Dmytro PALKO[1], Vira VIALKOVA[2]

Opiekun naukowy: Tetiana BABENKO[3]

MODELE INTELEKTUALNE DO CELÓW OCENY RYZYKA DLA BEZPIECZEŃSTWA CYBERNETYCZNEGO

Streszczenie: W artykule przeanalizowano istniejące podejścia do oceny ryzyka związanego z bezpieczeństwem informacji. Aby zaradzić brakom w metodach analizy i oceny ryzyka związanego z cyberatakami oraz dostosowywania się do danych innych niż liczbowe, w niniejszym opracowaniu proponuje się zastosowanie logiki neuro-rozmytej, która jest skuteczna w przypadkach niewystarczającej dokładności danych wejściowych.

Słowa kluczowe: ryzyko związane z bezpieczeństwem informacji, ocena ryzyka związanego z bezpieczeństwem informacji

INTELLECTUAL MODELS FOR CYBER SECURITY RISK ASSESSMENT

Summary: The article analyzes the existing approaches to information security risk assessment. To address the shortcomings of the methods of analyzing and assessing the risks of IS and adapting to non-numerical data, this study proposes to use neuro-fuzzy logic, which is effective in cases of insufficient accuracy of the input data.

Keywords: risk of information security, information security risk assessment.

1. Introduction

Nowadays the emergence of the Internet, and the transition to the information society, the problem of providing cybersecurity and building secure computing systems has become one of the most topical issues, due to the rapid development of computer and telecommunications technologies. However, the approaches used in these latter days have a number of disadvantages, mainly due to the narrow

[1] Taras Shevchenko National University of Kyiv, Faculty of Informaition Technology, Cyber Security and Information Protection: palko.dmytro@gmail.com
[2] Taras Shevchenko National University of Kyiv, Faculty of Informaition Technology, Cyber Security and Information Protection: veravialkova@gmail.com
[3] Dr.Sc., Professor of Cybersecurity and Information Protection department of the Taras Shevchenko National University of Kyiv, babenkot@ua.fm

specialization of certain cyber defense tools, insufficient level of their interaction, inadequate mechanisms for defining vulnerabilities, risk analysis and determining the level of security, monitoring of network conditions and adaptation to changing conditions. their functioning.

Information security management has become increasingly important in the operation of virtually any organization that applies modern technologies for the collection, storage and processing of information. This process is based on periodic analysis of information risks, which allows time to detect security threats and vulnerability of information system, implement appropriate measures to neutralize them and, consequently, constantly monitor the state of information security in the organization, given past experience, new threats and vulnerability. The prerequisite for the application of intelligent models is the presence of uncertainty due to lack of information or complexity of the system, as well as the availability of quality information about the system [1].

Today used many different methods of risk analysis information security, the main difference between them is the level of risk assessment scales used: quantitative or qualitative.

In quantitative methods, risk is assessed by numerical value. As an input, they usually use the accumulated statistical information on information security incidents for evaluation. However, the use of an unrepresentative amount of data leads to low adequacy of the assessment results.

Qualitative techniques are more common, but they use too simplistic scales, usually containing three levels of risk assessment (high, medium, low). The assessment is usually only based on expert surveys [2].

2. Main results of research

The existing approach to information security risk assessment can be described as follows: risk is a function of three input variables (probability of realization of threat, vulnerability, loss). Information security risk analysis is an ordered algorithm that consists of the same steps, each of which can be applied its own methods (Fig. 1).

Among the methods of risk assessment, as a rule, the following 3 groups are identified [3]:

1) statistical methods;
2) methods of expert assessments;
3) modeling methods.

As is well known, most common approaches to attack identification, risk identification, and assessment have poor accuracy and do not effectively counteract known zero-day attacks. Therefore, being developed by many different technologies to protect computer Books systems and networks that are based on the technology of data mining, including on the use of neural networks and the like. This has to do with the ability of neural network structures to solve complex tasks, as well as their capacity for learning, self-organization, and generalization.

This approach allows to obtain models that are able to quickly adapt to environmental conditions and allow to predict the development of the process on the basis of the generalization property.

However, even when using the device neural networks Esting information security risks are still many problems [4]:

1) incompleteness of information about the components of risk and the system under study;

2) the basis of the input data on the expert assessments, their lack of accuracy or incorrect presentation;

3) the complexity of creating a model information system and assessing its vulnerabilities;

4) the duration of the evaluation process and the relatively rapid loss of relevance of its results;

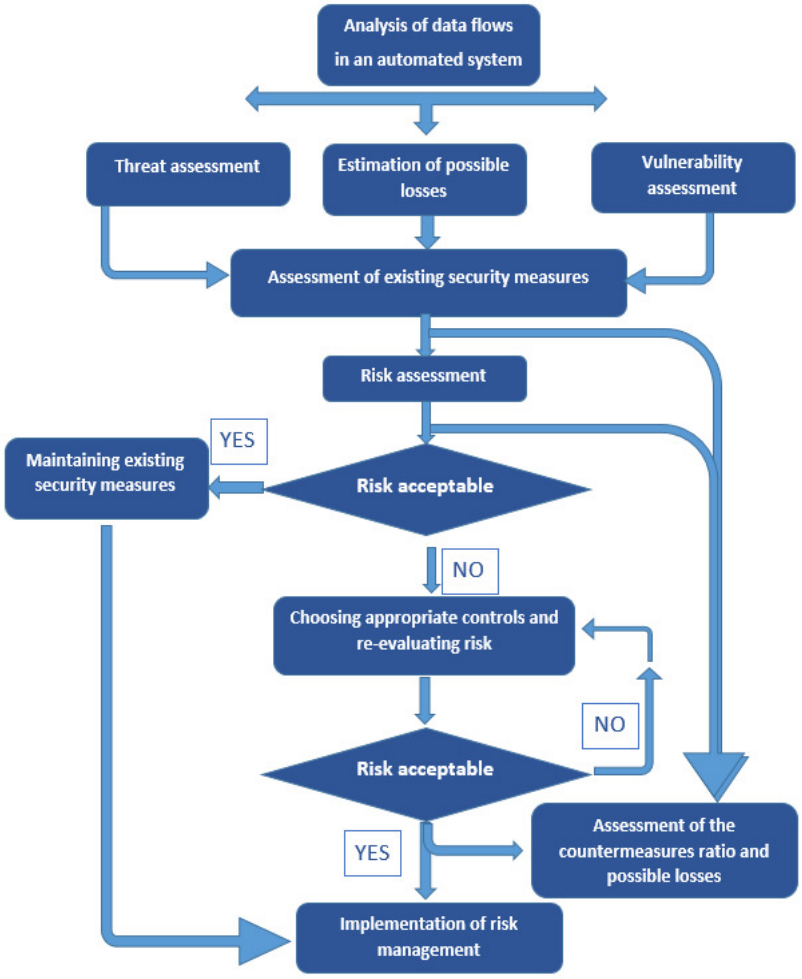

Figure 1. The process of information security risk analysis

Modern methods of managing risk are based on the use of probability of realization of threats, vulnerabilities and loss. However, in most cases, experts in the field of IB, based on their own experience, conduct the assessment in the form of verbal formulations, which are then linked to numerical values.

Such a mechanism to obtain estimates of risk significantly limits opportunities techniques in general, as factors risk (threat, vulnerability loss) are analyzed using heuristic methods of analysis, as a result can be obtained data different from each other if the examination conducted by various experts, and therefore confidence in the proposed expert assessment can be worn controversial nature [5].

In order to address the shortcomings of the methods of analyzing and assessing the risks of IB and adapting to non-numerical data, this study proposes to use neuro-fuzzy logic, which is effective in cases of insufficient accuracy of the input data. Sometimes it is difficult or difficult to quantify the components of the security system or even impossible to consider the input components in terms of fuzzy sets and linguistic variables.

In this study studied the possibility of synthesis models of neuro-fuzzy output for defined spare values and risk matches based on subjective evaluations of all levels of information security. As a tool in the simulation process, We were used fuzzyTECH - software for designing systems of fuzzy logic and neuro-fuzzy solutions and visualization of simulation results.

In the process of research used the algorithm of risk assessment [6]:

1. Sets of linguistic terms are defined to characterize the values of the input parameters (threat level, vulnerability level, damage level) and output (risk level).

2. Set a set of input parameters - damage, threats, vulnerabilities (i= 1…N), characterizing the levels of relevant indicators. In this case, the input parameters have quantitative [0… 1] or qualitative values, expressed in terms of linguistic variables.

3. There is an output risk parameter characterizing the level of risk.

4. Aset of rules of the form "If ..., then ..." is formulated, reflecting the relationship of the input parameters with the output.

5. There is a phasification of the input parameters - finding the values of the membership functions that correspond to the obtained values of the estimates of the input variables.

6. Aggregation is performed to determine the degree of truth of the conditions under each of the rules.

7. There is an accumulation of conclusions - finding the functions of belonging to the output parameter taking into account the aggregation.

8. The output parameter is dephased.

The neuro-fuzzy model of information risk analysis contains 3 inputs and 1 output variable with 5 linguistic terms and 125 fuzzy rules. As input variables, the model uses values of three risk factors in the range [0, 1] that have been described by linguistic terms - "very low", "low", "medium", "high", "very high" (presented in Table. 1).

The model training procedure combined backpropagation methods with the idea of competitive learning. After calculating the output of the model by direct propagation algorithms, the output error is determined by comparing the system output to the reference sample samples.

The error and gradient calculation for all samples, as well as the system parameters are updated after each complete iteration.

Table 1. Levels of scales when assessing risk factors

Level of scale	Threats	Loss	Vulnerabilities
very_low	The event almost never happens	Minor losses of material and resources that are rapidly recovering or have little effect on reputation	A vulnerability that can be ignored
Low	The event is rare	More tangible loss of tangible assets, and a significant impact on reputation	Minor vulnerability that is easy to resolve
medium	An event is quite possible under certain circumstances	There is sufficient loss of tangible assets and a significant impact on reputation and interests	Mild vulnerability
high	Most likely, the event will occur in the implementation of the attack	Significant damage to reputation and interests, which may threaten the further efficiency of the AU	Serious vulnerability, which can be eliminated but is costly
very_high	The event will occur when the attack is implemented	Destructive consequences and inability to continue	A critical vulnerability that is difficult to resolve

To educate and evaluate the adequacy of the neuro-fuzzy model, retrospective data were divided into training, control, and test samples. The training data is a numerical matrix of dimension m × (n + 1), in which the number of rows m corresponds to the volume of the training sample, the first n columns - the values of the input variables of the model, and the last - the value of the output variable Table 2 presents a piece of training sample data.

Table 2. Fragment of the training sample

damage	threats	vulnerabilities	risk
0.35	0.39	0.63	0.5 1
0.63	0.19	0.26	0.4 7
0.6	0.62	0.89	0. 78
0.26	0.19	0.92	0. 81
0.33	0.77	0.39	0.5 6
0.23	0.38	0.17	0.2 5
0.43	0.3	0.33	0. 39

Upon completion of the training process, the average relative error for the control and test sample did not exceed 7%, indicating a high level of accuracy.

3. Conclusion

The analysis of the results obtained in the process of modeling results allows to conclude that the synthesized model adequately describes the processes occurring in the information system while ensuring the adequacy of the expert assessments by calculating the coefficient of concordance and during refinement it can be applied in the process of information security audit.

REFFERENCES

1. ASTAKHOV A. M.: The art of managing information risks.M.: DMK Press, 2010.
2. MING-CHANG LEE: Information Security Risk Analysis Methods and Research Trends: AHP and Fuzzy Comprehensive Method, in the International Journal of Computer Science & Information Technology (IJCSIT). 6(2014)1, 29-45.
3. VELIGURA A.N.: About the choice of methodology for information security risk assessment. Information security. 4(2008), 16-17.
4. ROT A.: IT Risk Assessment: A Quantitative and Qualitative Approach // Proceedings of the World Congress on Engineering and Computer Science. 2008. 1073-1078.
5. LEONENKO A.V.: Fuzzy modeling in MATLAB and fuzzyTECH. St. Petersburg.: BHC - St. Petersburg , 2005.
6. STOVBA S.D.: Designing fuzzy systems with MATLAB. M.: Horyachaya Linia -Telekom, 2011.