Adv Syst Sci Appl 2024; 04; 44-56
Published online at https://ijassa.ipu.ru.

Possibilities of Assessing Information Security Risks Using Fuzzy Logic and Econometrics Methods

Alexander Kozlov[1], Nikolai Noga

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Abstract. This paper provides a brief overview and analysis of the most commonly used methods for assessing information security risks in various complex systems. These methods involve creating specific models that help assess, manage, and predict the possible occurrence of adverse situations related to information security. These methods help to make decisions aimed at minimizing the potential damage that could occur because of external attacks on information resources or other threats that exploit existing vulnerabilities. However, they provide an acceptable result in risk assessment in conditions where it is possible to quantify the parameters on which the risk depends. In conditions of high uncertainty, for example, when determining the dependence of risk on subjective factors, the use of these methods can lead to high errors. As a rule, risk assessment is associated with a high degree of uncertainty of parameter values and their mutual influence on the information security risk level. The method proposed by the authors, based on the combined use of fuzzy logic and regression analysis, makes it possible to assess the information security risk in conditions of uncertainty in complex information systems with a network structure. This method also allows you to identify the parameters that most influence the risk level, which in turn enables you to protect information resources while optimizing the cost of implementing an effective protection system. Using this method and basing on the predicted risk level values, allows you to plan events to improve the level of information system protection in both the short and long term.

Keywords: information security, information leakage, multifactorial risk assessment, fuzzy logic, econometrics, production rules, regression equation, multiple determination coefficient, subjective factors, uncertainty conditions.

INTRODUCTION

Currently, the global digitalization of the economy and the rapid development of digital technologies have led to the fact that the digital world has penetrated into all areas of human activity. Now a person can no longer refuse digital services. Moreover, these services, in addition to new opportunities, carry new risks and threats. It is difficult for an untrained person to understand all the nuances of the digital world. Attackers are ready to take advantage of this, always striving to benefit themselves in any unclear situation. The situation regarding countering hacker attacks on any information resources of companies, as well as on critical information infrastructures, has become extremely complicated. In particular, such attacks intensified during the global pandemic, when the staff of many companies had to switch to remote operation. The activity of such attacks has also increased because of the aggravation of the political situation in the world. Confidence in software and hardware manufactured abroad has significantly decreased. Thus, it became necessary to solve the problem of switching to domestic hardware and software.

Both large companies and small businesses face these challenges. We should note that, in the first half of 2024, the number of data breaches worldwide decreased by 19.2% compared to the same period in 2023. However, in Russia, during this time, there was an increase of 10.1% in the number of information breaches compared to the first half of 2023. Additionally, an all-

[1] Corresponding author: alex4590alex@yandex.ru

time high in the number of compromised personal data records was recorded in Russia during this period. Overall, approximately 1 billion personal data records were compromised in the first half of 2024, which is a 33.8% increase from the first half of 2023 [1].

It is worth noting that the nature of attacks has also changed. Currently, attacks are mainly aimed at disrupting various systems of critical infrastructure, leading to complete failure and maximum damage. Additionally, these attacks try to achieve political objectives by creating a media impact. Although, in the past, attacks on companies were mainly carried out for the purposes of industrial espionage and theft of personal and financial data.

In order to protect the company's business and ensure its efficiency and continuity, it is important to assess both current and potential risks associated with its future activities. Risk refers to the impact of uncertain events on the achievement of goals [2]. Therefore, risk assessment involves considering these uncertainties.

There can be a wide range of risk factors. Typically, general factors are identified for a particular class of systems, but specific ones can also be considered, which are unique to each individual system. It is also possible to observe the impact of various factors over time, meaning that risk assessments you should perform regularly and iteratively, using up-to-date information. This process includes risk identification, analysis, and comparative evaluation [2].

Based on the above, it follows that the methodology used for risk assessment should be straightforward and allow for dealing with various uncertainties.

In this article, the authors propose a solution to the problem of information security risk assessment using fuzzy logic and regression analysis together. These methods allow for solving this problem under conditions of uncertainty in the dependence of various parameters within complex information structures. Additionally, these methods allow for the identification of parameters that have a significant impact on information security risks, as well as those that you can ignore in a given situation. This information you can use to plan measures for improving the information resource protection system, reducing potential damage, and minimizing potential risks in both the short and long term.

LITERATURE REVIEW

As mentioned above, the focus of attacks on organizations and their information resources has changed. Therefore, it is necessary to look for extraordinary solutions to identify information security risks in operated information systems. Based on these solutions, methodologies for assessing, analyzing, and managing information security risks have been developed. For example, well-established but still relevant methodologies include CORAS (a risk and threat modeling and system analysis tool), Octaves (for assessing critical threats, assets, and vulnerabilities), CRAMM (a Central Computer and Telecommunications Agency analysis and management risk method), Vulture 2006 (an integrated risk analysis and management tool for digital security information systems), RiskWatch, and the methodology from Microsoft [3-4].
A number of works have also been dedicated to identifying security threats. For example, work [5] introduces the recursive internetwork architecture and carries out a security risk assessment to identify threats at runtime and verify the correctness of built-in security tools. It also examines mitigation measures designed to combat attacks related to these threats.

Paper [6] presents a methodology for assessing information security risks when using cloud technologies in corporate information systems. This methodology uses fuzzy logic methods and takes into account the level of control over information resources. Also, the assessment of security risks in cloud computing is considered in [7] based on a two-stage procedure for quantifying the information security risk of cloud computing, aimed at calculating the coefficient of countering possible attacks and comparing the amount of damage with the value of the organization's assets. Work [8] shows that with each distributed access implemented

during data exchange. There is a security risk that is assessed using security metrics such as the average cost of failure and the multidimensional cost of failure.

In all the above–mentioned works, risk assessment is the assessment of the threats, as well as the vulnerabilities that these threats are implemented through, and, finally, the damage resulting from their implementation.

Risk analysis and management include the construction of models that adequately show the occurrence of adverse conditions, taking into account various parameters that characterize these risks. In addition, these models help decision-making in order to reduce the potential damage that attacks on information resources can caused or the implementation of other threats through existing vulnerabilities, as well as the development of preventive measures to increase the security of the informatization object.

The above methods provide an acceptable result for risk assessment in situations where it is possible to quantify the parameters that affect risk. However, in conditions of high uncertainty, such as when determining the dependence of risk on subjective factors [9] that may affect business goals, including information security, these methods may not be sufficient. Factors such as the level of wages and professionalism can affect the outcome of a project, and incorrect management decisions based on incomplete information can lead to negative consequences such as increased costs, delays, or even the loss of the business due to information leakage.

Methods based on fuzzy logic have been shown to be effective for risk assessment in these situations, as they allow for the consideration of uncertain parameters and their impact on risk [10]. These methods can help identify potential risks and develop strategies to mitigate them, ensuring that businesses can achieve their goals while minimizing the risk of negative outcomes.

In [11], a methodology is proposed that allows assessing risk under conditions of uncertainty regarding the interrelation between factors such as the level of threats, presence of vulnerabilities, potential damage, amount of various resources, costs for creating and maintaining a system, and partial loss of control over one's information resources when using cloud-based structures.

In [12], the authors propose a risk analysis process combining qualitative and quantitative methods. The focus is on developing a methodology for assessing and analyzing risks and vulnerabilities in the context of security risk management. The authors carry out the risk assessment in four stages. At the first stage, a mathematical model is built in accordance with the results of risk identification. At the second stage, basic information or data available as a result of expert assessments are obtained. A mathematical method is chosen to quantify the information. At the third stage, a suitable model and analysis methods are selected. At the last stage, the risk level is determined in accordance with certain criteria. At the same time, the focus of the study is on the analysis of risk assessment using fuzzy logic.

In [13], to assess cyber threats to digital instrumentation and I/O systems, the authors proposed a risk assessment method based on multi-fuzzy systems that assesses the risk of cybersecurity depending on the following risk factors: the overall capabilities of the attacker, the probability of success of the attack and the consequences of the attack. This method involves the operation of three fuzzy inference system. The first system evaluates the overall capabilities of the attacking attacker, the second system evaluates the overall probability of success of the attack, and the third system evaluates the risk level based on the impact of the attack and the results of the first two systems.

In [14], a technique was proposed to solve the problem of detecting a set of critical nodes by fragmenting the graph as a result of removing a set of vertices with a certain power in such a way that the residual graph has minimal pairwise connectivity at a user-defined power value. Traditional optimization algorithms are unable to find the optimal set of vertices in large graphs with thousands or millions of nodes due to the high computational costs associated with them. The authors used the within-method method in a greedy algorithm to quickly identify a set of

critical vertices. The proposed algorithm can be easily extended to vertex- and edge-weighted variants to detect a critical node.

In [15], a method is presented that allows us to estimate the predicted value of information security risk for complex systems and its confidence interval, using regression analysis and fuzzy logic techniques, in terms of how risk depends on various factors such as the cost of resources and the threat level. This method is used to implement a risk assessment process under uncertainty.

The use of a fuzzy model in [16] allowed for a more flexible processing of inaccurate information security risk factors, and enabled us to switch to a quantitative representation of qualitative characteristics. The fuzzy model and methods proposed in this work can be used to assess both specific types of information security risks for an ERP system (Enterprise Resource Planning), as well as the general information security risk for an ERP. At the same time, the Fuzzy Logic Toolbox in the Matlab system [17] was used to implement the methodology described in [11] and [16].

The paper [18] provides a detailed description of the information risk analysis algorithm, as well as the specifics of the implementation of the information risk assessment stage in the overall analysis process. The authors proposed a technique based on a neuro-fuzzy network that ensures the adequacy of the assessment and applicability to non-numeric input data using the example of the implementation of a neuro-fuzzy network in the Matlab environment.

MATERIALS AND METHODS

In this paper, we propose a methodology based on fuzzy logic and regression analysis for determining the factors that significantly influence information security risk. This methodology allows us to identify a set of critical parameters that need to be taken into account when evaluating information security risks, as well as a list of less significant parameters that can be ignored. Additionally, this approach enables us to address the problem of identifying critical nodes in networked information systems [19].

It is clear that an increase in the number of parameters under consideration and their ranges of values can significantly complicate the construction of a table of production rules. The proposed methodology makes it possible to solve the problem of managing the security of a network information system with a large number of parameters, as well as with a different number of ranges of values for each parameter included in the consideration.

According to our methodology, the risk level we can represent as a function of n parameters.

$$R = R(P_1, P_2, ..., P_n) \tag{3.1}$$

where $P_i, i = 1,...,n,$ – parameters that affect the risk level.

It is assumed that the quantitative values of all these parameters are in the range [0; 1], and the qualitative ones take values, for example, low, medium and high.

The goal of this task is to identify the linguistic variables that have the most significant impact on the value of a specific output variable, which in turn determines the level of risk. In order to accomplish this, we will develop a multiple regression model. The ranges of values for these variables will be established based on expert assessments, such as those obtained through the Delphi method or other group-based techniques [20-22]. We will use mathematical statistical methods to process the estimates provided by the experts.

The authors of this paper did not intend to provide a detailed description of the process for processing expert estimates, such as assigning relative weights to factors considered during the evaluation process. A detailed description of this procedure you can find in [21]. For simplifying the coordination of results from expert assessments in this example, we will assume that experts provided estimates that were close in value and the concordance coefficient was close to one.

It is proposed to use fuzzy logic methods to determine the risk level [11, 15, and 19]. This makes it possible to conduct a multifactorial risk assessment. The following features we should consider when using these methods:

- the impact of the parameters under consideration can be very uncertain due to their possible influence on each other, i.e. there is no clarity on the issue of the mutual correlation of these parameters and, accordingly, their possible redundancy;
- the accuracy of the assessment depends on the quality of the formation of production rules;
- and how well and in detail are the terms of variables described , indicating the ranges of their values for each linguistic variable.

To calculate the coefficients of the regression equation, we use the least squares method (OLS) However, since the obtained coefficients of the equation are incomparable, in order to be able to compare them [23] and arrange the parameters according to the degree of influence on the risk level, we build the regression equation on a standardized scale. As a result, it is possible to rank coefficients according to the degree of influence on the risk level and, having determined redundant parameters, i.e. parameters that have little effect on the risk level, exclude these parameters from the equation.

The solution scheme provides that we perform the task in several stages.

Step 1. Conducting a Survey of the Informatization Facility

According to the results of the survey of the informatization object, a set of n factors (parameters) is determined, on which the information security of this object (the operated information system) depends. In relation to information systems, risk factors we can divide into four main types, which include:

1. Economic parameters: the value of the asset (information resource), the level of dependence of the main production on the information infrastructure, the level of costs for the information component in general and means of ensuring cybersecurity in particular, the level of potential damage and others.
2. Technical parameters, which include: the level of vulnerabilities in the hardware and software complex of the operated system; the level of technological independence from imported developments (the ratio of the number of software and hardware developed and manufactured in Russia to the total number of tools used at the informatization facility), the level of technical protection of information and others.
3. Organizational parameters: the level of isolation of the facility from external systems, the level of use involved external specialists, the organization level of the information system protection (availability and execution of instructions and regulations, availability of access controls, availability of security administrators, availability of contractors) and others.
4. Subjective parameters: the employees and administrators qualification level, the employee's salary level, the level of "turnover" of personnel and others [9].

According to the results of the examination of the informatization object, other risk factors may be established. The combination of such factors will represent the result of the risk identification.

Step 2. Creating a Fuzzy Knowledge Base

From the set of n parameters obtained at the first step, m parameters (from five to eight) are selected, representing all four of the above groups, on which, according to experts, the information security of the object depends largely. If we take into account all n parameters, then the table of production rules may take an unacceptably large size, which will be quite difficult to process. The selected parameters (linguistic variables) are normalized so that their values are within the range from 0 to 1.

All input linguistic variables P_i are evaluated each on their own scale, both at a qualitative level ("low", "medium" and "high") and in quantitative form in the range from 0 to 1. The boundaries of the terms of these linguistic variables are set by expert assessments, in tabular

form. Similarly, we assign values for the output parameter of the risk level *R*: "low", "acceptable", "high" and "critical". Then we form a table of production rules, each row of which is a combination of the values of the input parameters. In this case, each line corresponds to a certain value of the output parameter. Moreover, we do this for any possible combination of values of the selected parameters. Further, the averaged quantitative values are correlated with the qualitative values of the variables [19]. We get many data for the possibility of using regression analysis methods.

Step 3. Building a Regression Equation

Using Matlab, MS Excel or other software products with options for working with regressions, we build a linear model of multiple regressions from (3.1) with explanatory variables introduced into consideration for simplicity of presentation:

$$R = a_0 + \sum_{i=1}^{m} k_i P_i + \varepsilon \tag{3.2}$$

where a_0 and $k_i, i = 1,...,m$, incomparable coefficients that can be found using OLS. Next, from equation (3.2), we proceed to the equation on a standardized scale and find standardized coefficients.

Step 4. Analysis of the Result Obtained

We build the coefficients in ascending order to determine the variables on which the information security risk depends largely. In addition, we examine the variables for their dependence on each other.

To assess the quality of the constructed model, one of the important indicators in the construction of regression is calculated - the coefficient of multiple determination to assess the joint influence of variables. The higher it is, the greater the impact of the selected variables on the risk level *R*.

With a low value of the multiple determination coefficient (less than 0.6), variables that have little effect on the risk value with the lowest value of the ki coefficients are excluded from consideration. We return to step 2, where instead of the excluded parameters, we select other parameters from the set n and repeat the above steps until the value of the multiple determination coefficient is at least 0.8.

We can complete the iterative process when we make sure that any new combination of parameters does not significantly increase the coefficient of multiple determinations. In this case, we will assume that the parameters that maximally affect the value of the information security risk level of the information object in question have been found. Applying further Fisher's *F*-criterion and Student's *t*-criterion [23], one can also verify the statistical significance of both the regression coefficients and the resulting regression equation as a whole at a certain level of significance.

After determining the parameters that have the maximum impact on the level of information security risk, we implement the stage of comparative risk assessment. We compare the value of the risk level obtained with the above parameters with its acceptable value. If necessary, we develop a set of measures aimed at ensuring such values of the established parameters, at which the value of the risk level lies within acceptable limits.

It should be borne in mind that the cost of providing the required values of the desired parameters should not exceed the possible losses from the implementation of potential threats.

Example

As an illustration of the proposed methodology, let us consider an example of determining the factors that most affect the level of risk in the network information infrastructure. We select

five linguistic variables from the four groups indicated above in the first step. These are the following factors:

- the value level of the information asset of the network node,
- the level of wear of technical means in the node,
- node load level,
- the level of organization of node protection,
- the qualification level of the node staff.

Of the five indicators selected above, three indicators increase the level of risk with increasing values, and two indicators (the level of organization of node protection and the qualification level of the node staff), on the contrary, reduce the level of risk. Production rules are much easier to form when the indicators are unidirectional, i.e. the risk level increases (decreases) with an increase (decrease) in the values of each of the considered indicators. In order to comply with these conditions, while leaving the essence of the indicators unchanged, we will change the level of organization of node protection to the level of node vulnerability, and the level of staff qualification to the level of staff incompetence. That is, we are now considering the following factors:

- the value level of the information asset of the network node – x_1,
- the level of wear of technical means in the node – x_2,
- node load level – x_3,
- the vulnerability level of the node – x_4,
- the level of incompetence of the node staff – x_5.

Now, increasing the value of any of the selected indicators leads to an increase in the level of risk.

We believe that the values of all these variables take both quantitative values in the range [0; 1] and qualitative values (for example, low, medium, high).

Now the risk level from (3.2) we can represent as follows:

$$R = a_0 + \sum_{i=1}^{5} k_i x_i + \varepsilon \tag{3.3}$$

where ε is an error that includes the influence of factors unaccounted for in this equation, as well as random errors and measurement features.

To determine the coefficients in (3.3), we use OLS. As mentioned above, the coefficients obtained are incomparable. In order to be able to compare coefficients and build parameters according to the degree of influence on risk, a regression equation is constructed on a standardized scale [23]:

$$t_R = \sum_{i=1}^{5} l_i t_{x_i} \tag{3.4}$$

$l_i,\ i = 1,...,5,$ - standardized coefficients and $t_R, t_{x_i},\ i = 1,...,5$ - standardized variables such that

$$t_R = \frac{R - \overline{R}}{\sigma_R},\ t_{x_i} = \frac{x_i - \overline{x_i}}{\sigma_{x_i}},\ \overline{t_R} = \overline{t_{x_i}} = 0,\ \sigma_{t_R} = \sigma_{t_{x_i}} = 1,\ i = 1,...,5.$$

The boundaries of the terms of the above-mentioned linguistic variables, as indicated above, are set based on expert assessments, in tabular form.

We should note that the influence of the above-mentioned factors might be quite uncertain due to their possible influence on each other.

In this example, we will assume that the ranges of terms values of the above linguistic variables with the replacement of qualitative values with the corresponding averaged quantitative values correspond to those shown in Tables 3.1 to 3.5.

Table 3.1. The value level of an information resource (asset) in a node (x_1).

	Level (Fuzzy variable)	*The ratio of the cost of information stored in the node to the cost of all information in the system*	*The boundaries of the term "Asset value level"*	*The average value of the term*
1	Low	Up to 10%	0-0.11	0.05
2	Middle	From 10% to 40%	0.07-0.45	0.21
3	High	More than40%	0.4-1.00	0.7

Table 3.2. The level of wear of technical means (x_2)

	Level (Fuzzy variable)	*The ratio of the age of the equipment to the maximum service life (8 years)*	*The boundaries of the term "Technical means wear level"*	*The average value of the term*
1	Low	The equipment has been in operation for less than 2 years	0-0.21	0.1
2	Middle	The equipment is operated from 2 to 6 years	0.15-0.65	0.4
3	High	The equipment has been in operation for more than 6 years, there are signs of physical and moral deterioration	0.6-1.0	0.8

Table 3.3. Node load level (x_3)

	Level (Fuzzy variable)	*The ratio of the volume of passing information to the maximum throughput of the node's communication equipment*	*The boundaries of the term "Node load level"*	*The average value of the term*
1	Low	There is a large margin for node throughput	0-0.28	0.14
2	Middle	The node works stably and can withstand peak loads	0.2-0.6	0.4
3	High	The node is loaded and may not withstand peak load	0.55-1.0	0.725

Table 3.4. The vulnerability level of the node to security threats (x_4)

	Level (Fuzzy variable)	*Availability of an information security system*	*Boundaries of the term "Node vulnerability level"*	*The average value of the term*
1	Low	The node is protected by certified an information security system, and constantly monitored for employee actions and external attacks.	0-0.31	0.15
2	Middle	Additional protective equipment has been installed at key positions of the node, and periodic monitoring is carried out.	0.2-0.7	0.45
3	High	There is only an information security system, which is usually included in the hardware and software	0.6-1.0	0.8

Table 3.5. The level of incompetence of security staff (x_5)

	Level (Fuzzy variable)	*Work experience in the field of information security and data protection*	*The boundaries of the term "The level of incompetence of employees"*	*The average value of the term*
1	Low	Work experience of more than 10 years	0-0.25	0.125
2	Middle	Work experience from 2 to 10 years	0.2-0.7	0.45
3	High	Work experience up to 2 years	0.6-1.0	0.8

To form the production rules, we will also need the following Table 3.6 of changes in the level of risk.

Table 3.6. The output variable is Risk (R)

	Level	*Measures taken*	*The boundaries of the term "Risk"*	*The average value of the term*
1	Insignificant	Ignoring the risk	0-0.2	0.1
2	Acceptable	The operation of the system is possible, it is necessary to plan additional measures to protect the node.	0.16-0.5	0,33
3	High	Only limited operation of the node is possible, urgent measures must be taken to reduce the risk	0.45-0.65	0,55
4	Critical	The node impossible to work	0.6-1.0	0,8

The task of creating a fuzzy knowledge base at the risk assessment stage we propose to solve using the MATLAB software package (fuzzy logic option) [17]. Thus, we obtain the production rules in the form of Table 3.7.

Table 3.7. Production rules

	x_1	x_2	x_3	x_4	x_5	R
1	low	low	low	low	low	insignificant
2	low	low	low	low	middle	insignificant
…						
100	middle	low	high	low	high	high
…						
242	high	high	high	high	middle	critical
243	high	high	high	high	high	critical

Next, using the averaged values from Tables 3.1-3.6, we obtain from Table 3.7 of the production rules a fuzzy knowledge base in quantitative form in Table 3.8.

Table 3.8 Fuzzy knowledge base, production rules

	x_1	x_2	x_3	x_4	x_5	R
1	0,05	0,1	0,14	0,15	0,125	0,1
2	0,05	0,1	0,14	0,15	0,45	0,1
…						
100	0,21	0,1	0,725	0,15	0,8	0,55
…						
242	0,7	0,8	0,725	0,8	0,45	0,8
243	0,7	0,8	0,725	0,8	0,8	0,8

The values of the standardized coefficients are found using the MATLAB system [17] (regression option) or MS Excel. Thus, equation (3.4) takes the following form:

$$t_R = 0,3529t_{x_1} + 0,2169t_{x_2} + 0,2588t_{x_3} + 0,2954t_{x_4} + 0,1791t_{x_5}. \quad (3.5)$$

Now we calculate the coefficient of multiple determination to assess the combined effect of these five parameters on the value of the risk level

$$M^2_{Rx_1x_2x_3x_4x_5} = 0,826$$

This means that 82.6% of the risk level value we can explain by the parameters included in the regression equation and gives an assessment of the quality of the constructed model. To verify the statistical significance of the obtained regression equation as a whole and the regression coefficients at a certain level of significance, you can use Fisher's *F*-criterion and Student's *t*-criterion. Ranking in ascending order of standardized coefficients from (3.5) shows that in the given example, the parameter - the level of incompetence of the node staff – has the least influence on the overall risk value, and the value of information assets has the greatest influence.

RESULTS AND DISCUSSION

Based on the results of our example, we can conclude that in order to reduce the risk of information security of an information system, first, you should pay attention to the proper distribution of stored information assets between network nodes.

Having established a list of parameters on which the level of information security risk of the node mainly depends, the next step is to rank the nodes of a complex network by criticality based on these parameters, which will identify the most critical nodes of the network [19]. Node ranking can be carried out using multi-criteria optimization methods, for example, the method of Borda counts or the method of determining the number of elements in the upper and lower boundary sets [24-25].

As a result of conducting an information security risk assessment using the proposed methodology, the following decisions can be made:

- no additional measures are required, an acceptable level of risk is provided;
- urgent measures are required to eliminate critical vulnerabilities;
- to continue the risk analysis taking into account additional parameters, both external and internal;
- adjust the goals to reduce the level of risk and costs.

In addition, in the case of any of the above decisions, it is necessary to take into account the need for constant monitoring of the risk level, since any change in the external or internal conditions of the information infrastructure may affect its significance.

The authors note that one of the results of using this technique is the formation of a knowledge base using production rules. At the same time, instead of values in qualitative form (values, as indicated above, "low", "medium", "high", etc.), data are formed in quantitative form, as in [19. 26] by averaging the values of terms of linguistic variables.

Based on the knowledge base obtained, using regression analysis methods, it is possible to build the regression model in such a way as to minimize the number of explanatory variables that significantly affect the value of the output variable "risk level". At the same time, it is possible to study the dependence of explanatory variables on each other, i.e. their mutual correlation. Moreover, according to the results of the study, the proposed methodology makes it possible not to take into account variables that have little effect on the risk level.

The formed knowledge base simplifies significantly the monitoring of the risk level, since this requires making changes to it only according to those parameters whose values have changed during the period under review.

The proposed methodology has been tested in solving the problem of determining critical nodes in information systems with a complex network structure [19], as well as the problem of assessing insurance risks in the insurance of information systems, resources and their supporting infrastructures [26]. In both cases, the methodology showed a good result in assessing risk under conditions of uncertainty.

This technique has a certain versatility and you can use it not only to solve problems of assessing information security risks, but also to solve other problems in conditions of uncertainty. Thus, in [27], using the methods of multicriteria optimization used in the methodology, the parameters that most affect the demographic situation in various regions were determined.

CONCLUSION

The methodology we have considered, which combines the use of fuzzy logic and regression analysis techniques to assess information security risk, enables us to calculate predicted values of risk levels in conditions of uncertainty, taking into account various factors that may influence risk, including subjective factors such as the level of employee qualification, staff turnover within the company, and others.

This technique provides an opportunity to determine the parameters that most affect the value of the risk level, which allows companies to pay special attention to measures to counter threats in the most dangerous areas of activity and minimize the cost of measures to protect their information resources.

The proposed methodology you can apply to any complex network structure in both public and private companies with an extensive network of regional divisions (branches). At the same time, this technique allows not only to determine the level of risk, but also, in the case of additional multicriteria optimization methods, to find critical nodes and evaluate the effectiveness of branches in companies with a distributed structure [19, 28].

This technique you can also apply in the field of cyber insurance or risk insurance in the field of information technology [26]. Its application expands the possibilities of using information risk insurance tools in various insurance companies and makes it relatively easy to determine the level of insurance risk in conditions of uncertainty and greatly facilitate the underwriting procedure.

A similar technique based on the combined use of regression analysis and ranking methods [27], you can use to conduct a comparative analysis of the demographic situation in various regions of Russia with the determination of indicators that have the greatest impact on the demography of the region. This makes it possible for regional authorities to make long- and medium-term forecasts for the formation of a policy in the field of demography. We should note that this indicates some university of the proposed methodology.

The versatility of the methodology allow you to apply it in many areas of management from small companies and enterprises to the public sector, i.e. where it is required to assess various risks under conditions of uncertainty.

Currently, different companies increase using artificial neural networks in economics and business [29]. Neural networks are used, among other things, to increase the level of attacks detection on computer networks, as well as to determine the level of information security risk [30-33]. World statistics show that attackers have recently been massively organizing attacks on neural networks. Therefore, the creation and operation of such networks is also associated with certain risks, most of which lie in the area of uncertainty. The authors believe that the proposed methodology allows for the assessment of information security risk when using neural networks and artificial intelligence systems based on them.

REFERENCES

1. Information Leaks in the World and Russia for the First Half of 2024: An Analytical Report, [Online]. Available https://www.infowatch.ru/analytics/analitika/utechki-informatsii-v-mire-i-rossii-za-pervoye-polugodiye-dve-tysyachi-dvadtsat-chetvertogo-goda.pdf [In Russian]

2. GOST R ISO 31000-2019. (2019). *The National Standard of the Russian Federation. Risk Management. Principles and Guidelines*, [In Russian]. [Online]. Available www.gost.ru

3. Razumnikov, S. V. (2014). Analysis of the Possibility of Using OCTAVE, RiskWatch, CRAMM Methods to Assess IT Risks for Cloud Services, *Modern Problems of Science and Education*, **1**, 247-248, [In Russian].

4. Baranova, S. (2015). Methods of Analysis and Assessment of Information Security Risks, *Bulletin of the S.Y. Witte Moscow University. Series 3. Educational Resources and Technologies,* **1**(9), 73-79, [In Russian].

5. Asgari, H., Haines, S. & Rysavy, O. (2018). Identification of Threats and Security Risk Assessments for Recursive Internet Architecture, *IEEE Systems Journal,* **12**(3), 2437-2448.

6. Kozlov, A. D. & Noga, N. L. (2019). Information Security Risks of Corporate Information Systems when Using Cloud Technologies, *Risk Management*, **3**, 31-46, [In Russian].

7. Tsaregorodtsev, A. V., Zelenina, A. N. & Savelyev, V. A. (2017). Two-Stage Procedure for Quantifying the Risk of Information Security of Cloud Computing, *Modeling, Optimization and Information Technology*. **4**(19), [In Russian].

8. Shirisha, R., BalaRaju, M. & Ramana, N. (2019). Security Measures in Distributed Approach of Cloud Computing, *Advances in Intelligent Systems and Computing*, **768**, 19-30.

9. Kozlov, A. D. & Noga, N. L. (2021). About Some Risks Associated with Subjective Factors and the Methodology for their Assessment, *Review of Business and Economics Studies*. **3**, 94-102.

10. GOST R 58771-2019 (2020). *The National Standard of the Russian Federation. Risk Management. Risk Assessment Technologies*, [In Russian], [Online]. Available www.gost.ru

11. Kozlov, A. D. & Noga, N. L. (2020). Some Method of Complex Structures Information Security Risk Assessment in Conditions of Uncertainty, *Proc. of the 13th International Conference "Management of Large-Scale System Development (MLSD)"* (Moscow, Russia).

12. Choudhary, R. & Raghuvanshi, A. (2012). Risk Assessment of a System Security on Fuzzy Logic, *International Journal of Scientific & Engineering Research*, **3**(12).

13. Hany, S. (2015). Cyber Security Risk Assessment Using Multi Fuzzy Inference System, *International Journal of Engineering and Innovative Technology*, **4**(8), 13-19.

14. Ventresca, M. & Aleman, D. (2015). Efficiently Identifying Critical Nodes in Large Complex Networks, *Computational Social Networks*, **2**(6), 3-16.

15. Kozlov, A. D. & Noga, N. L. (2021). Applying the Methods of Regression Analysis and Fuzzy Logic for Assessing the Information Security Risk of Complex Systems, *Proc. of the 14th int. conf. Management of Large-Scale System Development (MLSD)* (Moscow, Russia).

16. Kornev, L. V. (2021). A Fuzzy Model for Assessing Information Security Risks and Maintaining the Security Level of ERP Systems, *A Young Scientist*, **27**(369), 48-54, [In Russian].

17. Matlab R2022b, [Online]. Available: https://matlab.sooftware.com/

18. Buldakova, T. I. & Mikov, D. A. (2015). Implementation of the Risk Assessment Methodology Information Security in the Matlab Environment, *Cybersecurity Issues*, **4**(12), 53-61.

19. Kozlov, A. D. & Noga, N. L. (2023). Methodology for Determining the Most Critical Nodes of Network Information Infrastructures in Order to Ensure Information Security, *Information Technologies,* **29**(6), 296-306, [In Russian].

20. Beshelev, C. L. & Gurvich, F. G. (1980). *Mathematical and Statistical Methods of Expert Assessments. 2nd Edition*. Moscow, Russia: Statistics, [in Russian].

21. Danelian, T. Y. (2015). Formal Methods of Expert Assessments, *Bulletin of the UMO*, **1**, 183-187, [In Russian].

22. Gudkov, P. A. (2008) *Methods of Comparative Analysis. Study guide*. Penza, Russia: Publishing house of Penza State University, [In Russian].

23. Eliseeva I. I., et. al. (2003). *Econometrics*. Moscow, Russia: Finance and Statistics, [in Russian].

24. Aleskerov, F., Ersel, H. & Yolalan, R. (2004). Multicriteria ranking approach for evaluating bank branch performance, *International Journal of Information Technology & Decision Making*, **3**(2), 321-335.

25. Podinovsky, V. V. (2022). *Multi-Criteria Decision-Making Tasks: Theory and Methods of Analysis: University Textbook*. Moscow, Russia: Yurayt, [In Russian].

26. Kozlov, A. D. & Noga, N. L. (2023). Methodology for Determining the Parameters that Most Affect Insurance Risks in the Field of Information Technology, *Insurance Business,* **12**, 17-26, [In Russian].

27. Kozlov, A. D. & Noga, N. L. (2023). The Comparative Assessment Methodology of the Demographic Situation in the Regions, *Proc. of the 16th International Conference "Management of Large-Scale System Development (MLSD)"* (Moscow, Russia).

28. Kozlov, A. D. & Noga, N. L. (2022). The Selection of the Comparative Evaluation Method of the Effectiveness of the Insurance Company Branches, *Proc. of the 15th International Conference "Management of Large-Scale System Development (MLSD)"*, (Moscow, Russia).

29. Gorodnova, N. V. (2021). Application of Artificial Intelligence in the Business Sphere: Current State and Prospects, *Issues of Innovative Economy*, **11**(4), 1472-1492, [In Russian].

30. Dyachenko, R. A., Chastikova, V. A. & Lyakh, A. R. (2022). Implementation of evasion attacks on neural networks and methods of their prevention, *Scientific Works of the Kuban State Technological University*, **5**, 68-77, [In Russian].

31. Xu, X., Chen, J., Xiao, J., Gao, L., Shen, F., et al. (2020). What Machines See is Not What They Get: Fooling Scene Text Recognition Models with Adversarial Text Images, *IEEE Xplore*, 12304-12314.

32. *SmartEngines How Invulnerable is Artificial Intelligence?*, [Online]. Available https://habr.com/ru/companies/smartengines/articles/528686/, [In Russian]

33. Ermakov, S. A. & Bolgov, A. A. (2022). Risk Assessment Using a Neuro-Fuzzy System, *Information and Security*, **25**(4), 583-592, [In Russian].

