Fuzzy model of cybersecurity risk assessment in digital economy

Nazarov Dmitry Mikhailovich
Institute Of Management And Information Technologies
Ural State University of Economics
slup2005@mail.ru

Samatov Konstantin Mikhailovich
Institute Of Management And Information Technologies
Ural State University of Economics
k.samatov@gmail.com

Begicheva Svetlana Viktorovna
Institute Of Management And Information Technologies
Ural State University of Economics
begichevas@mail.ru

Abstract **– The article deals with fuzzy model of risk assessment of cybersecurity based on fuzzy model of poorly formalized factors. Therefore, there was assessed general impact in critical information infrastructure facilities by using technology of Mamdani-type logical inference. Model of poorly formalized factors, which are described "source of threats", is used to cybersecurity risks for critical information infrastructure facilities modelling.**

Keywords: poorly formalized factors, peer review, fuzziness index, cybersecurity, source of threats, critical information infrastructure facilities, new industrialization, digital economy.

I. INTRODUCTION:

Considering the genesis of world civilization gradually, it can be claimed that the beginning of XXI century is kind of line between the ending life cycle of industrial age of human society development and the starting one of post-industrial phase. Except that, search for theoretical-economic meaning is just getting started. A key concept of new phase of civilization development is "knowledge-based economy" and new industrialization, based on trends such scientific and technical progress, globalization of investment and information flows, innovations, intellectual capital, networks and communication development and also arising with that cybersecurity risks. In the 1960's it was first spoken about "knowledge-based economy" as a new economies, which created tectonic shifts in the national and world economies. It (knowledge-based economy) was based on intellectual activity, which was founded on data, information and information technology. The new industrialization is naturally one of the display in global economy and the result of "knowledge-based economy". The modern age is a logical follow-up to concept "knowledge-based economy" and new industrialization, which is taking the form of new economic phenomenon – "digital economy", based on digital transformation of the society towards the integration of digital technologies into all areas of economy. Information, along with the others resources such energy, financial, etc., is a key characteristic of economies, included the complete set of data, produced by economic units' business processes and activities as a whole, in a digital economy. In considering this resource, there are some important things to note. Firstly, information is almost along inexhaustible resource. Secondly, it is the product of intellectual activity of most skilled and creative economic units. Moreover, information, possessing unique storage property, facilitates the efficient using of all the rest resources. Finally, recycling of information not only does not reduce it- but it creates new forms of its using and presentation

Accordingly, information becomes an integrated economic resource, affecting the society comprehensively and economic agent's activity particularly. The diversityof information and its increasing value certainly warrant the building new formalized models of cybersecurity risks, which would raise information resource to the new level of security. It is particularly important in the new industrialization, which is a natural display in the digital economy. Whereas, the theoretical basis for the contribution of cybersecurity as a particular activity to Economic science, which permanently lead to the new knowledge-based approaches and the change of economy, needs extra research.

Federal Act of 26 July 2017 No.187-F3 "On the Security of the Critical Information Infrastructure of the Russian Federation" entered into force in January 2018, the main purpose of which is to secure the critical information infrastructure during the cyber-attacks against it [9,10].

It should be mentioned that, many other countries besides Russian Federation are concerned about the information infrastructure protection. In particular, the same legal document was adopted in the USA in 2013 (United States Presidential directive dated 12 February 2013 "Critical Infrastructure Security and Resilence"). However, the peak in this matter took place in 2017 and 2018. For instance, "Cyber

Security Low" was passed in PRC (People's Republic of China) in June 2017 and the similar laws became effective in Great Britain and Ukraine on May 10, 2018[15].

All the above general documents are focused on cyber-attacks protection against the critical information infrastructure, included the technical systems available to the public authorities and legal units that automates business-processes related to information-processing in particular in the technological process: information systems, automated control systems (technological processes), data transmission systems.

Therefore, most of public authorities and legal units in the sphere of influence of law on the protection of critical information infrastructure have been obliged to take a set of actions to the information protection, aimed at security of critical information infrastructure facilities since mid-2017 [13]. From this perspective, cybersecurity become the important part of new industrialization process.

Thus, the definition of cybersecurity risks and probabilities of their achievements are important in the election process of some action. Considering that the risks assessment is nearly always based on peer review, the concept of fuzzy modelling of poorly formalized systems becomes important in developing risks models.

Poorly formalized factors are such factors that possess systematic complexities therein, the rate of its influence on the rapid economic processes has not been fully explored and the impossibility of their analytical presentation significantly reduces the efficiency of economic system management.

For example, in deciding on implementing certain measures of information systems security it has to base on risks models prepared by experts (in most cases prepared by one expert), which are based on the assessment of source of threats, founded on an element of judgement greatly that prevent to mixed relevance to reality.

II. The description of poorly formalized factor as the linguistic variable

In view of the foregoing, it may be assumed that "sources of threats" are poorly formalized factors because the issue of assessment of the impact in sustainability of the cybersecurity system is not fully predictable and based on the subjective assessment of one or more experts.

The poorly formalized factors modelling is typically difficult due to the linguistic uncertainty of the concept under consideration in natural language. Consequently, such factors must be seen as the linguistic variable conceptually for modelling and structure definition. A linguistic variable is a variable, which has words or sentences in a natural or formal language as the value, not the numeric value. In this case, such variable may be associated with the term of natural language. The linguistic approach is the basis for fuzzy logic and approximate ways to reasoning for realistic modeling of complex management systems for which judgement, perception and human emotions influence on its behavior greatly. Using linguistic variables makes it possible to do describing the problem correctly researches of systems in natural terms for decision makers and experts [1].

Methods of fuzzy set theories allow to disentangle structure of poorly formalized factors, to define the relationship within the structure under consideration and if it is necessary to take into account its specificity towards the groups studied, and , the most importantly, to get the measuring technique.

In this part of paper, the author's understanding of poorly formalized factor is introduced as a linguistic variable in form of union of finite numbers of terms Ti:

$$PF \leftrightarrow T_1\left(A_1(t)\right) \cup T_2\left(A_2(t)\right) \cup \ldots \cup T_n\left(A_n(t)\right),$$
(1)

where

PF – a linguistic variable describing a poorly formalized factor

Ti(Ai(t)) – a term-set describing i-argument LV as a system (i=1..n);

Ai(t) – a system element describing a thing under consideration at the moment t (i=1..n).

Based on the foregoing, a poorly formalized factor (PF) "sources of threats" in general form can be represented as a hierarchical structure (pic. 1) [2].

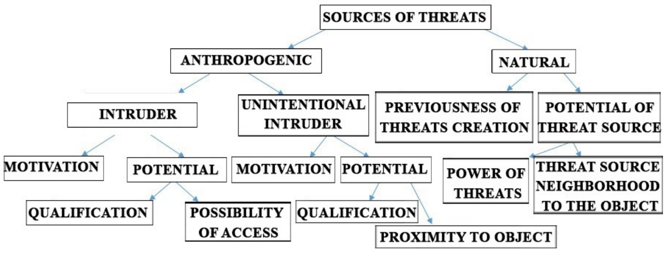

Fig. 1. Formalization of the factor "Sources of threats" in the form of an acyclic graph, each vertex of which is a fuzzy set (in the author's interpretation)

There is a typical hierarchical structure of the decomposition of the concept P in the form of an acyclic graph with nodes Ai(t) (i=1, 2, …, n). All nodes of the graph Ai(t) we will represent fuzzy sets. Let's consider the kth level of decomposition of the original concept of PF.Each component of this level Ak represented by sk partitions of A_{kj}:

$$A_k = \sum_{j=1}^{j=s_k} A_{kj} \qquad (2)$$

At this moment each partition of A_{kj} consists of n_{kj} components. Let's apply the simple notation of the fuzzy set A_{kj}:

$$A_{kj} = \sum_{i=1}^{i=n_{kj}} \{\mu(x\text{¿¿}i^{kj}) | x_i^{kj}\}\text{¿},$$
(3)

where x_i^{kj} (it's the ith element of the fuzzy set A_{kj}), $\mu(x\text{¿¿}i^{kj})\text{¿}$ (this function explains element owning to the fuzzy set $A_{kj}\text{¿}$.

Because of this analytical description of a poorly formalizable process two main problems arise: assessment of the depth of the acyclic graph and constructing of convolution models and obtaining summery of the poorly formalized factor.

A. Methods for obtaining an assessment of a poorly formalizable factor.

To get an aggregated grade that shows the rating of an object in the system, a linear convolution algorithm is usually used for the indicators that formalize this object. Based on the theory of fuzzy sets, aggregated result can be got from rules for constructing fuzzy-logical conclusions. Depending on which formulas we use in the structure of the fuzzy-logical inference, the following fuzzy inference algorithms used in various systems used: the Mamdani algorithm, the Larsen's algorithm, the Tsukamoto's algorithm, the Sugeno's algorithm, and the simplified algorithm. [3]

Let's watch for the Fuzzy Inference algorithm (by Mamdani).

1. We have m fuzzy sets that fit to the k-level partition, where p is the maximum grade of the researchable component of a poorly formalizable factor, E is the set of all components of a poorly formalizable factor.

$$M1 = \left\{ \frac{\mu_{11}(E)}{1} + \frac{\mu_{21}(E)}{2} + \ldots + \frac{\mu_{p1}(E)}{p} \right\}$$

$$M2 = \left\{ \frac{\mu_{12}(E)}{1} + \frac{\mu_{22}(E)}{2} + \ldots + \frac{\mu_{p2}(E)}{p} \right\}$$

…………………………………………………

$$Mm^{¿} \left\{ \frac{\mu_{1m}(E)}{1} + \frac{\mu_{2m}(E)}{2} + \ldots + \frac{\mu_{pm}(E)}{p} \right\} \qquad (4)$$

These sets we use as the base for aggregation as a sequence of membership function values of a single fuzzy set B:

α1= μ11(E) ∨ μ12 (E) ∨ … ∨ μ1m (E) =$\max_i \mu_{1i}(E)$

α2= μ21(E) ∨ μ22 (E) ∨ … ∨ μ2m (E)=$\max_i \mu_{2i}(E)$

……………………………………

αp= μp1(E) ∨ μp2 (E) ∨ … ∨ μpm (E)=$\max_i \mu_{pi}(E)$

(5)

where i=1,2, …,m.

Let's don't forget that operation disjunction (∨) in fuzzy logic is process of finding the maximum.

2. Let's transform the obtained membership function of the set B by the formula:

$$\mu_g(B) = \frac{\alpha_g}{\max_a \alpha_g, where\ g=1\ldots p}. \qquad (6)$$

We will get:

$$B = \left\{ \frac{\mu_1(B)}{1} + \frac{\mu_2(B)}{2} + \ldots + \frac{\mu_p(B)}{p} \right\}. \qquad (7)$$

Based on the obtained membership function it is possible to obtain a defuzzified value which express the aggregate expert opinions using the Mamdani's formula (center of mass method):[5,6]

$$x^{¿} = \frac{\sum_g g \cdot \mu_g(B)}{\sum_g \mu_g(B)},$$

(8)

where x* is a clear value of a variable fit to a logical association in the form of "if-then" which got as a result of the defuzzification procedure.

The procedure described above allows us to move from the k hierarchy level (see Pic. 1) to the k-1 level. This transition is based on the construction of a fuzzy k-1 level set which membership function are values.

$$\mu_{A_{k-1\,j}} = \mu_g(B), where\ j=1\ldots s_{k-1}$$

(9)

Considering the hierarchy (in general) of the representation of a poorly formalizable factor as a linguistic variable, the procedures and formula (2) which are described above allow moving from one level of hierarchy to another, obtaining and analyzing the membership functions of intermediate components. The advantage of this presentation is that it is possible to identify the most important components and discard irrelevant, and it's how to simplify further modeling. The results are universal and can be applied to formalize any poorly formalized factor.[7,8]

Methods of assessing the depth of an acyclic graph of a representation of a poorly formalizable factor.

Estimation of the depth of an acyclic graph of a representation of a poorly formalizable factor is associated with a semantic approach to the measurement of information. The most widely used here is the thesaurus approach, which links the semantic properties of information with the user's ability to formalize their knowledge of the object of study. A thesaurus is a collection of information that a user or system has. The essence of the thesaurus approach to the formalization of implicit factors can be represented as the implementation of the method of "hyponyms-hyperonyms".The idea of the method is to build interconnected hyperonym-hyponym associations. Hyperonym - a word with a broad meaning, expressing a common, generic concept, the name of the class (set) of objects (properties, attributes). Hyponyms are words that call objects (properties, attributes) as elements of a class (set). A hyperonym is the result of a logical operation of generalization or, in the mathematical sense, an addition to the set [6]. The method of hyponyms is based on the fact that any subject area can be described by a hierarchical dictionary of concepts, which is actually called thesaurus. The thesaurus is a tree of concepts for a given subject area, starting with the top, most general, and ending with the bottom, most specific, narrow concepts. Words (terms) in a thesaurus are usually related by the general-particular, the whole-part, etc.

Thus, the graph in Figure 1 of a poorly formalized factor "sources of threats" of information security is essentially a thesaurus, which, when customized for a specific object of protection and determining the level of probability of a particular source of threats, according to the methodology outlined in the previous section, will represent nothing more than a model of information security threats.

III. EXAMPLE OF ESTIMATION OF SOURCES OF THREAT FOR OBJECTS OF CRITICAL INFORMATION INFRASTRUCTURE

Critical Information Infrastructure (CII) is a set of objects (information systems, information and telecommunication networks, automated management systems) belonging to the subjects of critical information infrastructure (government bodies and institutions, legal entities and individual entrepreneurs) operating in one of 12 spheres of life activity: healthcare, science, transport, communications, energy, credit and financial sphere, fuel and energy complex, nuclear energy, defense, rocket and space, mining, metallurgical and chemical industries, as well as telecommunication networks used to organize the interaction of such objects.

In the process of developing information security threat models, the impact of a source of threats on the security of specific objects is assessed, it should be borne in mind that for each of the spheres of vital activity there will be a characteristic predominance of its critical information infrastructure objects with the most typical sources of threats [12].

So, in particular, when developing models of threats using the theory of fuzzy sets, for the objects of critical information infrastructure operating in the field of health, metallurgy and communications, the authors obtained the following results (Table 1).

TABLE 1. THE MOST TYPICAL SOURCES OF THREATS FOR HEALTH CARE, METALLURGY AND COMMUNICATIONS.

Field\Object	**Information Systems**	**Automated control systems**	**Information and telecommunication networks**
Public health	Autogenous: - medium potential external intruder; - low potential intruder. Natural: - fire; - earthquake.	isn't subject to the impact of the source of the threat (critical information is not processed)	Autogenous: - high potential external intruder; Natural: - flood; - hurricane; - heavy rains.
Metallurgical industry	Autogenous: - medium potential external intruder; - low potential intruder. Natural: - fire; - earthquake.	Autogenous: - low potential intruder; - high potential external intruder. Natural: - fire; - earthquake.	isn't subject to the impact of the source of the threat (critical information is not transmitted)
Connection	Autogenous: - medium potential external intruder; - high potential intruder. Natural: - fire; - earthquake.	Automatic control systems are not used	Autogenous: - medium potential external intruder. Natural: - flood; - hurricane; - heavy rains.

We investigated 5 objects of critical information structure in different spheres of economic activity, which were evaluated by 6 experts. According to the standard method, all 5 objects received a good rating. However, the applied standard technique did not allow to rank the evaluated objects and to answer the question of which of the CII objects is better protected. The algorithm proposed by the authors more accurately calculates the evaluation parameters of the CII object and allows evaluating the differences between the evaluated objects within the framework of the proposed strata. As a result, according to the proposed algorithm, 3 CII objects received a higher rank than others. In addition, the analysis of the results showed that the proposed method allows more accurate assessment of the sources of information security threats and does not contradict the approaches used in the methodology of industry regulators.

IV. CONCLUSION

Thus, taking into account the above, the following conclusions can be drawn:

1. The use of models of fuzzy sets, when developing models of information security threats, allows the use of a mathematical apparatus to analyze poorly formalized factors in order to identify the most significant (having the greatest impact) and increase the objectivity of expert assessments.

2. The dynamic nature of the linguistic variable, provided by the introduction of the variable t, which shows that some elements of the Ai (t) system change significantly over time, due to the development of scientific ideas about them; the presence of a hierarchical model of a linguistic variable; the thesaurus approach makes it clear the procedure for obtaining components, taking into account the ambiguity, practical understanding and interpretation of information security threats.

3. The advantages of such a presentation lie in the possibility of formally describing a poorly formalized factor, taking into account the purpose of the study using the apparatus of the theory of fuzzy sets. The algorithm of a fuzzy-logical conclusion we proposed for obtaining an aggregated estimate of a poorly formalizable factor is based on the idea of constructing a superposition of fuzzy membership functions, in the form of additive convolution technology of its components according to the Mamdani rule taking into account their degree of importance.

4. This approach was approved by the authors for building models of threats to the information security of critical information infrastructure facilities, due to the fact that there is no regulatory methodology to date. Taking into account the above, practical application of this method in methodological databases of industry regulators is possible.

References

[1] Alizadeh, A., Chehrehpak, M. (2017) Fuzzy group decision making model for identifying and ranking of success factors in fraud prevention in Iranian e-banking, *Journal of Information Technology Management(2)*, pp. 355-378

[2] Liu, F., Li, R., Li, Y., Yan, R., Saha, T. (2017) Takagi-Sugeno fuzzy model-based approach considering multiple weather factors for the

photovoltaic power short-term forecasting, *IET Renewable Power Generation*, 11(10), pp. 1281-1287

[3] Mahmoudi H., Aleenejad M., Moamaei P., Ahmadi R. (2016) Fuzzy adjustment of weighting factor in model predictive control of permanent magnet synchronous machines using current membership functions. 2016 IEEE Power and Energy Conference at Illinois, PECI 20167459225

[4] Nazarov D.M. (2017) Fuzzy model for assessment of causality of factors. Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017, 7970746, pp. 859-861

[5] Nazarov D.M. (2017) Fuzzy Model for Assessment of Causality of Factors in Collaborative Economy, Proceedings - 2017 IEEE 19th Conference on Business Informatics, CBI 20172,8012936, c. 28-31

[6] Norman T.J., Jennings N.R., Faratin P., Mamdani E.H. (2015) Designing and implementing a multi-agent architecture for business process management. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)1193, pp. 261-275

[7] Zadeh L.A., Abbasov A.M., Shahbazova S.N. (2015) Fuzzy-based techniques in human-like processing of social network data.International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems 23, pp. 1-14

[8] Zadeh L.A. (2015) Fuzzy logic - A personal perspective, Fuzzy Sets and Systems281, pp. 4-20

[9] Federal Act of 26 July 2017 No.187-F3 “On the Security of the Critical Information Infrastructure of the Russian Federation”, “Collection of Legislative Acts of the RF”, 31.07.2017, № 31 (Part I), p. 473.

[10] Decision No. 127 of the Government of the Russian Federation of 08 February 2018 “On approval of the rules for classified Critical Information Infrastructure facilities of Russian Federation as well as indicator list of important criteria of Critical Information Infrastructure facilities and its values ”, “Collection of Legislative Acts of the RF”, 19.02.2018, № 8, p. 1204

[11] GOST R ISO/IEC. 27001-2006. The Russian Federation national standard. Information technology. Cybersecurity methods and means. Cybersecurity management systems. Requirements, approved and introduced by Rostekh regulirovanie order of 27.12.2006 № 375-p.

[12] Samantov K.M. (2017) Complex restricted Information Security in corporate information system, *Journal “Information Security“*, № 2, pp. 34-37.

[13] Samantov K.M. (2018) CII: what to expect and what to do?, *Journal “Information Security”*, № 1, pp. 16-18.

[14] Skiba V., Kurbatov V. (2008) Guide to protection from internal threats of information security, *SPb.: Peter*, p. 320, Il.

[15] Harris Sean.CISSP All-in-One Exam Guide. CISSP. [Electronic source]. Retrieved from http://dorlov.blogspot.ru/2011/05/issp-cissp-all-in-one-exam-guide.html