Model of Assessment of Information-Psychological Influence in Social Networking Services Based on Information Insurance

Kateryna Molodetska^{1[0000-0001-9864-2463]}, and Yuriy Brodskiy^{2[0000-0002-6843-0192]},

Solomia Fedushko³*^[0000-0001-7548-5856]

¹Educational and scientific center of IT, Polissya National University, Zhytomyr, Ukraine ²Department of Computer Technologies and Systems Modeling, Polissya National University, Zhytomyr, Ukraine ³Lviv Polytechnic National University, 79013, Lviv, Ukraine

> kmolodetska@gmail.com¹, yubrodskiy@gmail.com² solomiia.s.fedushko@lpnu.ua³

Abstract. In the current context, social internet services play a leading role in various types of mass media. However, in the case of spreading destructive content with the aim of informational-psychological influence on the actors, social internet services can have a negative influence on the social and political processes in the country. The article summarizes and systematizes the effects created in virtual communities as a result of the implementation of threats to the information security of the state. We have originated the model for evaluating the informational-psychological impact on actors in the textual content of social networking services based on conditional entropy. The model takes into account not only the destructive informational-psychological impact from the content source in social internet services but also the increase of the destructive influence as a result of processing and further distribution of the content by other actors of virtual communities. This approach allows increasing the efficiency of detection of threats to the information security of the state in the information space of social networking services. The developed model can be used for the implementation of individual modules of the state information security system in order to automate the detection of threats in the information space and increase its efficiency.

Keywords: social networking service, virtual community, actor, information security, informational-psychological influence, conditional entropy, text content.

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). COAPSN-2020: International Workshop on Control, Optimisation and Analytical Processing of Social Networks

1 Introduction

The global development of information technology and the Internet have become systemic factors in the development of the information society. As a consequence, it has provoked changes in the system of strategic communication, where social networking services (SNS) have become the most popular means of communication [1-3]. Due to their high popularity, SNS can be used by the criminals to reach their own goals in the information space of services. Manipulation of public opinion, influence on the actors' freedom of choice, their emotional and mental state, discrediting the existing system of government in the state, etc. can be the result of information and psychological influence on SNS actors [4-9]. Such phenomena are widely spread not only in the national information space of Ukrainian SNS under the conditions of information war with the Russian Federation but all over the world. For instance, in January 2019, Russian media spread the information about Canada's imposition of sanctions on Russia as a result of the seizure of power in Canada by Ukrainian immigrants [10]. This is one of the numerous examples of a lengthy information campaign held by Russian state media to misrepresent Ukraine and its allies. Taking into account constant increase in the number of threats to the information security of the state in SNS, related to the dissemination of manipulative content, the problem of actualization of informational-psychological influence on citizens in the information space of services seems especially important. The development of technologies of latent information influence on the actors of virtual communities, the lack of universal and effective methods of detecting destructive influence determines the actuality of this work.

Analysis of recent research and publications [4, 11, 12–16] has shown that the methods of manipulating of public consciousness of SNS actors are constantly changing and improving. The existence of such a phenomenon is connected with the continuous development of information-psychological methods of influence on the actors in order to counteract its detection. In particular, the study [16] has found that the following are currently the most popular means of manipulating SNS:

- the use of labels in the posts of actors from the list of friends with whom the user is not personally familiar. This allows the content to be distributed among the actor's friends on his behalf;
- the purchase of virtual communities of actors that are popular in the SNS information space, ensuring that audience to be engaged in the spread of propaganda and campaign materials;
- blocking the profiles of actors, who are thought leaders, due to a large number of bot complaints. Thus, the functioning of the actor's page in the SNS is limited for up to 30 days;
- adding spam comments below the posts. That increases the size of the discussion and leaves only a few comments that are often unrelated;
- usage in the SNS of bots and trolls that have proven themselves to be the most effective tool for conducting a hybrid war. They are also used for information fighting at the level of state and local political groups;

- generation of virtual communities with limited access that unite like-minded people at the beginning of the creation and contribute to the effect of emotional resonance between actors. Being created, such communities disseminate destructive content that is fully perceived by its members which leads to manipulation of public opinion.
- usage of advertisement in SNS to spread unconcealed propaganda. Paid publications do not pass the required level of verification to detect fakes. As a result, it makes actors trust, capture, and distribute such content in the SNS.

Researchers [17] have also proved that the effects of provoking conflicts between actors in SNS are monotony, ambivalence and desensitization. Monotony is characterized by the usage of repeated calls, which makes actors stop responding to them and lose interest in current affairs in the society and the state. This effect is called "social fatigue".

Ambivalence is an effect associated with a disruption of sensible news perception, resulting in impulsive actions that are justified by encroachment on steady moral or cultural values. Decreasing the level of sensitivity and empathy as a result of ridiculing the manifestations of tolerance for violence or certain groups of the population has been called the desensitization effect.

On the other hand, despite the active development of information technology and psychological impact on actors in SNS, the development and improvement of methods for detecting such influence are significantly limited. This is connected, in particular, with the complexity of formalizing of the destructive content models and lack of consistent features of the mentioned content. We have conducted the critical analysis of publications devoted to the detection of manipulations in the information space which has shown that articles [18-22] propose formal models of virtual communities in SNS by describing them as an environment of information confrontation with characteristics of the audience, public importance, content, communication, state security allowing you to build a model of the management system for the protection of virtual information space. However, the use of expert polls to construct the models increases the level of the subjectivity of the results obtained. The proposed approach does not take into account the change in the characteristics of virtual communities with time. The article [23] proposes the method of detecting public opinion manipulations based on the intellectual analysis and consideration of information uncertainty in SNS text content. In this case, unconditional entropy is an integral indicator of the threat to the information security of the state in the SNS information space. However, increasing the efficiency of this task can be achieved by using conditional entropy "IF condition THEN - an event", which is a promising area for further research.

The article aims to increase the effectiveness of detecting informationalpsychological impact on SNS actors based on conditional entropy, which will allow taking into account not only the destructive informational impact from the source of content but also the increase of such an impact due to its processing and further distribution by other actors of virtual communities.

2 Models and methods

2.1 Content transformation with the informational-psychological impact as means for manipulating public opinion

Content distributed in SNS is created directly by actors and characterized by a diversity of forms and content. It reflects the subjective view of the author on events in the country, society and the world. Let us investigate the peculiarities of identifying signs of manipulation of public opinion in the textual content.

As researched in the publication [23], the connection between partial signs of information and psychological influence in the SNS is represented in the form of a hierarchy in a generalized form (Fig. [1]): F_1 – presence of a reference to a subjective point of view; F_2 – lack of argumentation; F_3 – share of interrogative sentences; F_4 – doubtful statements; F_5 – indicator of presence of exclamation sentences; F_6 – exclamations; F_7 – adverbs; F_8 – relative number of adverbs; F_9 – usage of emotional vocabulary; F_{10} – relative indicator of usage of the words to increase attention; F_{11} – indicator of the usage of the words to emphasize the responsiveness of events in content.

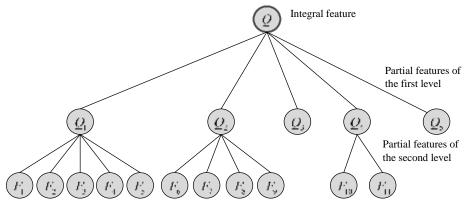


Fig. 1. Decision tree

According to the developed method [23], the signs $F_1 - F_5$ are used to establish the ambiguity of the facts Q_1 stated in the text content of the SNS $F_6 - F_9$ – its emotional colouring Q_2 , $F_{10} - F_{11}$ – manifestations of sensationalism Q_4 . The presence of features $F_1 - F_{11}$ in the content is determined on the basis of calculating the frequency of use of the given words-indicators. The hidden subject matter Q_3 and tone is connected with the usage of the methods of intellectual text analysis and machine learning. This introduces an indicator of the informational entropy of SNS text content, which determines the degree of uncertainty about the presence in the content of hidden information influence on the actor. Information entropy is calculated on the

basis of the number of detected words-indicators. It decreases with increasing frequency of their appearance and, accordingly, increases with low frequencies of manifestation of the given words [24–26].

Thus, textual content created by the participants of informational operations containing information-psychological influence on actors of virtual communities is characterized by some amount of incoming entropy H(x) or entropy of the content source. After being distributed in the SNS information space, the content of the virtual community actors is introduced. Then comes further distortion of the content due to the peculiarities of its perception by the consumers and its subsequent publication with the addition of subjective valuation judgments and conclusions.

Consequently, such content increases the level of information and psychological impact on SNS actors H(y/x), which will be characterized by conditional entropy. Then the loss of the meaning of the content after double distortion – first intentional which involves the participants of informational operation, and then conducted by the actors as the reaction to such content of directed meaning, is determined by the amount of entropy H(x/y). The process is presented in Fig. [2] in the form of a structural diagram.

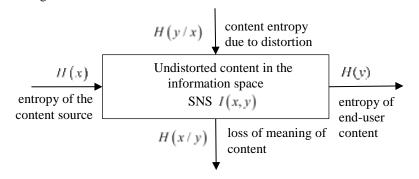


Fig. 2. The scheme of information processes in SNS in the conditions of manipulation of public opinion

Undistorted content in the SNS I(x, y) information space, which does not contain information and psychological influence on the actors, can be formalized in the form of the difference of the entropy of the content of the end consumer H(y) and the conditional entropy H(y/x), which characterizes the result of distorting the content by the actors, or through the entropy of the content source H(x) and conditional entropy H(x/y) which describes the loss of meaning of content as a result of its transformation by actors

$$I(x, y) = H(y) - H(y/x) = H(x) - H(x/y).$$
(1)

Therefore, the degree of uncertainty of the information-psychological impact on actors in the SNS text content is determined by the unreliability of the process of dissemination in the information space of virtual communities H(x/y) or by the level of additional distortion of the original content by the actors H(y/x).

2.2 Model of assessment of information-psychological influence

In general case, the scheme of interaction of actors, which is the source of content X_k , $k = \overline{1, n}$ in the SNS information space, and its consumers Y_l , $l = \overline{1, m}$ is presented in Fig. [3].

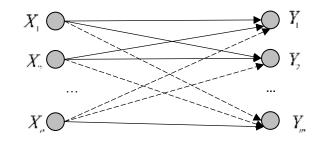


Fig. 3. Generalized scheme of content exchange between SNS actors

For convenience, let us limit ourselves to the assumption that each content source in the SNS and each consumer is connected by a network scheme, as shown in Fig. [4].

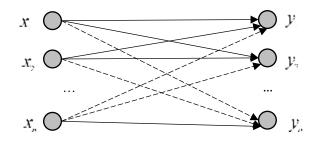


Fig. 4. A partial case of connection between the k -source and the l -content consumer

From the publication [23] it is known that while processing a large number of information messages in the SNS, the frequency of occurrence of the corresponding words-indicators approaches the probability of occurrence of signs of information and psychological impact. The model of such an influence should be described as a matrix of transient probabilities. The matrix will be constituted from n^2 elements, where n - the total number of signs of information and psychological influence on the actors. Each element of the matrix, depending on the content of the estimated information process in the SNS, represents a conditional probability $p(y_j / x_i)$ or $p(x_i / y_j)$. The

matrix under consideration describes the transformation of content into an SNS that spreads from the source X_k to the consumer Y_l and can take the following form

$$p(y/x) = \begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & \dots & p(y_n/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \dots & p(y_n/x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1/x_n) & p(y_2/x_n) & \dots & p(y_n/x_n) \end{bmatrix}.$$
 (2)

Note that the sum of the elements of each row is equal to one

$$\sum_{j=1}^{n} p(y_j / x_1) = \sum_{j=1}^{n} p(y_j / x_2) = \sum_{j=1}^{n} p(y_j / x_n) = 1.$$
(3)

If the matrix of transient probabilities is known, then the frequencies of occurrence of signs f_j^y of informational-psychological influence on the actors in the SNS at the output are explicitly determined by the frequencies of signs f_i^x in the content spread from the source, taking into account the transition probability $p(y_j / x_i)$

$$f_{j}^{y} = \sum_{i=1}^{n} f_{i}^{x} p(y_{j} / x_{i}), \qquad (4)$$

or expanded

$$f_{1}^{y} = f_{1}^{x} p(y_{1} / x_{1}) + f_{2}^{x} p(y_{1} / x_{2}) + \dots + f_{n}^{x} p(y_{1} / x_{n}),$$

$$f_{2}^{y} = f_{1}^{x} p(y_{2} / x_{1}) + f_{2}^{x} p(y_{2} / x_{2}) + \dots + f_{n}^{x} p(y_{2} / x_{n}),$$

$$\dots$$

$$f_{n}^{y} = f_{1}^{x} p(y_{n} / x_{1}) + f_{2}^{x} p(y_{n} / x_{2}) + \dots + f_{n}^{x} p(y_{n} / x_{n}).$$
(5)

Then the entropy of the content consumed by SNS actors can be written in a form of Shannon formula

$$H(y) = -\sum_{j=1}^{n} f_{j}^{y} \log_{2} f_{j}^{y}.$$
 (6)

Substituting equation (4) into expression (6) we obtain

$$H(y) = -\sum_{j=1}^{n} \sum_{i=1}^{n} f_{i}^{x} p(y_{j} / x_{i}) \log_{2} \sum_{i=1}^{n} f_{i}^{x} p(y_{j} / x_{i}).$$
(7)

The component of SNS content that determines the informational-psychological impact on the actors of virtual communities is represented as an expression of conditional entropy

$$H(y/x) = -\sum_{i=1}^{n} \sum_{j=1}^{n} p_i(x) p(y_j/x_i) \log_2 p(y_j/x_i) .$$
(8)

After substituting expressions (7) and (8) in equation (1) we obtain

$$I(x, y) = H(y) - H(y / x) = -\sum_{i=1}^{n} \sum_{j=1}^{n} f_{i}^{x} p(y_{j} / x_{i}) \log_{2} \sum_{i=1}^{n} f_{i}^{x} p(y_{j} / x_{i}) + \sum_{i=1}^{n} \sum_{j=1}^{n} p_{i}(x) p(y_{j} / x_{i}) \log_{2} p(y_{j} / x_{i}).$$
(9)

Assuming that the likelihood of content containing destructive information from the original source in the SNS approaches the frequency of occurrence of relevant signs of threat $p_i(x) \rightarrow f_i^x$, then expression (9) is simplified to the form

$$I(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_i^x p(y_j / x_i) \left[\log_2 p(y_j / x_i) - \log_2 \sum_{i=1}^{n} f_i^x p(y_j / x_i) \right].$$
(10)

To interpret the numerical values obtained by expression (10), we use normalized values I_{norm} for information about the presence or absence of informational-psychological influence on actors in the content of the SNS

$$I_{norm}(x, y) = \frac{I(x, y)}{I_{max}(x, y)}.$$
 (11)

where $I_{\text{max}}(x, y)$ is the maximum value of information.

Thus, the identification of informational-psychological influence on actors in the SNS can be determined by the proportion of undistorted content I(x, y) in the information space of virtual communities, which is compared to the permissible limit. The qualitative scale of threat information assessment of the psychological impact on actors in SNS was formed as a result of the computational experiment and generalization and adaptation of approaches to threat assessment in the field of informational security (Table [1]) [23, 28].

Table 1. Adapted interval scale

Threat class	Interval values of undistorted content $I_{norm}(x, y)$			
very high	0,00–0,20			
high	0,21–0,49			
considerable	0,50–0,74			
low	0,75–0,90			
very low	0,91–1,00			

3 Experiments

The textual content of the social network Facebook, as well as methods of interaction with the API FB service and the integrated development environment of MS Visual Studio have been used to evaluate the informational-psychological impact on the actors.

Determining the tonality of textual content is realized on the basis of the Bayesian multinomial naive method. Accordingly, the detection of hidden topics is realized through the usage of probabilistic latent-semantic indexing. As a result of calculating the entropy of partial signs of informational-psychological influence on public opinion, the SNS obtained the following numerical values given in Table [2].

Table 2. The calculated values of entropy

	H_{Q_1}	H_{Q_2}	H_{Q_3}	H_{Q_4}	H_{Q_5}
Value	0,30	0,40	0,34	0,52	0,50

A matrix P(Y|X) describing the transformation of the studied content, which is spread from source X_k to the consumer Y_l in the SNS, where n=5, takes the following form

$$P(Y \mid X) = \begin{bmatrix} 0,930 & 0,010 & 0,020 & 0,030 & 0,010 \\ 0,010 & 0,944 & 0,002 & 0,014 & 0,030 \\ 0,040 & 0,050 & 0,840 & 0,050 & 0,020 \\ 0,010 & 0,020 & 0,030 & 0,930 & 0,010 \\ 0,010 & 0,030 & 0,040 & 0,020 & 0,900 \end{bmatrix}$$

and the frequency of appearance of signs f_j^y , j = 5 informational-psychological influence on the actors after its distribution in the information space and distortion are equal to

$$f_i^y = \begin{bmatrix} 0,3071 & 0,4243 & 0,3299 & 0,5263 & 0,4771 \end{bmatrix}$$

According to the expression (9), information concerning the presence or absence of informational-psychological influence is equal to

I(x, y) = 2,5724 - 1,2165 = 1,3559, bit/feature.

The maximum value of information $I_{max}(x, y)$ in the absence of distortion of content H(y/x) = 0 acquires the value of 2.5734 bits/feature. Then the normalized value of the information is $I_{norm}(x, y) = 0,5271$. Thus, the level of threat to the informational security of the state in the textual content under study is characterized

by a level of "significant", which requires measures to counteract such destructive information influence on the actors in the SNS [4, 12].

4 Conclusions

We have developed the model for evaluating the informational-psychological impact on SNS actors in text content. The proposed approach uses the signs of manipulation of actors in the information space of virtual communities. The mentioned signs are summarized in the following such groups: doubtfulness of the presented facts; emotional colouring; the presence of hidden topics; sensationality and tone.

The suggested estimation model differs from the known by the emergence of information uncertainty, which is formalized as conditional entropy. This approach takes into account the informational-psychological impact incorporated in the text by the content creator, and the additional destructive impact of distorting that content when processing and disseminating it to other actors in virtual communities. The proposed model of assessment of informational-psychological impact provides the increasing the efficiency of monitoring the information space of SNS and identifying threats to the information security of the state. Thus, a general increase of the efficiency and speed of the system providing information security of the state in the SNS is achieved, which is an extremely urgent task for the modern society today.

References

- Oxley, A.: Security risks in social media technologies: Safe practices in public service applications. Elsevier (2013).
- Li, Na, Sajal K. Das.: Efficiently discovering users connectivity with local information in online social networks. Online Social Networks and Media 16, p. 100062 (2020). doi:10.1016/j.osnem.2020.100062.
- Haddow, G. D., Haddow, K. S.: Disaster communications in a changing media world. Butterworth-Heinemann (2013).
- Hryshchuk, R. Molodetska-Hrynhchuk, K.: Methodological foundation of state's information security in social networking services in conditions of hybrid war. Information & Security: An International Journal 41, pp. 55–73 (2018).
- Trach, O., Peleshchyshyn, A.: Functional-network model of tasks performance of the virtual communities life cycle directions. In: Proceedings of the 11th International Scientific and Technical Conference, CSIT 2016, pp. 108–110 (2016) doi:10.1109/STC-CSIT.2016.7589881
- Peleshchyshyn, A., Bandrovskyi, H.: Informational influence in social networks: dynamics modeling based on the system of linear equations. In: Proceedings of 2019 IEEE 14th International Conference on Computer Sciences and Information Technologies, CSIT Vol. 1, pp. 165–168 (2019, September) doi: 10.1109/STC-CSIT.2019.8929819.
- Korzh R., Peleshchyshyn A., Fedushko S., Syerov Y.: Protection of University Information Image from Focused Aggressive Actions. Advances in Intelligent Systems and Computing: Recent Advances in Systems, Control and Information Technology,

Proceedings of the International Conference SCIT 2016, vol. 543, pp. 104-110. Warsaw, Poland. Springer (2017).https://doi.org/10.1007/978-3-319-48923-0_14

- Korobiichuk, I., Snitsarenko, P., Katsalap, V., Hryshchuk, R.: Determination and Evaluation of Negative Informational and Psychological Influence on the Military Personnel Based on the Quantitative Measure. In: Proceedings of the 1st International Workshop on Control, Optimisation and Analytical Processing of Social Networks (COAPSN-2019) 2392 pp. 66–78 (2019).
- Peleschyshyn, A., Holub, Z., Holub, I.: Methods of real-time detecting manipulation in online communities. In: Computer Sciences and Information Technologies - Proceedings of the 11th International Scientific and Technical Conference, CSIT 2016, pp. 15–17 (2016). doi:10.1109/STC-CSIT.2016.7589857
- 10. MediaSapiens: https://ms.detector.media/ethics/manipulation
- Hryshchuk, R., Molodetska, K., Syerov, Y.: Method of improving the information security of virtual communities in social networking services. In: CEUR Workshop Proceedings 2392 pp. 23–41 (2019)
- Hryshchuk, R., Molodetska, K.: Synergetic control of social networking services actors' interactions. In: Recent Advances in Systems, Control and Information Technology 543 pp. 34–42 (2017) doi:10.1007/978-3-319-48923-0_5
- Cheng, Chun, Yun Luo, and Changbin Yu: Dynamic mechanism of social bots interfering with public opinion in network. Physica A: Statistical Mechanics and its Applications p. 124163 (2020) doi:10.1016/j.physa.2020.124163.
- Etesami, S., Rasoul, et al.: Evolution of Public Opinion under Conformist and Manipulative Behaviors. IFAC-PapersOnLine 50.1 pp. 14344–14349 (2017) doi: 10.1016/j.ifacol.2017.08.2012.
- Luwei Rose L., Schmierbach M., Yu-Leung Ng.: Willingness to follow opinion leaders: A case study of Chinese Weibo. Computers in Human Behavior 101 pp. 42–50 (2019) doi: 10.1016/j.chb.2019.07.005.
- 16. Independent publishing TEXTY.org.ua (in Ukrainian) http://texty.org.ua/pg/article/editorial/read/85655/Tehnika_manipulaciji_Visim_brudnyh_ pryjomiv_jaki_vykorystovujut.
- 17. Virtual warfare: how social networks are used to stir up conflict: https://konkurent.in.ua/publication/40122/virtualna-viyna-yak-vikoristovuut-sotsmerezhidlya-rozpaluvannya-konfliktiv/.
- Kereselidze, N.: Combined continuous nonlinear mathematical and computer models of the information warfare. International Journal of Circuits, Systems and Signal Processing 12, pp. 220–228 (2018)
- 19. Molodetska K., Tymonin Yu., Melnychuk I.: The conceptual model of information confrontation of virtual communities in social networking services. International Journal of Electrical and Computer Engineering **10**(1) pp. 1043–1052 (2020).
- Hryshchuk R., Molodetska K., Tymonin Yu.: Modelling of conflict interaction of virtual communities in social networking services on an example of anti-vaccination movement. In: Proc. of the International Workshop on Conflict Management in Global Information Networks CMiGIN 2019 2588 pp. 250–264 (2019).
- Zakharchenko A., Maksimtsova Y., Iurchenko V., Shevchenko V., Fedushko S.: Under the conditions of non-agenda ownership: Social media users in the 2019 Ukrainian presidential elections campaign. CEUR Workshop Proceedings, vol. 2392, pp. 199-219. Lviv, Ukraine (2019).
- 22. Syerov Yu., Fedushko S., Loboda Z.: Determination of Development Scenarios of the Educational Web Forum. Proceedings of the XIth International Scientific and Technical

Conference (CSIT 2016), pp. 73-76. Lviv (2016). https://doi.org/10.1109/STC-CSIT.2016.7589872

- Molodetska-Hrynchuk, K.: Methodology of detection of the public opinion manipulations in the social networking services (in Ukrainian). Information security 4(24) 80–92 (2016).
- Thuy, N. N., Wongthanavasu, S.: On reduction of attributes in inconsistent decision tables based on information entropies and stripped quotient sets. Expert Systems with Applications 137 pp. 308–323 (2019) doi: 10.1016/j.eswa.2019.06.071.
- Wang, Y., Yongming, L.: Bayesian entropy network for fusion of different types of information. Reliability Engineering & System Safety 195 106747 (2020) doi: 10.1016/j.ress.2019.106747.
- Li, DongDong, et al.: Information entropy based sample reduction for support vector data description. Applied Soft Computing 71 pp. 1153–1160 (2018) doi: 10.1016/j.asoc.2018.02.053.
- Dai, Jianhua, et al.: Uncertainty measurement for interval-valued decision systems based on extended conditional entropy. Knowledge-Based Systems 27 pp. 443–450 (2012) doi: 10.1016/j.knosys.2011.10.013.
- Hryschuk, R., Mamarev, V.: Optimization method of the input data flow dimensionality for the information protection systems (in Ukrainian). Information security 2(8) pp. 27– 34 (2012).