computing and Simulation Conference (HPCS), Leipzig, Germany.
[12] Trunfio P., et al.: Peer-to-Peer resource discovery in Grids: Models and systems. Future Gener. Comput. Syst. 23/2007, 864–878.
[13] Voorsluys W., et al: Introduction to Cloud Computing. Cloud Computing. John Wiley & Sons, Inc., 2011.

Ph.D. Beley Olexander Igorovych
e-mail: tiger_oles@i.ua, Oleksandr.I.Belei@lpnu.ua

Since 1998 he has worked as a software engineer and system administrator for various companies. Since 2000 he has been engaged in information technology, design, programming, database development and web pages, modeling and control of complex processes and systems, mining systems and forecasting data
He is the author of more than 100 scientific works, co-author of 6 monographs, 20 teaching and methodological works, 3 manuals. He participated in more than 50 international scientific-practical conferences.

ORCID ID: 0000-0003-4150-7425

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8014

MULTICRITERIA SELECTION OF THE OPTIMAL DESIGN OPTIONS OF TELECOMMUNICATION FACILITIES

Valeriy Bezruk, Daria Chebotareva, Yulia Skoryk
Kharkiv National University of Radio Electronics, Department of the Information and Network Engineering

Abstract. *The peculiarities of the application of methods of multicriteria optimization in the choice of optimal design options of telecommunication facilities taking into account a set of quality indicators are considered. Examples of multicriteria analysis and the choice of optimal variants for various telecommunication facilities, in particular, the variants of radio network construction of a mobile communication system, the structure and methods of routing in multiservice networks, ad-hoc networks, sensor and actuator networks, as well as mobile communication technologies of the 4th generation are provided.*

Keywords: telecommunication facilities, designing, multicriteria optimization, optimal design options

WIELOKRYTERIALNY WYBÓR OPTYMALNYCH OPCJI PROJEKTOWYCH DLA TELEKOMUNIKACJI

Streszczenie. *Rozważono osobliwości zastosowania metod optymalizacji wielokryterialnej w wyborze optymalnych opcji projektowych urządzeń telekomunikacyjnych z uwzględnieniem zestawu wskaźników jakości. Podane zostały przykłady analizy wielokryterialnej i wyboru optymalnych wariantów dla różnych obiektów telekomunikacyjnych, w szczególności wariantów budowy sieci radiowej systemu komunikacji mobilnej, struktury i metod trasowania w sieciach wielousługowych, sieciach ad-hoc, sieciach czujników i urządzeń wykonawczych, a także technologie komunikacji mobilnej czwartej generacji.*

Słowa kluczowe: urządzenia telekomunikacyjne, projektowanie, optymalizacja wielokryterialna, optymalne opcje projektowe

Introduction

Telecommunications facilities significantly affect the user service quality of infocommunication services. Therefore, there are requirements for the strict consideration of technical and economic requirements that are evaluated by the values of the relevant quality indicators, when choosing the best variant from a given set of feasible options of telecommunication facilities. This determines the necessity of using multicriteria optimization methods in choosing optimal design options for telecommunications facilities [1–6].

In this article, practical features of the application of multicriteria optimization methodology for solving various tasks of planning and designing of communication networks with consideration of the set of quality indicators are considered. Examples of multicriteria choice of optimal variants of communication networks in the tasks of nominal planning of cellular communication networks, selection of optimal variants of the structure of the data transmission network, choice of optimal routing in a multiservice communication network, selection of the preferred option of the routing protocol in ad-hoc networks and in the wireless sensor and actuator network, the choice of a prevailing telecommunication technology in mobile networks were provided [7].

1. Methodology of multicriteria optimization for the selection of optimal design solutions

In the work [7], a methodology of multicriteria optimization for the choice of optimal design solutions based on a set of quality indicators is presented. It includes the implementation of the following design procedures: the formation of a certain set of acceptable design solutions, evaluation of design options for a set of quality indicators and their representation in the criteria space of their assessments, setting of vector optimality criterion that takes into account a set of quality indicators, as well as the chose of design options optimal for the chosen optimality criterion. In the case of the usage of the unconditional advantage criterion (ADC), the optimal design variants correspond to Pareto-optimal estimates in the criterion space. When finding a subset of the Pareto-optimal design options worse design decisions are certainly excluded. The resulting Pareto-optimal design options are incomparable between themselves by ADC. In order to further narrow the Pareto subset and select a single design option for its further practical implementation, it is necessary to additionally introduce the conditional advantage criterion (CAC) using additional information from experts about the preference. There are a number of formalized methods for considering this information when constructing a formalized procedure that implements the selection of a single design option by CAC. In particular, it is a method of analyzing hierarchies, a lexographic method, a method of successive concessions, methods based on the utility theory and the theory of fuzzy sets.

2. Multicriteria optimization at the stage of nominal cellular communication network planning

Practical peculiarities of the application of multicriteria optimization methodology at the stage of nominal planning of cellular networks of mobile communication (CNC) were considered.

Finding the optimal design options for CNC with nominal CNC planning, considering the aggregate of quality indicators, includes the following stages:

- setting the initial set of network options that differ by data on the service area, the allocated frequency line, the number of users and others;
- selection of a set of permissible options taking into account restrictions on the structure and parameters of the network, as well as restrictions on the values of quality indicators;
- selection of a subset of Pareto-optimal network options using ADC;
- analysis of the obtained Pareto-optimal network options, evaluation of their multidimensional potential characteristics and multidimensional diagrams of the exchange of quality indicators;
- selection of a single network design variant from Pareto-optimal subset using the specified CAC.

In the example, a set of permissible CNC options was formed, determined by different data on the planned number of subscribers in the network, size of the serviced territory, subscribers activity, allocated frequency line, cluster size, transmitter power of base stations (BS), allowable call blocking, percentage of the time

of communication quality deterioration. As a result of the calculations, an initial (nominal) frequency-territorial plan has been developed and a subset of permissible options of CNC have been formed. An example of the nominal CNC planning was considered and it was possible to select the following quality indicators: k_1 – probability of error, k_2 – network capacity, k_3 – number of base stations in the network, k_4 – efficiency of radio frequency spectrum use, k_5 – probability of blocking, k_6 – coverage area. For each design version of the CNC, estimates of the values of the specified quality indicators were found, their rationing to maximum values was carried out and brought to a comparative form. The finding of a subset of the optimal variants of CNC was carried out in the criteria space V of assessments of the vector of entered quality indicators $\vec{k} = (k_1, k_2, k_3, k_4, k_5, k_6)$.

An initial set of 100 options of construction of CNC was formed. In the criterion space of the corresponding vector estimates $\vec{k}$, a subset of optimal estimates for the binary relation of a non-strict preference $\geq$ in the space of estimates V, which is also called a subset of Pareto-optimal estimates $P(V) = opt_{\geq} V$, was identified. The inclusion of vector estimates in this subset $\vec{k}^0 \in P(V)$ was performed if and only if there were no other vector estimates $\vec{k}$ for which vector inequality would hold $\vec{k} \geq \vec{k}^0$.

Such a criterion for the choice of optimal solutions is called an unconditional advantage or Pareto criterion. The detected subset of Pareto-optimal estimates corresponds to the subset of Pareto-optimal variants of SMS.

According to the given formalized selection procedure, in the considered example of the 100 design SMS options, 71 Pareto-optimal CNC construction options were selected. At the same time, 29 certainly the worst by the Pareto criterion options for CNC construction were rejected.

For the obtained Pareto-optimal design options, the following properties are characteristic:

1) All variants of SMS, which do not belong to the Pareto-optimal set are definitely worse.
2) All Pareto optimal versions of the SMS cannot be recognized as definitely worse or better than another version of the Pareto subset. This means that they are all not comparable to the Pareto criterion.
3) Each Pareto-optimal version of the SMS is responsible for the potentially possible value of each of the quality indicators $\vec{k}$, which can be achieved with fixed, but arbitrary values of other quality indicators. This is the property of the agreed multiplier optimum. The set of such optimal values of quality indicators is multidimensional potential characteristics of the system (MPC).

In order to choose from the found subset of Pareto-optimal variants a single design variant of CNC, the CAC was used as a weighted sum of values of the selected quality indicators $k_p = \sum_{i=1}^{6} c_i k_i$. Provided minimum k_p with $c_i = 1/6$ from the obtained Pareto-optimal subset a unique variant of the construction of the CNC was selected, which is characterized by the following data: the number of subscribers in the network is 30,000; area of the served territory is 320 sq·km; subscriber activity – 0.025 Earl; bandwidth – 4 MHz; admissible probability of call blocking – 0.01; percentage of time of deterioration of communication quality – 0.07; allowable density service – 94 act·sb/sq·km; cluster size – 7; the number of base stations in the network – 133; number of subscribers, serviced by one BS – 226; the efficiency of usage of the radio frequency spectrum – $1.614 \cdot 10^{-4}$ act·sb/Hz; telephone loading – 3.326 Earl; probability of error – $5.277 \cdot 10^{-7}$.

3. The selection of optimal design options for data network, taking into account a set of quality indicators

The peculiarities of the application of the methodology of optimization for the selection of optimal project version of the data network with packet switching, taking into account the set of quality indicators were considered. The quality indicators which are determined by the time of delivery and the possibility of packet loss in the framework of the datagram transmission of messages were defined. These quality indicators are interconnected and are antagonistic, that is, with the improvement of the value of one of the indicators, the other quality indicator deteriorates. Such a task of a data transmission network designing is relevant for practical applications that are critical to the timely delivery of messages, in particular, in video and voice messaging systems, banking terminal systems, alarm systems, troubleshooting systems in communication networks.

In the research, a mathematical model of a fully connected data transmission network topology was constructed. In the structure of mathematical model of the network, simulators of message sources, packaging messaging procedures in packages and their transmission via communication channels, routing and servicing procedures at switch nodes, error simulators in communication channels were introduced. The sources of messages with the Poisson law of distribution and different intensities of the flow of applications were modeled. Also, modeling of various delays in the transmission of packets associated with the final rate of signal propagation in communication channels, the fixed channel bandwidth, as well as the time that packets stay in the queue for transmission via the communication channels was carried out. Different variants of network operation were implemented that differed in the disciplines of packet data service in queues, routing methods for packet transmission, and the size of the transport connection window.

According to this problem thirty six variants of the network data transmission were given. As a result of simulation modeling, for each network option, estimates of quality indicators were found: average delivery time of packets $k_1 = \overline{T}$ and average probability of loss of a message $k_2 = \overline{P}$. In this case, the obtained admissible set of variants of the work of the network, which was presented in the criterion space of quality indicators estimates normalized to the maximum values. Here you can see a subset of Pareto-optimal network options by excluding definitely the worst options by the Pareto criterion.

Among the obtained Pareto-optimal variants of the network the only variant with the use of CAC was chosen – provided the minimum of the scalar function of value in the form of the resulting quality indicator $k_p = c_1 k_1 + c_2 k_2$. For the case $c_1 = 0.4$, $c_2 = 0.6$ the only variant of the work of the network that was characterized was chosen: the discipline of service flow of applications – in a random manner by the method of routing – uniform in accordance with the weight, the size of the "window" of transmission – equal to 8.

4. Optimal routing in multiservice communication network, taking into account the set of quality indicators

Multiservice communication network is a complex system with a set of elements and in order to provide high quality services of different types of traffic the actual task is optimal routing based on set of indicators. The peculiarities of applying of methodology of multicriteria optimization during the routing planning in such communication networks were considered.

The following multicriteria problem of routing was formulated. The set of acceptable solutions X (routes) in the final graph of the network $G = (V, E)$ was given, where V – the set of nodes, E – the set of connection lines. On a sat X vector target function is given function $\overline{F}(x) = (F_1(x),..., F_v(x),..., F_m(x))$ the components of which define the value of corresponding indicators of routs' quality k_v. The indicators of routs quality are connected and are antagonistic. The given task is the task of selection of the optimal routs options by the aggregate of quality indicators.

The solution of this problem is a subset of Pareto-optimal routing options, which corresponds to the agreed optimum of individual target functions $F_1(x),..., F_v(x),..., F_m(x)$ that determine the set of service quality indicators.

The choice of optimal routes taking into account a set of quality indicators means selecting a subset of Pareto-optimal routing options. The route option $\tilde{x} \in X$ is Pareto-optimal, if there is no other route $x^* \in X$ for which irregularities $F_v(x^*) \le F_v(\tilde{x}), v = 1,..., m$ are performed, and at least one of them is strict. When comparing the routes for this vector criterion, the benefits of the set of acceptable options exclude definitely worse route options and remain Pareto-optimal route variants among themselves. The subset of Pareto-optimal route options corresponds to the Pareto optimum of quality indicators, that is, the minimum possible value of one of the quality indicators at a possible change in the values of another quality indicator.

The found Pareto-optimal route variants are incomparable from the point of view of the ADC – the Pareto criterion. The resulting subset of Pareto-optimal route options can be used to organize multi-route routing, in particular, in MPLS technology. This approach allows to implement load balancing and traffic control and provides a given quality of service for a set of quality indicators.

5. The choice of the prevailing option for routing protocol in ad-hoc network based on hierarchy analysis method

Hierarchy analysis method (HAM) is based on obtaining and formalized processing of information from experts in the form of matrices of pairwise comparisons of the importance of quality indicators, as well as design options with respect to each quality indicator $\left[a_{ij} \right]$. From a mathematical point of view, the processing of matrices of pairwise comparisons is reduced to the calculation of the principal eigenvectors of the matrix, which, after a certain rationing, becomes the vector of priorities of the elements of comparisons at a certain hierarchy level. Calculation of the estimates of the components of the main eigenvector of the matrix of pairwise comparisons is determined by a ratio

$$V_j = \sqrt[n]{a_{ij}}, \ i = \overline{1, n},$$

and assessments of the vector components of the priorities of quality indicators are determined by the ratio

$$P_j = \frac{V_n}{\sum_{j=1}^{n} V_n}, \ j = \overline{1, n}.$$

The obtained components of the priority vectors are used to calculate the components of the vector of global priorities according to the ratio

$$C_j = \sum_{i=1}^{N} P_i P_{ij}, \ j = \overline{1, N},$$

where P_{ij} – the calculated priority vectors of the design options with respect to each quality indicator.

The number of the corresponding maximum components of the obtained vector of global priorities of the design variants in relation to each quality indicator C_j corresponds to the prevailing design variant.

The features of the HAM application were considered for the selection of a single prevailing option for routing protocol in ad-hoc networks, taking into account the set of quality indicators [4]. A set of routing protocols included the following protocols: DSDV, OLSR, WRP, AODV, DSR. As protocols' quality indicators, the main characteristics of the protocols were selected, in particular k_1 – convergence time, k_2 – memory, k_3 – management. These quality indicators are interconnected and competing. According to the hierarchy analysis method the matrix of paired comparisons of the importance of the quality parameters of routing protocols was constructed. Subsequently, paired comparisons of alternative variants of routing protocols with respect to selected quality indicators were performed and corresponding matrixes of pair comparisons were obtained. As a result of processing of received matrices, eigenvectors and vectors of priorities were calculated. Using these data, the components of the global priority vector were calculated. The prevailing option of the routing protocol in ad-hoc networks corresponds to the maximum value of the component of the global priority vector, taking into account the introduced quality indicators. This is the OLSR routing protocol, which is based on the Dijkstra's algorithm.

6. The selection of the prevailing routing protocol in the wireless sensor and actuator network

At present, wireless sensor and actuator networks (WSAN) have been actively developed, which are effectively used to solve applied tasks of distributed information collection on monitored parameters in monitoring networks. Such networks are homogeneous, self-organizing, peer, with cellular topology, whose nodes are capable of retransmission of information. Autonomous sources of the structural elements consumption of such networks in the form of batteries impose strict energy efficiency constraints on all routing protocols used in sensor networks. Therefore, for WSAN, routing tasks are important, in particular, the search of routes that are optimal in terms of energy efficiency and the search of routes that provide maximum network lifetime. The development of WSAN led to the emergence of a large number of routing protocols. Let's consider the features of the prevailing routing protocol selection in the WSAN by the hierarchy analysis method.

A set of protocols was specified, including the following protocols: Sensor Protocols for Information via Negotiation (SPIN), Direcied Diffusion, Rumor Routing, Low Energy Adaptive Clustering Hierarchy (LEACH), Threshold Sensitive Energy Efficient Sensor Network Protocol (TEEN), Power-Efficient Gathering in Sensor Information Systems (PEGASIS, Self-OrganizingProtocol (SOP), Geographic Adaptive Fidelity (GAF), Geographicand Energy and Aware Routing (GEAR), Sequential Assignment Routing (SAR), Stateless Protocol for Real-Time Communication in Sensor Networks (SPEED).

To select a single prevailing routing protocol in WSAN, taking into account the aggregate of quality indicators, a hierarchy analysis method was used to obtain expert data on pair comparisons of the importance of quality indicators and routing protocols in WSAN in relation to these quality indicators. These data were processed and the values of components of the global priority vector C_j were obtained. According to the hierarchy analysis method, the prevailing WSAN option is the GEAR routing protocol, which corresponds to the maximum component of the global priority vector.

7. The selection of the prevailing technology in the 4th generation mobile network

For comparative analysis, the following technologies of mobile communication were selected: HSPA (release 7 and separate amendments of the release 8), WiMAX (release 1.5) and LTE. At the same time, the following quality indicators for the specified mobile communication technologies were chosen: k_1 – spectral efficiency, k_2 – radius of action, k_3 – data transfer rate.

The quality indicators of these technologies are interconnected and competing, which determines the need to apply multicriteria optimization methods, in particular, the peculiarities of using the method of analysis of hierarchies for the choice of the prevailing option of mobile communication technology were considered. The matrix of paired comparisons of quality indicators was obtained, for which the components of the main eigenvector and the priority vector were calculated. Further, paired comparison of technologies in relation to selected quality indicators was conducted. After processing of the corresponding matrix of paired comparisons, the corresponding main eigenvectors and priorities vectors were obtained, which were used to calculate the components of the vector of global priorities C_j. The prevailing LTE technology with a data rate of 75 Mbit/s, a spectral efficiency – 1.57 bits/Hz/s and a radius of base stations – 5 km corresponds to the maximum value of the global priority vector components.

8. Conclusions

This article shows how to use a multicriteria approach when planning and designing optimal telecommunication facilities taking into account the totality of quality indicators. A number of examples that illustrate the practical features of applying the methods of multicriterion optimization when choosing the optimal design options for various types of communication networks were considered.

References

[1] Granat J., Wierzbicki A. P.: Multicriteria analysis in telecommunications. Proceedings of the 37th Hawaii International Conference on System Sciences, 2004, 1–6.

[2] Mishra A. R.: Advanced Cellular Network Planning and Optimisation. 2G/2.5G/3G Evolution TO 4G. John Wiley & Sons Ltd, London 2007.

[3] Grosan C., Abraham A., Hassainen A.: Designing resilient networks using multicriteria metaheuristics. Telecommunation System 40, 2009, 75–88.

[4] Saaty T.: Theory and Applications of the Analytic Network Process. Decision Making with Benefits, Opportunities, Costs and Risks. RWS Publications, Pittsburgh 2005.

[5] Lee H., Shi Y., Nazem S. M. et al.: Multicriteria hub decision making for rural area telecommunication networks. European Journal of Operational Research 133, 2001, 483–495.

[6] Carlos De Morais Cordeiro: Ad Hoc & Sensor Net works: Theory and Applications. Agrawal World Scientific 1, 2006, 641.

[7] Bezruk V. M., Bukhanko A. N., Chebotareva D. V., Varich V. V.: Multicriterion optimization in telecommunication networks planning, designing and controlling. Telecommunications Networks. Current Status and Future Trends, Dr. Jesús Ortiz (Ed.), InTech, 2012, 251-274.

Prof. Valeriy Bezruk
e-mail: valeriy_bezruk@ukr.net

Doctor of Technical Sciences, Professor, Head of the "Information and Network Engineering Department" of the Kharkiv National University of Radio Electronics. In 1973 Valery Bezruk graduated from the Radio Engineering Department of the Kharkiv Institute of Radio Electronics on the specialty "Radio Engineering". In 2005 Valeriy Bezruk defended his doctoral dissertation.
Area of scientific interests: statistical methods of processing and recognition of signals, mathematical modeling and multicriteria optimization of systems.
Valery Bezruk has over 200 publications, including 5 monographs and 9 textbooks.

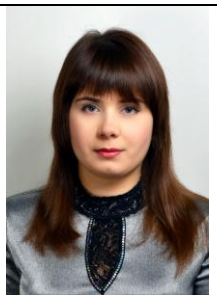
ORCID ID: 0000-0003-2349-7788

Ph.D. Daria Chebotareva
e-mail: dariia.chebotarova@nure.ua

Candidate of technical sciences, Docent of the "Information and Network Engineering Department " of Kharkiv National University of Radio Electronics.
In 2005 Daria Chebotareva graduated from the Telecommunications and Measuring Technology Department of Kharkiv National University of Radio Electronics on the speciality “Information Communication Networks”. In 2012 Daria Chebotareva defended the thesis for the Candidate's Degree.
Area of scientific interests: planning and multicriteria optimization of communication networks. Daria Chebotareva is the author of over 50 publications, including 21 articles, 3 monographs and 1 textbook.

ORCID ID: 0000-0002-8442-8412

Ph.D. Yulia Skoryk
e-mail: yuliia.skoryk@nure.ua



Candidate of technical sciences, Docent of the "Information and Network Engineering Department" of Kharkiv National University of Radio Electronics.
In 2008 Yulia Skoryk graduated from the Telecommunications and Measuring Technology Department of Kharkiv National University of Radio Electronics on the speciality “Information Communication Networks”. In 2014 Julia Skoryk defended the thesis for the Candidate's Degree.
Area of scientific interests: multicriteria optimization of telecommunication facilities. Yulia Skoryk is the author of more than 30 publications, including 16 articles, 2 monographs.

ORCID ID: 0000-0003-1729-1003

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8016

DEVELOPMENT OF THE METHOD OF INDIRECT STEGANOGRAPHIC DATA HIDING IN THE CONTAINER IMAGE CONTOUR

Volodymyr Barannik[1], Oleg Shatun[1], Dmitriy Barannik[2], Veronika Kobtseva[2]
[1]Ivan Kozhedub Kharkiv National University of Air Force, Faculty of Automated control systems and aviation ground support,
[2]Kharkiv National University of Radio Electronics, Department of information and network engineering

Abstract. *The article discusses issues related to the use of methods of digital steganography for information security in systems of critical appointment. The advantages of using data embedding in an image container are shown. The main disadvantages of the existing methods of embedding in the image container are given. The issues of JPEG image compression for digital steganography are described. The allocation of stable regions in the attacking effects on the basis of the moving mask is proposed. The mathematical apparatus for masking images by the Sobel method is shown. The indirect steganography method of hiding data in blocks which contain information about the circuit is developed.*

Keywords image-container, image contour, discrete cosine transformation, indirect method

OPRACOWANIE METODY POŚREDNIEGO STEGANOGRAFICZNEGO UKRYWANIA DANYCH W INFORMACJI O KONTURZE

Streszczenie*. Artykuł dotyczy konieczności stosowania cyfrowych metod steganograficznych do ochrony informacji w systemach o krytycznym znaczeniu. Pokazano zalety używania osadzania danych w kontenerze w postaci obrazu. Przeanalizowano główne wady większości istniejących systemów steganograficznych. Rozważono problemy kompresji obrazów JPEG występujące przy steganografii cyfrowej. Zaproponowano przydział stabilnych obszarów w oparciu o przesuwną maskę. Przedstawiono aparat matematyczny do maskowania obrazów metodą Sobela. Opracowano pośrednią metodę steganograficzną polegającą na ukrywaniu danych w blokach zawierających informacje o konturze.*

Słowa kluczowe: kontener obrazu, kontur obrazu, dyskretna transformacja kosinusowa, metoda pośrednia

Introduction

In the modern world, information and telecommunication systems (ITS) are widely used for data transmission to ensure the efficiency and quality of information. In the operation of state agencies to enhance economic, defence and social impacts of the use of the system critical applications (SCA). The functioning of the SCA is characterized by the presence of an attacker who can carry out active actions [7, 8]. This can lead to a breach of confidentiality, leakage of important information and lead to significant material and human losses. Therefore, the protection of information in data transmission systems is one of the most important problems of modern science.

1. Analysis

Cryptographic data protection is used to encrypt information resources in the UPC, which provides reliable protection against unauthorized access [4, 10]. Since cryptography only encrypts, not hides, the enemy can intercept the encrypted message and distort it. Therefore, an alternative to cryptography is the use of digital steganography [9, 12].

Digital steganography (DS) is a direction of steganography that hides data in computer files that have analog origin. The most developed and common methods of DS are data embedding in the image-content (IC). This is due to the following reasons:

- distribution of images in the information space;
- the presence in the image of high redundancy, which can potentially be used for embedding information;
- relatively large capacity of the steganographic channel using IC;
- unfavourable human vision to minor distortions in color and brightness.

Methods of embedding data in the IC are divided into direct and indirect [9].

When hiding information in direct methods, bits are embedded in the images, which leads to distortion [1, 6, 11, 13, 14] and reduced container resistance to attacks. In indirect data hiding, information is embedded by creating dependencies between certain elements of the image [17]. These methods are more complex due to the presence of mathematical calculations, but are more resistant to attacks [3, 15].

The Editor reserves the right to editorial redaction of the submitted texts.

2. Problem formulation

Analysis of recent publications [2, 3, 6, 7] showed that the use of indirect methods of hiding data in images provides reliability and reliability of hidden data. The concept of indirect data concealment is based on the creation of image element dependency.

This occurs most often after carrying out a discrete cosine transform (DCT).

The use of prep when embedding information due to the fact that the existing image, the transfer process technology with JPEG loss, part of which is prep.

The generalized analysis of data embedding methods in IC showed the following disadvantages:

- existing methods do not meet the requirements for the steganographic container capacity;
- the methods used are unstable to known attacks on the steganographic system;
- visual analysis of low-level imagery;
- methods have a low probability of correct data extraction.

The above mentioned disadvantages lead to a decrease in the resistance of steganosystems and the probability of loss of information when transmitting through channels [18].

Thus, it is necessary to use a method of hiding data, which will increase the resistance of the container to attacks and increase the probability of correct data extraction.

The basis of this approach is proposed to apply blocks that will be resistant to attacks and will not be distorted during transmission. Therefore, the purpose of the article is to develop a method of hiding data in stable blocks of the image contour, namely elements that contain information about the contour.

3. The allocation of persistent fields based on moving mask

For digital images, the most useful semantic loads are the contours of objects. Contours are lines that pass at the boundaries of homogeneous regions. Elements $\{z_{ij}\}$ of the spatial-temporal representation of the image, whose values do not exceed a certain threshold, form homogeneous areas [19]. This is set by the following condition:

$$|z_{\max} - z_{\min}| \le 1, \tag{1}$$

where $z_{\max}$ is the largest element of the image area; $z_{\min}$ is smallest element of the image.

Existing image compression algorithms reduce redundancy and, thus, introduce the smallest distortion that, while maintaining a high level, provides a disadvantageous human vision to distortion [6, 16]. Therefore, to identify areas resistant to compression effects, it is necessary to use methods of image contour selection, for their further use in steganographic embedding.

The contours of the images are formed at the boundaries of homogeneous areas of the image [5]. In order to determine the belonging of the elements of the spatial representation of the image to a homogeneous area at the same time checking the presence of the contour, the following conditions must be met:

- belonging of the image element to a homogeneous area is given by the condition:

$$|z_{i,j} - z_{i\pm1,j\pm1}| \le 1, \tag{2}$$

$i = \overline{1,x},\ j = \overline{1,y}.$

- belonging of the image element to the adjacent homogeneous region is given by the condition:

$$|z_{i,j} - z_{i\pm1,j\pm1}| > 1, \tag{3}$$

$i = \overline{1,x},\ j = \overline{1,y}.$

Most often used in practice for the selection of image contours and gradient methods.

The most common way to find contours is to process the image *Z* with a sliding mask *K*. The mask *K* is a square matrix with coefficients {k}. The process of image processing *Z* based on the matrix *K* is called filtering or masking and is given by the following functionality $f(\bullet)$:

$$M = f(Z,K), \tag{4}$$

where M is image obtained from the processing of the image *Z* based on the mask K. The filtering process is based on the gradual spatial movement of the filter mask from the element to the image element. The value of the $m_{i,j}$ (filter response) element is calculated using the values of the previous and subsequent elements in two-dimensional space. In this case, the value of the element mi, j of the image M, obtained as a result of masking is calculated by the formula:

$$m_{i,j} = \sum_{\xi=i-1}^{i+1} \sum_{\tau=i-1}^{\tau+1} z_{\xi,\tau} \cdot k\,. \tag{5}$$

The Sobel operator is proposed to be used as a method of image contour selection. This operator is most often used in practice and has the following form:

$$m_{i,j} = \sqrt{G_i^2 + G_j^2}\,, \tag{6}$$

$$G_i = K_i \cdot m_{i,j} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \cdot m_{i,j}; \tag{7}$$

$$G_j = K_j \cdot m_{i,j} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \cdot m_{i,j}. \tag{8}$$

where K_i and K_j are operators that determine the increment value of an image element horizontally and vertically, respectively; G_i and G_j are blocks of the image each element of which contains approximate values of derivatives horizontally and vertically, respectively.

Thus, the use of a sliding Sobel mask allows identifying the contours of objects at the boundaries of homogeneous areas [9]. It is proposed to use this technology to identify areas that will be used in steganographic embedding.

4. Embedding data in areas resistant to compression attacks

We will formulate the requirements for the developed method.

This method should ensure the reliability of hiding information in images, embedding a relatively large amount of information and resistance to distortion. Therefore, it is proposed to build the view steganographic embedding, by indirect modification of the elements of the image blocks that contain contour information.

Step 1. You must select the contours of the image to embed the data.

When selecting contours, it was found that the elements that are located on the positions of the contours, the detection that uses the image mask have the following properties:

1) a limited number of elements that contain contour information, that is, the area of values for embedding information is limited by the width of the contour. In most cases, the contour is not wide (Fig.1);

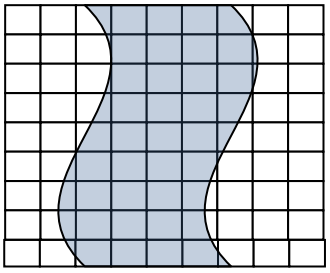


Fig. 1. The width of the contour image

2) through a large number of contours in the image, for the successful integration of data in steganographic transformations, it is necessary to take into account the different directions of the contour of the image: vertical and horizontal (Fig. 2)

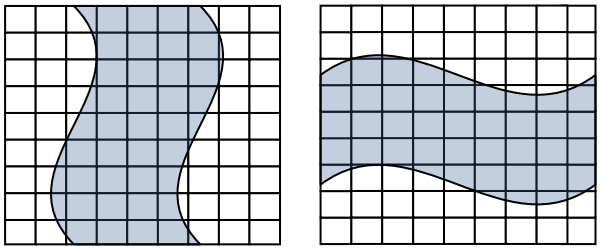


Fig. 2. The direction of the contour depicted (left: vertical position, right: horizontal position)

Step 2. Consider the possible options for selecting the matrix of contour elements for embedding information, taking into account these features.

It is necessary to consider that the allocation of a smart block to be embedded in the contour image, it can get items that are not part of the contour elements.

In this case, when sending a message, the data contained in the IC may be damaged. That is, when choosing a block with dimensions of 4×4 or 5×5 and embedding information in it, the messages will go beyond the contour, which does not meet the requirements for the developed method (Fig. 3).

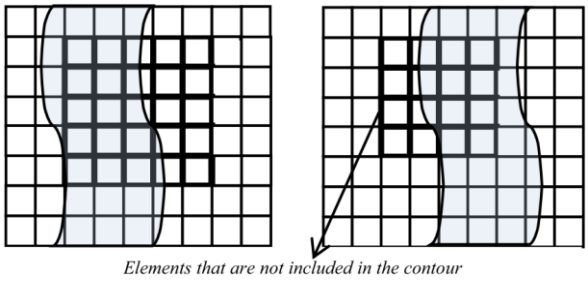


Fig. 3. Example of selection of blocks for embedding

It follows that the most optimal size of the contour for embedding is a matrix of dimension 3×3 elements (Fig. 4).

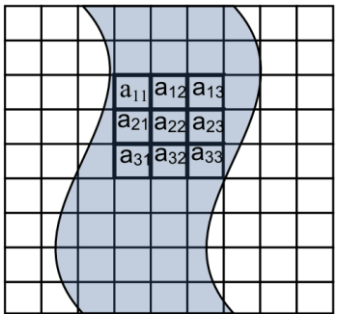


Fig. 4. Choice of the block for embedding data

By selecting the appropriate block and selecting the matrix a 3×3 it is necessary for further calculations to determine the following elements:

- no variability in the process of embedding (reference elements);
- elements that are modified in the process of embedding.

To select the elements that will not change during the implementation process, you must do the following:

In matrix A, the maximum and minimum elements should be defined.

These elements will remain unchanged for the selection of the interval on which the data is embedded.

$$A = \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}. \tag{9}$$

For example, $a_{12} = a_{min}$, $a_{32} = a_{max}$.

Calculate the width of the interval d. This value is required to calculate the interval boundaries and is characterized by the dynamic range of elements in the block matrix. Determined by the formula:

$$d = \frac{a_{\max} - a_{\min}}{3} \tag{10}$$

Step 3. At this stage, the interval boundaries are calculated. These elements are used for comparison. These elements are not modified in the process of data embedding – reference elements, they are used in the process of creating a dependency in indirect embedding. Since the range of values must be divided into 3 intervals, it is necessary to determine the 4 reference elements. These elements are calculated by the formula:

$$z_i = a_{\min} + d(i-1) \tag{11}$$

where z_i is a reference element for comparison, i = [1,4].

Step 4. The items that are selected for embedding of the matrix A, which will be involved in embedding distributed according to intervals. These elements are allowed to change in the process of embedding data modificam elements. The modifier elements must be placed in ascending order and their belonging to the intervals between the reference elements $\{z_i\}$, $i = \overline{1,4}$.

Step 5. It is necessary to calculate the average value of the SJ modifier elements that fall within the interval between two reference elements $\{z_i\}$, $i = \overline{1,4}$. These values will be used for further embedding of the data using the modifier elements.

The calculation of this value is necessary to create a dependency on the embedding of the data. This value is calculated by the formula:

$$S_j = \frac{\sum_{m=1}^{n} a_m}{n}, \tag{12}$$

where j is number of the interval j = [1,3]; a_m is matrix modifier element m = [1, n]; n is number of elements in the interval.

Step 6. To implement indirect embedding of information in the image contours, it is necessary to calculate the reference coefficient K.

This coefficient is constant, that means, it will not change from the modification of the block elements. This coefficient is calculated for each embedding block separately, takes the values $0 < k \le 1$ and is calculated by the formula:

$$k = \frac{z_4}{z_2 + z_3} \tag{13}$$

Data embedding will be performed by modifying the elements of the block according to the rule:

$$b = \begin{cases} 0, \ if\ H > k; \\ 1, \ if\ H \le k. \end{cases} \tag{14}$$

where b is the bit to be embedded in this loop block; H is the coefficient of comparison, and is calculated by the formula:

$$H = \frac{S_3}{S_1 + S_2} \tag{15}$$

If the condition is met, the data injection is complete. That is, it is assumed that the elements have values that satisfy the embedding condition.

Consider cases where the embedding condition does not hold.

Failure to comply with the rules of the embedding of H should increase or decrease for the condition of embedding. Then the rule of embedding will have the following form:

$$b = \begin{cases} 0, if\ H = \dfrac{S_3 + x}{(S_1 - x) + (S_2 - x)} > k; \\ 1, if\ H = \dfrac{S_3 - x}{(S_1 + x) + (S_2 + x)} \le k. \end{cases} \tag{16}$$

where x is modification factor $x \in [-255; 255]$.

Implementation will be the selection of the values of H thus, to ensure the performance of condition 16.

The values of the modified elements a'_{ij} of the block contain information about the contour and is calculated by the following formula:

$$a_i' = a_i + x. \tag{17}$$

The value of x is calculated by the formula:

$$x = \begin{cases} \dfrac{H(S_1 + S_2) - S_3}{1 + 2H}, \ for\ b = 0; \\ \dfrac{S_3 - H(S_1 + S_2)}{1 + 2H}, \ for\ b = 1. \end{cases} \tag{18}$$

We rewrite the finding formula 17 taking into account x:

$$a_i' = \begin{cases} a_i + \dfrac{H(S_1 + S_2) - S_3}{1 + 2H}, \ for\ b = 0; \\ a_i + \dfrac{S_3 - H(S_1 + S_2)}{1 + 2H}, \ for\ b = 1. \end{cases} \tag{19}$$

After finding the values of the modific elements, we perform steganographic embedding of the data in the blocks according to rule 14.

Extracting data occurs after receiving the stegoimage by image analysis, and comparison of the values of H and k in the blocks.

5. Conclusion

1) Reviewed current trends in the information security, it is proposed to apply the methods of digital steganography to protect your data for increased protection and reliability in the transmission of information. Using a container image is the most promising way to hide data.
2) Analysis of existing methods of data embedding in the image container showed that these methods have a low probability of correct data extraction, unstable to existing attacks and have a small steganographic bandwidth.
3) A method for allocating the blocks of the outline of the image based on the moving mask. These blocks are resistant to compression attacks and cause minor distortions in the image, which allows you to use the image mask for steganographic data hiding.
4) The method of indirect steganographic data hiding in the image contours is developed. This method allows hiding bits in image blocks with high probability of correct extraction of embedded data. The developed method is resistant to known active attacks and steganographic analysis by the enemy.

References

[1] Ablamejko S., Lagunovskij D.: Obrabotka izobrazhenij: teh-nologija, metody, primenenie. Amalfeja, Minsk 2000.
[2] Barannik V. V., Alimpiev A. M., Bekirov A., Barannik D. V., Barannik N. D.: Detections of sustainable areas for steganographic embedding. Proc. East-West Design & Test Symposium (EWDTS), 2017, 555–558 [doi: 10.1109/EWDTS.2017.8110028].
[3] Barannik V., Barannik D., Bekirov A., Lekakh A.: Steganographic method based on the modification of regions of the image with different saturation. Proc of 14th International Conference: Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering, 2018, 542–545 [doi: 10.1109/TCSET.2018.8336260].
[4] Barannik V., Krasnorutsky A., Larin V., Hahanova A., Shulgin S.: Model of syntactic representation of aerophoto images segments. Proc of. XVI-th Intern. conf.: Modern Problems of Radio Engineering, Telecommunications and Computer Science, 2018, 974–977 [doi: 10.1109/TCSET.2018.8336356].
[5] Christophe E., Lager D., Mailhes C.: Quality criteria benchmark for hiperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 43(9), 2005, 2103–2114.
[6] Gonzalez R., Woods R.: Tsyfrova obrobka zobrazhen. Tekhnosfera, Kiev 2018.
[7] Gribunin V., Okov I., Turincev I.: Tsifrovaya steganografiya. Solon-Press, Moscow 2018.
[8] Grundmann M., Kwatra V., Han M., Essa I.: Efficient hierarchical graph based video segmentation. IEEE CVPR, 2010.
[9] Konakhovich G. M, Puzyrenko A.: Computernaya steganographiya. Teoriya i praktika. To Press, Kiev 2016.
[10] Melnik A. M.: Informatsiyni systemy ta merezhi. Bulletin "Lviv Polytechnic" 673, 2010, 365–374.
[11] Miano J.: Compressed image file formats: JPEG, PNG, GIF, XBM, BMP. Addison-Wesley Professional, New York 1999.
[12] Miano J.: Formats and image compression algorithms in action. Triumph, Kiev 2013.
[13] Pratt W., Chen W., Welch L.: Slant transform image coding. Proc. Computer Processing in Communications. Polytechnic Press, New York 1969.
[14] Sindeev M., Konushin A., Rother C.: Alpha-flow for video matting. Technical Report, 2012.
[15] Stankiewicz O., Wegner K., Karwowski D., Stankowski J., Klimaszewski K., Grajek T.: Encoding mode selection in HEVC with the use of noise reduction. Proc. International Conference on Systems, Signals and Image Processing (IWSSIP), Poznan, 2017, 1–6.
[16] Wallace G.: Overview of the JPEG (ISO/CCITT) – Still image compression: image processing algorithms and techniques. Processing of the SPIE 1244, 1990, 220–233.
[17] Wallace G.: The JPEG Still Picture Compression Standard. Communication in ACM 34(4), 1991, 31–34.
[18] Wang S., Zhang X., Liu X., Zhang J., Ma S., Gao W.: Utility-Driven Adaptive Preprocessing for Screen Content Video Compression. IEEE Transactions on Multimedia 19(3), 2017, 660–667.
[19] Yudin O., Boiko Y., Frolov O.: Organization of decision support systems for crisis management. Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T), 2015, 115–117 [doi; 10.1109/INFOCOMMST.2015.7357287].

Oleg Shatun
e-mail: shatun_oleg@ukr.net

Cadet of Ivan Kozhedub Kharkiv National University of Air Force. Active participant of scientific conferences, computer science olympiads. Winner of the competition of students scientific works in the field of cyber defense. Scientific interests: systems, technologies of transformation, encoding, defence and information transfer.

ORCID ID: 0000-0002-5323-2580

Dmitriy Barannik
e-mail: d.v.barannik@gmail.com

Student of the Kharkiv National University of Radioelectronics. Scientific interests: systems, technologies of transformation, encoding, defence and information transfer, semantic processing of images. He has 15 scientific publications, including a monograph and scientific articles.

ORCID ID: 0000-0002-7074-9864

D. Sc. Volodymyr Barannik
e-mail: vvbar/off@gmail.com

Doctor of engineering sciences, professor, chief of department of the Ivan Kozhedub Kharkiv National Air Force University. Scientific interests: systems, technologies of transformation, encoding, defence and information transfer, semantic processing of images.

ORCID ID: 0000-0002-2848-4524

M.Sc. Veronika Kobtseva
e-mail: Veronika.Kobtseva@nure.ua

In 2018 she graduated from Kharkiv National University of Radio Electronics from the Master of Education. Active participant of scientific conferences. Deals with data protection issues. Scientific interests: systems, technologies of transformation, encoding, defence and information transfer

ORCID ID: 0000-0002-0914-5156

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8021

STEGANOGRAPHY METHODS USED IN ATTACKS ON INFORMATION AND COMMUNICATION SYSTEMS

Anna Romanova, Sergiy Toliupa
Taras Shevchenko University of Kyiv, Faculty of Information Technologies

Abstract. *An analysis of steganography methods that are can be potentially used as instruments in attacks on information and communication systems is presented. The possible solutions to ensure resilience to such attacks are presented.*

Keywords: steganography, TEMPEST, covert channel, information protection

ZASTOSOWANIE METOD STEGANOGRAFICZNYCH DO ATAKÓW W SYSTEMACH INFORMACYJNO-KOMUNIKACYJNYCH

Streszczenie. *W artykule został przedstawiony przegląd istniejących i potencjalnie dostępnych technik steganograficznych, które mogą zostać użyte jako narzędzia do ataków na systemy informacyjne i komunikacyjne. Podano możliwe sposoby zapewnienia ochrony przed takimi atakami.*

Slowa kluczowe: steganografia, TEMPEST, ukryty kanał transmisji, ochrona informacji

Introduction

Cryptography is widely used as one of the most efficient and approbated methods of critical information resources protection. Nevertheless, in particular cases it might be more effective to hide the communication channel itself instead of making the information within it unreadable. Such a practice, namely – concealing data within unsuspicious, innocent-looking containers – is called steganography.

While being primarily considered a means of information security assurance, steganography can be used with ill intentions, as well. In fact, several high-tech attacks are based on the hidden data transmission, and contemporary methods of counteraction do not provide satisfactory level of resilience to those. The main advantage of steganography becomes the main source of threat – the channels of the attack, not to mention information about the attacker left in the channels, cannot be identified due to the nature of the method. In other words, attacks become invisible, as does the transmission channel. Even the fact of trespassing cannot be easily proven.

The purpose of this article is to conduct an analysis of attacks that are carried out with the use of steganography methods as their basis, and are directed against information and communication systems. Both existing and potential methods are presented.

1. Steganography as a means of hiding information

1.1. 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 lc-long elements.
- Modified container s (also: package, steganogramme) is the one that contains a secret message.
- Steganographic algorithm means two transforms, a direct F: M x B x K → B and an inverse one F–1: B x K → M.
- Steganographic system (also: steganosystem) is a totality of messages, secret keys, containers and transforms that connect them [8, 13].

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 steganogramme detection is called *steganoanalysis*.

1.2. 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) [6].

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 [10].

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 [12].

Methods of embedding data within an image container are [5, 8]:

- 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 correspondsto 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 8×8 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 [10].

Audio steganography [8]:

- 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 [4].

Linguistic steganography [8]:

- 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 [10].
- Using similarly looking Unicode and ASCII characters [2, 6].
- Using non-printable Unicode characters [2].
- Creating a pattern of deliberate errors and/or marked corrections [6].

Some other methods:

- Converting a file so that it has the statistical characteristics of another one [6].
- Format steganography.
- Blog-steganography. Secret data is added as commentary pin boards on social networks [10].

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 [1].

There are also several perspective steganography methods, the use of which is still limited, but nevertheless possible [11]:
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. 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. The network steganography uses communication protocols' control elements and their functionality to hide information inside. 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.
2) The use of stream containers. As mentioned above, by the type of access to the data one can distinguish fixed and stream containers. 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. 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.

2. Steganography methods used as instruments for attacks

2.1. Format steganography

Perhaps, the easiest and the most well-known way, which is actually a steganography method, is using legitimate features of file formats to carry hidden malicious software within their structure. A file of every format contains specific fields, which ensure that the former will be processed correctly on the target computer. Some of these fields are optional, or more strictly – information that they contain is not vital for the file. Thus, changing data bits in these fields most probably will not lead to errors while operating with the file. Such characteristics make these formats perfect containers [11].

A vivid example is a virus Win95.CIH – specific malware which is embedded in *.exe files by using Portable Executable format features. This format includes a lot of additional data which are grouped according to their functions. Every group gets its own section in the file structure, and the size of the sections is predefined. If they are not entirely filled with data, it means the file contains a lot of spare space. For example, the first section is only for the PE header, so a big part of the virus uses it as a covert container [3].

2.2. Soft Tempest

In fact, there are a lot of ways to covertly transmit necessary information to the target system. Not only harmless files but also network protocols can be used as efficient containers within the attacker's steganography system. Nevertheless, necessary means depend on the final objective of the attack. If the goal is to steal data, there is need for both an inward and an outward information flow. Getting information into a system is important. A more interesting question, though, is how to get the stolen data out without raising suspicion of a legitimate user.

While operating, every electronic device (including those inside a computer) gives off compromising emanations – electromagnetic emanations, which can be demodulated and accordingly processed to illegitimately get the critical information from them. These are called TEMPEST emanations after an American standard on the matter.

Contemporary TEMPEST-based attacks tend to become more and more sophisticated as the countermeasures are being continuously enhanced, as well. Systems are contaminated with the malicious software which then conducts the search of necessary information (key data, passwords, specific files etc) and induces the leak through TEMPEST emanation. For example, if reception of the signal is the one from the monitor, then the information will be, say, amplitude modulated and sent as a visual picture to the monitor. The obvious disadvantage is that such an activity cannot be missed by an operator and will be deemed highly suspicious, which, on its part, will lead to finding and neutralizing the virus.

M. Kuhn and R. Anderson conducted a series of experiment in which they shown a possible solution [9]. The human eye is more sensitive to low-frequency than to high-frequency vibrations, while TEMPEST receivers work vice versa. What is more, any devices primarily perceive luminosity in a linear way, while humans are more sensitive for the dark colours. This difference in sensitivity perception can be used to embed a message in the emanation and make it invisible to an unsuspecting user. The suggested method is to control and modify monitor dithering patterns. Pixels of two colours put in a check pattern are seen as a uniform colour, on the one side; on the other side, they create a high-frequency signal, which is best received by TEMPEST equipment with the following use of gamma-correction. Basically, the target computer is programmed so that it acts as a radio transmitter and emits a compound TEMPEST signal: a legitimate user observes one picture, and the attacker receives another – embedded – one on the monitor of his/her TEMPEST receiver.

The only suggested method of counteraction, which is specific enough for this very type of attack, is still based on using the difference in perception sensitivity between humans and devices. TEMPEST fonts are designed with top 30% of the Fourier transform of the signal removed, which is most probably not noticed by a human eye, but makes it impossible to receive a strong TEMPEST signal [9]. Nevertheless, special equipment with necessary parameters (enhanced sensitivity to low-frequency emanations) might be designed, which will make the use of such fonts ineffective.

2.3. Acoustic emanations as containers

Electromagnetic fields are not the only by-product of the computer systems operation. A. Shamir and E. Tromer published the results of their research, in which they showed that computer emit high-pitched noise while operation, due to vibration in some of their electronic components [7].

A series of experiments conducted by the scientist revealed that acoustic emanations can provide a potential attacker with information about what kind of software is currently running on the target system, as well as leak data on security-related parameters and computations. For example, loops of CPU instructions were highly distinguishable, and different RSA keys appeared to induce different sound patterns. To extract individual keys, the technic of acoustic cryptoanalysis was presented (applicable to GnuPG's implementation of RSA). According to the results, it takes about an hour to extract full keys from a target computer, irrespectively to their models and manufacturers. The key piece of equipment used for the attack is a microphone, and that of a mobile phone was demonstrated to be enough. Apart form acoustics, the scientists demonstrated a low-bandwidth attack, based on the same principles. The main difference was that the attacker had to get the leakage from ends of VGA, Ethernet, USB or other cables [7].

If electromagnetic emanations can be used as containers in steganography systems, acoustic waves can be, too. The first case could be based on the nature of sound perception itself – the classical steganography technic. Human hearing systems cannot distinguish slight variations in an acoustic flow. Here, any known method, mentioned above (Least Significant Bit, Echo-signal use etc) can be used to embed stolen information in parasitic sounds, emitted by the target computer. The second possible scenario is similar to the use of emanations in Soft Tempest. Sound dithering is a widely used method in music digital processing. The principle is the following: any piece of musical record might contain extensive frequency transitions that are too slow and smooth. This is where so called quantization noise can appear. If the level of frequency fluctuation is insignificant, the processing software simplifies the sound by removing the frequencies that exceed some medium limit. To cope with such a situation, special noises are generated and gradually added to the record. In music processing, this technic allows to achieve a natural sound lost during quantization.

It is possible to suggest, that the same technic can be used in attacking steganography systems. The noise emitted by a computer is quite stable. It is not foiled by fan system noise, as critical acoustic signals appear to be mostly above 10 kHz, while a typical fan noise along with other noises lie in a much lower frequency band [7]. Task-switching is not a problem either, as it is the tasks that carry distinguishable acoustic spectral signatures. The same can be said about several computers working simultaneously in a closed space: they can be told apart using different sound patterns, as their depend on specific hardware, temperatures inside and outside the system, humidity, and other conditions. Thus, it acoustic emanations seem to be a sufficient container, while dithering can be accordingly modified and applied as an embedment method.

The only suitable countermeasure seems to be the use of sound dampening equipment that can diminish the level of high-frequency leakage. As for means of active protection, strong wide-band noise source can serve for masking the critical data signals. Rough-scale behaving algorithms are another solution: despite somewhat diminishing the level of performance, they can thwart side-channel attacks by shuffling the signal and making it thus useless for the attacker [7]. In addition, electronic components of the system should be those of the highest quality, designed to reduce the level of acoustic and any other leakage.

Nevertheless, at this point, efficiency of such protection methods is rather relevant, as sound-proving degrades other performance features along with being quite expensive. At the same time, due to the need of ventilation, there are still open parts in the cases, so their structure has to be constructed to shuffle outcoming noises very efficiently.

3. Conclusion

Stenganography is a powerful means of information protection. Nevertheless, it has to be also regarded as an instrument for a potential attacker, with all of the advantages turned threats.

Compromising emanations of different physical nature are invisible and can only be noticed with the use of special equipment. Using steganography technics for the attacks ensures that the fact of using those emanations is efficiently hidden, and the system operations remains unsuspicious. This is exactly why there is need to consider technics described above a real threat for information and communication systems, and to join academic and technical potential to develop cost-effective and technically efficient counteracting means.

References

[1] Agranovskiy A. V., Balakin A. V., Gribunin V. G., Sapozhnikov S. A.: Steganografiya, tsifrovyye vodyanyye znaki i steganoanaliz. Vuzovskaya kniga, Moscow 2009.

[2] Ali A. E.: A New Text Steganography Method by Using Non-Printing Unicode Characters. Eng& Tech. Journal 28 (1), 2010, 72–83.

[3] Chekhovskyy S.: Sovremennye metody skrytoy peredachy ynformatsyy putem prohrammnoho upravlenyya yzluchenyem komp′yutery. Pravove, normatyvne ta metrolohichne zabezpechennya systemy zakhystu informatsiyi v Ukrayini. 2003.

[4] Echo Data Hiding (html): http://www.slidefinder.net/a/audio_steganography_echo_data_hiding/24367218.

[5] Ferreira A. M.: An Overview on Hiding and Detecting Stego-data in Video Streams. University of Amsterdam. System & Network Engineering – Research Project II, 2015.

[6] Fridrich J., Goljan M., Soukal D.: Searching for the Stego Key. Proc. SPIE 5306, 2004 [doi: 10.1117/12.521353].

[7] Genkin D., Shamir A., Tromer E.: RSA key extraction via low-bandwidth acoustic cryptoanalysis. Tel Aviv University. 2013.

[8] Konakhovich G. F., Puzyrenko A. Yu.: Computer steganography. Theory and practice with Mathcad. MK-Press, Kyiv 2006.

[9] Kuhn M. G., Anderson R.: Soft Tempest: Hidden data transmission using electromagnetic emanations. University of Cambridge, Computer Laboratory, New Museum Site, 1998.

[10] League C.: An overview of digital steganography, particularly within images, for the computationally curious. Long Island University 2015: https://www.youtube.com/watch?v=-7FBPgQDX5o.

[11] Romanova A., Toliupa S.: Perspective steganographic solutions and their application. Proceedings of the VII Inter University Conference Engineer of XXI Century, volume 2, 2017, 269–278.

[12] 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.

[13] Zorin Ye.Ye., Chichvarin N.V.: Steganografiya v SAPR. Uchebnoye posobiye. MGTU im. N.E. Baumana. Moscow.

Anna Romanova
e-mail: anitraromanova@gmail.com

Since 2015, has been a student of the Department of Cybersecurity and Information Protection. Has published articles in European scientific journals, as well as took part in conferences, devoted to scientific and applied topics. Academic interests include steganography, TEMPEST, system security, and critical information infrastructure protection. Is a certifies SearchInform DLP-System Specialist.

ORCID ID: 0000-0003-1403-6322

Prof. Toliupa Sergiy
e-mail: tolupa@i.ua

Scientific and practical interests are related to such areas as intelligent control systems, the direction of improving the efficiency of information technology, information security systems, cybersecurity and cyber defense. He is the author of 5 monographs and over 150 scientific and methodological works, 16 textbooks and manuals.

ORCID ID: 0000-0002-1919-9174

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8022

BUILDING INTRUSION DETECTION SYSTEMS BASED ON THE BASIS OF METHODS OF INTELLECTUAL ANALYSIS OF DATA

Serhii Toliupa, Mykola Brailovskyi, Ivan Parkhomenko
Taras Shevchenko Kyiv National University, Faculty of Infirmation Security

Abstract. *Nowadays, with the rapid development of network technologies and with global informatization of society problems come to the fore ensuring a high level of information system security. With the increase in the number of computer security incidents, intrusion detection systems (IDS) started to be developed rapidly.Nowadays the intrusion detection systems usually represent software or hardware-software solutions, that automate the event control process, occurring in an information system or network, as well as independently analyze these events in search of signs of security problems. A modern approach to building intrusion detection systems is full of flaws and vulnerabilities, which allows, unfortunately, harmful influences successfully overcome information security systems. The application of methods for analyzing data makes it possible identification of previously unknown, non-trivial, practically useful and accessible interpretations of knowledge necessary for making decisions in various spheres of human activity. The combination of these methods along with an integrated decision support system makes it possible to build an effective system for detecting and counteracting attacks, which is confirmed by the results of imitation modeling.*

Keywords: intrusion detection systems, attacks, fuzzy logic, neural networks

BUDOWA SYSTEMÓW WYKRYWANIA ATAKÓW NA PODSTAWIE METOD INTELIGENTNEJ ANALIZY DANYCH

Streszczenie. *W chwili obecnej szybki rozwój technologii sieciowych i globalnej informatyzacji społeczeństwa uwypukla problemy związane z zapewnieniem wysokiego poziomu bezpieczeństwa systemów informacyjnych. Wraz ze wzrostem liczby incydentów komputerowych związanych z bezpieczeństwem nastąpił dynamiczny rozwój systemów wykrywania ataków. Obecnie systemy wykrywania włamań i ataków to zazwyczaj oprogramowanie lub sprzętowo-programowe rozwiązania automatyzujące proces monitorowania zdarzeń występujących w systemie informatycznym lub sieci, a także samodzielnie analizujące te zdarzenia w poszukiwaniu oznak problemów bezpieczeństwa. Nowoczesne podejście do budowy systemów wykrywania ataków na systemy informacyjne jest pełne wad i słabych punktów, które niestety pozwalają szkodliwym wpływom na skuteczne pokonanie systemów zabezpieczania informacji. Zastosowanie metod inteligentnej analizy danych pozwala wykryć w danych nieznane wcześniej, nietrywialne, praktycznie użyteczne i dostępne interpretacje wiedzy niezbędnej do podejmowania decyzji w różnych sferach ludzkiej działalności. Połączenie tych metod wraz ze zintegrowanym systemem wspomagania decyzji umożliwia zbudowanie skutecznego systemu wykrywania i przeciwdziałania atakom, co potwierdzają wyniki modelowania.*

Słowa kluczowe: systemy wykrywania włamań, ataki, logika rozmyta, sieci neuronowe

Introduction

With the rapid development of network technologies and global informatization of society to the fore problems of ensuring a high level of information system security. With the increase in the number of computer security incidents, started to be developed rapidly intrusion detection systems (IDS).

Traditionally, intrusion detection systems (IDS) are classified according to two characteristics: detection method and system level at which the protection is performed. All developers of intrusion detection systems and organizations that use IDS should understand and study their classification in order to choose the best solutions for information security systems. In the study of various aspects of taxonomy and the application of various options we can achieve a higher level of information system security [8].

1. Main part

Today intrusion detection systems are usually software or hardware-software solutions, that automate the process of control occurring in the information system or network (IDS) also independently analyze these events in search of signs of security issues [4, 7, 11]. Since the number of different types and methods of organizing unauthorized intrusions into foreign networks has increased significantly in recent years and intrusion detection systems have become a necessary component of the security infrastructure of most organizations.

Detection of intrusions has been an area of active research for several decades. There are a large number of different methods and approaches for identifying remote network attacks. To protect the information system such common means and methods are used: corporate network security policy; firewalls; router level protection; network audit; intrusion detection systems; procedure for responding to identified attacks, etc. (Fig. 1).

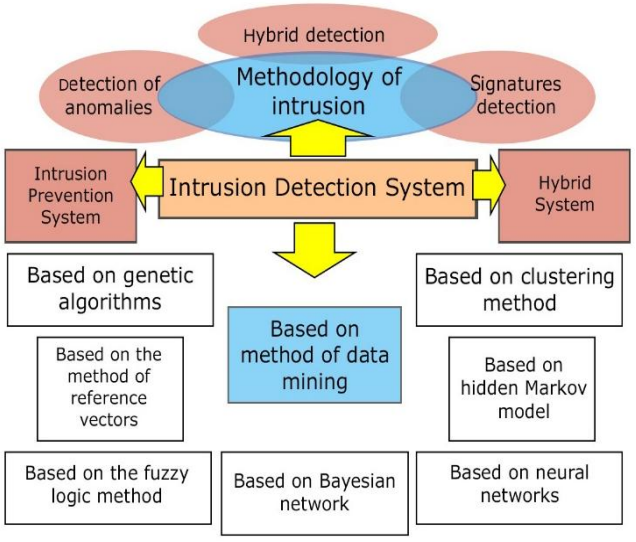


Fig. 1. Classification features intrusion detection systems

In their work, intrusion detection systems are guided not only by network traffic and *a* variety of rules, but also by auditing the system, different audit logs, operating system performance indicators, etc. There are also intrusion detection systems that allow not only to detect the fact of the intrusion into the system, but also to minimize the consequences by breaking the network connection, blocking suspicious user activity or even the administrator.

The most effective way to prevent unauthorized use of information systems and network resources is to support multi-level protection, when firewalls, intrusion detection systems, auditing systems, security policies, and other types of protection are used together.

The most general structure of an intrusion detection system, developed by a group of researchers CIDF (Common Intrusion Detection Framework) [14] is presented in Fig. 2.

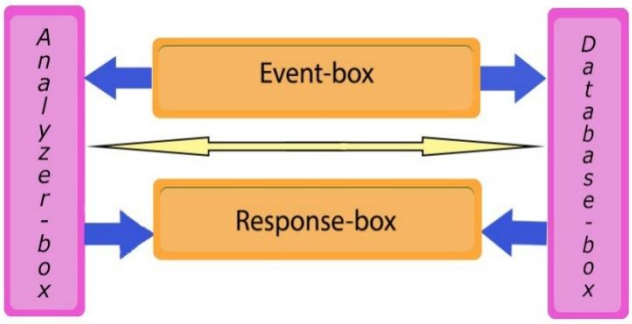


Fig. 2. General structure of an intrusion detection system

Event-box (Sensor) – analyzes data for processing and decision making by the analyzer [12]. The Data may contain names of controlled parameters, their features and values. The sensor can convert data to the required format or to reduce the amount of data transmitted.

Analyzer-box – makes a decision about the presence or absence of signs of an attack or anomaly based on data from the sensors. As part of the analysis, this block can perform the functions of filtering, normalizing, transforming, and correlating data. When an attack is detected, the analyzer block can add a description of the detected attack to the source data. The analyzer unit may have a multi-level system.

Database-box – contains a set of decisive rules and a semantic description of attacks, as well as accumulation of information from sensors. The data may be in text files, a database, etc.

Response-box – informs the administrator about a fixed attack, and in the case of an intrusion detection system creates an active response. Intrusion detection systems track activity in real time and quickly implement actions to prevent attacks. Possible measures are: blocking traffic flows in the network, resetting connections, issuing signals to the operator. Also, intrusion detection systems can perform packet defragmentation, TCP packet sequencing to protect against packets with modified sequence numbers and confirmation.

Network attack detection systems collect information from network traffic packets, system logs and system performance indicators. Traditional systems of detecting network attacks are based on the signature approach: with the help of a set of rules or signatures, which are formed by experts and placed in the database of decision rules, all possible scenarios and features of attacks are described. In this approach, there are many known shortcomings. Using signature analysis it is impossible to detect new types of attacks, because the base of decision rules does not contain information about the corresponding attack. The process of analyzing signatures for distributed attacks is extremely difficult. In addition, the databases of decision rules of popular intrusion detection systems are practically publicly available, so the perpetrator can test the possibility of hiding the attack.

The listed problems of the signature search compel experts to look for alternative ways to organize protection against network attacks. One of the popular areas of research is the use of various data mining methods in network attack detection systems. One of the popular areas of research is the application of various methods of intelligent data analysis in systems for detecting network attacks [2, 13].

Data mining – a set of methods for detecting previously unknown, non-trivial, practically useful and accessible interpretations of knowledge necessary for decision-making in various spheres of human activity. The basis of these methods is the assumption that all legitimate activity in the system can be represented as a mathematical model. Mining methods used to detect network attacks have one of the following aims: detection of violations; detection of anomalies.

The first method is simulate attacks and use classification tools. The other method is about simulating normal behavior and searching for exceptions.

Using data mining technique to detect network attacks, the following problems can be identified: considered data by intrusion detection systems have a high dimension and volume the requirement of data processing in real time; a large amount of noise and inconsistencies in the processed data and those which cause inadequate response of data mining methods.

Analyze IDS based on data mining techniques. One such method is to detect an attack using the Hidden Markov Model (HMM) [10]. The Hidden Markov Model (HMM) is a statistical model [6] where the system is modeled as a Markov process with unknown parameters. The objective of the method is to evaluate the hidden parameters that are based on the observed parameters. Event sequences collected from normal operating systems used as a training sample to evaluate the parameters of the Hidden Markov Model (HMM). After investigating the hidden Markov model, probabilistic estimates are used as threshold values for identifying network anomalies in test data.

Detection of attacks using Bayesian networks. The Bayesian network is a model that encodes probabilistic relationships between variables. The main method of using Baessian networks implies independence among the attributes. Several variants of the use of Bayesian networks have been proposed to detect network anomalies [15]. Most of the methods are aimed at forming conditional\to form conditional dependencies between attributes using complex Bayesian networks. Bayesian methods are often used in the procedure for the classification and localization of false positives. Bayesian networks may be effective in some cases to detect intrusions or predict intruder behavior, but in the general case, the accuracy of this method depends on the assumptions related to the behavior of the target system model. Thus, any significant deviation from the assumptions will reduce the accuracy of the detection.

Detection of attacks using clustering methods. The methods of clustering group data into clusters based on the similarity of objects. Most clustering methods begin with the choice of a central point for each cluster, and the set of elements is distributed among the clusters. After that, the centers are adjusted, and the elements are redistributed. Clustering allows you to study and identify anomalies without requiring a plurality of classes or types of anomalies. That is, to identify anomalies using clustering methods, there is no need for a training set. Clustering is widely used to detect network anomalies [16]. Detection of unknown network attacks is most often based on clustering methods. Homogeneous groups with similar characteristics or clusters are formed by splitting a set of elements without any marks. It is extremely important for a system to correctly identify the clusters in order to keep them from emissions as far as possible. The ultimate goal of these methods is to determine the degree of deviation of emissions from clusters. By simply comparing with a threshold values, emissions with a high degree of deviation from clusters are marked as anomalies.

Detection of attacks using the support vector machine. This method is one of the most popular classification methods. The method constructs an optimal hyperplane in the characteristic space: $w \times x - b = 0$, that separates normal and abnormal elements [3, 5]. As a result, the problem can be reduced to quadratic programming: $min \parallel \omega \parallel \frac{2}{H}/2 + c\sum_{i=1}^{N}\xi_i$, when $c_i(wx_i - b) \geq 1 - \xi_i$, $1 \leq i \leq N$, where ξ_i is the magnitude of the error in the objects x_i. The parameter c is a compromise between the accuracy of the description of the model, is determined by the magnitude of the error $\sum_{i=1}^{N}\xi_i$ and the possibilities of the model to generalization, that is, the value of the limit $1/\parallel \omega \parallel \frac{2}{H}$.

Detection of attacks using neural networks. The interest in artificial neural networks is caused by the fact that the human brain produces computing operations in a completely different way than a conventional digital computer. Neural networks provide a variety of tools for a variety of applications: clustering of data, drawing of signs, diminishing of dimension, etc [1].

Detection of attacks using genetic algorithms. Genetic algorithms represent a computational model based on the principles of evolution and natural selection. With this approach, the problem that needs to be solved is transformed into an environment that uses the chromosomal data structure. Chromosomes have evolved over many generations, using operations such as choice, recombination, and mutation [9]. In the task of detecting network anomalies, the record chromosome contains genes that correspond to attributes such as services, flags, number of hits, etc.

Detecting attacks using rules of fuzzy logic. Fuzzy network intrusion detection systems use a variety of fuzzy rules to determine the probability of specific or general network attacks. A fuzzy set can be formed to describe the traffic on a particular network. Work [9] describes a method for constructing classifiers that use fuzzy associative rules that are used to detect an intrusion into a network. Fuzzy sets of association rules are used to describe normal and anomalous classes. Belonging of record to particular class is determined by the appropriate metric. Fuzzy association rules are formed based on the usual training samples. A test sample is classified as normal if the indicator generated by the set of rules is above a certain threshold value. Samples with a lower rate are considered abnormal.

Consider the possibility of using a group of data mining methods for designing a network attack detection system.

According to the previously presented subtasks associated with the identification of network attacks, we can distinguish several groups of data mining methods. At the core of most IDS is the classification process, which forms a conclusion about the fixation of an attack or anomalous behavior. Currently, there are numerous studies on the identification of network attacks. The basis of these studies are such techniques as neural networks, decision trees, association rules, genetic algorithms, and many others.

As a result of the analysis of a large number of studies, the support vector machine has been selected as a classifier. This method shows one of the best indicators of attack detection and has ample opportunities for internal configuration.

The support vector machine finds patterns that are on the borders between two classes, which are called support vectors. In the case of linear inseparability, the kernel mechanism is used, which translates the initial data into a large-dimensional space, in which there is a linear boundary between the classes. The following types of nuclei were most widely used: polynomial, radial-base and sigmoidal nuclei. Due to the high computational complexity, most studies are limited to using radial-base kernel.

Other functional tasks solved by the network attack detection system are mainly aimed at improving the quality of detection, speed of the system, unification and performance of other auxiliary tasks.

Learning principal component analysis is a complex computational task. In addition, the quality of the classification essentially depends on the internal settings of this method, individual for various types of attacks and training sets. In this regard, there is a need to simplify the procedure for preparing a classification model. The solution to this problem is to reduce the size, including the rejection of noise and emissions, as well as the breakdown of the training set into parts – the implementation of the clustering procedure.

The main purpose of the methods of reducing the size is to search for a smaller space in which the internal properties of the source data are stored. These methods allow you to define the set of most important parameters to detect a specific attack. The choice of a particular size reduction method is strongly depends on training data. For the problem that is solved, the method of main components is the most promising.

The purpose of the cluster analysis is to partition data sets into groups in such a way as to minimize the differences between the elements of the same group and maximize the differences between the elements of different groups. The main methods of cluster analysis are divided into hierarchical and non-hierarchical.

Hierarchical methods allow us to construct an optimal cluster structure, but have an exponential dependence on the number of records. Due to the large size of the training data arrays, the use of hierarchical clustering is impossible for all data arrays. But when considering fragments of training data containing records of individual attacks, the amount of information allows the use of hierarchical clustering.

Of the non-hierarchical methods of cluster analysis, iterative methods, which are more universal than hierarchical, but have one serious drawback, are the most commonly used – the need for a priori knowledge of the number of clusters, which greatly complicates their automation.

For the problem of forming the composition of the detection modules, both hierarchical methods and iterations are selected. The hierarchical ones are applied to a subset of attacks from the training set, iterative allows clustering the entire training set, the initial number of clusters and their centers are determined based on the hierarchical clustering subset of attacks.

Fuzzy logic is the generalization of classical formal logic and set theory. Instead of the values of lie and truth, the membership function of an element in the fuzzy set whose values are in the interval [0; 1]. For unclear logic, there are general logical operations based on operations with fuzzy sets. The current base of fuzzy rules allows you to compare the detected anomaly with a set of known attacks. The modular architecture allows you to build an over-the-top model, increasing the likelihood of an attack, and the fuzzy conclusion allows you to reduce the number of false positives.

Fuzzy logic allows you to solve the problem of having the same records with different marks in the training data: "questionable" network packets are a set of anomalous with some probability.

A side application of fuzzy logic is is the construction of clusters that intersect, and an extension of the support vector method. Of course it is ineffective to resist invasion and attacks based on only one of the data mining methods, so it is necessary to approach this issue comprehensively and to build an intellectual system to counteract the invasion. When constructing such an expert system it is suggested to choose a fuzzy model. This is due to the fact that much of the information about the causes and source of attacks can be obtained only by expert or in the form of heuristic descriptions of processes. To determine the sources of attacks, the security system must be represented by a model of the information network on which it is oriented. This model divides the problem of moving information between computers through a network environment into a number of levels of smaller and easier to solve subtasks. Each of these subtasks is solved using a single network level. Therefore, the primary task of a security specialist can be represented by the decomposition of security tasks at individual levels of the network.

Imagine a separate level of security in the form of a nonlinear object with a set of input variables $\{x_i\}, i = \overline{1,n}$ and one output variable *y*: $y = f_y(x_1, x_2, ..., x_n)$.

As input variables are selected signs of sources of attacks. The output variable y is an indicator of the degree of network-level state capability.

The model uses the following assumptions and limitations:

- input variables $\{x_i\}$ within the same level are independent,
- at each level of the network, separate network functions are isolated.

Building an integrated intelligent decision support systems to determine intrusions based on data mining methods should contain a set of functional components that allow to maximize automation and speed up the development of managing actions when chaning the security situation.

References

[1] Bankovic Z., Stepanovich D., Bojanic S., Nieto-Taladris O.: Improving network security using genetic algorithm approach, Computers and Electrical Engineering, 33(5-6)/2007, 438–451.

[2] Barsegyan A. A., Kupriyanov M. S., Stepanenko V. V., Kholod I. I.: Technologies of data analysis: Data Mining, Visual Mining, Text Mining, OLAP, SPb. BHV, Petersburg 2007.

[3] Bhattacharyya D. K., Kalita J. K.: Network Anomaly Detection. A Machine Learning Perspective, CRC Press, 2014.

[4] Brailovskyi M. M., Pogrebna T. V., Ptakhok O. V.: Essential requirements for the construction and safety of next-generation networks. Telecommunication and Information Technologies 2/2014, 41–49.

[5] Brailovskyi N. N., Ivanchenko E. V., Khoroshko V. A.: Diagnostics of information space protection systems" Information protection. Special issue 2014, 59–67.

[6] Ghahramani Z.: An Introduction to hidden Markov models and Bayesian networks. International Journal of Pattern Recognition and Artificial Intelligence 15/2001, 9–42.

[7] Koboseva A. A., Machalin I. O., Khoroshko V. O.: Analysis of the security of information systems. DUIKT, Kiev 2010.

[8] Pavlov I. M., Toliupa S. V., Nishchenko V. I.: Analysis of Taxonomy of Attack Detection Systems in the Context of the Current Level of Information Systems Development. Modern Protection of Information 4/2014, 44–52.

[9] Tajbakhsh A., Rahmati M., Mirzaei A.: Intrusion detection using fuzzy association rules. Applied Soft Computing 9(2)/2009, 462–469.

[10] Tereikovskiy I., Toliupa S., Parkhomenko I., Tereikovska L.: Markov Model of Normal Conduct Template of Computer Systems Network Objects. 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET-2018.

[11] Toliupa S. V, Borisov I. V.: Methodology of evaluation of the complex system of information security at the object of information activity. Scientific and Technical Journal "Modern Information Protection" 2/2013, 43–49.

[12] Toliupa S. V, Parkhomenko I. I., Konovalenko A. D.: Analysis of vulnerabilities of local wireless networks and ways to protect them from possible attacks. Journal of the Engineering Academy of Ukraine 3/2017, 72–76.

[13] Toliupa S. V., Parkhomenko I. I.: Multilevel hierarchical models of information security systems. Proceedings of the II International scientific and practical conference Trends in the development of corvergent networks: decision of the post: NGN, 4G, 5G. Kyiv 2016, 111–114.

[14] Valdes A., Skinner K.: Adaptive model-based monitoring for cyber attack detection. Proc. of the Recent Advances in Intrusion Detection, Toulouse, France, 2000, 80–92.

[15] Valdes A., Skinner K.: Adaptive model-based monitoring for cyber attack detection. Proc. of the Recent Advances in Intrusion Detection. Toulouse 2000, 80–92.

[16] Yang H., Xie F., Lu Y.: Clustering and classification based anomaly detection. Fuzzy Systems and Knowledge Discovery 4223/2006, 1082–1091.

Prof. SerhiiToliupa
e-mail: tolupa@i.ua

Scientific and practical interests are related to such areas as intelligent control systems, the direction of improving the efficiency of information technology, information security systems, cybersecurity and cyber defense. He is the author of 5 monographs and over 150 scientific and methodological works, 16 textbooks and manuals.

ORCID ID: 0000-0002-1919-9174

Prof. Mykola Brailovskyi
e-mail: bk1972@ukr.net

Scientific and practical interests are related to such areas as information and communication systems and networks, systems of technical protection of information, cyber defense. He is the author of over 40 scientific and methodological works.

ORCID ID: 0000-0002-3031-4049

Prof. Ivan Parkhomenko
e-mail :parkh08@ukr.net, parkh08@gmail.com

Scientific and practical interests are related to such areas as information and communication systems and networks, systems of technical protection of information, cyber defense. He is the author of over 80 scientific and methodological works.

ORCID ID: 0000-0002-9197-2600

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8023

OVERVIEW OF APPLICATIONS OF WIRE MEDIUM IN RADIO ENGINEERING MEANS

Mykola Khobzei[1], Dmytro Vovchuk[1], Magdalena Michalska[2]
[1]Yuriy Fedkovych Chernivtsi National University, Department of Radio Engineering and Information Security
[2]Lublin University of Technology, Institute of Electronics & Information Technology, Lublin, Poland

Abstract. *This paper presents an overview of the main possible applications of wire medium in different field of radio engineering. Wire media possesses the unique electromagnetic properties. There is a number for approaches for manufacturing the wire media of different shapes from microwave frequency up to optical one. The wire media that consists of parallel metallic rods finds application for narrow and broadband power radiation in photovoltaic devices, imaging, endoscopy and many others. The wire metastructures such as hyperlens and brush are applicable to the antennas of wide frequency range.*

Keywords: wire media, superlens, hyperlens, brush, metamaterial

PRZEGLĄD ZASTOSOWAŃ KONSTRUKCJI Z PRZEWODÓW W URZĄDZENIACH RADIOWYCH

Streszczenie. *W niniejszej pracy dokonano przeglądu głównych możliwych zastosowań struktur z przewodów w różnych dziedzinach inżynierii radiowej. Struktury z przewodów posiadają unikalne właściwości. Istnieje wiele podejść do wytwarzania struktur o różnych kształtach i zastosowaniach od zakresu częstotliwości mikrofalowych do optycznych. Struktury składające się z równoległych przewodników są wykorzystywane do transmisji fal wąskopasmowych i szerokopasmowych fal elektromagnetycznych w urządzeniach fotowoltaicznych, transferu obrazów, endoskopii i wielu innych. Konstrukcje z przewodów, takie jak hiperłącza oraz z losowym rozmieszczeniem przewodów, mogą być wykorzystywane do projektowania anten działających w szerokim zakresie częstotliwości.*

Słowa kluczowe: struktury z przewodów, supersoczewka, hipersoczewka, struktura z losowym rozmieszczeniem przewodów, metamateriały

Introduction

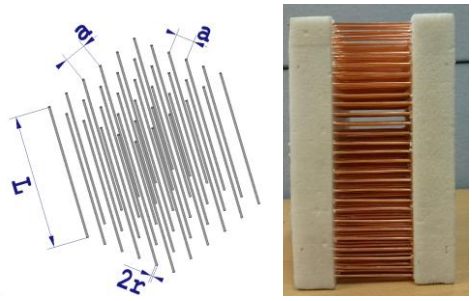
Metamaterials is the class of artificial materials with the negative values of permittivity and/or permeability that cannot be found in nature [6, 13]. Such effect is explained by purely internal characteristics that can depend on the dimensions of metamaterial elements, their shape, distances between them and others.

Wire media is a unique class of metamaterials that can be simply manufactured for the wide frequency range from radio and microwave frequencies up to optical and higher [14]. It is possible due to the number of approaches for production of WM of different dimensions that include mechanical assembling of metallic wires into arrays and fabrication of micro- and nanostructured WM which operate in THz, IR and visible frequencies. Some of such technologies are anodic electrochemical etching, anodization of aluminum and influence of ion beam on track membranes of different dielectrics. As a result, there is a possibility to manufacture the porous matrixes with thickness up to 2 μm and diameter of pore 20 nm [2, 7, 17]. It became the reason of intensive studying WM. There are many works where such metamaterials are presented with different shapes [1, 6, 14]. The most popular of them (superlens, hyperlens and brush) and their applications are considered in this chapter based on the last scientific investigations. Therefore, goal of the paper is an overview of modern applications of WM for different radio engineering devices and telecommunication means as well as discussion of further possible usage of such metamaterial.

1. Parallel Metallic Wire Structure

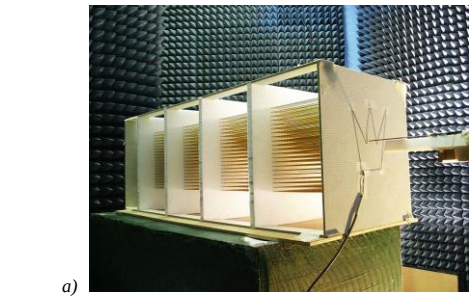
The structure that consists of arranged parallel metallic wires with the finite values of wires length *L*, their diameter *2r* and lattice period *a* included into dielectric matrix is called superlens (Fig. 1 – CST model and experimental sample) and finds the number of applications [3, 14, 19].

First of all, WM is known as a mean of imaging. As is shown in earl works it is possible at Fabri-Perot resonances of WM in different frequency ranges [4, 5]. One of examples is depicted in Fig. 2 [5]. An antenna that has shape of the wished picture is located in front of the structure in the distance not more than the value of lattice period of WM. There were two experiments for both 894 and 1034 MHz. As one can see from comparison of electric field distributions at the back side of WM (Fig. 2b – right top and bottom), such device can operate only at Fabry-Perot resonance and the right bottom picture is vague.

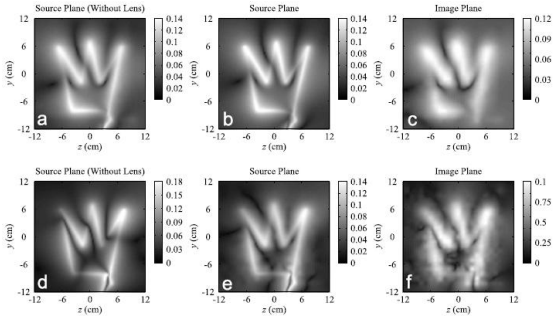


a) *b)*

Fig. 1. Wire media that consists of parallel metallic wires with length L, diameter 2r and lattice period a: CST model (a) and experimental sample (b) [19]



a)



b)

Fig. 2. Wire media of microwave frequency range (a) that can transfer images at Fabri-Perot resonance (b) [4]

There was a restriction within the previously described approach for EM wave transfer because such application of WM is limited by general dimensions of the structure. The first idea of broadband power transfer was proposed in [11] as an effective WM multilayer between the emitter and photovoltaic cell as shown in Fig. 3. The experimental investigation as a proof of broadband power transfer was carried out in [19] where waveguide ports were placed one in front of another by their apertures in some distance. The air gap between ports was filled out by parallel metallic slab that allows enhancing transmission of EM power in a few times in comparison with the case when the transmission through the free space [10, 19]. Those studies opened new possibilities of WM in a number of applications including control radiative heat flows in thermophotovoltaic devices, sensing, effective multimode telecommunications and others.

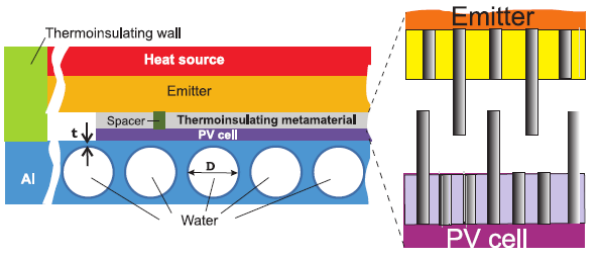
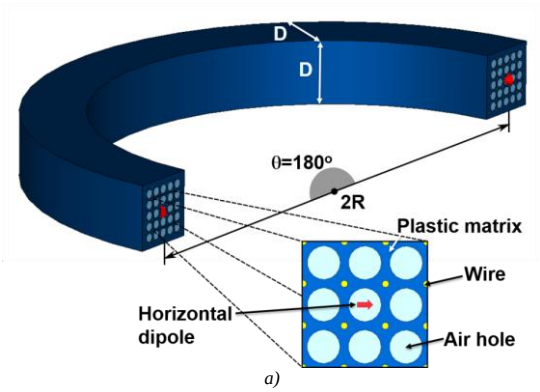
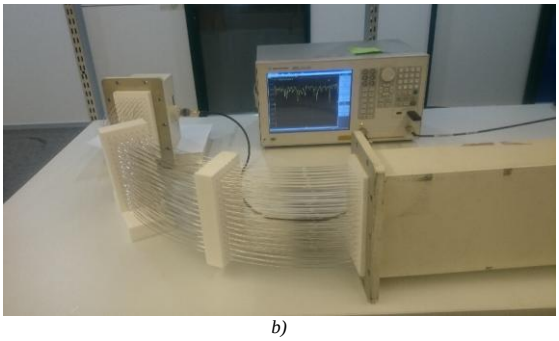


Fig. 3. A micro-gap thermophotovoltaic system enhanced by metamaterial with parallel metallic wires [11]

The broadband effect of power transfer became an initial point for the development of the means for endoscopy. In [16, 20] the WM that consists on parallel wires is base for the suggested device. The endoscope contains WM included into the porous polyethylene matrix that can operate in microwave and THz frequency ranges. The investigation in [16, 20] shows that the endoscope keeps its transmission properties under the bending with different values of angle even more than 180° for the optically long structure 2 cm for the frequency range 0.5÷1 THz (for the bending radius R approx. 6.3 mm). Moreover, such device demonstrates functionality when it is folded with radius $R \approx \lambda$ at the frequencies below 0.75 THz [20].



a)



b)

Fig. 4. A model of THz endoscope that consists of parallel metallic wires included into porous dielectric matrix (a) and its microwave experimental setup (b) [16, 20]

The experimental investigation of the suggested WM endoscope in microwave range has confirmed its stable work for the bending with different angles that were equal 90°, 180° and more. The calculated loss dispersion for the straight and bent endoscope is shown in Fig. 5 [20]. The averaged values <P> are 0.38, 0.4 and 0.41 that correspond on the different angles demonstrate the effectiveness and the huge role of WM in this scientific field.

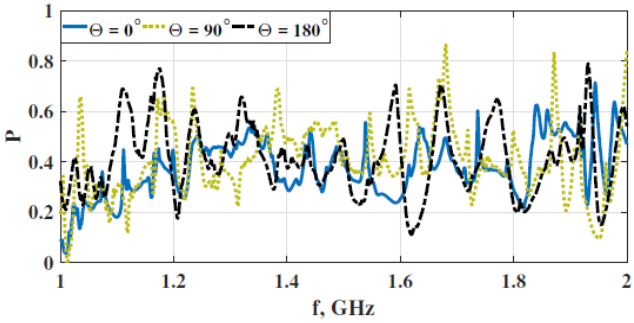


Fig. 5. Dispersion of power loss for straight endoscope as well as the bent on with angles 90 and 180° that obtained by the microwave experiment [20]

2. Tapered Wire Media

The WM structure that consists of array of tapered metallic wires, so-called hyperlens, was suggested to solve a problem of enhancement of propagation EM modes. In Fig. 6 the hyperlens was considered as a structure that allows to fix the noise oscillations and enhance the power of very weak radiation [9]. It is possible due to the shape of suggested structure because the input and output aperture are different and the enhancement factor depends on the relation of their dimensions.

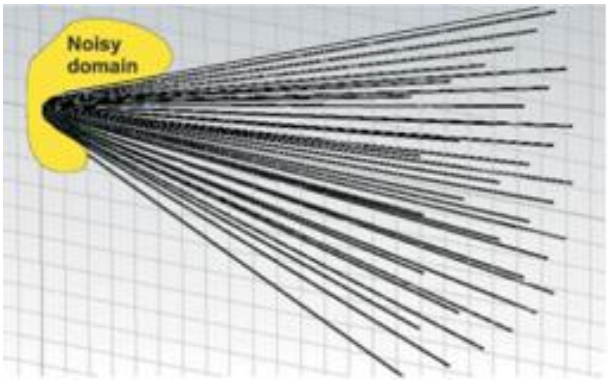


Fig. 6. A model of tapered metamaterial [9]

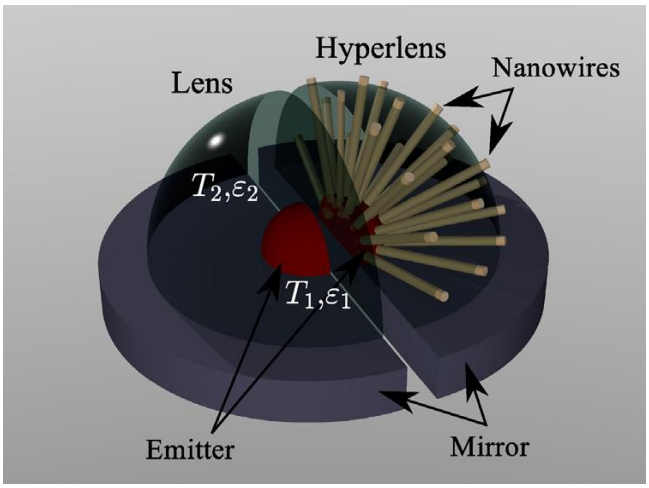


Fig. 7. A hemispherical structure for the super-Planckian thermal radiation from a hot emitter modified by tapered metamaterial [15]

In [15] such metastructure was suggested to modify the usual lens with the aim to enhance thermal radiation of hot body of infrared frequency range. The principle was that the usual hemispherical lens was filled out by tapered wires and the radiative element (emitter) was located as is shown in Fig. 7. The hyperbolic metamaterial of the hyperlens converts emitter's near fields into propagation modes which effectively radiate from the hyperlens surface.

3. Tapered Wire Media

Wire medium can be used for modification of different antennas to radiate evanescent EM waves at the frequencies that do not correspond to the resonance frequencies of modified antenna. Of course, the radiation properties of such modified antennas depend on the features of WM configuration. The work [8] gives an excellent description of this. There, the broadband radiation was the goal of investigation and the irregular metallic structure – WM brush, was suggested based on the properties of parallel and tapered metallic structures. The synthesis from the regular WM structure (superlens) and hyperlens to the brush is presented by Fig. 8 [8]. The superlens (Fig. 8a) can support the narrowband radiation at the Fabri-Perot resonance frequencies (Fig. 8d – red solid plot). At the same time, a hyperlens (Fig. 8b) provides radiation not only at resonances, but also between of them, because values of Parsell factor there are more than 1 (Fig. 8d – green dotted plot). However, the dispersion of Parsell factor is not smooth. Therefore, the irregular WM structure (Fig. 8c) was designed as a combination of regular one and hyperlens. It is because the dimensions are the same with first, but the whole structure is as a volume of a huge number of local hyperlens with different lengths, tilts of angles and locations. It makes an effect of reradiating and as a result the dispersion of Parsell factor is smoothed throughout the investigation frequency range (Fig. 8d – pink dash-dotted plot).

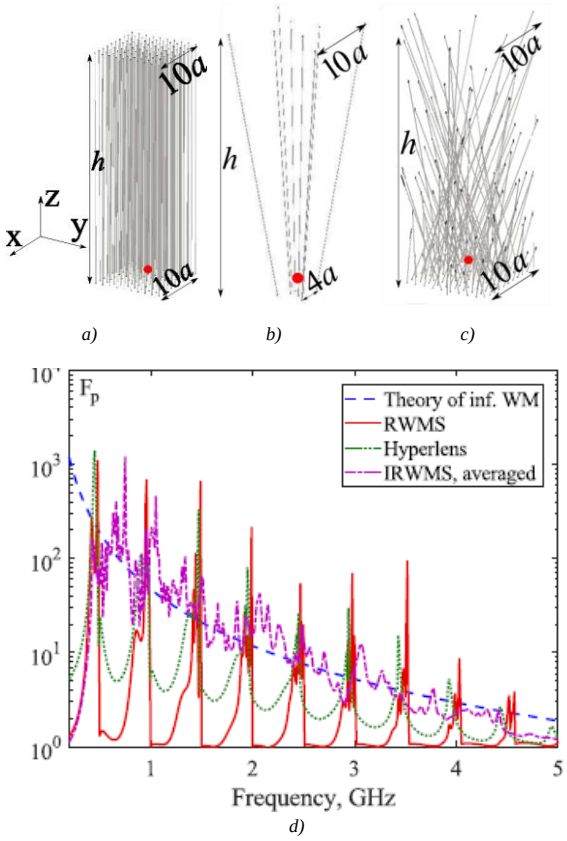


Fig. 8. A synthesis of wire medium of different shapes from the parallel (a) and tapered (b) ones the irregular structure (c) and the dispersion of Parcell factor when the point electromagnetic wave source is modified by those metamaterials (d) [8]

The experimental investigation was performed for previously manufactured samples of parallel and randomized aluminum WM [8]. The external dimensions (length, width and high) of both were the same. The experimental setup that consists of Network Analyzer for displaying of the results, loaded by inductance diaphragm waveguide port and WM sample inserted into the port is shown in Fig. 9a.

In the case when the open waveguide port radiates into the free space without any WM the values of S_{11}-parameters are approximately equal -5 dB (Fig. 9b – blue dotted plot). If the waveguide port is modified by regular WM structure the radiation is possible at the resonance frequency (Fig. 9b – black dotted plot). The broadband radiation in the frequency range from 1.1 up to almost 1.4 GHz was achieved with using irregular WM metamaterial (Fig. 9b – red solid plot).

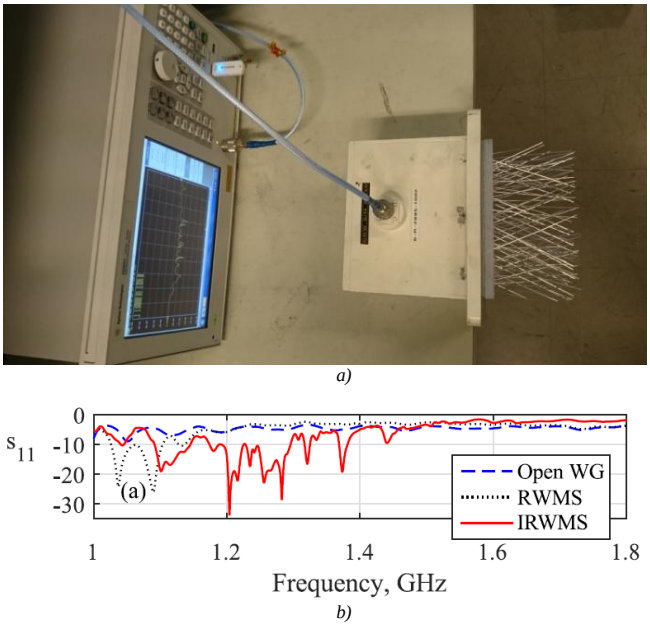


Fig. 9. An experimental setup for the investigation of waveguide port radiation that loaded by irregular WM (a) and S_{11}-dispersions when the port radiates into free space (blue dotted plot), through the parallel WM (black dotted plot) and irregular WM (red solid plot) (b) [8]

4. Conclusions

The WM metamaterials is a prospective class of unusual materials with unique property (negative value of permittivity) that cannot be found in nature. However, such materials can be synthesized in laboratory. The flexibility of WMs includes the possibility of their manufacturing of different dimensions for the application in very wide frequency range from microwaves up to infrareds [14]. It is possible due to different methods from clearly mechanical to chemical approaches [2, 7, 17].

The most popular structures are WM that consists of parallel metallic wires, tapered one and with irregular placement of the wires. Each of them possesses the different properties. For example, the first structure is effective for the narrow and broadband power transfer. It has found its application, first of all, for the imaging. Further development and investigation of WM have shown the possibility of using of parallel metallic structures for the sensing, thermophotovoltaic, endoscopy and many others. When the possibility of operation in wide frequency range was proven, the outlook of WM application was changed. The first endoscopes that based on the WM were narrowband [12]. However, the suggested in [20] WM endoscope can operate in wide THz frequency range that is experimentally proven for microwaves. Moreover, such endoscope is very resistant to the bending with different values of radius including the radius that approximately equal the wavelength.

The WM with tapered wires shown the possibility to use this class of metamaterials for EM wave radiation. Since the tilt of wires is defined the dispersion of Parcell factor is not smoothed and there are the picks at the resonant frequencies. Therefore, in [8] the new kind of wire metamaterials was proposed, so-called WM brush. The randomized arrangement of wires allows creating a huge number of local hyperlenses with different values of wires, angles and their positions. After such modification the dispersion of Parcell factor became smoothed.

Despite the fact that the metastructure with parallel wires is narrowband it does not mean that this structure is strongly only radiates at resonance frequency. For instance, in [20] was presented the results where the parallel metallic structure can be used for modification of antenna that gives the possibility of radiation in some frequency range. This range, of course, is not very wide, but it is controllable by WM parameters and can find a number of applications in antennas technologies.

The authors believe this paper can expend the horizons in WM development and its applications as well as will be useful for the scientists who work in metamaterial area and find the new engineering solutions for the development for the different radio engineering means.

References

[1] Albooyeh M.: Electromagnetic Characterization of Metasurfaces. PhD Thesis at Aalto University, 2015.

[2] Atrashchenko A.V., Katz V. N., Ulin V. P., Evtikhiev V. P., Kochereshko V. P.: Fabrication and optical properties of porous InP structures. Physica E 44/2012, 1324–1328.

[3] Belov P.A., Marques R., Maslovski S. I., Nefedov I. S., Silverinha M., Simovski C., Tretyakov S. A.: Strong spatial dispersion in wire media in the very large wavelength limit. Phys. Rev. B. 67/2003, 113103c(1–4).

[4] Belov P.A., Zhao Y., Sudhakaran S., Alomainy A., Hao Y.: Experimental study of the subwavelength imaging by a wire medium slab. Appl. Phys. Lett. 89/2006, 459–462.

[5] Belov P.A., Zhao Y., Tse S., Ikonen P., Silveirinha M. G., Simovski C. R., Tretyakov S. A., Hao Y., Parini C.: Transmission of images with subwavelength resolution to distances of several wavelengths in the microwave range. Phys. Rev. B. 77/2008, 193108 (1–4).

[6] Capolini F.: Theory and Phenomena of Metamaterials. Taylor and Francis Group, LLC, 2009.

[7] Föll H., Leisner M., Cojocaru A., Carstensen J.: Macroporous semiconductors. Materials 3/2010, 3006–3076.

[8] Kosulnikov S. Yu., Mirmoosa M. S., Vovchuk D. A., Tretyakov S. A., Glybovski S. B., Simovski C. R.: Enhancement of Radiation With Irregular Wire Media. IEEE Transactions on Antennas and Propagation 64(12)/2016, 5469–5474.

[9] Kosulnikov S., Filonov D., Glybovski S., Belov P., Tretyakov S., Simovski C. Wire-Medium Hyperlens for Enhancing Radiation From Subwavelength Dipole Sources. IEEE Trans. Antennas Propag. 63/2015, 4848–4856.

[10] Kosulnikov S., Vovchuk D., Nefedov I., Tretyakov S., Simovski C.: Broadband Power Transfer Through a Metallic Wire Medium Slab. URSI International Symposium on Electromagnetic Theory (EMTS), Espoo 2016, 596–599.

[11] Mirmoosa M.S., Simovski C. R.: System Analysis of Micron-Gap Thermophotovoltaic Systems Enhanced by Nanowires. Photonics and Nanistructures – Fundaments and Applications 13/2015, 20–30.

[12] Radu X, Garray D., Craeye C.: Towards a wire medium endoscope for MRI imaging. Metamaterials 3/2009, 90–99.

[13] Sihvola A.: Metamaterials in electromagnetic. Physical Review Letters 85(18)/2000, 3966–3069.

[14] Simovski C., Belov P.A., Atrashchenko A.V., Kivshar Yu. S.: Wire metamaterials: Physics and applications. Advanced Materials 24/2012, 4229–4248.

[15] Simovski C., Maslovski S., Nefedov I., Kosulnikov S., Belov P., Tretyakov S.: Hyperlens makes thermal emission strongly super-Planckian. Photonics and Nanistructures – Fundaments and Applications 13/2015, 31–41.

[16] Simovski C., Vovchuk D., Kosulnikov S.: Power Vortices in Wire-Medium Endoscopes (Invited). 12th International Congress Metamaterials, Espoo 2018, 315–317.

[17] Ulin V. P., Konnikov S. G.: Electromechanical pore formation mechanism in III-IV crystals. Semiconductors 41/2007, 832–844.

[18] Vovchuk D., Khobzei M., Zhadan I.: Properties of Antennas Modified by Wire Media. Proceedings of 1st International Conference on Informatioon and Telecommunication Technologies and Radio Electronics (UkrMiCo), Odessa 2018.

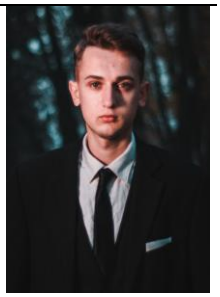
[19] Vovchuk D., Kosulnikov S., Nefedov I., Tretyakov S., Simovski C.: Multi-Mode Broadband Power Transfer through a Wire Medium Slab. Progress In Electromagnetics Research (PIER) 154/2015, 171–180.

[20] Vovchuk D., Kosulnikov S., Simovski C.: Unusual eigenmodes of wire-medium endoscopes: impact on transmission properties. Optics Express 26(14)/2018, 17988–18005.

Mykola Khobzei
e-mail: hobze7@gmail.com

Bachelor student at Radio Engineering and Information Security Department of Yuriy Fedkovych Chernivtsi National University. Research field includes wire metamaterials for the transmission and radiation. Author of more than 5 publications.

ORCID ID: 0000-0002-9101-8569



Ph.D. Dmytro Vovchuk
e-mail: dimavovhuk@gmail.com

Teaching assistant (Ph.D., Candidate of Science) at Radio Engineering and Information Security Department of Yuriy Fedkovych Chernivtsi National University. Research field includes metamaterials for the radio devices applications. Author of more than 10 publications.

ORCID ID: 0000-0001-7108-8984

M.Sc. Magdalena Michalska
e-mail: magdalena.michalska@pollub.edu.pl

Ph.D. student, Institute of Electronics and Information Technology, Lublin University of Technology
Recent graduate Warsaw University of Technology The Faculty Electronics and Information Technology. Since 2017Graduate Student Member, IEEE Students or IEEE Graduate Students, Poland Section Branch Lublin University of Technology. Research field covers medical image processing, 3D modelling, optoelectronics, spectrophotometry.

ORCID ID: 0000-0002-0874-3285

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8025

IMPROVEMENT OF FIDELITY OF MOVING OBJECTS CLASSIFICATION IN GUARD SIGNALING COMPLEXES WITH SEISMIC SENSORS

Bohdan Volochiy[1], Mykhailo Zmysnyi[1], Leonid Ozirkovskyy[1], Volodymyr Onyshchenko[2], Yuriy Salnyk[2]
[1]Lviv Polytechnic National University, Department of Theoretical Radio Engineering and Radio Measurement,
[2]Hetman Petro Sahaidachnyi National Army Academy, Scientific Centre

Abstract. *The effectiveness of guard signaling complexes (GSC), when there is an important validity of the classification of moving objects (MO), is evaluated by the following indexes: probability of GSC task execution; probability of partial fulfillment of the task; probability of user's "deception". Accordingly, the performance indicators of the GSC, in turn, depend on the indexes of the functionality of its constituents: probability of fixation of moving object by seismic sensor, probability of correct classification of MO type and probability of receiving radio signal by the system of receiving and displaying information (SRDI). The article describes a discrete-continuous stochastic model of of GSC reaction to moving object crossing control zone, in which three seismic sensors are installed. Majority principle of identifying the type of moving object was used on the receiving part of the complex. A comparative analysis of the effectiveness of guard signaling complexes using one, two and three sensors in control zone are carried out.*

Keywords: seismic sensor, guard radio electronic complex, efficiency indexes

POPRAWA JAKOŚCI KLASYFIKACJI OBIEKTÓW RUCHOMYCH W SYSTEMACH ALARMOWYCH Z WYKORZYSTANIEM CZUJNIKÓW SEJSMICZNYCH

Streszczenie. *Skuteczność systemów alarmowych, w przypadkach, gdy ważna jest dokładność klasyfikacji poruszających się obiektów, ocenia się za pomocą następujących wskaźników: prawdopodobieństwo wykonania zadania; prawdopodobieństwo częściowej realizacji zadania; prawdopodobieństwo "oszukiwania" użytkownika. W związku z tym, wskaźniki jakości działania systemów alarmowych zależą od wskaźników funkcjonalności ich komponentów: prawdopodobieństwa ustalenia poruszającego się obiektu za pomocą czujnika sejsmicznego, prawdopodobieństwa poprawnej klasyfikacji rodzaju poruszającego się obiektu oraz prawdopodobieństwa odbioru sygnałów radiowych przez system odbioru i wyświetlania informacji. Artykuł przedstawia dyskretno-ciągły model stochastyczny reakcji systemu alarmowego na przekroczenie strefy kontrolnej przez poruszający się obiekt, w której zainstalowane są trzy czujniki sejsmiczne. Identyfikacja poruszającego się obiektu odbywa się na wyjściu systemu na podstawie zasady większości. Porównano skuteczności systemów alarmowych z wykorzystaniem jednego, dwóch i trzech czujników sejsmicznych w strefie kontrolnej.*

Słowa kluczowe: czujnik sejsmiczny, system alarmowy, wskaźniki efektywności

Introduction

In the phase of system design of guard signaling complex (GSC) it is necessary to perform research of its effectiveness with different versions of its implementation. Corresponding research must form the requirements for using GSC constituents. A future GSC must detect moving objects (MO) by seismic sensors, perform MO classification by seismic signals and transmit messages by radio channel from autonomous systems of detection, object classification and transmitting radio signals (DOCTRS) to the system of receiving and displaying information (RDI).

Seismic sensors (SS) [11, 12] have widespread application in designing guard systems.

The main advantages of SS use: operational conditions specify covert SS layout in the ground; a principle of detecting moving objects by SS is passive which does not imply energy emission to the environment [3, 4].

The GSC effectiveness (Fig. 1) is assessed by the following factors: Probability of GSC task execution; probability of partial fulfillment of the task; probability of user's "deception". Above-mentioned factors of GSC effectiveness, on their part, depend on functionality factors of its constituents: probability of fixation of moving object by seismic sensor, probability of correct classification of MO and probability of delivering message about moving object in the system of receiving and displaying information. Probability of GSC task execution implies the situation when the MO, detected (fixed) in control zone, is correctly classified. Probability of partial fulfillment of the task implies the situation when the MO is detected (fixed) in control zone, but is not classified. Probability of user's "deception" implies the situation when the MO, detected (fixed) in control zone, is classified incorrectly. In all cases the message about MO with a certain probability is delivered to the user.

The effectiveness of guard signaling complex with installing one or two SSs in control zone is examined in works [7–9].

In order to perform comparative research it is necessary to develop a model of GSC reaction to MO crossing control zone with three seismic sensors. To improve the fidelity of MO classification it is suggested to use majority principle of taking decisions {2 from 3} [2, 5] on the receiving part of GSC. So, actual is the task of developing a model of GSC reaction to crossing control zone by moving object where there are three seismic sensors.

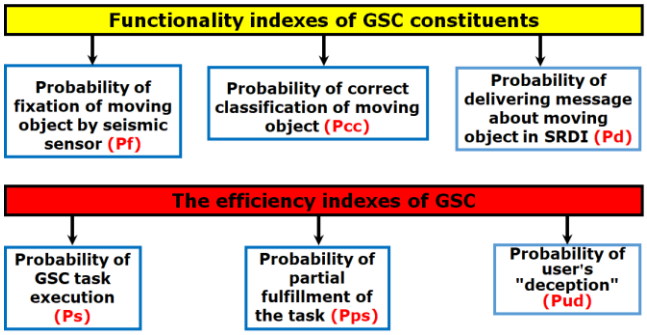


Fig. 1. Effectiveness factors of guard signaling complex and functionality factors of its constituents

The developed model allows conducting analysis of GSC effectiveness of seismic sensor sensitivity (probability of MO detecting), of effectiveness of classification method (probability of correct classification), and of effectiveness of the system of transmitting radio signals (probability of receiving radio signal). In addition, the development of the model will show the advantage of GSC using the majority principle of taking decisions. Thus, the object under study is GSC reaction to MO crossing of control zone where there are three SSs.

The task of developing a method for classification of MO with the use of signals from seismic sensors is complex. Therefore, it is necessary to look for such principles of constructing of GSC (technical solutions) in order to reduce the requirements to the classification method (to the value of the index functionality of the classification method). In this regard, two questions arises. How much can the value of the index of the functionality of the classification method be reduced, if:

The instead of one SS, to install two SSs in the control zone?

The instead of two SSs, to install three SSs in the control zone and use the majority principle for decision-making on the type of MO according to the rule of voting "2 of 3" on the receiving side?

Experimental researches have been carried out by the method of computer simulation using the software tool ASNA [12] and the formalized representation of the research object in the form of a structural automaton model [12]. ASNA software performs the following functions:

1. The development of a model in the form of a graph of states and transitions is carried out on the basis of a structural and automatic model.
2. Forms a system of Kolmogorov-Chapman differential equations on the basic of a state graph.
3. The result of solving the system of differential equations is the distribution of probabilities of staying in states.
4. Defines reliability, functionality and efficiency indexes using the necessary states.

Therefore, it is necessary to develop a structural automaton model of GSC reaction to MO crossing control zone. Structurally automatic models of the reaction of the GSC with the installation in the control zone of one and two SSs are developed and presented by the authors in articles [7, 9]. This article presents the structural automatic model of the GSC reaction with the establishment in the control zone of three SDs and with the majority principle for decision-making on the type of MO with the rule of voting ″2 of 3″.

1. Principle of GSC functioning with three seismic sensors

The GSC include three seismic sensors with autonomous systems DOCTRS, system of receiving and displaying information. The block diagram of the GSC is shown in Fig. 2. In the system of receiving and displaying information the majority principle with the election rule ″2 from 3″ is used for taking decision about correct classification of MO type. GSC reaction to crossing control zone by MO is presented in the following way. An object is moving across control zone.

Around the control zone there are three seismic sensors (SS1, SS2, SS3) which have to react to MO appearance in control zone. Each sensor with a defined probability can detect or not detect the moving object. That is, all three sensors, only two sensors or only one sensor could react to the moving object. It is also possible that none of the sensors could detect the moving object. It is specified by several factors, such as different distance of MO movement from the sensor, ground condition, specific character of landscape of the control zone, way of MO moving, etc. After reaction to MO, the sensor's autonomous system DOCTRS starts to perform classification of the object. Classification may be correct or incorrect. After classification procedure the autonomous system transmits a message about MO type to system of receiving and displaying information. However, the message can be delivered or not delivered. It should be noted that the majority element will be able to send a correct message about MO type only in that case if there are signals with correct classification from three or two autonomous systems DOCTRS. If there is no signal from one of SSs, and two other signals come with correct and incorrect MO classification, then in RDI there is a message "MO type is not determined".

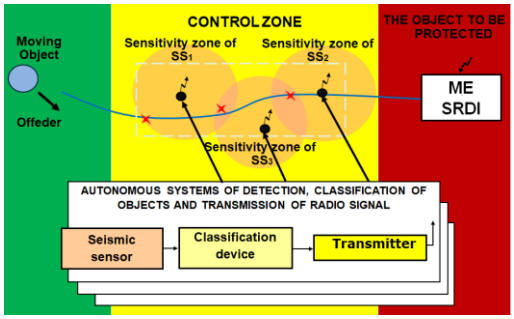


Fig. 2. Layout of three seismic sensors in control zone near the MO route and structure of autonomous system DOCTRS for each of them

2. Development of the structural automaton mode model of GSC reaction to MO crossing control zone with majority principle of taking decision about correct classification of MO type

The following procedures that form the GSC behavior in the process of crossing control zone by moving object are taken into account in the developed model.

Procedure 1. Detecting of moving object by seismic sensor.

A moving object may be detected or not detected, however, autonomous system DOCTRS with SSs in control zone, is in good order and ready to work. A moving object may not be detected in the following cases: it passed a seismic sensor at a safe distance; a moving object used special equipment that cannot be detected or was wearing special uniform; unsuitable place of seismic sensor location.

Procedure 2. Classification of moving object.

Alternative events are inherent in classification procedure, that is, it can be performed correctly or incorrectly. The error in classification may be caused by unsound method of processing seismic signal in autonomous system DOCTRS. Message with the result of MO classification is delivered to the system of transmitting radio signals.

Procedure 3. Delivering a message with information about MO to RDI.

The process of delivering radio signal about MO may be successful or not. Failure of message delivery to RDI may be caused by conditions of radio-wave transmission, presence of radio interference of natural and man-caused character.

Procedure 4. Taking decision about type of moving object in RDI using majority principle.

Correct decision in RDI is taken in the following situations:

1. When a signal from all three sensors came to ME, though MO was classified correctly.
2. A signal with correct classification came to ME from the first and second sensors. The third sensor either did not detect MO, or there was no message from it with correct classification, or a message was delivered with incorrect classification.
3. A signal with correct classification came to ME from the first and third sensors. The same situation is described in paragraph 2 for a signal from the second sensor.
4. A signal with correct classification came to ME from the second and third sensors. The same situation is described in paragraph 2 for a signal from the first sensor.

In order to develop structural automaton model [1] of GSC reaction to crossing control zone by moving object it is necessary to work out a reference graph of states and transitions [10]. To do this it is necessary to set values of system parameters and specify basic events which represent all processes and procedures that are included in the algorithm of its behavior. And also internal and external processes with which every GSC channel interacts during all operation period. It is also necessary to substantiate the components of state vector that will represent a state of GSC reaction.

With the help of presenting GSC procedures make a list of events taking place in the given complex. Events must be presented in pairs showing the beginning and ending of corresponding procedure. Pairs of events are shown in Table 1. Events which correspond to procedure ending are basic events for developing the model.

As far as duration of MO classification is less than the time of MO being in control zone, its value is taken equal to null. That is why events 1, 3, 5 will be considered to be basic ones, and events 2, 4, 6 will be brought into coincidence with them respectively. On that basis we will use the following basic events in developing structural automaton model:

- Basic event 1 ″Fixation of moving object by seismic sensor 1″ and brought into coincidence with it basic event 2 ″Ending of procedure of moving object classification by seismic sensor 1″.

- Basic event 3 ″Fixation of moving object by seismic sensor 2″ and brought into coincidence with it basic event 4 ″Ending of procedure of moving object classification by seismic sensor 2″.
- Basic event 5 ″Detection of moving object by seismic sensor 3″ and brought into coincidence with it basic event 6 ″Ending of classification procedure of moving object by seismic sensor 3″.
- Basic event 7 ″Ending of procedure of receiving 3-rd message by majority element RDI″.

Table 1. Presentation of pairs of events that fix the beginning and ending of procedures which form the GSC reaction to MO crossing control zone

No.	Event-beginning	Event-ending	Average duration of procedure
1	Beginning of procedure of detecting moving object in control zone	Basic event 1: detecting of moving object by seismic sensor 1	$1/\lambda_1$
2	Beginning of procedure of MO classification detected by seismic sensor 1	Basic event 2: ending of procedure of classification of moving object detected by seismic sensor 1	0
3	Beginning of procedure of detecting moving object in control zone	Basic event 3: detecting of moving object by seismic sensor 2	$1/\lambda_2$
4	Beginning of procedure of MO classification, detected by seismic sensor 2	Basic event 4: ending of procedure of MO classification, detected by seismic sensor 2	0
5	Beginning of procedure of detecting moving object in control zone	Basic event 5: detecting of moving object by seismic sensor 3	$1/\lambda_3$
6	Beginning of procedure of MO classification detected by seismic sensor 3	Basic event 6: ending of procedure of MO classification, detected by seismic sensor 3	0
7	Beginning of obtainment of three messages to the input of the majoritarian element of the system RDI	Basic event 7: ending of obtainment of three messages to the input of the majoritarian element of the system RDI	$1/\lambda_{1\text{-}3}$

In the model of GSC reaction to MO crossing control zone with majority principle of taking decisions about correct classification of MO type the following parameters are presented:
P_f – probability of fixation of moving object by seismic sensor,
P_{cc} – probability of correct classification of moving object,
P_d – probability of delivering the message about moving object,
λ_1 – intensity of event ″Detecting of MO by seismic sensor 1″,
λ_2 – intensity of event ″Detecting of MO by seismic sensor 2″,
λ_3 – intensity of event ″Detecting of MO by seismic sensor 3″,
$\lambda_{1\text{-}3}$ – intensity of event ″Operation of majority element″.

Let's present state vector of the system under study with the following components: V1 – shows the state of the first seismic sensor SS1 with classification result; takes values V1 = {0; 1; 2; 3}, where 0 – initial state, 1 – sensor fixed MO and classification device identified MO type correctly, 2 – sensor fixed MO, but MO classification was incorrect, 3 – sensor did not fix moving object. V2 – shows the state of the second seismic sensor SS2 with classification result; takes value V2 = {0; 1; 2; 3}, where 0 – initial state, 1 – sensor fixed MO, and device identified MO type correctly, 2 – sensor fixed but classification of MO was incorrect, 3 – sensor did not fix moving object. V3 – shows state of the third seismic sensor SS3 with classification result, takes value V3 = {0; 1; 2; 3}, where 0 – initial state, 1 – sensor fixed MO type, classification was correct, 2 – sensor fixed MO but classification was incorrect, 3 – sensor did not fix moving object. V4 – shows the result of receiving a message with information about moving object; takes value V4={0; 1; 2}, where 0 – initial state, 1 – a message is delivered, 2 – a message is not delivered. V5 – shows a state of carrying out a task and takes value V5 = {0; 1; 2}, where 0 – initial state, 1 – a task is carried out, 2 – a task is nor carried out.

Model of GSC reaction to crossing control zone with three SSs and with majority principle of taking decision about MO type is presented in the form of graph of states and transitions shown in article [12].

Formalized representation of a research object in the form of a structural automaton model, using the software ASNA shown in the Fig. 3–5.

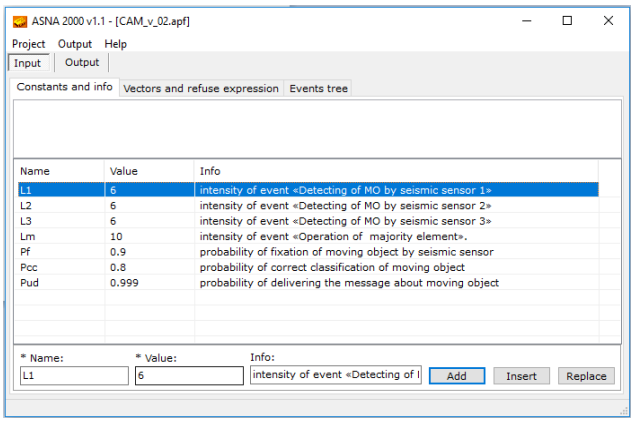


Fig. 3. Constants and info

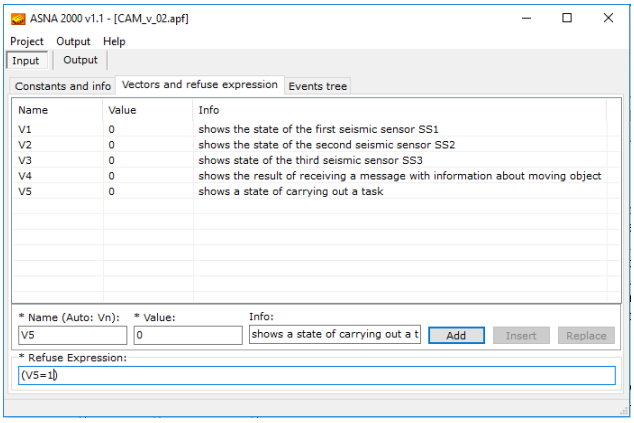


Fig. 4. Vectors and refuse expression

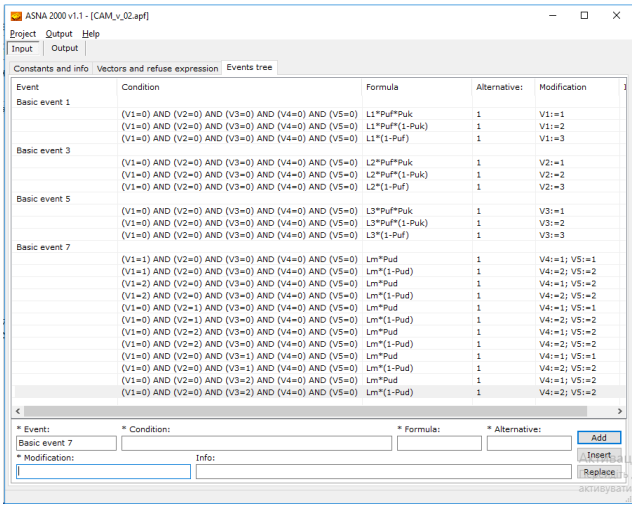


Fig. 5. Events tree

3. Comparative research of GSC effectiveness with three versions of seismic sensors layout in control zone

Comparative research was carried out with the following effectiveness factors of GSC constituents:

- Probability of fixation of moving object by seismic sensor – $P_f = 0.8$.
- Probability of correct classification of MO – $P_{cc} = 0.8$.
- Probability of delivering message about MO in RDI – Pd = 0.999. Research results are shown in Fig. 6.

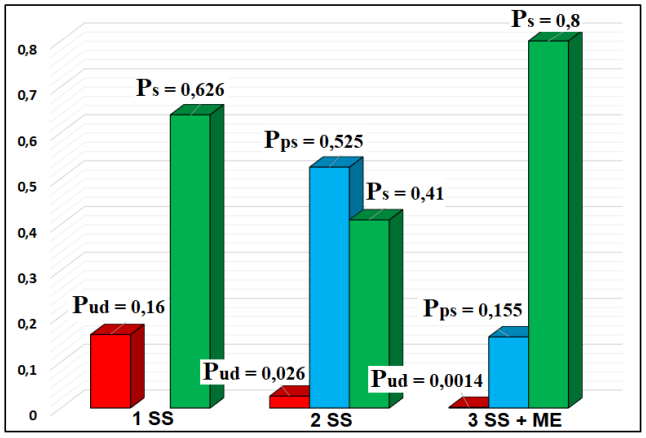


Fig. 6. Comparing of GSC effectiveness factors with the use of one, two or three seismic sensors in control zone

The results of the research confirmed the effectiveness of using GSC with three seismic sensors in control zone with the majority principle of taking decisions about MO type. The use of GSC with three SSs in control zone and with the majority principle of taking decisions about MO type in comparison with GSC with one or two SSs lessens the probability value of the user "deception" (Pud) by two orders (GSC with 1 SS) and by one order for GSC with two SSs.

The use of GSC with three SSs in control zone and majority principle of taking decision about MO type in comparison with GSC with two SSs and taking decision about MO type by signals coincidence lessens probability of partial fulfillment of the task 3 fold.

4. Conclusion

Unlike GSC with one or two seismic sensors in control zone, GSC with three seismic sensors in control zone and use of majority principle of taking decisions in the system of receiving and displaying information, provides better fidelity in identifying a type of moving object. In this case it is not necessary to raise requirements to effectiveness in the method of classifying moving objects.

References

[1] Fedasiuk D. V., Volochiy S. B.: Method of development of structural automaton models of discrete continuous stochastic systems. Radioelectronic and computer systems 6(80)/2016, 24–34.
[2] Koren Israel, Krishna C. Mani: Fault tolerant systems. Morgan Kaufmann Publishers is an imprint of Elsevier, 2007.
[3] Pricon. Technical Information – Resource access mode http://www.signalsecurity.gr/html/pdf/brochures/psicon_brochure.pdf [1.10.2018].
[4] Quantum multichannel seismic-acoustic system – Resource access mode: http://qtsi.com/wpcontent/ [1.10.2018].
[5] Shooman M. L.: Reliability of Computer Systems and Networks: Fault Tolerance, Analysis, and Design. John Wiley & Sons, Inc., New York, 2002.
[6] Volochiy B. Y.: Assessment of potential capabilities of guard signaling complex using seismic sensors. IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT) 2018, 435–441.
[7] Volochiy B. Y.: Modelling of reaction of guard signalling complex to moving object appearance by seismic sensors placed in far and close control zones. Military technical collection. Army Academy 1(10)/2014, 7–13.
[8] Volochiy B. Y.: Modelling the reaction of guard signaling complex on appearance of moving object when seismosensors are deployed in far and close control zones. Modern problems of radio engineering, telecommunications and computer science: proceedings of the International Conference TCSET'2014, Publishing National University "Lviv Politechnic", Lviv 2014, 252–254.
[9] Volochiy B. Y.: Research of the dependence of the guard signaling complex on the location of seismic sensors. Eastern-European Journal of Eenterprise Technologies 2/9(68)/2014, 54–60.
[10] Volochiy B. Y.: Technology of modelling algorithms of information systems behavior. NU "Lviv Politechnika", Lviv 2004.
[11] Zvezhynskii S. S.: Perimeter concealed seismic detection means. Special equipment 2/2004, 20–28.
[12] Zvezhynskii S. S.: Problem of choosing perimeter detection means. BDI 4(44)/2002, 36–41.

Prof. Bohdan Volochiy
e-mail: bvolochiy@ukr.net

Professor of Theoretical Radio Engineering and Radio Measuring Department of Lviv Polytechnic National University Experience of teaching in higher education for is over 40 years. Author of 280 publications, including 4 monographs, 3 textbooks, 5 inventions. He has prepared 4 Doctors of philosophy (Ph.D.). Research interests are: theory and practice of system design of radioelectronic information systems.

ORCID ID: 0000-0001-5230-9921

Ph.D. Mykhailo Zmysnyi
e-mail: zmysnyim@gmail.com

Senior Lecturer of Theoretical Radio Engineering and Radio Measuring Department of Lviv Polytechnic National University He has been teaching in higher educational establishments for over 5 years. The author of over 30 scientific publications. Research interests include the development of methods and tools for modeling functional and reliability behavior of fault-tolerant systems with majority structure.

ORCID ID: 0000-0002-3384-6139

Ph.D. Leonid Ozirkovskyy
e-mail: l.ozirkovsky@gmail.com

Associate Professor of Theoretical Radio Engineering and Radio Measuring Department of Lviv Polytechnic National University.He has been teaching in higher educational establishments for over 18 years. The author of 93 scientific publications and 32 scientific and methodical works, including 2 monograph, 5 textbooks. He has prepared 3 Doctors of philosophy (Ph.D.). Research interests include the development of methods and tools for modeling functional and reliability behavior of information systems

ORCID ID: 0000-0003-0012-2908

Volodymyr Onishchenko
e-mail: onishchenkovolodymyr@gmail.com

Senior researcher of Scientific Centre of Hetman Petro Sahaidachnyi National Army Academy.
Area of research: development and improvement of theoretical fundamentals of the structure analysis and synthesis, characteristics and mathematical software for systems (complexes, samples) of equipment and their components.

ORCID ID: 0000-0002-9616-838X

Ph.D. Yuriy Salnyk
e-mail: jurasalnik@ukr.net

Senior researcher, head of research department of Scientific Centre of Hetman Petro Sahaidachnyi National Army Academy Area of research: system engineering design and application of radio-electronic means of remote monitoring and security systems.

ORCID ID: 0000-0002-0772-6811

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8028

DEVELOPMENT OF THE REMOTE-PILOTED VEHICLE ALGORITHMIC SUPPORT AND ON-BOARD NAVIGATION COMPLEX STRUCTURE

Mykola Mykyjchuk, Volodymyr Markiv
Lviv Polytechnic National University, Department of Measuring Information Technologies

Abstract. *The article dwells upon the peculiarities of remote piloted vehicles of on-board navigation complex. It is highlighted that total number of the remote-piloted vehicles have increased and as a result the question of their integration into common space with manned aircraft becomes urgent. It is possible only when specified quality of the remote piloted vehicle movement parameters have been determined, including accuracy and interference immunity. It has been highlighted that remote-piloted vehicle equipment is subject to stringent requirements for minimization of cost, mass and size characteristics and power consumption, which are often mutually contradictory, and their implementation in general leads to deterioration of accuracy. The problem of ensuring accuracy and noise immunity when using the general-purpose element base becomes urgent.*

Keywords: remote-piloted vehicle, on-board navigation complex, inertial navigation complex, algorithmic support

ALGORYTMICZNE SYSTEMY WSPARCIA BEZZAŁOGOWYCH STATKÓW POWIETRZNYCH I CECHY STRUKTURY POKŁADOWYCH SYSTEMÓW NAWIGACYJNYCH

Streszczenie. *W artykule poruszono zagadnienie pokładowych systemów nawigacyjnych w bezzałogowych statkach powietrznych. Należy podkreślić, że liczba bezzałogowych statków powietrznych wzrosła obecnie na tyle, że integracja bezzałogowych oraz załogowych statków powietrznych we wspólnej przestrzeni powietrznej stała się ważnym i palącym problemem. Jest to możliwe jedynie pod warunkiem uzyskania najwyższej jakości parametry ruchu bezzałogowych statków powietrznych, z odpowiednią dokładnością i odpornością na zakłócenia. W artykule podkreślono, że wyposażenie bezzałogowych statków powietrznych jest pod ciągłą presją w zakresie minimalizacji kosztów, przystosowania do produkcji masowej, minimalizacji wielkości oraz zużycia energii, które często są ze sobą sprzeczne i implementacja których prowadzi do pogorszenia dokładności nawigacji oraz odporności na zakłócenia. Tym samym, zagadnienie zapewnienia precyzji oraz odporności na zakłócenia podczas użytkowania elementów ogólnego przeznaczenia w pokładowych systemach nawigacyjnych w bezzałogowych statkach powietrznych stało się palącym problemem.*

Słowa kluczowe: bezzałogowy statek powietrzny, pokładowy system nawigacyjny, bezwładnościowy system nawigacyjny, wsparcie algorytmiczne.

Introduction

The theoretical and practical value of the remote-piloted vehicles use is concerned with the development of methods of improvement control quality based on different methods and approaches. They will help to provide better quality of remote-piloted vehicles control.

Robotic technologies are increasingly used in various spheres of life. At the same time, the market of robotic aircraft is developing very dynamically. As the total number of unmanned aerial vehicles grows, it becomes urgent to integrate them into a common space with manned aircraft, which can only be solved if the specified quality of remote-piloted vehicles movement parameters is achieved, including accuracy and noise immunity. Operational standards are under development by authorized civil aviation authorities and are likely to repeat similar requirements for onboard equipment of the civilian aircraft.

Currently, international and domestic markets do not offer specialized serial navigation systems for unmanned aircraft, primarily of the small and middle classes, which meet the requirements for safety of flights in the general airspace. Tight requirements for cost minimization, weight and size characteristics and power consumption are often mutually contradictory, and their performance in general leads to a deterioration in accuracy and noise immunity. The developers of the remote-piloted navigation equipment are faced with the problem of ensuring accuracy and noise immunity when using the general-purpose element base [1, 16, 20].

To achieve this goal it is important to develop the concept of constructing of the on-board navigation complex and to increase noise immunity with variable structure.

It is necessary to investigate problem of development of algorithmic software and hardware solutions that provide increased accuracy of the flight and the peculiarities of the on-board navigation complex.

The basic requirements for accuracy and range of measurement parameters of orientation and navigation of on-board navigation complexes meet requirements of the inertial satellite navigation system.

1. Peculiarities of on-board navigation complex construction

Consistent improvement of techniques, technologies, production development and increased efficiency dramatically changes value of orientations. The main problem concerning remote-piloted vehicles is precise control and data transmission at long distances, because remote-piloted vehicles are under the influence of a variety of factors during the flight:

- weather (temperature, wind direction, humidity),
- the level of radio barriers,
- relief of the region,
- atmospheric pressure.

Due to the fact that it is difficult to evaluate these factors the accuracy of remote-piloted vehicles metrology supply, processing speed and data transmission are very important.

It is extremely important to improve the control and localization of remote-piloted vehicles by improving metrological measurements and the reduction of the metrology errors.

In navigation system there is the place for the movement of controlled objects, which are described in relation to a predetermined reference system. The motion of controlled object in any reference system is always purposeful (the trajectory, speed, acceleration and direction of travel are calculated). In real conditions, the various forces and moments disturb (deterministically and randomly) the moving object [4, 7]. Therefore, the navigation process in the systems of traffic control has the following stages [1, 14, 15]:

- determination of the required law of motion of the object;
- measurement of the characteristics of the actual motion of the object;
- comparison of the parameters of the actual position of the object with the given one and the definition of the deviation;
- the development of control signals (commands) and their transmission to the executive authorities in order to fulfill the prescribed law of motion.

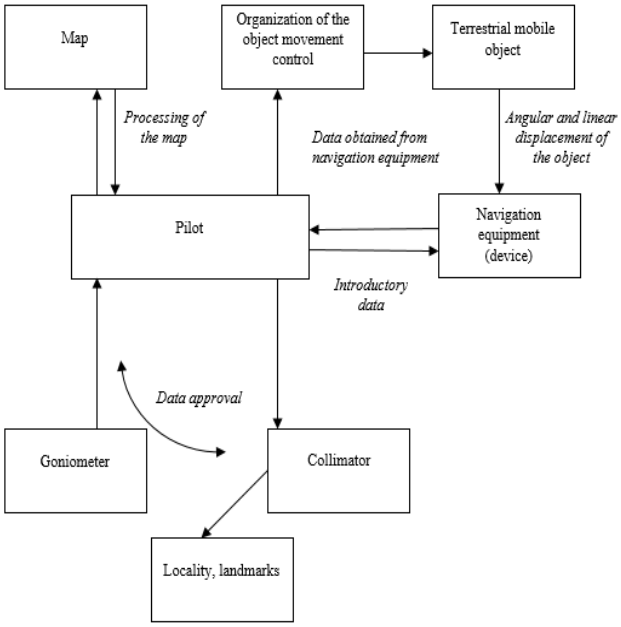


Fig. 1. General scheme of navigation control of terrestrial moving object

General scheme of navigation control of terrestrial moving object is depicted in Fig. 1. This process is as following [14]:

- the commander (driver), using sighting and angle devices, orientation on the ground and the map, prepares the initial data, which is involved into the navigation equipment;
- in the process of motion of an object, navigation equipment automatically determines its linear and angular displacements and produces navigation information that the commander (the driver) takes when choosing a further route of the movement.

For navigational control at the present stage, various technical means and physical principles are used. To obtain navigation information in order to control the movement of the center of mass of an object on its board, various technical means of navigation are set-up.

Specificity of remote-piloted application without terrestrial support under low-altitude flight, while reducing the visibility of signals of navigation satellites of global navigation satellite systems, complicates the problem of accuracy and provision of interference immunity [13, 17].

There are two possible ways of the solution of the problem [14]:

- use of the equipment used in manned aircraft (advantage of this approach is the use of waste products and technologies, disadvantage is ignoring the remote-piloted vehicle specificity that makes practically impossible to use it as a part of the remote-piloted vehicle of small and middle classes);
- creation of specialized navigation systems of small and middle-class remote-piloted vehicle in which it is possible to apply inexpensive general-purpose sensors (possibility to maintain algorithms software, and apparatus core of remote-piloted on-board navigation complex based on to the inertial navigation system).

It is necessary to substantiate the concept of construction, development of software and algorithmic support and hardware solutions to improve accuracy, as well as to study the properties of on-board navigation complex with high interference immunity and with variable structure for remote-piloted vehicles [1, 6, 19].

At the first stage of development of on-board navigation complex it is necessary to justify its structure in accordance with the requirements for the accuracy of determining orientation and navigation parameters, to propose the concept of the development procedure, operation modes, to develop the structure of algorithms for complex information processing, orientation and navigation algorithms, to justify the choice of the hardware of on-board navigation complex.

The basic requirements for the accuracy and range of measurements of orientation and navigation parameters of the on-board navigation complex correspond to the requirements of the inertial-satellite navigation system [14].

The traditional approach to the design of the remote-piloted vehicles involves the following:

- selection of measurers,
- development of orientation and navigation algorithms based on measurer values,
- development of algorithms for complex processing of information.

Such an order is used because of the small volume of sensors and high cost of products. However, the reverse order is more flexible and productive:

- requirements,
- modes of operation,
- measurements,
- development of algorithms for complex processing of information,
- development of orientation and navigation algorithms based on measurer values,
- selection of measurers.

It is more convenient due to the lack of mandatory requirements for remote-piloted vehicle navigation equipment and, accordingly, the absence of the need for certification, which gives advantages in the development of algorithmic support and choice of measurers.

When creating remote-piloted vehicle, it is necessary to take into account such critical factors as ensuring the accuracy levels established by the technical specification while minimizing the cost, weight and size characteristics and energy consumption. From this point of view, the optimal method for constructing on-board navigation complex is integration into single complex of sensors and systems with the integration of measurement information. The core of on-board navigation complex should be built on the basis of the free-form inertial navigation system. In order to ensure the piloting tasks, the on-board equipment includes system of air signals [1, 21, 22].

On the basis of the air signals system and magnetic compass air course counting is performed, which together with the inertial calculation allow to obtain comprehensive solution in an autonomous mode (without the use of GNSS). It is important to include in the on-board navigation complex receiver of GNSS signals. Thus, the ideology of constructing of the on-board navigation complex initially consists in the integration of measurements from the sensors and systems that make up its structure. On-board navigation complex directly consists from inertial sensors, GNSS and magnetic compass receivers and also from interface with air signal system [9, 14, 21].

Specific types of sensors and systems are selected in accordance with the requirements of software and algorithmic support of on-board navigation complex.
It is based on the sequence of stages:

- requirements,
- operating modes,
- measurements,
- development of algorithms for complex processing of information,
- development of orientation and navigation algorithms (based on measurer values and selection of measurers),
- structure of on-board navigation complexes.

There are various types of remote-piloted vehicles control system (Fig. 2) [10, 11, 15]:

- manual control when remote-piloted vehicle is under the control of a pilot,
- automated control when the control is made on the basis of the telemetric data,
- automatic control when the control is made on the basis of on the system chosen parameters.

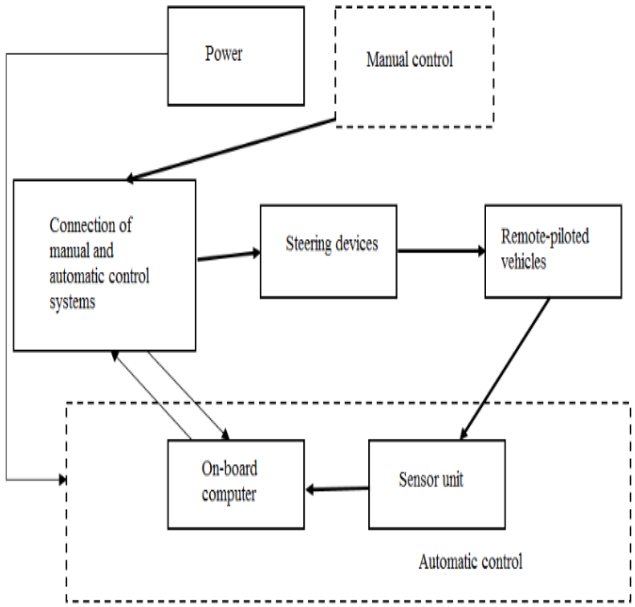


Fig. 2. Control system structure of remote-piloted vehicle

Figure 2 depicts the overall structure of the control system for drone. In particular, the sensor unit contains an inertial module, a tri-axial magnetometer, a receiver of satellite navigation signals, receivers of static and dynamic pressure, an ultrasound altimeter.

2. On-board navigation complex modes

The navigation complex is the set of on-board measuring tools and calculators, allowing to determine the location and speed of the aircraft relative to the Earth. None of the existing navigation meters can completely solve these problems, since each of them separately does not provide the necessary accuracy, noise immunity or reliability.

The tasks solved by the navigation complex are diverse. Among them, one of the most important is the reckoning of the path, providing a continuous measurement of the object coordinates. The main disadvantage of the number systems is the deterioration of the accuracy of determining the coordinates with an increase in the working time.

Therefore, to obtain the required accuracy, the numerical coordinates must be continuously or periodically corrected on the basis of information from radio engineering meters, i.e., comprehensive data processing should be carried out.

The basis of navigation complex is inertial navigation system on a gyro-stabilized platform. It measures both the angular position of the aircraft, and the components of acceleration and speed. The speed of the aircraft is also measured by using the airspeed sensor, which is part of the air signal system. The height and speed of its change are measured with the aid of radio altimeter. The signals of these devices are processed in computing device, which is part of a distributed on-board computing system [8, 18].

The on-board navigation complex has three main modes: "Initial parameters", "Navigation", "Attitude-and-heading reference" and four auxiliary modes: "Test Control", "Deviation", "Calibration", "Axis Coordination" (Fig. 3) [14].

Modes "Deviation", "Calibration" and "Coordination of axes" can not work simultaneously and outside the mode "Attitude-and-heading reference". In the "Test-control" mode, the other modes don't work.

In the "Navigation" mode on-board navigation complex generates and outputs navigation parameters with the required level of accuracy and in full volume. For successful operation of the regime, stable reception of GNSS signals and/or air signal system is required. In the absence of reception of GNSS and/or air signal system during a time interval of a predetermined duration, on-board navigation complex switches to the operation mode of "Attitude-and-heading reference". In this mode, according to the navigation parameters, there are signs of failure, and the orientation parameters are determined with increased errors [5, 6].

In the "Attitude-and-heading reference" mode, the angular position of the remote-piloted vehicle relative to the vertical is determined by the measurements of the accelerometers. The error in determining the angles in this mode depends on the flight mode of the remote-piloted vehicle, the maximum accuracy is achieved with a flight close to a straight uniform. In the modes of intensive maneuvering of the remote-piloted vehicle, an error is accumulated, the magnitude of which depends on the duration and intensity of maneuvering. After the reduction of the effect of accelerations and angular velocities, the errors decrease. When start receiving GNSS signals again, the "Navigation" mode is restored. The proposed structure and logic of switching modes fully meets the requirements of the on-board navigation complex and makes it easier to conduct routine maintenance [1, 2, 3, 14].

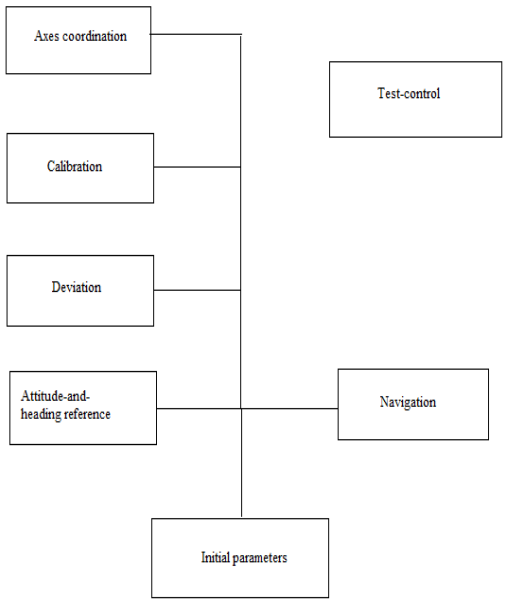


Fig. 3. On-board navigation complex operation modes structure

It is necessary to determine the structure of on-board navigation complex. It includes sensors and systems that allow constructing complex that meets the minimum requirements for determining the orientation and navigation parameters: the inertial module, the satellite navigation receiver GPS and magnetic compass (Fig. 4) [8, 1].

The output information of the inertial module are three projections of the absolute angular velocity of rotation and three projections of the apparent acceleration onto the orthogonal axes of the coordinate system. At the same time, it is required to provide a range of measured angular velocities.

The choice of the on-board navigation complex calculator is carried out taking into account the requirements for the interaction interfaces and the functional load [14, 18]:

- control, monitoring and signaling functions for all elements (basic and optional) that are part of the on-board navigation complex,
- collection, processing and transmission of data on internal and external data bases,
- functions for synchronizing the operation of sensors, systems and algorithms,
- implementation of algorithms for orientation, navigation and auxiliary algorithms,
- recording and storage in the non-volatile memory of initial and current settings of parameters of system.

Interfaces of on-board navigation complex interaction should provide reception and transfer of single discrete commands and data packets exchange with the on-board equipment and implemented navigation complex.

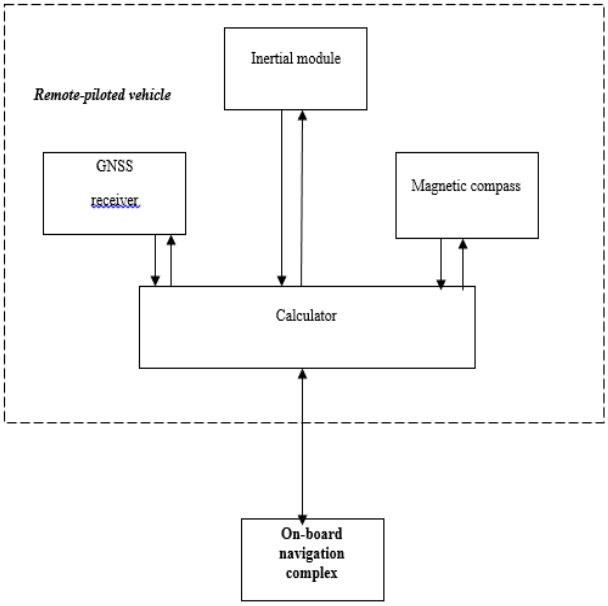


Fig. 4. General structure scheme of the navigation complex

The proposed structure allows to obtain the entire required volume of measurement information for the implementation of orientation and navigation algorithms. It is important to investigate in details the on-board navigation complex structure.

3. Conclusions

At the present stage of market economy development remote-piloted vehicles are intensively used in various spheres of scientific life. But there are some hindrances during their movement. Therefore, it is necessary to increase the height of the automatic flight and thus reduce the efficiency of the accomplishment of the tasks.

It is important to analyze the conception of constructing on-board navigation complex, that is the following: determination of requirements, operating modes, selection of measurement types, development of algorithms for complex processing of information, development of orientation and navigation algorithms based on readings of measurers, the choice of measurers.

Structures of algorithmic support of complex processing of information, solving problems of navigation and orientation of the on-board navigation complex are presented.

Hardware structure of the on-board navigation complex, which allows implementing the proposed algorithmic support structure, is proposed.

References

[1] Austin R.: Unmanned aircraft systems UAVs design, development and Deployment. John Wiley & Sons Ltd, West Sussex 2010.
[2] Bond L.: Overview of GPS Interference Issues. GPS Interference Symp., Volpe National Transportation System Center 1998, 28–32.
[3] Brown A.K., Yan Lu: Performance Test Results of an Integrated GPS/MEMS Inertial Navigation Package. ION GNSS 17th International Technical Meeting of the Satellite Division, Long Beach, CA, 2004.
[4] Forssel B., Olsen T.: Jamming Susceptibility of Some Civil GPS Receivers. GPS World 1/2003, 54–58.
[5] Grewal M.S., Weill L.R., Andrews A.P.: Global Positioning Systems, Inertial Navigation, and Integration. John Wiley & Sons, Inc., New York 2001.
[6] Key E.: Technique to Counter GPS Spoofing. Int. Memorandum, MITRE Corporation, 1995.
[7] Kim J.-H., Sukkarieh S.: Flight Test Results of GPS/INS Navigation Loop for an Autonomous Unmanned Aerial Vehicle (UAV). ION GPS, 24-27 September 2002, Portland, OR, 2002.
[8] Lawrence A.: Modern Inertial Technology (Navigation, Guidance, and Control). Springer-Verlag Inc., New York 1998.
[9] Markiv V.: Analysis of remote-piloted vehicles use and control system description. Computer sciences and information technologies 843/2016, 347–351.
[10] Markiv V.: Justification of remote-piloted vehicles use and metrology supply improvement. 5th Int. Scientific Conf. ICS-2016, 20–21.
[11] Martin M.: Non-linear DSGE Models and The Optimized Central Difference Particle Filter 2010, 2–45.
[12] Mukyichuk M., Markiv V.: Peculiarities of Remote-Piloted Vehicles On-Board Navigation Complex Construction. The II International Conference on Computational Linguistics and Intelligent Systems – CoLInS 2018. Lviv, 2018, 161–170.
[13] Mykyichuk M., Markiv V.: Metrology tasks of airphotoshooting by remote-piloted vehicle**.** Radioelectronics and telecommunications 874/2017, 57–61.
[14] Mykyichuk M., Markiv V.: Peculiarities of fractal analysis of remote-piloted vehicles recognition. VI International Scientific Practical Conference Practical Application of Nonlinear Dynamic Systems for Infocommunication, Chernivtsi 2017, 20–21.
[15] Mykyichuk M., Markiv V.: Peculiarities of the radio signals and hindrances in the navigation system of the remote-piloted vehicles. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOS 8(1)/2018, 40–43.
[16] Neitzel F., Klonowski J.: Mobile 3d mapping with a low-cost UAV system. Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-1/C22, 67–70.
[17] Roach D.: Dimensionality analysis of patterns: fractal measurements. Computers Geosciences 1993, 849–869.
[18] Salychev O.S.: Applied Inertial Navigation: Problems and Solutions, BMSTU, 2004.
[19] Sandau K.: Measuring fractal dimension and complexity – an alternative approach with an application, 1993, 164–176.
[20] Savage P.G.: Strapdown Analytics Part 1&2. Strapdown Associates, Inc., Maple Plain 2000.
[21] Tsui J. B.-Y.: Fundamentals of Global Positioning System Receivers. A Software Approach. John Wiley & Sons, Inc., New Jersey 2005.
[22] Winkler S., Schulz H.-W., Buschmann M., Vorsmann P.: Testing GPS/INS Integration for Autonomous Mini and Micro Aerial Vehicles. ION GNSS 18th International Technical Meeting of the Satellite Division. Long Beach 2005.

Prof. Mykola Mykyjchuk
e-mail: mykolam@ukr.net

Director of Computer Technologies, Automation and Metrology Institute, Professor at the Department of Measuring Information Technologies. Research field covers wide variety of methods and means of metrology supply improvement based on metrology indices, development of methods and means of operational control of industrial errors. Author of nearly 150 publications in this research area.

ORCID ID: 0000-0001-8343-3298

M.Sc. Volodymyr Markiv
e-mail: vovamarkiv6230@gmail.com

Ph.D. student at the Department of Measuring Information Technologies, specialization in 05.01.02 "Standardization, certification and metrology supply". Research field covers development of theoretical and methodical bases of the quality characteristics of products and services and the development of metrological supply for measurement and monitoring of certification. Author of nearly 20 publications in this research area.

ORCID ID: 0000-0003-0598-593X

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8032

METHOD OF SAFE LANDING THE EMERGENCY QUADROCOPTER

Bohdan Blagitko, Yuriy Mochulsky
Ivan Franko Lviv's National University, Faculty Electronics and Informational Technology, Department Radiofhysics and Computer Technology

Abstract. *By means of mathematical modelling, the main features of emergency landing of quadrocopter are determined. The results of simulate landings of the quadrocopter in cases of failure of one of four pairs of electric motor screw are given. The methods of safe landing of an unmanned quadrocopter in case of failure of one of four pairs of motor-screw are proposed. The basis of the proposed methods is to use a parachuting effect. Parachuting achieved by forced off the power of the motor, which is located at the opposite end of the same yoke as faulty motor. As a result, the vertical speed of the quadrocopter at the time of landing significantly decreases in comparison with the speed of free fall.*

Keywords: quadrocopter, 3-axis gyroscope, 3-axis accelerometer, 3-axis magnetometer

METODA BEZPIECZNEGO LĄDOWANIA AWARYJNEGO KWADROKOPTERA

Streszczenie. *Podstawowe cechy awaryjnego lądowania bezzałogowego kwadrokoptera zostały opisane metodami modelowania matematycznego. Zaproponowano metodę bezpiecznego lądowania bezzałogowego kwadrokoptera w przypadku awarii jednego z czterech par silnik-wirnik. Podstawą proponowanej metody jest zastosowanie efektu spadochronu. Efekt spadochronu jest osiągany poprzez przymusowe wyłączenia zasilania silnika elektrycznego, który znajduje się na przeciwległym końcu belki z uszkodzonym silnikiem. W rezultacie prędkość pionowa kwadrokoptera w momencie lądowania znacząco maleje w stosunku do prędkości swobodnego spadania.*

Słowa kluczowe: kwadrokopter, żyroskop 3-osiowy, 3-osiowy akcelerometr, 3-osiowy magnetometr

Introduction

In most articles on the problems of the quadrocopter flight, it is implicitly assumed that all four pairs of motor-screw and their control circuit are serviceable [5, 6, 8, 9]. In practice, often there are those or other faults in the sequence of links: the control circuit – the motor-screw. When a malfunction suddenly appears during the quadrocopter flight, then an emergency situation occurs.

In this article through mathematical modeling the basic features of quadrocopter emergency landings are considered. The results of mathematical modelling of quadrocopter emergency landings after the appearance of failure of one of the four electric motor-screws are presented.

1. Simulation of the quadrocopter landings in the event of the failure of one of the four electric motor-screw

The volume of the whole article should include an even number of pages. The last page should be filled at least 50%. The author should make efforts to fill a total volume of pages with formatted article. We propose to write the text directly on this template, it will help you to meet the editorial requirements. Please do not break the text – do not use non-breaking spaces, hard (manual) line breaks, etc. The Editor reserves the right to editorial redaction of the submitted texts.

2. Equations

The mathematical model of a quadrocopter in flight in the event that all four pairs of the electric motor screw and the control scheme of them are regular are considered as two essentially different mathematical models. Actually the mathematical model of the quadrocopter is the UAV the inputs of which are the values of the voltage supply of electric motors, and the outputs are the altitude and angles of the roll, pitch and yaw. The mathematical model of the PD regulator is a system in which inputs a measured value from the sensors of height, the angles of the pith, the roll and the yaw, and the microcontroller implements the PD algorithm and outputs the voltage values of the electric motors. The mathematical model of the quadrocopter in flight in the event that all four pairs of electric motor screw and the control scheme are useful, is described by the system of differential equations of the 16th order in the form of Cauchy [6].

In case of failure of one of the four pairs of electro-motor-screw quadrocopter, the following cases are possible: failure of the nasal steam motor-screw; failure of the tail-pair motor-screw: failure of the right pair of electric motor-screw; failure of the left pair motor-screw. The simulating of malfunctioning was performed by switching off the power of the faulty motor.

In order to solve the differential equations of the mathematical model of the quadrocopter, the implicit Runge-Kutta method TR-BDF2 was used (in the first stage of the equation integrated by the trapezium formula method, and in the second stage of the second order (differentiation formula back is used)).

The landing of the quadrocopter occurs after it has been suspended at a height of 50 m above the point with coordinates X = 0, Y = 0. The quadrocopter "nose" leads in the direction plus by axis X. In this case, the initial values of the angles of the pitch and the roll are zero, and the angle of yaw is 90^0. Simulation of the failure of a specific pair of electric motor screw is carried out by assigning the supply voltage of the corresponding motor to a value of 0.0 V for all moments of time **t** => **ta**, where **ta** is the moment of failure of a specific pair of the electric motor-screw of the quadrocopter.

In all articles [5, 6, 8, 9] about the simulations, the quadrocopter motions are considered as movements of the material point of a mass m. But the quadrocopter has a complex spatial structure with certain dimensions. In order to see not only the movement of the center of the masses of the quadrocopter but also the idea of its rotation during an emergency landing, the movements of the two points of the quadrocopter, the center of mass and the nasal engine, were simulated. On Fig. 1 and on Fig. 4 the dashed line shows the trajectory of the center of the masses of the quadrocopter, and the squares of the trajectory of the movement of the nasal engine.

The results of mathematical modeling of the process of landing a quadrocopter in the event of a failure of one of the four pairs of the electric motor-screw of the quadrocopter are given in the form of a graph of trajectories of motion in Fig. 1. Trajectory 1 reflects the process of landing a quadrocopter in the failure of the nose pair of electric motor screw, and trajectory 3 – in the event of failure of the tail pair the motor-screw.

As can be seen from Fig. 2, after the accidental disconnection of the nasal electric motor, the pitch angle begins to decrease (the quadrocopter "lowers" the nose). As shown in [8], the nasal and tail motors create the momentum of the forces that the quadrocopter rotates in relation to the vertical axis clockwise (if viewed from the bottom). The right and left motors rotate the quadrocopter counter-clockwise relative to this axis. Therefore, when the nasal engine fails, the quadrocopter starts to rotate counter-clockwise. The Pitch and Yaw create a gyroscopic moment of forces, which begins to rotate the quadrocopter relative to the longitudinal axis, that is, there is a roll. Having made a little more than one turn relative to the transverse axis, the quadrocopter starts to rotate in the opposite direction, and in the future the Pitch

ranges in the vicinity of -150º. As for the longitudinal axis, the quadrocopter makes almost five turns in the negative direction and falls to the ground by the right engine.

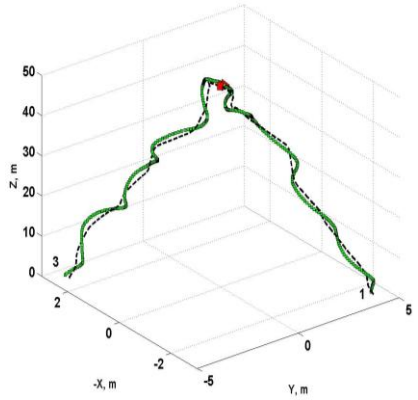
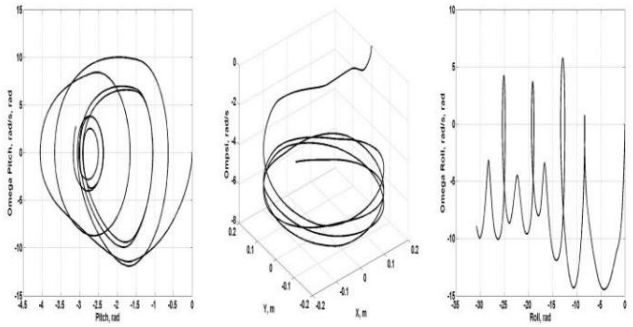


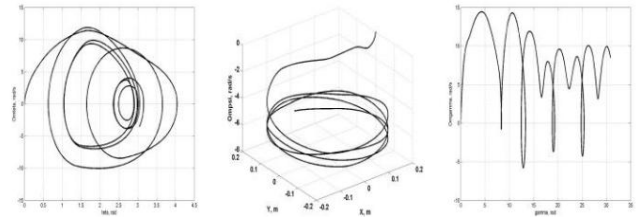
Fig. 1. Trajectories of the motion of the quadrocopter with the regulator: 1 – in case of failure of the nose pair motor-screw, 3 – in case of failure of the tail motor- screw



a) Pitch b) Yaw c) Roll

Fig. 2. Phase portraits of pitch, yaw and roll angles when the nose engine fails

Phase portraits of the angles of the Pitch, of the Yaw and of the Roll after having at the moment **ta =** 1 s rejected the tail engine, drawn on Fig. 3.



a) Pitch b) Yaw c) Roll

Fig. 3. Phase portraits of pitch, yaw and roll angles when the tail engine fails

As can be seen from Fig. 3, after the emergency shutdown of the tail motor, the pitch angle begins to increase (quadrocopter "lowers" the tail). As indicated above, the nasal and tail motors rotate the quadrocopter relative to the vertical axis in the same direction. But the angular Pitch speed changed the direction of rotation to the opposite. Therefore, the gyroscopic moment of forces causes the quadrocopter's Roll in the opposite direction. By making more than one turn in a positive direction relative to the transverse axis, the quadrocopter starts to rotate in the opposite direction, and in the future the pitch fluctuates in the vicinity of + 150º. With regard to the longitudinal axis, the quadrocopter makes almost five turns in a positive direction and falls to the ground on the left engine.

The results of mathematical modelling of the process of landing of the quadrocopter in the event of failure of one of the other pairs of the electric motor-screw quadrocopter are given in the form of a graph of trajectories of motion in Fig. 4. Trajectory 2 reflects the process of landing the quadrocopter with the failure of the right pair motor-screw, and trajectory 4 – in the case of failure of the left pair motor-screw.

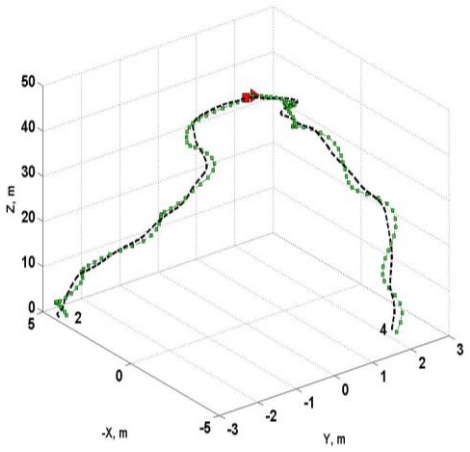
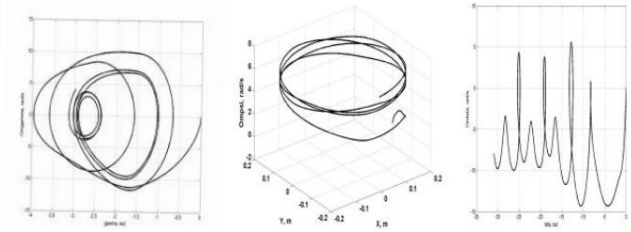


Fig. 4. Trajectories of the motion of the quadrocopter with the regulator: 2 – in case of failure of the right pair of electric motor- screw, 4 – in the failure of the left pair motor-screw

Phase portraits of the angles of the Roll, of the Yaw and of the Pitch after having at the moment **ta** = 1 s rejected the right engine are shown in Fig. 5.

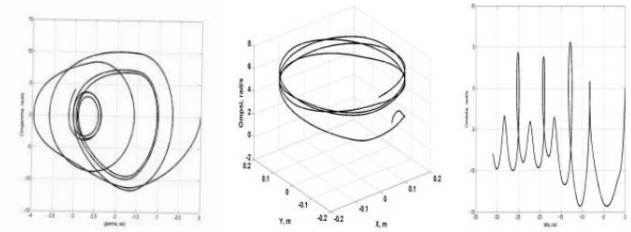


a) Pitch b) Yaw c) Roll

Fig. 5. Phase portraits of roll, yaw and pitch angles when the right engine fails

As can be seen from Fig. 5, after the trip of the right electric motor, the angle of the roll begins to increase (the quadrocopter "drops" to the right). The right and left engines create moment of forces, which rotates the quadrocopter relative to the vertical axis counterclockwise (if you look from below). Therefore, when the right engine fails, the quadrocopter starts to rotate clockwise. The Roll and Yaw create a gyroscopic moment of forces, which begins to rotate the quadrocopter relative to the transverse axis, that is, there is a pitch. Having made a little more than one turn relative to the longitudinal axis, the quadrocopter begins to rotate in the opposite direction, and in the future the roll varies in the vicinity of + 150º. As for the transverse axis, the quadrocopter makes almost five turns in a positive direction and falls to the ground on the nasal motor.

Phase portraits for changing the angles of the roll, yaw and pitch after having turned off the left engine at the time **ta** = 1 s, are shown in Fig. 6.



a) Pitch b) Yaw c) Roll

Fig. 6. Phase portraits of roll, yaw and pitch angles when the left engine fails

As can be seen from Fig. 6, after the emergency cut-off of the left electric motor, the angle of the roll begins to decrease (the quad-compressor "falls" to the left). At the same time, as a result of uncompensated moments of forces relative to the vertical axis, the quadrocopter starts to rotate clockwise. The gyroscopic moment of forces causes the quadrocopter Pitch. Having made a

little more than one rotation relative to the longitudinal axis, the quadrocopter starts to rotate in the opposite direction, and in the future the roll varies in the vicinity of -150º. As for the transverse axis, the quadrocopter makes almost five turns in the negative direction and falls to the ground on the tail engine.

It should be noted that the above-mentioned fall of the quadrocopter to the ground, when one of the engines fails, is valid only when a drop from a height of 50 m and zero initial angles of pitch and roll. Under other initial conditions, due to the rotation of the quadrocopter relative to all three axes, it is possible to fall to the ground in any position.

Emergency landing of a quadrocopter with failure of one of four pairs of electric motor-screw has significant disadvantages, namely: unpredictable landing point, unpredicted speed at the time of landing, unpredicted duration of fall and its nature, unpredicted landing angle, and the like.

2. Method of safe landing of quadrocopter with failure of one of the pairs of electric motor-screw

This article proposes a control method for the safe landing of a quadrocopter with a failure of one of the four pairs of electric motor-screw. The proposed method is to forcibly shut off the electric motor, which in paring with the already defective electric motor is located on the opposite end of the same console. This method uses the parachuting effect of the quadrocopter. Parachuting achieved by forced off the power of the motor, which is located at the opposite end of the same yoke as faulty motor As a result, the vertical speed of the quadrocopter at the time of landing significantly decreases in comparison with the speed of free fall.

For example, the simulation results of a safe landing of an emergency quadrocopter with a height of 50 m over a point with coordinates X = 0, Y = 0. Quadrocopter at the same time "hangs" motionless (the angles of the pitch and the roll are zero). The main parameters of the quadrocopter: mass – 468 g, distance from the center of mass to the engine – 22.5 cm, EMC battery – 11.1 V.

Phase portraits of the angle of Yaw for the proposed method of safe emergency landing of the quadrocopter are shown in Fig. 7 (in the of the failure of the nose engine and the instantaneous disconnection case of the tail engine) and in Fig. 8 (in the case of failure of the left engine and instant disengagement of the right motor).

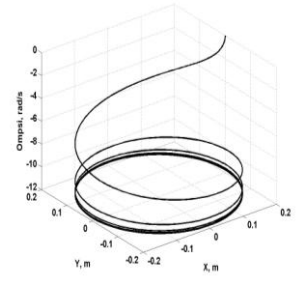


Fig. 7. Phase portrait of the angle of Yaw in the failure of the nose or tail engine

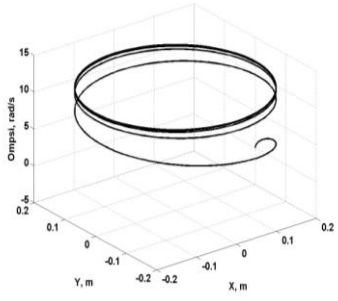


Fig. 8. Phase portrait of the angle Yaw at the failure of the right or left engine

Fig. 9 depicts a fragment of a safe emergency landing of a quadrocopter with a failure of the nasal engine and the instantaneous shutdown of the tail engine or with the cessation of the tail engine and the instantaneous disconnection of the nasal engine.

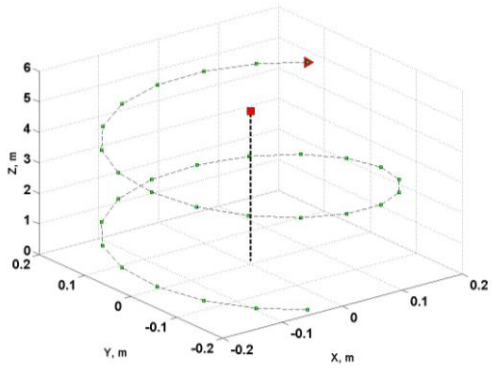


Fig. 9. Safe emergency landing of the quadrocopter at the failure of the nose or tail engine

Fig. 10 depicts a fragment of a safe emergency landing of a quadrocopter with a failure of the right engine and the instant disconnection of the left engine or the failure of the left engine and instant disabling of the right engine.

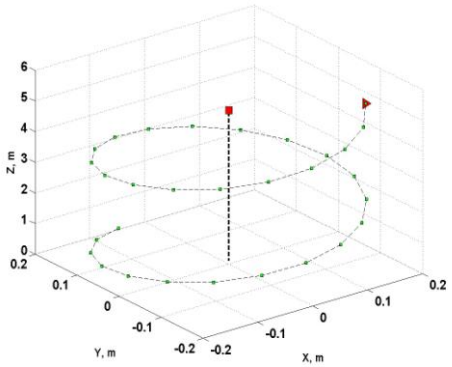


Fig. 10. Safe emergency landing of the quadrocopter with the failure of the left or right engine

For comparison: Fig.11 shows the process of landing a emergency quadrocopter in the event of a failure of one of the four pairs of motor-screw and the safe emergency landing of a quadrocopter with a failure of right engine and instantaneous shutdown of the left engine of a failure of left engine and instantaneous shutdown of the right engine.

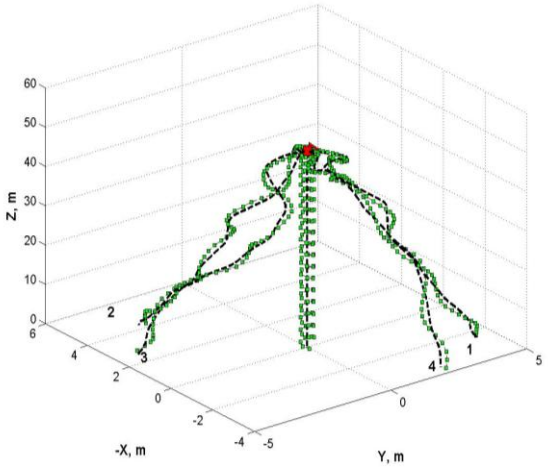


Fig. 11. Trajectories of the motion of the quadrocopter with the regulator: 1 – in case of failure of the nose pair motor-screw, 3 – in case of failure of the tail motor-screw, 2 – in case of failure of the right pair of electric motor-screw, 4 – in the failure of the left pair motor-screw, M-in case of safe emergency landing of the quadrocopter with the failure of the left or right engine

3. Discussion of the simulation results of the landing process of an emergency quadrocopter

When free fall all four engines are off, the quadrocopter reaches the ground in just 6 seconds. The vertical component of the landing speed at the same time is 16.5 m/s. In case of refusal of one engine and working three remaining during the time of reduction, there are moments of forces concerning the axes of pitch, roll and yaw.

Table 1 shows the simulation results of the landing process of an emergency quadrocopter in case of failure of one of the four motor-screw pairs. Due to this, the rate of change does not change monotonously and the vertical component of the landing speed decreases to 12.7 m/s. The roll and pitch change randomly during a fall and at the time of landing they take on different values depending on the fault.

Table 1. The simulation results of the landing process of an emergency quadrocopter in case of failure of one of the four motor-screw pairs

Engine failure	Pitch	Roll	Yaw	V_x, m/s	V_y, m/s	V_z, m/s
nasal motor	180º	40º	10º	0	1.47	-12.7
right motor	-47º	180º	160º	-1.05	0.59	-12.7
tail motor	180º	-33º	5º	0.03	-1.44	-12.7
left motor	28º	-170º	-180º	1.1	-0.6	-12.7

In the event of a faulty nasal or tail motor, the quadrocopter falls to the landing site with screws down (pitch = 180°) and with a roll of 40° and -33°, respectively. Non-zero values of the horizontal velocity component at the moment of landing are an additional danger V_y = 1.47 m/s and V_y = -1.44 m/s, respectively. Depending on the friction coefficient at the landing site, the quadrocopter will still travel some distance along the ground.

In the event of a faulty right or left motor, the quadrocopter falls to the landing site with screws down (roll = 180°) and with a pitch of -47° and 28°, respectively. Non-zero values of the horizontal velocity component at the moment of landing are an additional danger V_x = -1.05 m/s, V_y = 0.59 m/s and V_x = 1.1 m/s, V_y = -0.6 m/s, respectively. Depending on the friction coefficient at the landing site, the quadrocopter will still travel some distance along the ground.

Table 2 shows the simulation results of the landing process of an emergency quadrocopter in case of failure of one of the four motor-screw pairs and simultaneously with the engine that is out of order, turn off the power of the electric motor, which is located at the opposite end of the same traverse as the defective electric motor.

Table 2. The simulation results of the landing process of an emergency quadrocopter in case of failure of one of the four motor-screw pairs and simultaneously with the engine that is out of order, turn off the power of the electric motor, which is located at the opposite end of the same traverse as the defective electric motor

Engine failure	Pitch	Roll	Yaw	V_x, m/s	V_y, m/s	V_z, m/s
nasal motor	0º	0°	175°	0	0	-5.6
right motor	0º	0°	5°	0	0	-5.6
tail motor	0º	0°	175°	0	0	-5.6
left motor	0º	0º	5°	0	0	-5.6

If simultaneously with the engine that is out of order, turn off the power of the electric motor, which is located at the opposite end of the same traverse as the defective electric motor, then the time of fall is 11.7 s, and the vertical landing speed is 5.6 m/s. The horizontal components of the speed all the time will be zero, that is, the quadrocopter will fall vertically down. The roll and pitch during the fall are zero, that is, the quadrocopter will always land on the chassis. Due to the fact that in this case, not compensated moments of the forces of two opposite engines, the speed of rotation around the vertical axis at the time of landing will be 4.4 rad/s. This speed will create a distance of 22.5 cm from the center of mass of the acceleration 4.4 m/s^2.

4. Conclusion

The results of the conducted research determine the effectiveness and practical significance of the use of the proposed methods of safe landing of the quadrocopter, which in the flight unexpectedly refused one of the four pairs of propeller. To reduce the vertical component of the landing speed when the failure of one of the engines it is expedient as soon as possible to disconnect the voltage from the opposite engine, which is located on the opposite end of the console. In this case, the horizontal components of the speed are practically zero.

References

[1] Blagitko B., Mochulsky Y.: Emergency landing of quadrocopter in unmanned flight. Electronics and Information Technologies 8/2017, 137–142.
[2] Blagitko B., Mochulsky Y.: Mathematical modeling of quadrocopter stabilization in flight. Electronics and Information Technologies 3/2013, 96–107.
[3] Blagitko B., Zaiachuk I., Kit L., Mochulsky Y.: Modeling of the process of control of vertical take-off and landing of unmanned quadrocopter. Physical-mathematical modeling and information technologies 2/2012, 111–117.
[4] Blagitko B., Zaiachuk I., Mochulsky Y.: Mathematical modeling of the effect of sensor features on the flight process of an unmanned quadrocopter. Physical-mathematical modeling and information technologies 21/2015, 22–29.
[5] Huang H., Hoffmann G. M., Waslander S. L., Tomlin C. J.: Aerodynamics and control of autonomous quadrotor helicopters in aggressive maneuvering. IEEE International Conference on Robotics and Automation, May 2009, 3277–3282.
[6] Marcelo B.,.Sampaio R.C.B, Bounabdallah S., Perrot V., Siegwart R.: In-Flight Collision Avoidance Controller Based Only on OS4 Embedded Sensors. Journal of the Brazilian Society of Mechanical Sciences and Engineering XXXIV(3)/2012, 294–107.
[7] Mochulsky Y.: Matlab In Physical Research: tutor. – method. manual. VTS LNU them. Ivan Franko, Lviv 2004.
[8] Nartin P., Salaun E.: The True Role of Accelerometer Feedback in Quadrotor Control. IEEE International Conference on Robotics and Automation, Anchorage, May 2010, 1623–1629.
[9] Wissiere D., Bristeau P.-J., Martin A., Petit N.: Experimental autonomous flight of a small-scaled helicopter using accurate dynamics model and low-cost sensors. Proceeding of the 17th Word Congress The International Federation of Automatic Control. Seoul Korea 2008, 14642–14650.

Ph.D. Blagitko Bohdan
e-mail: blagitko@gmail.com

Associate-Professor of the Department Radiophysics and Computer Technology, Faculty Electronics and Informational Technology, Ivan Franko Lviv's National University.
His research and teaching interests are in areas microprocessor technique and systems, mainly concerned with mathematical modelling and simulation of processes in control dynamic systems, and computer-aided system analysis and design.

ORCID ID: 0000-0002-0516-9353

Ph.D. Mochulsky Yuriy
e-mail: mochulsky@ukr.net

Associate-Professor of the Department Radiophysics and Computer Technology, Faculty Electronics and Informational Technology, Ivan Franko Lviv's National University.
His scientific interests are mainly concerned with mathematical modelling and simulation of processes in dynamic radiophysical systems.

ORCID ID: 0000-0001-5488-4065

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8035

IMPROVEMENT OF METHODS OF MOTION COMPENSATION OF DYNAMIC OBJECTS MOVING IN VIDEO STREAM OF THE VIDEOCONFERENCING SYSTEM

Volodymyr Barannik[1], Mykola Dvorsky[1], Valeriy Barannik[2], Viktoria Himenko[2], Anton Sorokun[2]
[1]Ivan Kozhedub Kharkov National University of Air Force, Faculty of Automated Control Systems and Ground Support for Aviation Flights,
[2]Kharkov National University of Radio Electronics

Abstract*. Videoconferencing gives us the opportunity to work and communicate in real time, as well as to use collective applications, interactive information exchange. Videoconferencing systems are one of the basic components of the organization of manegment, ensuring, the timeliness and necessary quality management of the implementation of objective control over the solution of the tasks. The quality of the image and the time of transmission of video information is unsatisfactory for the quality control of the troops. Considered ways to increase the efficiency of management and operational activities, due to methods of compensation of motion, using technology to reduce the volume of video data for quality improvement.*

Keywords: videoconferencing, video stream, coding, motion compensation, troop control

POPRAWA METOD KOMPENSACJI RUCHU PORUSZAJĄCYCH SIĘ OBIEKTÓW DYNAMICZNYCH W STREAMIE WIDEO SYSTEMU WIDEOKONFERENCYJNEGO

Streszczenie. *Wideokonferencje dają możliwość pracy i komunikowania się w czasie rzeczywistym, a także korzystania ze zbiorowych aplikacji, interaktywnej wymiany informacji. Systemy wideokonferencyjne są jednym z podstawowych elementów organizacji zarządzania, zapewniając terminowość i niezbędne zarządzanie jakością w zakresie realizacji kontroli nad rozwiązaniem zadań. Jakość obrazu i czas transmisji informacji wideo jest niezadowalający dla kontroli jakości wojsk. Rozważono sposoby zwiększania efektywności zarządzania i działań operacyjnych, ze względu na metody kompensacji ruchu, z wykorzystaniem technologii zmniejszającej ilość danych wideo w celu poprawy jakości.*

Słowa kluczowe: wideokonferencje, strumień wideo, kodowanie, kompensacja ruchu, kontrola oddziałów

Introduction

Due to the rapid development of network and communication technologies, increased productivity of computers, and the need to process more and more information (both local and network and Internet), the role of software and hardware has increased. Remote access, distance education and management, as well as video conferencing facilities are experiencing a period of rapid growth. Bearing your appointment, ease and improve the effectiveness of interaction with both a computer and groups of people with computers integrated into the network.

Video conferencing gives us the opportunity to work and communicate in real time, as well as to use collective applications, interactive information exchange, videoconferences begin to be seen not only as something experimental, but also as a partial solution to the problem of automation of activity and gives a significant advantage over traditional solutions [16].

The basis of video streaming encoding algorithms is the redundancy of information and features of human perception of a visual image [9]. The image changes smoothly, small distortions when the image is restored become incomprehensible [13, 14]. So there were algorithms for encoding a video stream with a loss of quality. Such algorithms allow to efficiently encode the video stream, but do not guarantee accurate information recovery.

1. Videoconferencing

Videoconference is an area of informational technology that simultaneously provides two-way transmission, processing, transformation and representation of interactive information at a distance in real time with the help of hardware and software computer facilities.

Interaction in video conferencing mode is also called video conferencing session.

Videoconferencing is a telecommunication technology of interactive interaction between two or more remote subscribers, in which between them it is possible to exchange audio and video information in real time, taking into account the transfer of control data.

Table 1. Characteristics of values of average speed of stream of uncoded video information depending on spatial resolution and frame rate for videoconference complexes in departmental control systems

Levels of video quality	Number of rows	Vertical resolution	Frequency of frames per second	Average speed (Mbps)
CIF	320 – 352	240 – 288	24 – 30	66
SD	640	480 – 576	24 – 30	252
ED	720	480 – 576	50	500
HD	1280	720	50	1105
Full HD	1280 – 1920	720 – 1080	24 – 30; 50	1500; 2500
Advantage HD	1920 – 2048	1080	48; 60	2548; 3180

The presented assessments are based on the analysis of the main characteristics of telecommunication technologies used in the organization of videoconference in the system of troop control and assessments of the required volumes of video resource videoconference depending on the required spatial resolution of video and frame rates [1, 2]. With the growth of the volume of video information and the lack of provision of appropriate data on the productivity of technology for the transmission and processing of video information in the complexes of videoconference, there are conditions for violating the security categories of the video resource of the videoconference – accessibility and integrity.

The high quality of sound and full screen video, the operational exchange of documents and data (allowing video conferencing) gives this tool of distance communication the widest range of practical applications. To date, these systems have already been used in large companies, law firms, medicine and the armed forces.

After telephone conversations, the information received from the interlocutor of the audience, reaches, on average, up to 10% of the total amount of information.

With telephone conversations with the ability to exchange data, the amount of information collected can be increased to approximately 25%.

In case when it is possible during the conversation to visually monitor the interlocutor, it is possible to ensure assimilation of information up to 60%. But not only this statistics convinces us that the videoconference allows to provide a connection of a new level.

In practice, video conferencing is an indispensable assistant in the management and interaction of military units that are territorially diverse, advised, experienced in the management of troops, in providing moral and psychological support to troops in the zones of vision of hostilities, in telemedicine, in the transmission of audiovisual transmission from the battlefield. In this case, there is no need to send them to expensive business trips. Cost savings on business trips are not all advantages, the main thing – using video conferencing increases the efficiency of work.

Video conferencing can be arranged on wired and wireless communication channels [7, 20]. Leading channels differ from wireless data rates by cost of construction and ease of implementation.

Recently, Internet access systems based on the use of satellite systems have become widely distributed. Since transmission equipment is expensive, many companies use DirecPC systems. Most often when working with the Internet incoming traffic ("from the Internet") is much higher than the original. DirecPC systems use satellite antennas as receiving devices, but transmit information from users to the Internet via regular terrestrial communication channels. Conventional communications satellites are in stationary orbits at an altitude of approximately 35 kilometers above the Earth. New low-orbit satellites (or low-orbit satellites) operate in orbits close to the earth's surface, allowing them to receive signals even from low-power transmitters. Such satellites consume less power, and their launch and operation cost much less than traditional satellite communications systems.

The use of low-orbital satellites makes it possible to use communication services while in any part of the globe.

Technological implementation options, the telecommunication network usually contains a variety of hardware and software components that need to work together to transmit information. The various components of the network "communicate" with each other, following a series of rules that allows them to work all together. Such a set of rules that governs the process of transferring data between two network points is called a protocol. Each device on the network must correctly understand the protocol of another device.

The effectiveness of videoconference systems in this paper is considered as the efficiency of the transmission of video streams.

2. Methods motion compensation of dynamic objects moving in video stream

Methods of compression are divided into two groups: methods for loss of image quality and methods without loss of quality. Lossless image compression is a method where no loss of image quality is achieved compared to the original one. Unsafe image is mathematically identical to its original [18, 19]. Lossless compression typically provides lower compression ratios than compression loss.

Existing compression methods do not ensure the effective operation of objective video control systems [8, 11]. This is manifested in the delayed display of information on the operator's screen, in the low quality of the information received [12].

Motion compensation is one of the main algorithms used in the processing and compression of video data. The algorithm uses the similarity of neighboring frames in a video sequence and finds the motion vectors of individual parts of the image (usually blocks of 16 × 16 and 8 × 8). Use of compensation allows compression to increase the degree of compression by eliminating redundancy in the form of matching parts of frames. It is used not only for compression, but also for filtering video, changing the frame rate.

Motion compensation is one of the most important component part of the MPEG 1 and MPEG 2 standards. The method of compensation motion prediction can significantly reduce the temporal redundancy of video streams. If the next frame contains the shifted parts of the previous frame, then in this case, it is advantageous not to transfer the whole frame, but only information about the movement and changes of the shifted pixel.

The algorithms of compensation of motion can be carried out according to the following criteria:

- the analyzed element – a frame, blocks, or objects [10];
- type of movement – operations of parallel shifts, turns, scaling;
- measure of decision.

There are following methods of motion compensation:

- pixel method;
- object method;
- block matching method.

The pixel method is one of the earliest methods of motion compensation. Compensation is performed separately for each pixel of the frame, which is considered a class of transformations – linear shifts.

The approach is based on the assumption that the brightness can be approximated by a linear function from the point position in the frame. This assumption is valid only for a relatively small neighborhood of this point, which significantly reduces the scope of this method and allows him to correctly evaluate only small shifts. This restriction can be overcome by evaluating not the shear vector itself, but its difference with some prediction vector, which is located closer to the desired vector than zero [17]. In the general case, when the motion can make dozens of pixels, the vector of displacement is searched with the help of an iterative algorithm – at each step there is a refinement of the value found in the previous step. As an initial approximation, we can take the shear vector for the same point found when processing the previous frame. This method has a number of serious shortcomings, as a result of which at present it is purely theoretical interest and is practically not used anywhere. Its main drawbacks are high complexity, low accuracy and a large amount of information.

The object method is a precursor to the method of matching the blocks, where as a compensation element a rectangular block of the frame acts and the compensation is a rectangular block of the frame and linear shifts are considered.

When block approach (Object method) for each block compensation is made regardless of adjacent blocks. Proceeding from the fact that the motion of the scene carries out objects, it is necessary that compensation for all the blocks of the same object gave the same result. This can be done if one or another object of the scene moves to each frame unit. After that you will be able to estimate the total compensation error for all blocks.

The segmentation process can take place regardless of the process of finding the parameters of the movement, either both, and the other can be defined within a single process, which is repeated iteratively. In the first case, the basis for segmentation is usually high-quality information, in the second segmentation is made taking into account the found parameters of motion, which then specify. Sometimes frame segmentation on objects is used after determining the displacement vector for individual blocks in order to correct the found vector field.

Thus, the proposed method is proposed as follows:

1) Accumulate information about the background in a temporary plane (subject to the time frame for the accumulation of information is a sharp change in the background part of the scene) and accumulation of it in the form of a single background for some sequence of personnel. In addition, similar accumulation of information about moving objects of the scene can be similar, taking into account the nature of their movement;
2) Conducting simultaneous construction of a difference image of a scene with a given threshold and segmentation of allocated objects of motion (for segmentation, it is proposed to use information about the eight binding boundary of the pixel of the image). In addition, during the segmentation segment is

filtered left after a different comparison with the specified threshold noise.

The use of the proposed method will improve the quality of motion compensation by smoothing random emissions into the field of motion vectors inherent in block methods. Accumulation of information in a temporary plane will reduce the noise level of the scene, as well as increase the degree of compression of the video.

The method of matching the blocks, more precisely, the class of methods, is the logical consequence of the pixel method of dynamic objects, eliminates most of its disadvantages, since the unit of compensation in it adopted a rectangular block (usually a square 16×16 pixels or smaller). The motion is also sought in the class of linear displacements; therefore, this motion describes a two-dimensional bias vector for each block.

The basic assumption of the method – for the time that passes between two consecutive frames, the objects in the scene and their location vary slightly. Then in the vicinity of any point in the frame, this change with a fairly high degree of accuracy can be approximated by parallel transfer of this neighborhood to a certain vector. In fact, the vast majority of ordinary video sequences satisfy this limitation, with the exception of areas of abrupt change in the scene, that is, the nature of the motion of objects can be considered almost everywhere continuous.

Various modifications to this approach differ in how the minimum function of the compensation error is in the whole area. Checking only a few points in the area can localize this minimum. The algorithm by which these points are selected is called a template.

Methods based on templates show a good rate of performance, but often find a local minimum instead of a global one. It is advantageous to note that the search for the vector of motion for each individual block does not depend on the results of the search in adjacent blocks and in the previous frame, which makes the method more effective in a very intense and complex movement [3, 4].

After analyzing these methods, their main advantages and disadvantages, which are presented, are revealed.

Pixel method. Disadvantages: high complexity, low accuracy, large amount of information.

Object method. Advantages: high-noise method. Disadvantages: high computational complexity.

Blocks comparison method. Advantages: high accuracy Disadvantages: not a high processing speed

To increase the efficiency of the method of matching the blocks using different templates for finding similar blocks in the next frames.

For optimal work of videoconferencing in the control system and operational communication, the development of a new template – six-point.

The schematic of the six-point search is as follows. First, six blocks are considered. Then, if one of the blocks is the best candidate among others, then the search center shifts there, and the distance between the points shrinks two times [5, 6]. Such actions are repeated as long as the distance between the blocks becomes equal to one.

$$K_{j+1}(x_1;y_1) = K_j(x_0;y_0), \quad (1)$$

where K_{j+1} is an interpolated frame; $(x_0;y_0)$ is coordinates of the center of the current frame; $(x_1;y_1)$ is center coordinates of the interpolated frame.

Formula (1) represents the coordinate of the center of two neighboring frames.

The calculation of the upper left point:

$$K_{j+1}(-4x_1;(y_1+4)) = \frac{[K_j(x_0;y_0) + K_j(-4x_0;(y_0+4))]}{2}, \quad (2)$$

Top right point calculation:

$$K_{j+1}(4x_1;(y_1+4)) = \frac{[K_j(x_0;y_0) + K_j(4x_0;(y_0+4))]}{2}, \quad (3)$$

Left point calculation:

$$K_{j+1}(-6x_1;y_1) = \frac{[K_j(x_0;y_0) + K_j(-6x_0;y_0)]}{2}, \quad (4)$$

Right-pointcalculation:

$$K_{j+1}(6x_1;y_1) = \frac{[K_j(x_0;y_0) + K_j(6x_0;y_0)]}{2}, \quad (5)$$

Lower left point calculation:

$$K_{j+1}(-4x_1;(y_1-4)) = \frac{[K_j(x_0;y_0) + K_j(-4x_0;(y_0-4))]}{2}, \quad (6)$$

Lower right point calculation:

$$K_{j+1}(4x_1;(y_1-4)) = \frac{[K_j(x_0;y_0) + K_j(4x_0;(y_0-4))]}{2}, \quad (7)$$

According to the formulas, the following cohorts of each of the six points are calculated. These calculations are repeated until the distance between the blocks becomes equal to one.

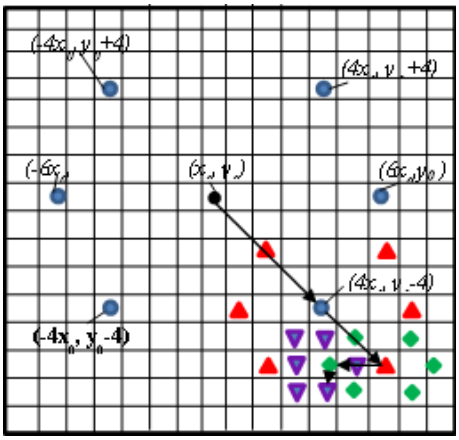


Fig. 1. Possible ways of convergence of the six-point algorithm of search

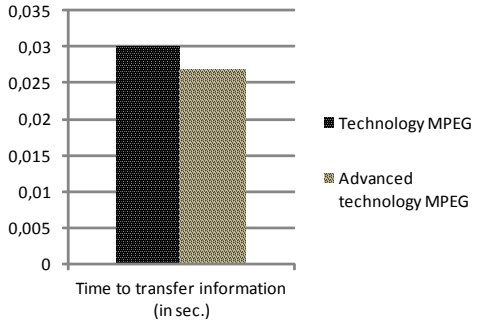


Fig. 2. Comparative characteristics of information transmission time (in sec.) of MPEG technology and after its improvement

After comparative characteristics of the time of information transmission of standardized MPEG technology before and after the improvement of the method of matching the blocks, the effectiveness of the application of the template can be seen.

3. Conclusions

Considered methods of motion compensation in a video data stream. Having analyzed the methods of encoding the stamping of a video data flow, it is interesting to study the possibility of further increasing the coefficient of coding for methods of series lengths due to the additional use of methods of compensation of motion.

As a consequence, it is proposed to improve the method of matching the blocks of motion compensation by simplifying the use of templates.

Video stream speed estimation is performed depending on the quality of video data required, from spatial resolution and frame rate. With the tendency of growth of volumes of video information and not providing the corresponding data volumes of the productivity of technologies of transmission and processing of video information in complexes videoconference – it is necessary to improve the coding methods. In future, it is necessary to improve the methods of motion compensation for other standardized technologies of encoding video information.

References

[1] Baccouch H., Ageneau P. L., Tizon N., Boukhatem N.: Prioritized network coding scheme for multi-layer video streaming. Proc. 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, 2017, 802–809.

[2] Bai X., Wang J.: Towards temporally-coherent video matting. Proceedings of the 5th international conference on Computer vision/computer graphics collaboration techniques – MIRAGE'11, 2011, 63–74.

[3] Barannik V., Krasnorutsky A., Larin V., Hahanova A., Shulgin S.: Model of syntactic representation of aerophoto images segments. Proc of. XVI-th Intern. conf.: Modern Problems of Radio Engineering, Telecommunications and Computer Science, 2018, 974–977 [doi: 10.1109/TCSET.2018.8336356].

[4] Barannik V., Ryabukha Yu., Barannik D., Podlesny S.: The Information Integrity Enhance in Telecommunication Systems with the Binomial Coding. 4th International Scientific-Practical Conference, 2017, 547–550 [doi: 10.1109/INFOCOMMST.2017.8246459].

[5] Barannik V., Ryabukha Yu., Tverdokhlib V., Dodukh A., Suprun O., Tarasenko D.: Integration the non-equilibrium position encoding into the compression technology of the transformed images. East-West Design & Test Symposium (EWDTS), 2017, 1–4 [doi: 10.1109/EWDTS.2017.8110030].

[6] Barannik V. V., Ryabukha Yu. N., Tverdokhleb V. V., Baranmk D. V.: Methodological basis for constructing a method for compressing of transformants bit representation, based on non-equilibrium positional encoding. 2nd IEEE International Conference on Advanced Information and Communication Technologies – AICT, 2017, 188–192 [doi: 10.1109/AIACT.2017.8020096].

[7] Christophe E., Lager D., Mailhes C.: Quality criteria benchmark for hiperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 43(9), 2005, 2103–2114.

[8] Ding Z., Chen H., Gua Y., Peng Q.: GPU accelerated interactive space-time video matting. Computer Graphics International, 2010, 163–168.

[9] Gonzales R. C., Woods R. E.: Digital image processing. Prentice Hall, New Jersey 2002.

[10] Grundmann M., Kwatra V., Han M., Essa I. Efficient hierarchical graph based video segmentation. IEEE CVPR, 2010, 85–91.

[11] Lee S. Y. Yoon J. C.: Temporally coherent video matting. Graphical Models 72, 2010, 25–33.

[12] Lezama J., Alahari K., Sivic J., Laptev I.: Track to the future: Spatio-temporal video segmentation with long-range motion cues. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, 256–289.

[13] Miano J.: Compressed image file formats: JPEG, PNG, GIF, XBM, BMP. Addison-Wesley Professional, New York 1999.

[14] Miano J.: Formats and image compression algorithms in action [Text]. Triumph, Kiev 2013.

[15] Pratt W. K., Chen W. H., Welch L. R.: Slant transform image coding. Proc. Computer Processing in communications. Polytechnic Press, New York 1969.

[16] Recommendations of the International Telecommunication Union ITU-T G.1010 "End-User multimedia QoS categories".

[17] Stankiewicz O., Wegner K., Karwowski D., Stankowski J., Klimaszewski K., Grajek T.: Encoding mode selection in HEVC with the use of noise reduction. International Conference on Systems, Signals and Image Processing (IWSSIP), 2017, 1–6.

[18] Wang S., Zhang X., Liu X., Zhang J. , Ma S., Gao W.: Utility-Driven Adaptive Preprocessing for Screen Content Video Compression. IEEE Transactions on Multimedia 19(3), 660–667.

[19] Zhang Y., Negahdaripour S., Li Q.: Error-resilient coding for underwater video transmission. OCEANS 2016 MTS/IEEE Monterey 2016, 1–7.

[20] Zheng B., Gao S.: A soft-output error control method for wireless video transmission. 8th IEEE International Conference on Communication Software and Networks (ICCSN), 2016, 561–564.

Mykola Dvorsky
e-mail: dmv06@i.ua

A cadet of Ivan Kozhedub Kharkov National University of Air Force. Soldier Dvorsky Mykola Vitaliiovich is a winner of the All-Ukrainian Students Competition of science-based works in the field of the telecommunications.

ORCID ID: 0000-0003-0497-5676

D.Sc. Volodymyr Barannik
e-mail: vvbar.off@gmail.com

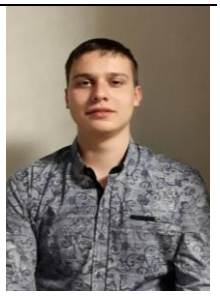
Doctor of technical sciences, professor.Head of the department of combat use and operation of the ACS since 2012. Scientific interests: information security. Publications: more than 550 scientific publications, including monographs, textbooks, teaching aids, scientific articles and patents for inventions.

ORCID ID: 0000-0002-2848-4524

Valeriy Barannik
e-mail: valera462000@gmail.com

Student of the Kharkov national university of Radio Electronics. Scientific interests: systems, technologies of transformation, encoding, defence and information transfer, semantic processing of images.

ORCID ID: 0000-0003-3516-5553



Viktoria Himenko
e-mail: vika.iv55@gmail.com

Position: graduate student of Kharkov National University of Radio Electronics. Research interests: information security, information coding. Publications: more than 10 scientific publications.

ORCID ID: 0000-0002-1337-2404

Anton Sorokun
e-mail: anton.sorokun@gmail.com

Ministry of Education and Science of Ukraine Kharkov National University of Radio Electronics

ORCID ID: 0000-0001-8469-641X

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8038

INFORMATION SYSTEM FOR ASSESSING THE PROFESSIONAL ACTIVITIES COMPLEXITY – THEORETICAL AND PRACTICAL ASPECTS OF IMPLEMENTATION

Oleg Zaritskyi
National Aviation University, Information Technologies Security Department, Educational and Scientific Institute of Information and Diagnostic Systems

Abstract. *The article deals with theoretical and practical issues of the implementation of professional activity analytical evaluation information technologies. The focus is on new methods for presenting and processing data in the activity description and evaluation tasks.*

Keywords: information system, profession activity estimation, job analysis, graphs theory, fuzzy set

SYSTEM INFORMACYJNY DO OCENY ZŁOŻONOŚCI DZIAŁALNOŚCI ZAWODOWEJ – TEORETYCZNE I PRAKTYCZNE ASPEKTY WDROŻENIA

Streszczenie. *Artykuł dotyczy teoretycznych i praktycznych zagadnień wdrażania technologii informatycznych do analitycznej oceny aktywności zawodowej. Nacisk kładziony jest na nowe metody prezentacji i przetwarzania danych w opisie i ocenie zadań.*

Słowa kluczowe: system informacyjny, szacowanie działalności zawodowej, analiza pracy, teoria grafów, zbiór rozmytych

Introduction

Nowadays, society is on the verge of the fourth industrial revolution. The development of engineering science and information technology (IT) in the field of genetics, artificial intelligence, robotics, nanotechnology, 3D printing technologies and complementary biotechnologies accelerates the processes of innovation in almost all industries [16, 17]. Changes in business models and their almost complete rebuilding will significantly affect the professional activity market in the next 5–10 years. Most of the characteristic indicators or markers of industrial transformation that affect the development of industry will have a significant impact on professional activity. They will induce the intensive emergence of new types of professional activity, the alternation or disappearance of existing ones and, consequently, will change the concepts of enterprise resource management, including human one.

Major changes will take place in the field of labour productivity, which is determined by a number of factors, including wage flexibility, which will increasingly be described by factor-based models and wage systems, which are formed according to the estimates of work complexity.

It should be noted that currently there are no generally accepted theoretical and methodological foundations as regards the formal description, modelling and classification of professional activities, and analytical assessment of the work complexity within such activities, which makes it almost impossible systematic collection of information and processing with the help of IT.

The problems of theoretical and methodological development of professional activity analytical evaluation are devoted by the works of scientists [3, 6–11, 13].

There are a certain number of structural models in the format of a simplified graphical description of the professional activity, which gives an idea of common factors set that describe the main functional areas in its content. The peculiarity of the considered models is their orientation to the simulation of certain types of professional activity, which significantly limits the possibility of their application for a wide range of occupations, that is, one of the general requirements for models – universality – is not fulfilled.

One way to solve the issue of universality is to define the minimum and the required number of factors of the model and their characteristics, which provide conditions of maintaining the professional activity content i.e. adequacy of the model. The second one is to increase the number of characteristics described by various scales, including linguistic variables, given by fuzzy sets, which leads to a complication of the model.

Obviously, there is an optimal solution to simultaneously ensure the adequacy and universality of the model in terms of minimizing its structural elements number.

The author examines the theoretical and methodological foundations of professional activity analytical evaluation in terms of the system of principles and methods of organization, construction theory and practice of analytical evaluation, as well as the doctrine of the system itself. The methodology involves a combination of research methods and strategies and also certain hypotheses and ideas that need to be clarified and confirmed by relevant research:

1. It is assumed, that in any profession it is possible to identify and assess a set of common factors that describe the main functional areas in its content.
2. In fact, it is possible to allocate the minimum and the sufficient number of structural elements to ensure the adequacy of the model.
3. It is supposed that the presentation of professional activity in the form of a graph will allow to investigate it as a system of interconnected operations and to calculate their weight characteristics.
4. It is assumed that the use of methods of fuzzy set theory to describe the weakly structured elements of the model (factors, characteristics), will allow to ensure its universality.

1. Conceptual model of professional activity

It should be noted that the existing conceptual models do not carry a meaningful load, but are used only for the formalization of the subject and object of research. The meaningful analytical evaluation of professional activity involves the presentation and processing of quantitative data using the theory of information and methods of mathematical apparatus and modelling.

Mathematical modelling involves the creation of an ideal structure, the so-called conceptual (content) model, but on a deeper level. Its development allows one to formalize with certain assumptions the main structural elements and relationships between them, with a subsequent description by means of mathematical expressions in the form of a formal or mathematical model. Content models are constructed using ready-made designs, developed in the form of hypotheses.

Professional activity belongs to the field of knowledge in which all theories are in a constant process of development and refinement, which greatly complicates the development of conceptual models that are classified depending on the depth of description and understanding of the formalized phenomenon and create a certain hierarchy. Based on fundamental research [5], which formalizes mathematical models for the analysis

of phenomena in the humanities, a meaningful model of professional activity can be built within hypotheses and phenomenological models that differ in the level of description and detail of structural elements.

The conceptual model is used to formalize the boundaries of the research object. It should also be further detailed concerning the transition directly to the subject of research within the framework of the corresponding structural and information models, detailed to a level of sufficiency and necessity for the description and analysis of professional activity – the level of entities and attributes.

Analytical evaluation of professional activity involves the allocation of the structural elements (entities) of the model's upper-level characteristics (Fig. 1) [21, 22], which, in turn, combine the characteristics of individual operations on certain features: educational-qualification level, methods and means of operations etc.

The proposed method of allocating structural elements of professional activity and combining them with a clearly defined feature in groups allows us to obtain a structural model of professional activity at a more detailed level in which the main emphasis is on the characteristics of entities and the connection between them as elements of a complex system (Fig. 1).

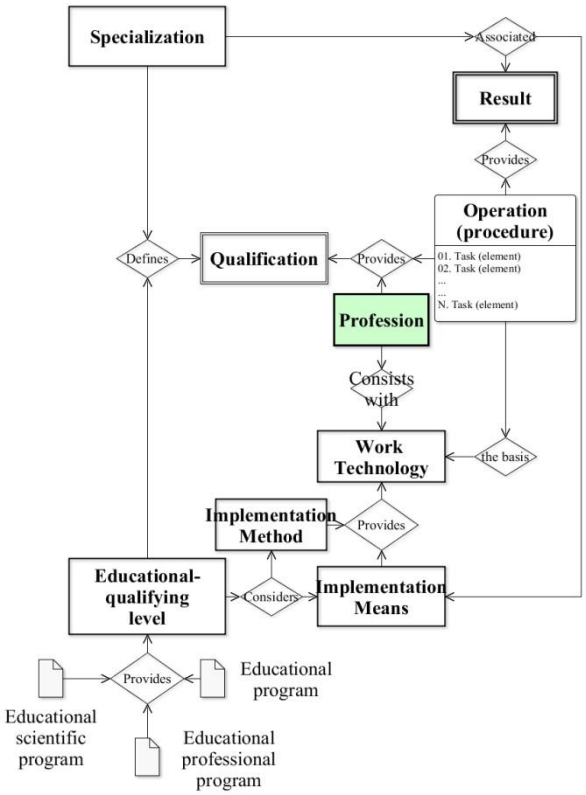


Fig. 1. Structural model of professional activity

Structural model is a system of domain elements – factors of the model (structural elements of the first level), links between them and the mechanism of functioning, which can be presented as a knowledge base, built on the rules, or mathematical expressions.

Thus, professional activities consist of the number of successive or parallel operations and procedures defined in the process of division of labor, which in general form the technology of work or management.

The technology of works (including management) is a rational sequence of operations and procedures (information, logical-thinking, settlement, organizational, etc.) that are performed by the executors using specific means and methods of production in order to influence the object of management.

The basis of the repetitive work technology is the operation (Fig. 1). It represents a homogeneous, logical, indivisible part of the process of work execution or management. The notions of "operation" and "procedure" are key ones in the technology of professional activity study. Along with this, the term "operation" has a broader meaning and covers any useful action of a person – production, technological, managerial, etc.

Operation is any action, measure (or system of events) that is united by a single plan and aimed at achieving an appropriate purpose. The complex of actions carried out within the framework of one operation may consist of several elements (tasks), therefore the operation can be considered as a set of elements (tasks) of labor processes performed by employees. The basis for the construction of operations should be, above all, the target completion of the performance.

Thus, the operation (task) is the basic element of professional activity, which was formalized in the form of a structural model for further information modeling within the framework of constructing a methodology for analytical evaluation of professional activity [14, 15, 20, 21].

2. A new approach in work presenting

The conceptual structure of the subject area was represented as a plurality of structural elements, that are individually described by the basic characteristics [21].

The analytical evaluation of professional activity involved the calculation of each operation weight within the profession (job) and rankings of occupations. Such an assessment is possible only with the use of a mathematical apparatus for describing information models developed by the author in order to obtain quantitative characteristics of work [19].

The main purpose of the mathematical modelling of professional activity is the study of the model entities (structural elements) in their description by means of mathematical relations. In general, the characteristics of the professional activity model structural elements can be represented by matrix *A*:

$$A\left(a_{ij}\right)_{\substack{i=\overline{1,n}\\ j=1,k}} = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk} \end{pmatrix}, \tag{1}$$

where (a_{ij}) is the evaluation of the entity attribute, n is the number of operations (tasks) within the professional activity, k is the number of attributes describing the nature of the model.

Each structural element of the professional activity model is calculated as the average of the estimates for each operation with the corresponding weighting factors ω_i:

$$s_{e_j} = \frac{1}{n}\sum_{i=1}^{n}\omega_i \tilde{a}_{ij}, \qquad \tilde{a}_{ij} = \frac{a_{ij}}{a_{ij_{\max}}}, \tag{2}$$

where $a_{ij_{\max}}$ is maximum value of the *i-th* attribute.

Based on the characteristics of the parameters, the model of professional activity is deterministic, discrete and stationary, since it describes the steady state of the system only for discrete values of its variables (attributes). For the first time, the calculation of weights of each operation was introduced into the model, which significantly reduced the impact of operations that occupy a small share in the total working time fund, on the overall professional activity assessment.

Since the weight coefficients in the mathematical model do not take into account the interconnections between the operations, the author proposes to consider the activity in terms of the theory of graphs, which will enable the calculation of coefficients that characterize each operation on certain features.

2.1. Professional activity graph

Professional activity is described by the non-oriented graph, which is given by matrices of adjacency and incidence, the analysis of which allows revealing the main characteristics of the operation: importance factors, ranking (Fig. 2, second and third digits) and clustering, which significantly affect the activity in general [4, 12].

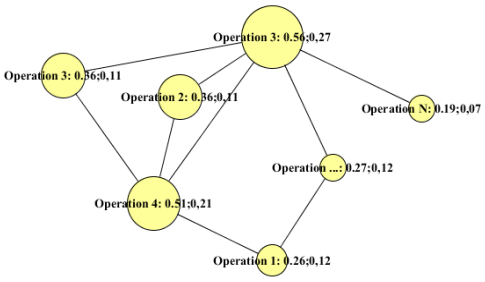


Fig. 2. Professional activity graph

The clustering factor reflects how much the operations tend to make a connection if they are connected through a third operation:

$$C_i = \frac{2\left|\left\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\right\}\right|}{k_i(k_i - 1)}, \tag{3}$$

where k_i is the number of operations on the set of adjacent operations: $N_i = \left\{v_j : e_{ij} \in E \vee e_{ji} \in E\right\}$. V is the set of operations, and E is the set of connections between them.

The weight of the operation on the basis of connections – the calculation of the weight of the operation on the basis of its connections within the activity – a recursive characteristic, which determines the impact of the operation on the activity as a whole:

$$x_i = \frac{1}{\lambda} \sum_{j \in G} a_{ij} x_j, \tag{4}$$

The importance of the operation in terms of the switching function (located on the shortest distance between the two other paths):

$$C_D(v) = \sum_{i \neq v \neq j \in V} \frac{\sigma_{ij}(v)}{\sigma_{ij}}, \tag{5}$$

where σ_{ij} is total number of the shortest paths between operations, $\sigma_{ij}(v)$ is the number of shortest paths that pass through the operation *v*.

The importance of the operation in terms of the minimum average distance to other operations:

$$C_L(v_i) = \sum_j \frac{1}{d(v_j, v_i)}, \tag{6}$$

where $d(v_j, v_i)$ is distance between operations v_j, v_i.

The calculated coefficients (3–6) are taken into account when weighting factors ω_i are calculated.

2.2. Representation of weakly structured characteristics of the professional activity model

Subject area of research – professional activity, expressed by models, described by entities in terms of their attributes and relationships between them. Within the scope of the research, a field of knowledge was defined – a model of knowledge about the subject field [18]:

$$N_p = \left(S_{p_d}; R_{a_p}; A^r \to A^s; R_{b_p}\right), \tag{7}$$

where S_{p_d} is description of the class of situations; R_{a_p} is the condition under which the products are activated; $A^r \to A^s$ is the core of products (7); R_{b_p} is condition (action) to be performed after the kernel is implemented:

$$P_t = \bigwedge_{r=1,r} A^r \to \bigwedge_{s=1,s} A^s, \tag{8}$$

where P_t is the designation of products; t is the number of products in the production model; A^r is the variable predicate (fact) from *r* variables, which is called parcel; A^s is a variable terminal predicate from *s* variables, which is called the conclusion.

Developing ideas presented by the fourth hypothesis, the author proposes to describe the weakly structured elements of the professional activity model by linguistic variables using the theory of fuzzy sets. Templates of presentation products in the expert system taking into account (7–8) and the syntax of the productive programming language CLIPS [18] are presented as designs:

$$P_t = \begin{pmatrix} defrule\left\langle N_p \right\rangle \left[S_{p_d}\right] \\ \left(R_{a_p}\right) \\ \left(\left\langle a_{ij} \right\rangle \left\langle value \right\rangle\right) \\ (...) \\ \left(\left\langle a_{ij} \right\rangle \left\langle value \right\rangle\right) \\ \Rightarrow \\ assert\left(\left\langle a_{ij} \right\rangle \left\langle value \right\rangle\right)\right) \\ R_{b_p} \end{pmatrix}, \tag{9}$$

The method of analysis of weakly structured data is directed at solving the problem of description, calculations and numerical representation of the operation characteristics, described by weakly structured data. The algorithmic part is given by the mathematical apparatus of the theory of fuzzy sets, the logic of the machine of logical output and the rules of the knowledge base in the form of a production model.

As a result of the logical output machine operation, linguistic variables were obtained (9), the transition to their numerical representation was realized due to the usage of standard methods in the theory of fuzzy sets [2].

3. Information system of professional activity evaluation

Analytical evaluation as information technology is considered from the point of view of the combination of data processing technology and expert systems, which represents a set of methods, production processes, software and hardware integrated with the purpose of collecting, processing, storing, disseminating, demonstrating and using structured information and analytical assessment with the purpose to make managerial decisions.

Taking into account the results of investigations of modern information systems functional structures [1], methods and tools for evaluating the work's complexity, theoretical approaches to formalize the tasks of professional activity evaluation with the help of expert systems, there was developed the structural scheme of intellectual IT of analytical assessment based on knowledge. The diagram of the calculation part of the information system is presented in (Fig. 3), it includes three main blocks: a block of graphical representation of the activity model, a block of representing operations in the form of linguistic variables, and directly a block of calculations.

The technology of professional activity analytical evaluation involves the formalization of specific tasks of assessment (classification) and the use of developed tools for the implementation of the evaluation system itself.

Experts (operators) enter data on aspects (factors) of work at the level of their attributes (Fig. 4).

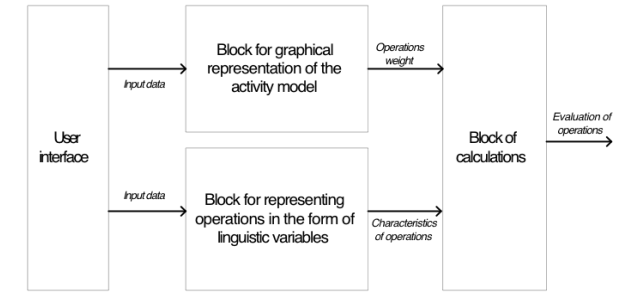


Fig. 3. Algorithmic part of the information system

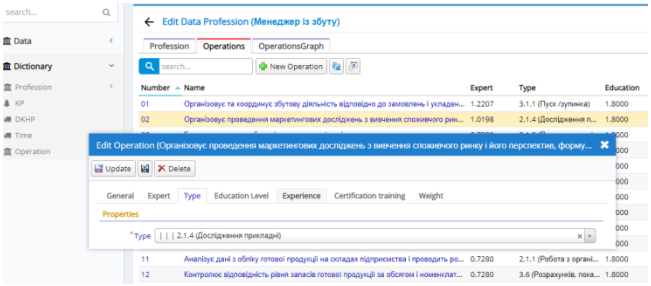


Fig. 4. Analysis and description of the model factors

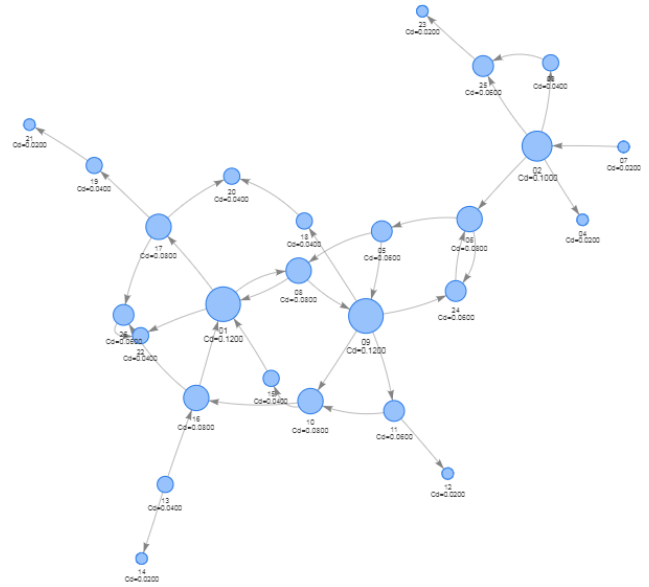
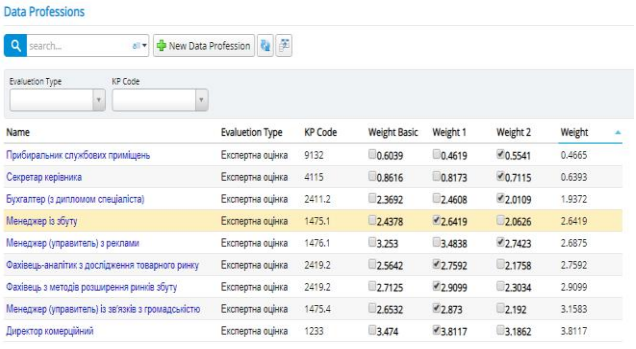


Fig. 5. Professional activity graph



Name	Evaluation Type	KP Code	Weight Basic	Weight 1	Weight 2	Weight
Прибиральник службових приміщень	Експертна оцінка	9132	0.6039	0.4619	0.5541	0.4665
Секретар керівника	Експертна оцінка	4115	0.8616	0.8173	0.7115	0.6393
Бухгалтер (з дипломом спеціаліста)	Експертна оцінка	2411.2	2.3692	2.4608	2.0109	1.9372
Менеджер із збуту	Експертна оцінка	1475.1	2.4378	2.6419	2.0626	2.6419
Менеджер (управитель) з реклами	Експертна оцінка	1476.1	3.253	3.4838	2.7423	2.6875
Фахівець-аналітик з дослідження товарного ринку	Експертна оцінка	2419.2	2.5642	2.7592	2.1758	2.7592
Фахівець з методів розширення ринків збуту	Експертна оцінка	2419.2	2.7125	2.9099	2.3034	2.9099
Менеджер (управитель) із зв'язків з громадськістю	Експертна оцінка	1475.4	2.6532	2.873	2.192	3.1583
Директор комерційний	Експертна оцінка	1233	3.474	3.8117	3.1862	3.8117

Fig. 6. Evaluation of profession

The block of the graphical representation of the activity represents the model in the form of interrelated operations and calculates the operations weight coefficients (3–6) (Fig. 5).

The algorithmic part of the method ensures the automatic formation of adjacency and incidence matrices on the basis of the introduced operations and the connections between them.

The calculation unit implements the developed algorithms (1, 2, 7–9) for quantitative estimation of operations (Fig. 6).

The peculiarity of the technology implementation is the need for the integration of data obtained using the developed algorithmic and production programming languages algorithms, i.e., different principles of data processing.

References

[1] CLIPS Architecture Manual. Version 5.1. Johnson Space Center, Washington 1992.
[2] CLIPS Reference Manual. Volume I. Basic programming guide. Version 6.22. Johnson Space Center, Washington 2004.
[3] Fine S. A., Cronshaw S. F.: Functional job analysis: A foundation for human resources management. Lawrence Erlbaum Associates Publishers, Mahwah 1999.
[4] Girvan M., Newman M. E. J.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99(12), 2002, 7821–7826.
[5] Gorban A., Kazantis N.: Model reduction and Coarse-Graining Approaches for multiscale phenomena. Springer-Verlag Heidelberg, Berlin 2007.
[6] Harvey R. J.: Research Monograph: Development of the Common-Metric Questionnaire (CMQ). Personnel Systems and Technologies Corporation, 1999.
[7] Harvey R. J.: The common-metric questionnaire (CMQ): A job analysis system. First edition. The Psychological Corporation, San Antonio 1991.
[8] McCormick E. J., Cunningham J. W., Gordon G. G.: Job dimensions based on factorial analyses of worker-oriented job variables. Personnel Psychology, 1967 [doi: 10.1111/j.1744-6570.1967.tb02442.x].
[9] McCormick E. J., Jeanneret P. R., Mecham R. C.: The development and background of the position analysis questionnaire (PAQ). Occupational Research Center Purdue University Lafayette, Indiana 1969.
[10] McCormick, E. J., Jeanneret, P. R., Mecham, R. C.: A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). Journal of Applied Psychology 56(4), 1972, 347–368.
[11] Miller B. Z.: Organization theory: study. INFRA-M, Moscow 2005.
[12] Pons P., Latapy M.: Computing communities in large networks using random walks. Computer and Information Sciences-ISCIS 2005, 284–293.
[13] Prigozhin A. I.: Methods of organization development. MCFER, Moscow 2003.
[14] Simon H. A.: A Behavioral Model of Rational Choice. Quarterly Journal of Economics 69, 1955, 99–118.
[15] Simon H. A.: Rational Decision-Making in Business Organizations. A Nobel Memorial Lecture – Economic Sciences, 1978, 343 – 371.
[16] The Future of Jobs Employment. Skills and Workforce Strategy for the Fourth Industrial Revolution – Global Challenge Insight Report. World Economic Forum, Cologny 2016.
[17] The Global Competitiveness Report 2016–2017 – Insight Report. World Economic Forum, Cologny 2017.
[18] Zaritskiy O., Pavlenko P., Tolbatov A.: Data representing and processing in expert information system of professional activity analysis. Proceedings of the 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, Lviv 2016, 718–720.
[19] Zaritsky O., Sudik V.: Development of a mathematical model of professional activity. Eastern European Journal of Advanced Technologies "Technological Center" 1/ 4 (79), 2016, 10–18.
[20] Zaritsky O.: Application of the basics of communication theory for the development of information systems for the modeling of professional activities. Bulletin of the Chernihiv State Technological University – Series "Technical Sciences" 1 (1), 2015, 94–98.
[21] Zaritsky O.: Functional modeling of professional activity basic elements within the model "Entity – Communication". Problems of Informatization and Management 2 (50), 2015, 70–75.
[22] Zaritsky O.: Modeling of human professional activity. Materials of abstracts V international scientific practice conf.: Integrated quality assurance of technological processes and systems, Chernihiv 2015, 202.

D.Sc. Oleg Zaritskyi
e-mail: olegzaritskyi@gmail.com

Associate Professor of the Information Technologies Security Department, Educational and Scientific Institute of Information and Diagnostic Systems, National Aviation University, Kiev, Ukraine. Research interests: information technology, artificial intelligence. Author of more than 50 scientific publications and theses of 20 scientific and technical conferences, 3 certificates of copyright in a work registration.

Scopus Author ID: 57189327322
ORCID ID: 0000-0002-6116-4426

otrzymano/received: 1.10.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8039

SUN TRANSITS IN GEO SATELLITE SYSTEMS IN THE ASPECT OF RADIO WAVES PROPAGATION

Jacek Łukasz Wilk-Jakubowski
Kielce University of Technology, Faculty of Electrical Engineering, Automatic Control and Computer Science, Department of Information Systems, Division of Computer Science

Abstract. *The article includes description of process of Sun transits, including indispensable information on the parameters employed to show their impact on these phenomena. On this basis the date, the start time and the end time and the duration time of Sun transits could be presented for a given location. The main considerations are pertinent to propagation studies. The part of results connected with Sun transits for the limited frequencies of K_u band (downlink) in satellite systems, depending on the antennae, is presented in this article.*

Keywords: wave propagation, technical parameters, thermal noises, sun fades, solar outages, solar interferences

ZJAWISKO INFLUENCJI SŁONECZNYCH W SATELITARNYCH SYSTEMACH GEO W ASPEKCIE PROPAGACJI FAL RADIOWYCH

Streszczenie. *Artykuł zawiera opis zjawiska influencji słonecznych, ukazując niezbędne informacje na temat zastosowanych parametrów, aby przedstawić ich wpływ na te zdarzenia. Na tej podstawie data, czasy początku i końca oraz czas trwania influencji słonecznych mogą zostać określone dla danej lokalizacji. Rozważania dotyczą badań propagacyjnych. W artykule dla łącza downlink zamieszczono część wyników, w zależności od zastosowanych anten, dla granicznych częstotliwości pasma K_u powszechnie wykorzystywanego w łączności satelitarnej.*

Słowa kluczowe: propagacja fal, parametry techniczne, szumy cieplne, zaniki słoneczne, blokady słoneczne, interferencje słoneczne

Introduction

Sun transits, also known as: Sun outages, Sun fades, solar outages or solar interferences, occur around the spring and autumn equinoxes when the overwhelming noise from the Sun swamps the intended signal from the satellite. These phenomena affect all satellite links from time to time (twice a year). At these times the day and night are about the same length. During vernal and autumnal equinoxes the Sun crosses the equator and traces an arc that is directly behind the geo-arc of satellites. The heat which is emitted by the Sun is becoming a strong source of radiation at all frequencies, including the microwave frequencies used to satellite communication (C band, X band, K_u band, and K_a band). During both equinoxes, the reception equipment should accept any interference received, including solar interference that may cause undesired operation. As a consequence, satellite reception can be momentarily interrupted, because the station is incapable of distinguishing between desired communication signal and energy from the Sun. Therefore due to interferences, this phenomenon disrupts the proper reception of geostationary satellite signals.

The effects of solar interferences range from momentarily degradation of satellite signals (increase of BER) to total destruction of the signal (a lack of communication between terminal and satellite). In practice, digital systems are a little more resilient in this respect than analog systems. To prevent equipment damage due to the radiated energy from the Sun being focussed into the unit (fortunately, it's rare with recent projects), this danger can be avoided by the use of appropriate temporary or permanent instruments (guards) to receiving stations.

1. Energy from the Sun

At the present time, it is possible to estimate the exact date, the start and the end time and the duration time of Sun transits. Their effects depend on the location of the earth station, satellite location above the equator and earth station equipment (e.g., antenna beamwidth of earth station or focal resolution, accuracy of antenna pointing and the station keeping accuracy of the satellite). Thus, the region of sky scanned for the Sun interferences is associated with antenna beamwidth of ground station. Furthermore, the effects of solar interferences depend on the size of the receive antenna, as well as a frequency of reception satellite signal due to the impact of all noise sources.

The Sun as a meaningful source of extraterriestrial natural noises contributes to temporarily interferences with the satellite signal. It is also of immense importance for the satellite links operating at frequencies below 2 GHz. At high frequencies of at least 2 GHz or more, the impact of Moon and nonthermal radiation are additionally visible (e.g., the constellations of Cassiopeia A, Cygnus or planetary nebula – Crab). The noise temperature of the Sun (as a powerful broadband transmitter of microwaves) is in excess tens of thousands of Kelvin. This thermal energy is strong enough to cause an outage when the Sun passes directly behind the satellite (when viewed from the Earth). The Sun moves further north, the Sun's alignment with the satellite and station on the ground moves ever so little. When the Sun, satellite and station on the ground are exactly aligned with each other, peak outage time appears. Therefore, the outage duration increases when the Sun becomes more aligned with the satellite and ground station. In practice, the Sun fades will appear for several days both before and after the peak. Further, the interferences will arise at roughly the same time each day and could repeat on a daily basis for one week or more. The interferences decrease moderately as the Sun starts moving away from the satellite and station alignment, until another solar outage when the Sun starts heading south (during autumnal equinox).

2. Experimental results

We can distinguish many factors affecting the reception of microwaves. Many articles include studies of the effects of rain. The other ones are connected with technical parameters (e.g., antenna aperture, antenna mispointing, dish efficiency, coupling loss, etc.). Sometimes it is necessary to compensate poor quality of signal via modifications of technical parameters. For this reason the knowledge about technical equipment is of great importance to the design and implementation of high-quality satellite systems. Currently, we can safely inspect and predict the Sun transits to estimate their impact on the quality of a satellite signal. Actual monitoring and measurements should be carried out in appropriate conditions. The predicted results for Lublin city will be presented in this article for the Eutelsat Hot Bird 13A satellite (13ºE) according to the limited (bounded) middle frequencies of the carrier from K_u band (downlink) and antenna aperture. Antenna aperture has been changed twice (from 0.6 m – it is the minimal recommended aperture for receiving signals in Poland, to 1.2 m – the doubled minimal value). The results of solar interferences for Geostationary Orbit Satellite (for vernal and autumnal equinoxes in year 2018) for the antenna aperture equalling 0.6 m and frequency equalling 10.7 GHz (lower bound of bandwidth) are presented in Table 1 and Table 2. The analogous results but for the frequency equalling 12.75 GHz (upper bound of bandwidth) are presented in Table 3 and Table 4. All solar outage events are listed in date order and grouped according to the early and late seasons.

Table 1. Vernal equinox in year 2018 for the frequency equalling 10.7 GHz and antenna aperture equalling 0.6 m

Data	Start	End	Outage window duration
22.02	12:23	12:27	00:04
23.02	12:19	12:31	00:12
24.02	12:17	12:33	00:16
25.02	12:16	12:34	00:18
26.02	12:15	12:34	00:19
27.02	12:14	12:35	00:21
28.02	12:14	12:35	00:21
1.03	12:13	12:35	00:22
2.03	12:13	12:35	00:22
3.03	12:13	12:34	00:21
4.03	12:14	12:33	00:19
5.03	12:14	12:32	00:18
6.03	12:15	12:31	00:16
7.03	12:17	12:29	00:12
8.03	12:19	12:26	00:07

Table 2. Autumnal equinox in year 2018 for the frequency equalling 10.7 GHz and antenna aperture equalling 0.6 m

Data	Start	End	Outage window duration
5.10	11:57	12:04	00:07
6.10	11:54	12:06	00:12
7.10	11:52	12:08	00:16
8.10	11:50	12:09	00:19
9.10	11:49	12:09	00:20
10.10	11:49	12:10	00:21
11.10	11:48	12:10	00:22
12.10	11:48	12:09	00:21
13.10	11:48	12:09	00:21
14.10	11:48	12:08	00:20
15.10	11:48	12:08	00:20
16.10	11:49	12:07	00:18
17.10	11:50	12:05	00:15
18.10	11:51	12:03	00:12
19.10	11:54	12:00	00:06

As we can conclude, according to the closest Sun's alignment within the prescribed time limits, during vernal equinox (February and March equinox) the maximum deph was between 27.02 and 4.03. Deph has a minimum value from 22.02 to 24.02 and from 6.03 to 8.03. It has a medium value between 25.02 and 26.02, similarly to 5.03. During autumnal equinox (October equinox) the maximum deph was between 10.10 and 14.10. Deph has a minimum value from 5.10 to 7.10 and from 17.10 to 19.10. It has a medium value between 8.10 and 9.10, the same as between 15.10 and 16.10.

As already noted, the results for the frequency equalling 12.75 GHz, in the case of antenna aperture equal 0.6 m, for vernal and autumnal equinoxes in year 2018 are shown in Table 3 and Table 4. For this signal frequency during vernal equinox (February and March equinox) the maximum deph was between 27.02 and 3.03. Deph has a minimum value from 23.02 to 25.02 and from 5.03 to 7.03. This parameter has a medium value 26.02, similarly to 4.03. During autumnal equinox (in October) the maximum deph was between 10.10 and 14.10, similarily as in the case of the earlier-mentioned frequency equalling 10.7 GHz. Deph has a minimum value from 6.10 to 8.10 and from 16.10 to 18.10. It has a medium value for two days: 9.10 and 15.10.

Table 3. Vernal equinox in year 2018 for the frequency equalling 12.75 GHz and antenna aperture equalling 0.6 m

Data	Start	End	Outage window duration
23.02	12:24	12:26	00:02
24.02	12:20	12:30	00:10
25.02	12:18	12:32	00:14
26.02	12:17	12:33	00:16
27.02	12:16	12:33	00:17
28.02	12:15	12:33	00:18
1.03	12:15	12:33	00:18
2.03	12:15	12:33	00:18
3.03	12:15	12:33	00:18
4.03	12:15	12:32	00:17
5.03	12:16	12:30	00:14
6.03	12:17	12:29	00:12
7.03	12:20	12:26	00:06

Table 4. Autumnal equinox in year 2018 for the frequency equalling 12.75 GHz and antenna aperture equalling 0.6 m

Data	Start	End	Outage window duration
6.10	11:58	12:03	00:05
7.10	11:54	12:06	00:12
8.10	11:53	12:07	00:14
9.10	11:51	12:08	00:17
10.11	11:50	12:08	00:18
11.10	11:50	12:08	00:18
12.10	11:49	12:08	00:19
13.10	11:49	12:08	00:19
14.10	11:49	12:07	00:18
15.10	11:50	12:06	00:16
16:10	11:51	12:05	00:14
17.10	11:52	12:03	00:11
18:10	11:55	12:00	00:05

On the basis of analysis carried out so far we can conclude that very important factor is the antenna aperture. This technical parameter is taken as the average diameter of a dish in metres to ascertain compliance with international engineering practice. The results indicate that the major diameter is a better choice than minor diameter. Against this background, the larger the antenna aperture, the shorter the duration and intensity of solar interferences and vice versa.

The solar interferences in year 2018 during both equinoxes for the same frequencies as before, but for the antenna aperture equalling 1.2 m are presented in tables from 5 to 8.

Table 5. Vernal equinox in year 2018 for the frequency equalling 10.7 GHz and antenna aperture equalling 1.2 m

Data	Start	End	Outage window duration
26.02	12:22	12:28	00:06
27.02	12:20	12:30	00:10
28.02	12:19	12:30	00:11
1.03	12:19	12:30	00:11
2.03	12:18	12:30	00:12
3.03	12:19	12:29	00:10
4.03	12:20	12:28	00:08
5.03	12:22	12:25	00:03

Table 6. Autumnal equinox in year 2018 for the frequency equalling 10.7 GHz and antenna aperture equalling 1.2 m

Data	Start	End	Outage window duration
8.10	11:58	12:01	00:03
9.10	11:55	12:03	00:08
10.10	11:54	12:04	00:10
11.10	11:53	12:04	00:11
12.10	11:53	12:04	00:11
13.10	11:53	12:04	00:11
14.10	11:53	12:03	00:10
15.10	11:54	12:01	00:07

The results in the case of frequency equalling 10.7 GHz indicate that, according to the closest Sun's alignment within the prescribed time limits, during vernal equinox (February and March equinox) the maximum deph was between 28.02 and 2.03. Deph has a minimum value from 4.03 to 5.03 and 26.02. It has a medium value within two days: 27.02 and 03.03. During autumnal equinox (October equinox) the maximum deph was between 11.10 and 13.10. Deph has a minimum value from 8.10 to 9.10, similarly to 15.10. It has a medium value for two days: 10.10 and 14.10.

The results for the upper bound of bandwidth (12.75 GHz) are presented in Tables 7 and 8 for vernal and autumnal equinoxes, respectively.

Table 7. Vernal equinox in year 2018 for the frequency equalling 12.75 GHz and antenna aperture equalling 1.2 m

Data	Start	End	Outage window duration
26.02	12:23	12:27	00:04
27.02	12:21	12:29	00:08
28.02	12:20	12:29	00:09
1.03	12:19	12:30	00:11
2.03	12:19	12:29	00:10
3.03	12:20	12:28	00:08
4.03	12:21	12:27	00:06

Table 8. Autumnal equinox in year 2018 for the frequency equalling 12.75 GHz and antenna aperture equalling 1.2 m

Data	Start	End	Outage window duration
9.10	11:57	12:02	00:05
10.11	11:55	12:03	00:08
11.10	11:54	12:04	00:10
12.10	11:54	12:04	00:10
13.10	11:54	12:03	00:09
14.10	11:54	12:02	00:08
15.10	11:56	12:00	00:04

These data indicate that during vernal equinox (February and March equinox) in year 2018 the maximum deph was between 28.02 and 2.03. Deph has a minimum value for two days: 26.02 and 4.03. This parameter has a medium value 27.02, similarly to 3.03. During autumnal equinox (in October) the maximum deph was between 11.10 and 13.10. Deph has a minimum value within two days: 9.10 and 15.10. It has a medium value also for two days: 10.10 and 14.10.

Generally during both equinoxes, with the increase in antenna aperture the number of days with outages significantly declined. This becomes particularly evident for both frequencies for all bounded days from Tables 5–8, in comparison to the first and the last days from Tables 1–4. As we can observe, the differences are equal several days for each bounded day in the frequency range covered. As might be expected, the increase in the antenna aperture leads to minimize the number of days with outages, which is desirable.

3. Results and discussion

This article presents the process of Sun transits. The article includes information on the parameters employed and their impact on these phenomena. One of them is antenna aperture. In practice, the duration of the solar interferences is inversely related to the antenna aperture and frequency of the satellite receive dish (the smaller the antenna aperture, the great outage window duration and intensity of the interferences).

The main scientific objective was to analyze the impact of solar interferences on the quality of satellite transmissions during vernal and autumnal equinoxes (twice a year). Solar outages may cause total loss of signal sometimes.

On this basis the date, the start time and the end time and the outage window duration for the limited frequencies of K_u band (downlink) in satellite systems, depending on the antenna aperture for the given location was presented. In this context the essential information could be used in practical applications. All data due to quite substantial repeatability may be accommodated by defining the terminal equipment over subsequent years, as soon as the links that should meet the quality requirements of signal by defining the specific parameters. Therefore, we can use the tabular data to achieve the forecast of solar outages for the future with quite good precision in the selected geographical area. Tests have shown that the predicted times of solar interferences are close to what is de facto viewed in practice. These data are also important to reduce the risk of interruption of the connection. So, another application could be present in the link budget analyses. Because of it, it is possible to use a transponder at an alternative time relative to solar outages as a concrete knowledge of when these events appear is meaningful. If we do wish more reliable data transfer connection at that time, it is possible to use an another satellite.

Apart from that, if we consider alongside the solar outages – another information, e.g., free space attenuation, degradation due to hydrometeors (especially rainfall), influence of the Earth atmosphere (especially total gaseous absorption) we could determine reasonably comprehensive link budget analyses. Therefore, the accurate results may be used to characterize the receiver by the minimum acceptable quality of signal which takes into account undesirable weather conditions and solar interferences.

4. Conclusions

The solar interferences during vernal and autumnal equinoxes may last several minutes either side of the peak each day. As can be seen from the analyses presented in Tables 1–8, the outage will last longer for the smaller the antenna. In this sense, determination of solar outage findings with satellite systems (for selected parameters) for planning and management in the city is significant. Fortunately, as may be expected the sustainable city for planning and management uses to determine the large satellite. It abundantly shows in many examples that we can use these data for manufacturers and users of satellite equipment. Moreover, the results can be applied to dynamical changes in transmission parameters (e.g. to improve energy efficiency).

These data can be useful for calculating signal quality in practice. Only if elderly satellites have lapsed into severe inclined orbits, results might be inaccurate as a sign for the future. So, the satellite's inclination should be check before measurements or predictions. The predictions of solar outages associated with a particular location and satellite combination (by using astronomical algorithm) can be applied to provide information about the days when energy from the Sun interferes with the excepted signals. Therefore, this knowledge may be useful to ensure the proper reception signals and minimize the risk of lack of communication (e.g., it is possible to book transponder time to relay information in good time – not when solar outage appears just to coincide with transmission of information on the downlink).

Data about solar outages can be useful for improve the knowledge about propagation of wave. Nevertheless, to put additional data into practice in a considerable way is a challenge requiring further efforts about modelling remote sensing, signal attenuation, noise increase or total signal degradation. Currently, the similar studies are lead many of numerical forecast models. It is planned to estimate the maximum safe distance of margin during solar outage events and conduct regression analysis.

References

[1] Baghsiahi H., Wang K., Kandulski W., Pitwon R., Selviah D.: Optical waveguide end facet roughness and optical coupling loss. Journal of Lightwave Technology 31(16)/2013, 2659–2668 [doi: 10.1109/JLT.2013.2271952].

[2] Benarroch A., García-del-Pino P., García-Rubia J. M., Riera J. M.: Derivation of rain attenuation from experimental measurements of drop size and velocity distributions. COST IC0802 (MCM3). Available online: http://www.tesa.prd.fr/cost/input_documents.pdf [05.04.2018].

[3] Boulanger X., Castanet L., Jeannin N., Lacoste F.: Study and modelling of tropospheric attenuation for land mobile satellite system operating at Ku and Ka band. COST IC0802 (MCM2). Available online: http://www.tesa.prd.fr/cost/input_documents.pdf [06.04.2018].

[4] Cetin M.: Consideration of permeable pavement in Landscape Architecture. Journal of Environmental Protection and Ecology 16(1), 2015, 385–392. Available online: https://docs.google.com/a/jepe-journal.info/viewer?a=v&pid=sites&srcid=amVwZS1qb3VybmFsLmluZm98amVwZS1qb3VybmFsfGd4OjNmY2FkYzYzN2Y1MWNlYjM [08.05.2018].

[5] Cetin M.: Determination of bioclimatic comfort areas in landscape planning: A case study of Cide Coastline. Turkish Journal of Agriculture-Food Science and Technology 4(9)/2016, 800–804.

[6] Cetin M.: Determining the bioclimatic comfort in Kastamonu City. Environmental Monitoring and Assessment 187(10)/2015, 640. Available online: http://link.springer.com/article/10.1007%2Fs10661-015-4861-3 [08.05. 2018].

[7] Cetin M.: Evaluation of the sustainable tourism potential of a protected area for landscape planning: a case study of the ancient city of Pompeipolis in Kastamonu. International Journal of Sustainable Development & World Ecology 22(6)/2015, 490–495.

[8] Cetin M., Sevik H.: Assessing potential areas of ecotourism through a case study in Ilgaz Mountain National Park, ed. L. Butowski. InTech. Rijeka 2016. Available online: http://www.intechopen.com/books/tourism-from-empirical-research-towards-practical-application/assessing-potential-areas-of-ecotourism-through-a-case-study-in-ilgaz-mountain-national-park [08.05.2018].

[9] Cetin M., Sevik H.: Evaluating the recreation potential of Ilgaz Mountain National Park in Turkey. Environmental Monitoring and Assessment 188(1)/2015, 52. Available online: http://link.springer.com/article/10.1007% 2Fs10661-015-5064-7 [08.05.2018].

[10] Cetin M.: Sustainability of urban coastal area management: a case study on Cide. Journal of Sustainable Forestry 35(7)/2016, 527–541 [doi: 10.1080/10549811.2016.1228072].

[11] Cetin M., Topay M., Kaya L. G., Yilmaz B.: Efficiency of bioclimatic comfort in landscape planning process: case of Kutahya. Turkish Journal of Forestry 1(1)/2010, 83–95.

[12] Cetin M.: Using GIS analysis to assess urban green space in terms of accessibility: case study in Kutahya. International Journal of Sustainable Development & World Ecology 22(5), 2015, 420–424 [doi: 10.1080/13504509.2015.1061066].

[13] Heiser D. A., Keyser R. B.: Microwave measurements for antenna randome maintenance and replacement. IEEE Conference Publications. Symposium on Antenna Technology and Applied Electromagnetics 1/1998, 501–506 [doi: 10.1109/ANTEM.1998.7861713].

[14] Ho Ch., Kantak A., Slobin S., Morabito D.: Atmospheric attenuation and noise temperature effects. The Interplanetary Network Progress Report 42(168)/2007, 1–22.

[15] Ho Ch., Kantak A., Slobin S., Asmar S.: Solar brightness temperature and corresponding antenna noise temperature at microwave frequencies. The Interplanetary Network Progress Report 42(175)/2008, 1–11.

[16] Huo X., Li D., Han X., Wang J.: Effects of structural and environmental parameters on the coupling loss of leaky rectangular waveguide in tunnel. Antennas & Propagation Conference (LAPC), 2015 [doi: 10.1109/LAPC.2015.7366022].

[17] Ippolito L. J.: Satellite communications. Systems engineering. Atmospheric effects, satellite link design and system performance. John Wiley & Sons. Chichester 2008.

[18] Johannsen K. G., Koury A.: The moon as a source for G/T measurements. IEEE transactions on aerospace and electronic systems AE-S10(5)/1974, 718–727.

[19] Knoch L. (red.): Systemy radiokomunikacji satelitarnej. Wydawnictwa Komunikacji i Łączności. Warszawa 1980.

[20] Lozano A., Tulino A. M.: Capacity of multiple-transmit multiple-receive antenna architectures. IEEE Transactions on Information Theory 48(12)/2002, 3117–3128 [doi: 10.1109/TIT.2002.805084].

[21] Marciniak M., Wilk J.: Czynniki tłumienia fal radiowych w atmosferze ziemskiej. Logistyka 4/2015, 6578–6588.

[22] Marciniak M., Wilk J.: Relationship between the quality coefficients signal and rainfall intensity. TRANSCOM 2015. 11-th European conference of young researchers and scientists 3/2015, 94–98.

[23] Marek M.: Wykorzystanie ekonometrycznego modelu klasycznej funkcji regresji liniowej do przeprowadzenia analiz ilościowych w naukach ekonomicznych. Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne. The role of informatics in economic and social sciences. Innovations and interdisciplinary implications, ed. T. Grabiński. Wydawnictwo Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach. Kielce 2013.

[24] Pardo E., Kapolka M., Kováč J., Šouc J., Grilli F., Piqué A.: Three-Dimensional Modeling and Measurement of Coupling AC Loss in Soldered Tapes and Striated Coated Conductors. IEEE Transactions on Applied Superconductivity 26(3)/2016 [doi: 10.1109/TASC.2016.2523758].

[25] Rec. P. 372-10: Radio noise. ITU-R. Genewa 2009.

[26] Rec. P. 837-6: Characteristics of precipitation for propagation modeling. Radio noise. ITU-R. Genewa 2012.

[27] Roddy D.: Satellite communications. McGraw-Hill. New York 2001.

[28] Szóstka J.: Fale i anteny. Wydawnictwa Komunikacji i Łączności. Warszawa 2006.

[29] Weitzen J. A.: Effects of polarization coupling loss mechanism on design of meteor scatter antennas for short- and long-range communication. Radio Science 24(4)/1989, 549–557 [doi: 10.1029/RS024i004p00549].

[30] Wilk-Jakubowski G.: Wpływ technologii informatyczno-komunikacyjnych na funkcjonowanie współczesnych społeczeństw. Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne. The role of informatics in economic and social sciences. Innovations and interdisciplinary implications, ed. T. Grabiński. Wydawnictwo Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach. Kielce 2011.

[31] Wilk-Jakubowski J.: Ocena wpływu źródeł szumów naturalnych na propagację fal radiowych. Autobusy. Technika, Eksploatacja, Systemy Transportowe 12, 2016, 1450-1453.

[32] Wilk-Jakubowski J.: Wpływ warunków klimatyczno-atmosferycznych na mechanizm propagacji fal radiowych w atmosferze ziemskiej. TTS. Technika transportu szynowego 12/2016, 266–268.

[33] Wilk-Jakubowski J.: Wpływ aktywności słonecznej na propagację fal radiowych. Autobusy: technika, eksploatacja, systemy transportowe 12/2016, 1454–1457.

[34] Zieliński R. J.: Satelitarne sieci teleinformatyczne. Wydawnictwa Naukowo-Techniczne. Warszawa 2009.

Ph.D. Jacek Łukasz Wilk-Jakubowski
e-mail: j.wilk@tu.kielce.pl

Kielce University of Technology, Ph.D. (technical) Faculty of Electrical Engineering, Automatic Control and Computer Science, Department of Information Systems.
Research interests: teleinformatic systems, wave propagation.

ORCID ID: 0000-0003-1275-948X

otrzymano/received: 12.09.2018 *przyjęto do druku/accepted:15.12.2018*

DOI: 10.5604/01.3001.0012.8041

INFLUENCE OF MOBILE ROBOT CONTROL ALGORITHMS ON THE PROCESS OF AVOIDING OBSTACLES

Piotr Wójcicki, Paweł Powroźnik, Kamil Żyła, Stanisław Grzegórski
Lublin University of Technology, Institute of Computer Science

Abstract. *This article presents algorithms for controlling a mobile robot. An algorithms are based on artificial neural network and fuzzy logic. Distance was measured with the use of ultrasonic sensor. The equipment applied as well as signal processing algorithms were characterized. Tests were carried out on a mobile wheeled robot. The analysis of the influence of algorithm while avoiding obstacles was made.*

Keywords: mobile robot, algorithms, collision avoidance

WPŁYW ALGORYTMÓW STEROWANIA ROBOTEM MOBILNYM NA PROCES UNIKANIA PRZESZKÓD

Streszczenie. *W artykule przedstawiono algorytmy sterowania robotem mobilnym. Wykorzystano algorytm oparty o sztuczną sieć neuronową oraz logikę rozmytą. Odległość od przeszkód mierzono za pomocą czujnika ultradźwiękowego. Scharakteryzowano zastosowane urządzenia oraz algorytmy przetwarzania sygnałów. Testy przeprowadzono na mobilnym robocie kołowym. Przeprowadzono analizę wpływu algorytmów sterowania podczas omijania przeszkód.*

Słowa kluczowe: robot mobilny, algorytmy, unikanie kolizji

Introduction

Controlling a mobile robot in the process of avoiding obstacles is a very important issue. The sensor system used to retrieve data should be adapted to the environment in which the location of obstacles may be unknown and variable. To adapt the robot's behaviour to any complex environmental dynamics without further human intervention, it is necessary to extract key information from the environment [3, 11]. The most commonly used and described detector elements use infrared light, radio waves or ultrasonic waves. Simplified versions of these systems require the use of markers that limit the work area of the robot. For infrared light based systems, reflective stickers are used, while radio systems use, for example, RFID stickers. In the case of ultrasonic waves, it is necessary to use a special microphone and ultrasound transmitters as markers [13]. Then, the route is determined by detecting a specific marker and choosing an alternative route. This approach prevents correct detection in the event of a sudden appearance of the object. More advanced systems are equipped with CCD cameras, Global Navigation Satellite System (GNSS) and image sensors based on CMOS technology [13]. Ultrasound sensors have found the greatest application due to their low cost and availability on the market. They are very effective in measuring distances, they also work well when detecting obstacles and exploring an unknown environment. They are successfully used in complicated maneuvers of passing or parking. In the applications presented in the literature [5–7], the authors most often use six or eight ultrasound sensors to capture information about obstacles around them. The information is then processed in order to e.g. avoid an obstacle, calculate a new route, determine the shape of the surface or update the map (if it is created).

The solution for measuring the distance from the obstacle during the manoeuvre was presented in the literature [7, 15]. An algorithm was then used that allowed the robot to maintain a safe distance from the obstacle while moving forward along the wall (mode T). A significant number of research papers using fuzzy logic navigation [1–2, 5–6, 10, 13] have been proposed. This approach requires adjustment of tuning parameters depending on the environment, which can be problematic. Another solution uses an artificial neural network for intelligent robot control, because it has the ability to educate itself and a high efficiency in the approximation of non-linear functions [8, 11].

The study defines goals: the main is to exam how particular algorithm affects the process of avoiding an obstacle. Another is whether it suits for the size of the room and it can detect an obstacle during drive without stopping. In this research, an analisys of the influence mobile robot control algorithms has been done. The article is organized as follows: Chapter 1 discusses the structure of the robot, in particular mechanism of uses one sensor. Chapter 2 presents selected algorithms used to control robot. Chapter 3 and 4 describe conducted experiment and results.

1. The structure of a mobile robot

A mobile wheeled robot consisting of several modules was used for the research. The body was designed in the Autodesk Inventor environment (Fig. 1) and made in 3D printing technology using ABS plastic. The robot has two wheels driven by separate DC motors and one support wheel without a drive. The measuring system consists of an HC-015 ultrasonic sensor placed on a movable head – a servo drive that allows the user to rotate left and right by a maximum of 90° to the front plane of the robot. This solution allows to reduce the number of sensors necessary for the correct location of obstacles in the case of moving, for example forward or along the wall.

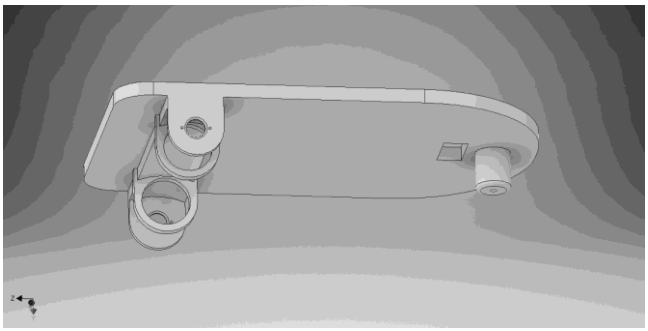


Fig. 1. The robot body design

There is no need to use an extensive system of six or eight sensors, just to rotate one by a given angle. An additional advantage is the smaller number of ports necessary to operate the sensor. The system centre is a Wi-Fi module based on the ESP8266-NodeMCU system. It is responsible for processing the signal received from the ultrasonic sensor, controlling the sensor head, calculating and generating a motor control signal. The NodeMCU Motor Shield based on the L293D chip was used for the motor control itself, which allows independent power supply and control of two DC motors at the same time. The last element is a power supply system consisting of three Li-Ion battery packs to supply respectively: DC motors (2 × 3.7V Li-Ion), servo drives (2 × 3.7V Li-Ion together with L7805CV voltage stabiliser) and the NodeMCU system (1 × 3.7V Li-Ion).

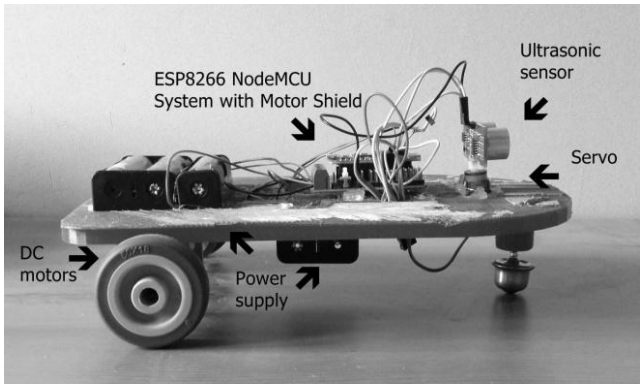


Fig. 2. The robot used in the research

2. Algorithms

This section presents proposed wheeled robot control algorithms in the process of bypassing obstacles. As an initial assumption it was taken for granted that until the obstacle was detected, the robot moved at a constant linear velocity. The main task was to travel the longest possible distance without colliding with an obstacle. Detection of an obstacle changed the route in a way depending on the specific algorithm. The distance reading was carried out using a sensor from five pre-set servo positions – two positions were responsible for the left-hand reading, one for the front reading and two for the right-hand reading. Due to the fact that the main goal was to move forward, the possibility of turning the servo drive in other directions was abandoned. Pulse Width Modulation (PWM) was used to control the motors, if a change in the direction of the movement occurred, the signal fill ratio was adjusted correspondingly. Thanks to this, it was possible to smoothly perform an evasive manoeuvre without stopping the robot. Algorithms using fuzzy logic and artificial neural network were used.

3. The first case – application of fuzzy logic

Fuzzy systems describe imprecise, unspecified deterministically or non-specific information. Knowledge about the system is stored in the form of fuzzy rules.

Stages of fuzzy system design:

- Fuzzification – transformation of input signals into qualitative values.
- Fuzzy application – determining the value of outputs in the field of quality based on inputs using a set of fuzzy rules.
- Creating a database of rules describing qualitative knowledge about the system in the form of fuzzy rules. This is a set of character instructions:

  IF <CONDITIONS> THEN <SELECTED ACTION>

The applied algorithm based on fuzzy logic consists of N steps. In the general case, it looks as follows:

- Start and initial initialisation.
- Initial distance readings from each servo position and direction selection.
- Driving forward.
- Check if an obstacle has been detected.
- If yes, distance classification.
- Pre-preparation for skipping manoeuvre.
- Change in the ratio of PWM signals controlling motors (left or right).
- Distance control when evading.
- A mirror-change in the ratio of PWM signals controlling motors (return to the route if possible).
- Checking the distance – choosing the direction.
- Driving forward.

In the initial phase of the algorithm preliminary initialisation is started and the possible directions of motion are checked. This direction is selected where the detected obstacle is farthest away (or no obstacle is detected). The robot then moves forward and every 100 ms the distance is measured from the fixed servo positions. In the absence of obstacles, the robot will move steadily forward. If the read distance is qualified to the 0.4–0.5 m range, the robot will start preparations for the skipping maneuver. Depending on the position of the servo from which the measurement was made, a different ratio of the signal filling to the motors would be adjusted. For example, if the distance read from position 2 (Fig. 3) is in the above-mentioned range, the robot will perform a slight right turn manoeuvre. Speed will be varied for the same distance detected by sensor from position 1 or 2.

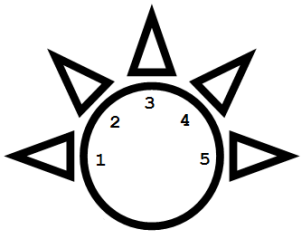


Fig. 3. Fixed sensor positions

While avoiding the obstacle, distance control is carried out to avoid collisions in case of a large obstacle (with a complicated shape). At the same time, a mirror turning manoeuvre is being carried out (if the first turn was to the right, the next turn would be to the left). After avoidance, another check of the movement possibility takes place – choice of the direction, then setting the robot and the movement is continued. The algorithm part is shown in Fig. 4. Three special cases of applying the algorithm have also been specified:

- Detecting an obstacle too late – it may not be possible to bypass it without stopping.
- Detecting large obstacles – it may not be possible to bypass them.
- Failure to detect an obstacle – collision.

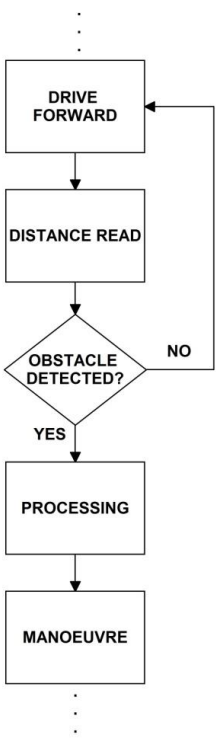


Fig. 4. Part of the control algorithm

4. The second case – the use of artificial neural network

In conducted research artificial neural networks have been created in Arduino IDE application. The mathematical formulation on neuron output is as follows [9]:

$$n = \sum_{i=1}^{N} x_i w_i + b$$

where: n – is the neuron output, x_i – is value of I input, w_i – is weight of I input, N – is number of inputs, b – indicates bias.

Algorithms based on an feedforwarded artificial neural network were used when planning path movements due to the possibility of approximating complex, non-linear dependencies, as well as learning and adaptation. The structure (Fig. 5) of the network used was as follows:

- the entrance layer consists of five neurons – corresponding to the next positions of the ultrasonic sensor
- the hidden layer consists of six neurons – responsible for processing input signals
- the output layer consists of two neurons – corresponding to the number of controlled motors

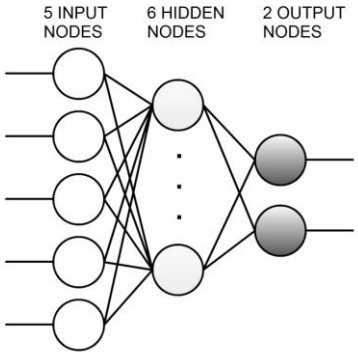


Fig. 5. The structure of used artificial neuron network

The structure of artificial neuron network during research was optimized. Starting from a small number of neurons in the hidden layer, computed as geometric mean of the number of inputs and outputs, through the number of neurons determined based on Kolmogorov's Theorem, to networks containing several dozen neurons in the hidden layer.

The first stage of the algorithm is training the network and relies on such a selection of parameters (weights) that will allow linking the input signal with the output signal – that there is a dependence of the output from the input for each case. The scales determine the degree of influence of the information carried by the input signal on the final result which is the activation of the motor. Network learning is necessary each time after starting, because the selected weights are not saved on any durable medium and in the case of turning off the power, the learning procedure should be performed again. Thanks to this solution, the robot has the opportunity to train a different reaction to a given event. This process takes place iteratively.

Due to the fact that the network is taught every time the robot is started, the learning time has a significant impact on the operation of the device. It was important to find a solution that would allow fast and effective learning. The learning took place until the number of epoques equals to 150 or the maximum permissible error equals to 0.015 was reached. Weight values taken by individual neurons before starting learning process were designated randomly. Figure 6 compares learning times for a few standardly used modifications of the backpropagation algorithm (GD), i.e.: with momentum (GDM), with adaptive learning rate (GDA), momentum and adaptive learning rate (GDX) and scaled conjugate gradient (SCG) [4, 12, 14, 16].

The effectiveness of passing the obstacles by the robot for each of the training algorithms was at a similar level (the difference was about 12%). Therefore, in the further stage of the research, a network taught by gradient descent with momentum and adaptive learning rate backpropagation algorithm was used.

After training the network, it can begin to be used in the process of avoiding obstacles. Each input neuron corresponds to a specific sensor position (Fig. 3) – depending on the value received at the input, the robot decides whether and how to change the speed of a given motor, causing e.g. a turn. As in the case of the fuzzy logic algorithm, the measurement is carried out in five directions while driving. The robot has the task to drive in the direction determined by the farthest obstacle detected.

5. Experiment

The research involved the use of a mobile circular robot with a built-in programmable ESP8266 NodeMCU system. The software was created in C/C++ using the integrated development environment – Arduino IDE. The test involved uploading software in two versions – first based on the fuzzy logic algorithm and the second based on the artificial neural network.

The test used obstacles made of paper of various shapes and sizes. The algorithm tests were performed in the laboratory in area 3 m × 4 m. The test scenario assumed that robot moved forward until the obstacle was detected. The evasive maneuver and drive forward were proceed.

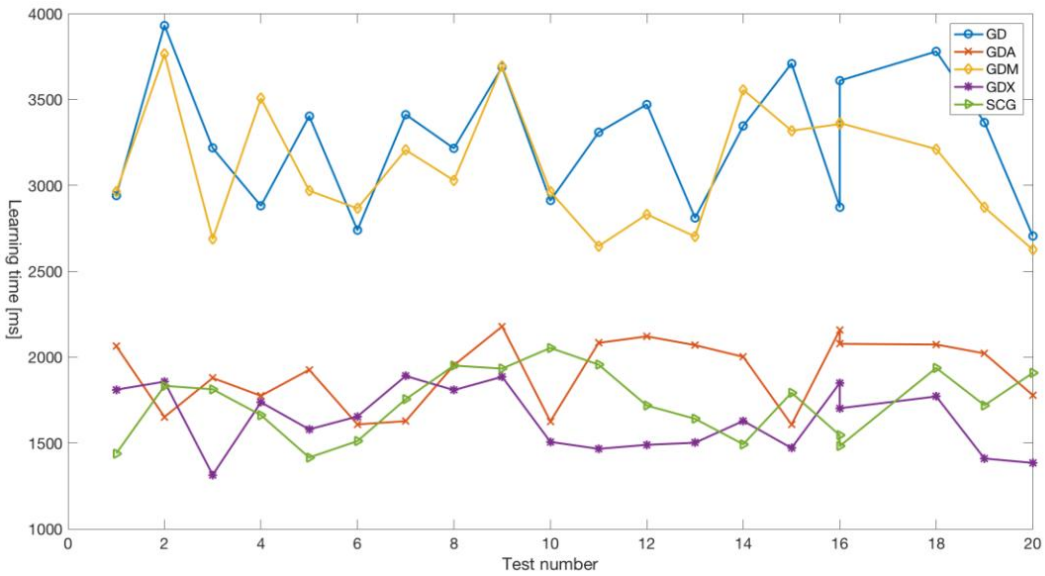


Fig. 6. Learning algorithms comparison

6. Discussion

As a result of the experiment, the analysis of the impact of mobile robot control algorithms in the process of avoiding obstacles was carried out. The algorithms based on fuzzy logic and using an artificial neural network were used.

In the case of the fuzzy logic algorithm, the calibration of parameters is necessary depending on the size of the room in which the test is carried out. Adjusting the speed of moving robot during the maneuver is necessary and depends on how much distance robot can overcome after the obstacle has been avoided. In addition, if the room size is small in relation to the number and size of obstacles it may be impossible to avoid without stopping. This situation was not considered in this work.

The next case was the use of an artificial neural network. Due to the fact that the set network weights were not permanently saved anywhere, it was necessary to train each time after turning on the power again. For this reason, the process of evading the obstacle looked different each time. Comparison of learning times for several modifications of the back propagation algorithm has been shown in figure 6.

In the case of both algorithms, there was a situation of insufficient adjustment of the robot's moving speed in relation to the sensor driving servo. It was possible that the sensor did not detect an obstacle (blind spot) and caused a collision.

7. Summary

An experimental system of checking the impact of robot control algorithms on the process of passing obstacles has been developed and tested. The work presents the algorithms used. The main conclusions include:

The implementation process of the algorithm based on fuzzy logic is much simpler and takes less time, but requires initial calibration, which may cause difficulties if the robot is used in too large or too small areas.

The process of network training takes place each time after starting, it may cause a different reaction to the same situations.

It is very important to adjust the speed of the robot's movement to the velocity of the servo positioning the sensor – it may happen that the obstacle is not detected because the robot moves too fast (or changes the position of the sensor too slowly).

Further work is planned to use wireless communication to signal the position of the obstacle and replace ultrasonic sensor with an optical. In addition, it will be possible to create a map of a given room.

References

[1] Adib A., Masoumi B.: Mobile robots navigation in unknown environments by using fuzzy logic and learning automata. Artificial Intelligence and Robotics (IRANOPEN), 2017, 58–63 [doi: 10.1109/RIOS.2017.7956444].

[2] Bajrami X., Dërmaku A., Demaku N., Maloku S., Kikaj A., Kokaj A.: Genetic and Fuzzy logic algorithms for robot path finding. 5th Mediterranean Conference on Embedded Computing (MECO), 2016, 195–199 [doi: 10.1109/MECO.2016.7525739].

[3] Boujelben M., Ayedi D., Rekik C., Derbel N.: Fuzzy logic controller for mobile robot navigation to avoid dynamic and static obstacles. 14th International Multi-Conference on Systems, Signals & Devices (SSD), 2017, 293–298 [doi: 10.1109/SSD.2017.8166963].

[4] Hammed A. A., Karlik B., Salman M. S.: Back-propagation algorithm with variable adaptive momentum. Knowledge-Based Systems 114, 2016, 79–87 [doi: 10.1016/j.knosys.2016.10.001].

[5] Handayani A. S., Dewi T., Husni N.L., Nurmaini S., Yani I.: Target tracking in mobile robot under uncertain environment using fuzzy logic controller. 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2017, 1–5 [doi: 10.1109/EECSI.2017.8239079].

[6] He K., Sun H., Cheng W.: Application of fuzzy neural network based on T-S model for mobile robot to avoid obstacles. 7th World Congress on Intelligent Control and Automation, 2008, 8282–8285 [doi: 10.1007/978-3-540-88513-9_120].

[7] Khan S., Ahmmed M. K.: Where am I? Autonomous navigation system of a mobile robot in an unknown environment. 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016, 56–61 [doi: 10.1109/ICIEV.2016.7760188].

[8] Mazare A., Ionescu L., Lita A., Serban G., Ionut M.: Mobile system with real time route learning using Hardware Artificial Neural Network. 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2015, 45–48 [doi: 10.1109/ECAI.2015.7301250].

[9] McCulloch W., Pitts W.: A logical calculations of the ideas in nervous activity. Bulletin of Mathematical Biophisics 5, 1943, 115–133.

[10] Mohammad S. H. A., Jeffril M. A., Sariff N.: Mobile robot obstacle avoidance by using Fuzzy Logic technique. IEEE 3rd International Conference on System Engineering and Technology, 2013, 331–335 [doi: 10.1109/ICSEngT.2013.6650194].

[11] Panigrahi P.K., Ghosh S., Parhi D.R.: A novel intelligent mobile robot navigation technique for avoiding obstacles using RBF neural network. International Conference on Control, Instrumentation, Energy and Communication (CIEC), 2014, 1–6 [doi: 10.1109/CIEC.2014.6959038].

[12] Powroźnik P., Czerwiński D.: Effectiveness comparison on an artificial neural networks to identify Polish emotional speech. Przegląd Elektrotechniczny 07/2016, 45–48 [doi: 10.15199/48.2016.07.08].

[13] Stączek P.: Digital signal processing in ultrasonic based navigation system for mobile robots. ITM Web Conf. 15, 2017 [doi:10.1051/itmconf/20171505008].

[14] Tiwari S., Naresh R.: Comparative study of backpropagation algorithms in neural network based identification on power system, International Journal of Computer Science and Information Technology 5(4), 2013, 93–107 [doi: 10.5121/ijcsit.2013.5407].

[15] Wu T. F., Tsai P. S., Hu N. T., Chen J. Y.: Use of Ultrasonic Sensors to Enable Wheeled Mobile Robots to Avoid Obstacles. Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2014, 958–961 [doi: 10.1109/IIH-MSP.2014.240].

[16] Yong L., Yang F., Hui L., Si-Wen Z.: The Improved Training Algorithm of Back Propagation Neural Network with Selfadaptive Learning Rate, International Conference on Computational Intelligence and Natural Computing, 2009, 73–76 [doi:10.1109/CINC.2009.111].

M.Sc. Piotr Wójcicki
e-mail: p.wojcicki@pollub.pl

Piotr Wójcicki graduated from the Lublin University of Technology, where he now works as assistance in the Institute of Computer Science. His previous research is related to microelectronics, mobile robotics, IoT and applied computer science.

ORCID ID: 0000-0002-0522-6223

Ph.D. Paweł Powroźnik
e-mail: p.powroznik@pollub.pl.

Paweł Powroźnik graduated from the Maria Curie-Skłodowska University in Lublin. Ph.D. degree recived in 2018 at Lublin University of Technology. Currently, he works as an assistant in Institute of Computer Science, Lublin University of Technology. His previous research is related to digital signal processing technology, wavelet transform and artificial neural networks.

ORCID ID: 0000-0002-5705-4785

Ph.D. Kamil Żyła
e-mail: k.zyla@pollub.pl

Employee of the Institute of Computer Science at the Lublin University of Technology. His research interests include: model-driven engineering, software engineering and mobile technologies.

ORCID ID: 0000-0002-6291-003X

D.Sc. Stanisław Grzegórski
e-mail: s.grzegorski@pollub.pl

Stanisław Grzegórski graduated from the Maria Curie-Skłodowska University in Lublin. He was the Head of Institute of Computer Science at the Lublin University of Technology. He is an associate professor there. His main scientific interests include computer science, numerical methods, algorithms and IT solutions developing.

ORCID ID: 0000-0001-7640-6195

otrzymano/received: 13.11.2018 *przyjęto do druku/accepted: 15.12.2018*

DOI: 10.5604/01.3001.0012.8046

MODIFICATIONS OF MAXIMUM POWER POINT TRACKING P&O METHOD FOR PHOTOVOLTAIC PANEL

Piotr Kozierski[1,2], Adam Owczarkowski[3], Marcin Lis[3], Dariusz Horla[2]
[1]Poznan University of Technology, Faculty of Computing, Institute of Automation and Robotics, Division of Signal Processing and Electronic Systems
[2]Poznan University of Technology, Faculty of Electrical Engineering, Institute of Control, Robotics and Information Engineering
[3]Spirvent sp. z o.o., Pokrzywno street 4A, 61-315 Poznan

Abstract. *Perturbation and Observe method for maximum power point tracking is presented in this paper. Three method modifications have been proposed, which allow satisfying tracking efficiency, even for very fast and noisy irradiance changes.*

Keywords: tracking, photovoltaic panels, perturbation and observe, MPPT

MODYFIKACJE METODY P&O ŚLEDZENIA MAKSYMALNEGO PUNKTU MOCY DLA PANELU FOTOWOLTAICZNEGO

Streszczenie. *W artykule przedstawiono metodę zaburzania i obserwacji P&O do śledzenia maksymalnego punktu mocy. Zaproponowano trzy modyfikacje metody, dzięki którym efektywność śledzenia jest zadowalająca, nawet dla bardzo szybkich oraz zaszumionych zmian irradiancji.*

Słowa kluczowe: śledzenie, panele słoneczne, zaburzanie i obserwacja, MPPT

Introduction

Thanks to Maximum Power Point Tracking (MPPT) methods it is possible to obtain a maximum power of device under given conditions. These methods are used mainly when device model is too complex, when too many variables should be taken into account (including cases in which these variables are difficult to measure), and also when satisfying results must be obtained with relatively low computing costs. MPPT methods are applicable among others in energy generation from renewable energy sources, e.g. from wind [1, 13] or sunlight [10,11].

There are many different MPPT methods, but Perturbation and Observe (P&O) is the most often used method, mainly due to the simplicity of implementation while achieving quite good tracking results [2], also for changing weather conditions. This method is described in Section 1 in more detail.

The all methods one can divide into direct and indirect [14]. The indirect methods require some knowledge about working device, its characteristics, et cetera. Therefore, they require a certain "initial work", however the algorithms are in fact quite simple and fast. Short Circuit Current method, Open Circuit Voltage method, Curve Fitting method and Look-up Table method belong to such algorithms.

The second group of techniques includes methods, which do not require any additional information about used devices, and moreover, they can work in variable climatic conditions. This group includes such algorithms as Differential method, previously mentioned P&O method, Conductance Incremental method [3, 6] and Forced Oscillations method.

The additional subgroup – Hill Climbing Techniques – was specified in [4]. In these methods the device operation point is shifted in the direction, which increases the device output power. P&O and Conductance Incremental methods belong to this techniques.

A separate algorithms group are methods, which based on the neural networks, fuzzy logic [12] or genetic algorithms [8].

In the presented work the authors are focused on the P&O method and few modifications to this MPPT algorithm were proposed.

In Section 1 P&O method principle of operation was described. In the second section the model of photovoltaic panel was described – based on this model further simulations were performed. The third and fourth sections contain simulation results and descriptions of all modifications. In the last section the whole paper was summarized.

1. Perturbation and Observe (P&O) method

As it was mentioned earlier, P&O method is very commonly used. The algorithm is easy to implement – in subsequent steps, the voltage of PV generator is forced to change, and then it is observed whether the power of the generator has increased or decreased [14]. If the power has increased, voltage changes are continued in the same direction, and if the power has decreased, the changes direction should be set to the opposite.

The algorithm, which appears in literature [7, 9] and presents operation principle of the P&O method is presented in Fig. 1, whereas the authors propose a pseudocode (see Algorithm 1), for better readability.

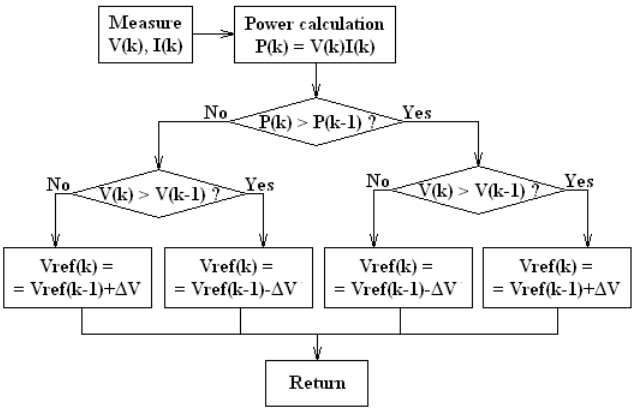


Fig. 1. Scheme of operation principle of the P&O method

Algorithm 1 – Principle of operation of the P&O method
1. Set the voltage changes direction $d_{ir} = 1$ and changes step ΔV.
2. Measure voltage $V^{(k)}$ and current $I^{(k)}$ at k-th time step. Calculate power $P^{(k)}=V^{(k)} \cdot I^{(k)}$.
3. If $P^{(k)} < P^{(k-1)}$, then $d_{ir} = - d_{ir}$.
4. $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} + d_{ir} \cdot \Delta V$.
5. Go to step 2.

However, this technique has two disadvantages – oscillations near the optimal value and low tracking quality, when irradiance (power of light per unit area [W/m^2]) grows rapidly [2].

In the paper [9] the modified P&O method was presented. In this algorithm every second step the voltage is not changed and the power change resulting from changes in atmospheric

conditions is checked. In the next step it is decided if the changes direction is proper. Also the second modification was proposed, in which this verification of direction is made less frequently, every third step.

In the work [7] authors proposed modification, which introduces hysteresis and step of changes autotuning mechanism. Also in the article [2] one has focused on the adaptation varieties of P&O, which allow changes of the step ΔV size.

In this paper the authors also proposed few modifications of the P&O method and thanks to this the improvement of MPPT method was obtained.

2. Model of Photovoltaic Panel

The model of photovoltaic panel (PV), which was used in the research, can be described by the equations [2]

$$I = I_{PV} \cdot N_p - I_{D1} - I_{D2} - \frac{V + N \cdot I \cdot R_s}{N \cdot R_p}, \tag{1}$$

$$I_{D1} = I_{d1} \cdot N_p \cdot \exp\left(\frac{V + N \cdot I \cdot R_s}{a_1 \cdot V_{T1} \cdot N_s}\right), \tag{2}$$

$$I_{D2} = I_{d2} \cdot N_p \cdot \exp\left(\frac{V + N \cdot I \cdot R_s}{a_2 \cdot V_{T2} \cdot N_s}\right), \tag{3}$$

$$I_{PV} = \left(I_{PV\_STC} + K_I\left(T - T_{STC}\right)\right) \cdot \frac{G}{G_{STC}}, \tag{4}$$

$$I_{d1} = I_{d2} = \frac{I_{sc} + K_I\left(T - T_{STC}\right)}{\exp\left(\frac{V_{oc} + K_V\left(T - T_{STC}\right)}{V_T}\right) - 1}, \tag{5}$$

$$N = N_s / N_p, \tag{6}$$

where: I – PV current, V – PV voltage, N_p – number of panels connected in parallel, N_s – number of panels connected in series, R_s – series resistance, R_p – parallel resistance, a_1 and a_2 – ideality factor for the first and second diode, V_T – thermal voltage of the diodes, G – irradiance, T – temperature, K_I – current temperature coefficient, I_{sc} – short circuit current, V_{oc} – open collector voltage, K_V – voltage temperature coefficient, STC – refers to values in Standard Test Conditions.

As one can easily see, after substitutions, calculated in (1) PV current value is also on the right side of the equation and in power. To solve this equation one should use W Lambert function or any numeric techniques.

The characteristics, which were obtained from the model of photovoltaic panel were presented in Fig. 2–3. The results of the studies presented later in the article were performed using this model.

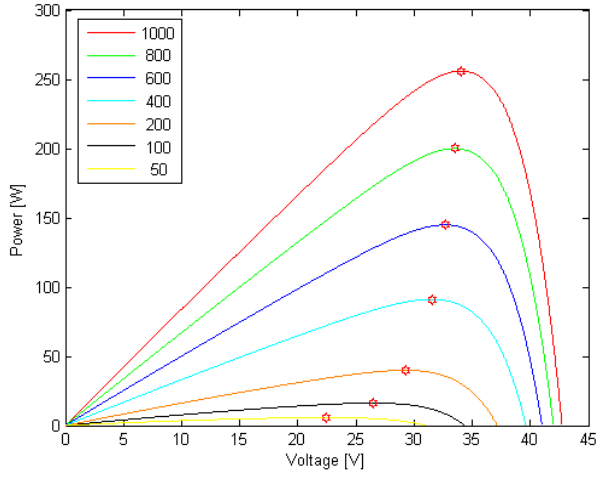


Fig. 2. Characteristics of power from voltage for different irradiance values; points in which power value is the highest are marked by red stars

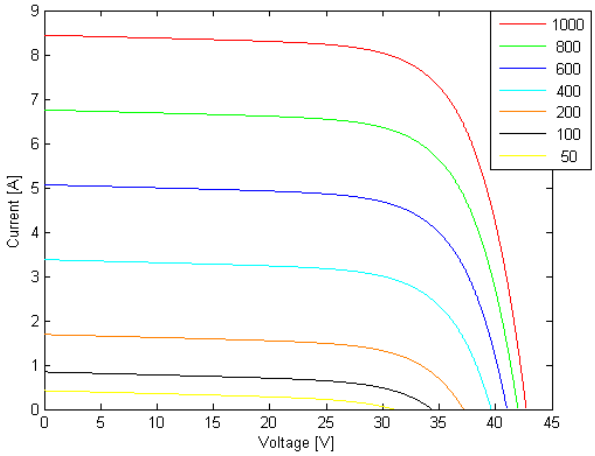


Fig. 3. Characteristics of current from voltage for different irradiance values

3. Performed simulations

It was assumed in the performed simulations that reference voltage $V^{(k)}_{\text{ref}}$ is equal to the measured value $V^{(k+1)}$, and current $I^{(k)}$ is calculated based on the equations (1–6). Thanks to this approach it was possible to focus entirely on the P&O algorithm and on the proposed modifications.

The theoretical course of irradiance was proposed in the studies

$$G = 500\sin\left(\frac{t\pi}{2}\right) + 550\left[\frac{W}{m^2}\right] \tag{7}$$

and also theoretical course of temperature

$$T = 57.5\sin\left(\frac{t\pi}{30}\right) + 12.5\left[^{\circ}\text{C}\right] \tag{8}$$

for time t varying from 0 to 60 seconds. Based on the equations (1–6) the waveform of maximum power point was obtained – see Fig. 4. As one can see, a fifteenfold brightening and dimming of sunlight within a minute was assumed. It means that the time between minimum and maximum irradiance is only 2 seconds.

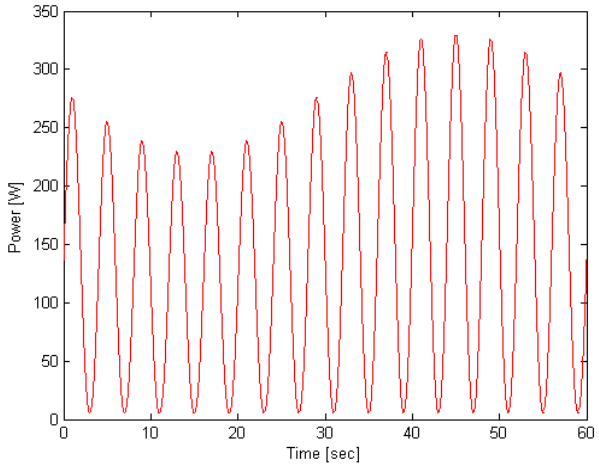


Fig. 4. The waveform of the maximum power point for a given model and weather conditions

To compare the quality of MPPT methods the efficiency was calculated [2, 5, 6]

$$\eta = \frac{\int P\,dt}{\int P_{\max}\,dt} = \frac{\sum_{k=1}^{N_k} P^{(k)}}{\sum_{k=1}^{N_k} P^{(k)}_{\max}}, \tag{9}$$

where N_k is number of time steps in whole simulation (the length of time step was set on 0.5 ms, what gives 120 thousands of time steps during 60 seconds of simulation), and $P^{(k)}_{\max}$ is a maximum

power, which can be obtained from PV in given weather conditions at k-th time step.

It was assumed that the time step of MPPT method is equal to 2.5 ms, the same time was proposed in [9]. Between subsequent MPPT steps it was assumed that voltage is constant and equal to the last calculated value.

The simulation results for conditions described above, with changes step ΔV = 0.05 [V], are presented in Fig. 5, and the efficiency was equal to η = 0.831 (and calculating from the second period of irradiance, to ignore the effect of the initial power increase, η = 0.8454 was obtained).

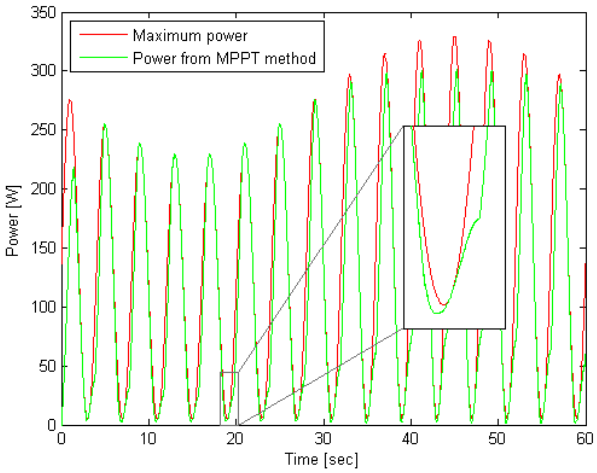


Fig. 5. Characteristics of power in time – simulation results of the basic P&O method (Algorithm 1); obtained efficiency was equal to η = 0.831 (η = 0.84543 calculating from the second period)

4. Modifications of P&O method and obtained results

In the Fig. 5 one can see enlarged chart fragment. It can be conclude that the highest power loses in basic P&O method (Algorithm 1) are during increase of the irradiance – power obtained from MPPT method grows slower and slower, and at some point "something switches" and rapid growth starts. What is going on? To the "switch" moment the voltage still decreases, because voltage reduction by ΔV has lower impact on obtained power than the irradiance G growth.

Accordingly, the modification was proposed – every period of time (specifically it was proposed every 50 ms, i.e. every 20 steps of MPPT method) the change of voltage is tried in the opposite direction. If the power growth in subsequent step will be higher than in current step, then the direction will be changed to the opposite. And if no – voltage changes should be continued in the right direction, but the wrong attempt should be made up by triple ΔV step. Remaining 19 MPPT steps should be carried out according to the standard P&O method. Principle of operation is presented by pseudocode in Algorithm 2.

Algorithm 2 – The first P&O modification

1. Set the voltage changes direction d_{ir} = 1, multiplicity d_{mult} = 1 and changes step ΔV = 0.05.
2. Measure voltage $V^{(k)}$ and current $I^{(k)}$ at k-th time step. Calculate power $P^{(k)}=V^{(k)}\cdot I^{(k)}$.
3. If modulo(k/20) == 0, then perform step 4:
   4. If $P^{(k)} > P^{(k-1)}$ and $P^{(k)}-P^{(k-1)} > P^{(k-1)}-P^{(k-2)}$, then $d_{ir} = -d_{ir}$.
   4a. Otherwise d_{mult} = 3.
5. If modulo(k/20) == 19, then $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} - d_{ir}\cdot\Delta V$.
5a. Otherwise perform steps 6-7:
   6. If modulo(k/20) ≠ 0 and $P^{(k)} < P^{(k-1)}$, then $d_{ir} = -d_{ir}$.
   7. $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} + d_{ir}\cdot\Delta V\cdot d_{mult}$.
8. Set d_{mult} = 1; go to step 2.

Using algorithm with the first modification the efficiency η = 0.967 was obtained (η = 0.988 calculating from the second period).

Afterwards, fragment before minimum value in Fig. 5 was improved. To understand, where is the problem, the voltage waveform was checked – it is shown in Fig. 6.

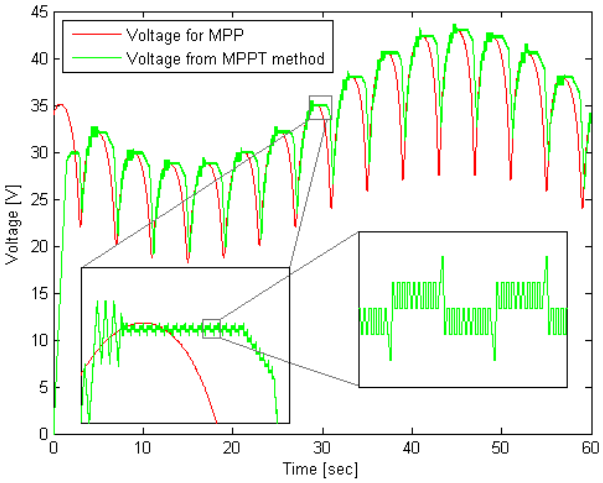


Fig. 6. Characteristics of voltage in time – simulation results of P&O method with the first modification (Algorithm 2); it was obtained η = 0.967 (η = 0.988 calculating from the second period)

As one can see, the changes direction d_{ir} has opposite sign in almost every MPPT step. Every 20 MPPT steps one can see impact of the first modification; however, after that d_{ir} changes are continued. It is because of the power decrease, which is caused by the irradiance decreasing. And the impact of irradiance is much bigger than the impact of voltage changes.

Therefore, the second modification of P&O method was proposed – it checks whether change of direction occured 5 or more times in a row. If yes, then it is checked how much the power was decreased in the latest steps. If alternately there is higher and lower decrease, the direction which provides lower decrease is taken and multiplicity is set on d_{mult} = 3. In such a way voltage decreasing is forced even during irradiance decreasing. Precise operation of principle was presented in Algorithm 3.

Algorithm 3 – The second P&O modification

1. Set the voltage changes direction d_{ir} = 1, multiplicity d_{mult} = 1, changes counter d_{zm} = 0 and changes step ΔV = 0.05.
2. Measure voltage $V^{(k)}$ and current $I^{(k)}$ at k-th time step. Calculate power $P^{(k)}=V^{(k)}\cdot I^{(k)}$.
3. If modulo(k/20) == 0, then perform step 4:
   4. If $P^{(k)} > P^{(k-1)}$ and $P^{(k)}-P^{(k-1)} > P^{(k-1)}-P^{(k-2)}$, then $d_{ir} = -d_{ir}$ and $d_{zm} = d_{zm} + 1$.
   4a. Otherwise d_{mult} = 3 and d_{zm} = 0.
5. If modulo(k/20) == 19, then $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} - d_{ir}\cdot\Delta V$.
5a. Otherwise perform steps 6-11:
   6. If $d_{zm} < 5$, then perform steps 7-8:
      7. If modulo(k/20) ≠ 0 and $P^{(k)} < P^{(k-1)}$, then $d_{ir} = -d_{ir}$ and $d_{zm} = d_{zm} + 1$.
      7a. Otherwise d_{zm} = 0.
      8. $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} + d_{ir}\cdot\Delta V\cdot d_{mult}$.
   6a. Otherwise perform steps 9-11:
      9. If $P^{(k-1)}-P^{(k-2)} > P^{(k-2)}-P^{(k-3)}$ and $P^{(k-1)}-P^{(k-2)} > P^{(k-4)}-P^{(k-5)}$ and $P^{(k-3)}-P^{(k-4)} > P^{(k-2)}-P^{(k-3)}$ and $P^{(k-3)}-P^{(k-4)} > P^{(k-4)}-P^{(k-5)}$, then $d_{ir} = -d_{ir}$ and d_{mult} = 3.
      10. If $P^{(k-1)}-P^{(k-2)} < P^{(k-2)}-P^{(k-3)}$ and $P^{(k-1)}-P^{(k-2)} < P^{(k-4)}-P^{(k-5)}$ and $P^{(k-3)}-P^{(k-4)} < P^{(k-2)}-P^{(k-3)}$ and $P^{(k-3)}-P^{(k-4)} < P^{(k-4)}-P^{(k-5)}$, then d_{mult} = 3.
      11. $V^{(k)}_{\text{ref}} = V^{(k-1)}_{\text{ref}} + d_{ir}\cdot\Delta V\cdot d_{mult}$ and d_{zm} = 0.
12. Set d_{mult} = 1; go to step 2.

Using P&O algorithm with the second modification, the efficiency η = 0.9782 was obtained (calculating from the second period η = 0.9993). It can be seen that algorithm with modifications works very well; however, it is not robust for any device noise or slight (in comparison to the general changes)

irradiance fluctuations. It is due to the fact that algorithms are based on values differences in specific time moments. If any "noise" will be higher than changes step ΔV, then all these modifications will be have insignificant influence on obtained quality of tracking.

Therefore, the irradiance waveform was changed – the Gauss noise was added

$$G = 500\sin\left(\frac{t\pi}{2}\right) + 550\left[\frac{W}{m^2}\right] + N\left(0,\sigma^2\right) \quad (10)$$

and the third P&O modification was proposed. This modification change the ΔV step based on the power from last 50 ms (21 MPPT steps). Specifically, this value should be equal to

$$\Delta V = \max\left[0.05\cdot\left(\log_{10}\left(\mathrm{var}\left[P^{(k-20)},\ldots,P^{(k)}\right]\right)+3\right);0.05\right]. \quad (11)$$

The operation principle was proposed in Algorithm 4.

Algorithm 4 – The third P&O modification

1. Set the voltage changes direction $d_{ir} = 1$, multiplicity $d_{mult} = 1$, changes counter $d_{zm} = 0$, changes step $\Delta V = 0.05$ and step of the last calculation of changes step $k_{\Delta V} = 0$.
2-11. Steps 2-11 are identical as in Algorithm 3.
12. If $k>100$ and $|P^{(k-20)} - P^{(k)}| < 0.2$ and $k - k_{\Delta V} > 1000$ and $V^{(k)} > 10$, then perform steps 13-15:
    13. $k_{\Delta V} = k$.
    14. Calculate power variance v_P from $P^{(k-20)}$ to $P^{(k)}$.
    15. If $v_P < 0.01$, then $\Delta V = 0.05$.
    15a. Otherwise $\Delta V = 0.05\cdot(\log_{10}(v_P) + 3)$.
16. Set $d_{mult} = 1$; go to step 2.

As one can see few conditions were introduced, among others one which does not allow to change ΔV more than once every half a second. The efficiency of all algorithms for different noise standard deviations σ is presented in Table 1

Table 1. Efficiency results of all presented MPPT methods for different noise variance (results in brackets were calculated from the second period)

σ	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
0	0.8310 (0.8454)	0.9670 (0.9880)	0.9782 (0.9993)	0.9783 (0.9993)
0.2	0.8860 (0.9020)	0.9636 (0.9846)	0.9736 (0.9950)	0.9739 (0.9952)
1	0.9628 (0.9855)	0.9641 (0.9866)	0.9665 (0.9890)	0.9746 (0.9957)
3	0.9450 (0.9746)	0.9451 (0.9745)	0.9420 (0.9730)	0.9709 (0.9944)
5	0.9181 (0.9534)	0.9260 (0.9599)	0.9310 (0.9653)	0.9675 (0.9909)
10	0.8730 (0.9118)	0.8799 (0.9177)	0.8734 (0.9136)	0.9514 (0.9847)

5. Summary

The operation principle of P&O MPPT method was presented in the article, and also three P&O modifications were proposed – thanks to them algorithm has a good tracking quality, even at rapid and noisy irradiance changes.

Acknowledgements

Research was cofinanced by The National Centre for Research and Development as part of the Intelligent Development operational program 2014-2020, operation 1.1/suboperation 1.1.1, application number POIR.01.01.01-00-0468/17.

References

[1] Abdullah M. A., Yatim A. H. M., Tan C. W.: A study of maximum power point tracking algorithms for wind energy system. In Clean Energy and Technology (CET), 2011 IEEE First Conference, 321–326.
[2] Ahmed J., Salam Z.: A modified P&O maximum power point tracking method with reduced steady-state oscillation and improved tracking efficiency. IEEE Transactions on Sustainable Energy 4/2016, 1506–1515.
[3] Chung T. M., Daniyal H., Sulaiman M. H., Bakar M. S.: Comparative study of P&O and modified incremental conductance algorithm in solar maximum power point tracking. In 4th IET Clean Energy and Technology Conference (CEAT) 11/2016, 5.
[4] Gaur P., Verma Y. P., Singh P.: Maximum power point tracking algorithms for photovoltaic applications: A comparative study. In Recent Advances in Engineering & Computational Sciences (RAECS), 2nd International Conference 12/2015, 5.
[5] Hohm D. P., Ropp M. E.: Comparative study of maximum power point tracking algorithms. Progress in photovoltaics: Research and Applications 1/2003, 47–62 [doi: 10.1002/pip.459].
[6] Hussein K. H., Muta I., Hoshino T., Osakada M.: Maximum photovoltaic power tracking: an algorithm for rapidly changing atmospheric conditions. IEE Proceedings-Generation, Transmission and Distribution 1/1995, 59–64.
[7] Jung Y., So J., Yu G., Choi J.: Improved perturbation and observation method (IP&O) of MPPT control for photovoltaic power systems. In Photovoltaic Specialists Conference, Conference Record of the Thirty-first IEEE 01/2005, 1788–1791.
[8] Lasheen M. A., Bendary F. M., Sharaf M., El-Zoghby H. M.: Maximum power point tracking of a wind turbine driven by synchronous generator connected to an isolated load using genetic algorithm. In Smart Grid Conference (SASG), Saudi Arabia 12/2014, 1–9.
[9] Liu C., Wu B., Cheung R.: Advanced algorithm for MPPT control of photovoltaic systems. In Canadian Solar Buildings Conference, Montreal 08/2004, 20–24.
[10] More D. N., Naik A. V., Kumar V. C.: Maximum power point using P&O technique for photovoltaic system. In Technological Innovations in ICT for Agriculture and Rural Development (TIAR) 04/2017, 134–137.
[11] Seyedmahmoudian M., Horan B., Rahmani R., Maung Than Oo A., Stojcevski A.: Efficient photovoltaic system maximum power point tracking using a new technique. Energies 147(9)/2016, 1–18.
[12] Surma P.: Porównanie metod MPPT paneli fotowoltaicznych (P&O, IC, Fuzzy Logic) w środowisku Matlab Simulink. Przegląd Elektrotechniczny 1/2014, 66–69.
[13] Van T. L., Nguyen D. Q., Duy V. H., Nguyen H.: Fast Maximum Power Point Tracking Control for Variable Speed Wind Turbines. In International Conference on Advanced Engineering Theory and Applications 12/2017, 821–829.
[14] Zaremba A., Rodziewicz T., Wacławek M.: Algorytmy śledzenia punktu mocy maksymalnej (MPPT) w systemach fotowoltaicznych. Proceedings of ECOpole 2/2012, 805–810 [doi: 10.2429/proc.2012.6(2)112].

M.Sc. Eng. Piotr Kozierski
e-mail: piotr.kozierski@gmail.com

Third year full-time third degree student in Poznan University of Technology, Faculty of Computing. Is interesting in normal and whispery speech recognition using Kaldi toolkit. Additionally research area is associated with state estimation of nonlinear, multivariable plants and with particle filters.

ORCID ID: 0000-0001-8777-6132

Ph.D. Eng. Adam Owczarkowski
e-mail: adam.owczarkowski@spirvent.pl

Ph.D. in control and robotics from 2017 (at Poznan University of Technology) He conducted research on dynamics of two-wheeled vehicles with reaction wheel. Is interested in scientific research project realization associated with modern theory of non-linear systems control, informatics and robotics.

ORCID ID: 0000-0002-1398-6484

M.Sc. Eng. Marcin Lis
e-mail: marcin.lis@spirvent.pl

Over the years, he worked for such companies as Energo and Spirvent in positions related to research and development activities. Earlier, he devoted himself to scientific research on the use of neural networks in control.

ORCID ID: 0000-0001-7324-2383

D.Sc. Eng. Dariusz Horla
e-mail: dariusz.horla@put.poznan.pl

Is working in Institute of Control, Robotics and Information Engineering at Poznan University of Technology (PUT). Ph.D. in 2005, Post-doctoral degree in 2013 in technical sciences on Faculty of Electrical Engineering at PUT. His area of interest is windup compensation, optimal control, adaptive control, optimization theory and linear matrix inequalities.

ORCID ID: 0000-0002-9456-6704

otrzymano/received: 14.09.2018 *przyjęto do druku/accepted: 15.12.2018*

LISTA RECENZENTÓW WSPÓŁPRACUJĄCYCH Z CZASOPISMEM W ROKU 2018

prof.	Ageiev Dmytro	Kharkiv National University of Radio Electronics (Charków, Ukraina)
prof.	Barabash Oleg	State University of Telecommunications (Kijów, Ukraina)
prof.	Barylo Grygoriy	Narodowy Uniwersytet "Politechnika Lwowska" (Lwów, Ukraina)
prof.	Bezruk Valeriy	Kharkiv National University of Radio Electronics (Charków, Ukraina)
dr	Bitchenko Oleksandr	Ukrainian Engineering and Pedagogical Academy (Charków, Ukraina)
dr	Boyko Oksana	Danylo Halytsky Lviv National Medical University (Lwów, Ukraina)
prof.	Brauer Hartmut	Technische Universitat Ilmenau (Ilmenau, Niemcy)
dr	Braunig Sylvia	Technische Universität Ilmenau (Ilmenau, Niemcy)
dr	Brygilevych Volodymyr	Państwowa Wyższa Szkoła Techniczno-Ekonomiczna w Jarosławiu
prof.	Chaplyha Vyacheslav	Lviv Institute of Banking University (Lwów, Ukraina)
prof.	Chevardin Vadym	Kyiv National University of Construction and Architecture (Kijów, Ukraina)
prof.	Cieśla Antoni	Akademia Górniczo-Hutnicza w Krakowie
prof.	Cieślik Sławomir	Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy
prof.	Czerwiński Dariusz	Politechnika Lubelska
prof.	Dorosz Dominik	Akademia Górniczo-Hutnicza w Krakowie
prof.	Druzhinin Volodymyr	National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” (Kijów, Ukraina)
dr	Gorbulyk Volodymyr	National Technical University “Kharkiv Polytechnic Institute” (Charków, Ukraina)
dr	Grela Jakub	Akademia Górniczo-Hutnicza w Krakowie
dr	Grygorovych Victor	Drohobych Ivan Franko State Pedagogical University (Drohobycz, Ukraina)
dr	Haliuk Serhii	Yuriy Fedkovych Chernivtsi National University (Czerniowce, Ukraina)
prof.	Hlaponin Yurii	Kyiv National University of Construction and Architecture (Kijów, Ukraina)
prof.	Hotra Zenon	Lviv Polytechnic National University (Lwów, Ukraina)
prof.	Ivakhiv Orest	Lviv Polytechnic National University (Lwów, Ukraina)
prof.	Kaidan Mykola	Lviv Polytechnic National University (Lwów, Ukraina)
prof.	Klymash Mykhailo	Lviv Polytechnic National University (Lwów, Ukraina)
prof.	Konakhovich Georgiy	National Aviation University (Kijów, Ukraina)
prof.	Koroliov Dmitri	Institute of Telecommunications and Global Information Space (Kijów, Ukraina)
prof.	Koshevyy Vitaliy	National University “Odessa Maritime Academy” (Odessa, Ukraina)
prof.	Kotyra Andrzej	Politechnika Lubelska
dr	Kovaluk Zakhar	Chernivtsi Division of I. M. Frantsevich Institute of Materials Science Problems (Czerniowce, Ukraina)
dr	Kushnir Mykola	Yuriy Fedkovych Chernivtsi National University (Czerniowce, Ukraina)
dr	Kusznier Jacek	Politechnika Białostocka
prof.	Liashenko Sergey	Kharkiv Petro Vasylenko National Technical University of Agriculture (Charków, Ukraina)
prof.	Liubchyk Vitalii	Khmelnytskyi National University (Chmielnicki, Ukraina)
prof.	Lukin Vladimir	Kharkiv Aviation Institute (Charków, Ukraina)
dr	Melnychyn Andriy	Lviv Ivan Franko State University (Lwów, Ukraina)
prof.	Mosorov Volodymyr	Politechnika Łódzka
prof.	Mykytyuk Zinoviy	Lviv Polytechnic National University (Lwów, Ukraina)
dr	Naumchuk Oleksandr	National University of Water and Environmental Engineering (Równe, Ukraina)
prof.	Osadchuk Alexander	Vinnitsa National Technical University (Winnica, Ukraina)
prof.	Osuch Tomasz	Politechnika Warszawska
prof.	Ozirkovskyy Leonid	Lviv Polytechnic National University (Lwów, Ukraina)
prof.	Pałka Ryszard	Zachodniopomorski Uniwersytet Techniczny w Szczecinie
prof.	Peleshchak Roman	Drohobych Ivan Franko State Pedagogical University (Drohobycz, Ukraina)
prof.	Politanskyi Leonid	Yuriy Fedkovych Chernivtsi National University (Czerniowce, Ukraina)
prof.	Politanskyi Ruslan	Yuriy Fedkovych Chernivtsi National University (Czerniowce, Ukraina)
prof.	Polozhaenko Sergiy	Odessa National Polytechnic University (Odessa, Ukraina)
dr	Rolek Jarosław	Politechnika Świętokrzyska
prof.	Ruzhentsev Mykola	Kharkiv Aviation Institute (Charków, Ukraina)
prof.	Ryabova Zoya	University of Educational Management of NAES of Ukraine (Kijów, Ukraina)
prof.	Semenko Anatoliy	State University of Telecommunications (Kijów, Ukraina)
prof.	Sharonova Natalia	National Technical University “Kharkiv Polytechnic Institute” (Charków, Ukraina)
prof.	Sikora Jarosław	Politechnika Lubelska
prof.	Sivyakova Galina	Karaganda State Industrial University (Temyrtau, Kazachstan)
dr	Slynko Vasyl	Institute of Materials Science Problems of NAS of Ukraine (Kijów, Ukraina)
prof.	Stepanov Mykola	National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” (Kijów, Ukraina)
prof.	Subach Ihor	National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” (Kijów, Ukraina)
prof.	Toepfer Hannes	Technische Universität Ilmenau (Ilmenau, Niemcy)
prof.	Tomashevskyi Valentyn	National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” (Kijów, Ukraina)
dr	Trokhimchuck Petro	Lesya Ukrainka Eastern European University (Łuck, Ukraina)
prof.	Udovenko Serhiy	Simon Kuznets Kharkiv National University of Economics (Charków, Ukraina)
dr	Vakhnyak Nadiia	Institute of Semiconductor Physics of NAS of Ukraine (Kijów, Ukraina)
prof.	Vasiuta Konstiantyn	Ivan Kozhedub Kharkiv National Aviation University (Charków, Ukraina)
prof.	Volochiy Bohdan	Lviv Polytechnic National University (Lwów, Ukraina)
dr	Yegoshyna Hanna	Yuriy Fedkovych Chernivtsi National University (Czerniowce, Ukraina)
dr	Yevseiev Serhii	Simon Kuznets Kharkiv National University of Economics (Charków, Ukraina)
prof.	Yuzevych Volodymyr	National Academy of Science of Ukraine (Kijów, Ukraina)
prof.	Zachko Oleh	Lviv State University of Life Safety (Lwów, Ukraina)
dr	Zagorodna Nataliya	Narodowy Uniwerystet Techniczny w Tarnopolu im. I. Puluja (Tarnopol, Ukraina)
prof.	Ziółkowski Marek	Technische Universität Ilmenau (Ilmenau, Niemcy)

International Interdisciplinary PhD Workshop

In 2019 the International Interdisciplinary PhD Workshop will take place in Wismar (Germany) at the Hochschule Wismar: University of Applied Sciences: Technology, Business and Design from 15th May to 17th May. The goal is to gather PhD students and postdocs in order to share knowledge and discuss problems related to their research and scientific interests. The workshop enables the participants to gain valuable experience that will reflect in their professional research careers. Importantly, the event also provides the opportunity to integrate with the scientific community and develop informal contacts. The session chairs are among the most renowned experts in the fields covered by the workshop. Thus, attending the event is the only way to meet these specialists and possibly ask some intricate questions.

The scope of the workshop includes, but is not limited to, the following engineering areas:

- Automation
- Civil Engineering
- Communications and Signal Processing
- Mechanical and Production Engineering
- Electrical Engineering
- Computer Science
- Maritime Studies and Transport Operations

Important Dates

Full paper submission due: 28.01.2019

Notification to authors: 25.03.2019

Camera-ready paper due: 15.04.2019

Contact

iiphdw2019@hs-wismar.de

André Sandmann
Hochschule Wismar
Fakultät für Ingenieurwissenschaften / EuI
Philipp-Müller-Straße 14
23966 Wismar (Germany)

https://www.hs-wismar.de/iiphdw2019/