An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security[1]

Sergey A. Glushenko
Senior Lecturer, Department of Information Systems and Applied Computer Science
Rostov State University of Economics
Address: 69, Bolshaya Sadovaya Street, Rostov-on-Don, 344002, Russian Federation
E-mail: gs-gears@yandex.ru

Abstract

This article explains the importance of applying risk assessment in the implementation of information security systems. It is considered the most common risk assessment procedure and entails application of fuzzy logic theory for this purpose. The paper describes the proposed fuzzy production model (FPM), which defines seven input linguistic variables describing risk factors, four output linguistic variables that characterize different areas of information security risks, as well as four base rules.

It is noted that the FPM is the first approach to the subject area and requires optimization to minimize the model's output errors. The most common methods of optimization of fuzzy models parameters are examined, and the advantages of applying methods based on neuro-fuzzy networks (NFN) are justified.

The article describes the process of converting fuzzy model elements, such as unit fuzzification, rule base unit and unit defuzzification, into fragments of the neural network. The result of this process is a neuro-fuzzy network corresponding to the fuzzy model.

Formation of the developed NFN is based on an adaptive neuro-fuzzy inference system (ANFIS), using the specialized Neuro-Fuzzy Designer package of MATLAB software. The model was trained by a hybrid method which represents a combination of the methods of least squares and backpropagation. The result of this process is optimization (setting) the parameters of membership functions of input linguistic variables.

Application of neuro-fuzzy modeling made it possible to obtain a more appropriate fuzzy production model which is able to conduct linguistic analysis of the risks of an organization's information security. The information obtained with its help allows IT managers to determine risk priorities and to develop effective action plans to reduce the impact of the most dangerous threats.

Key words: risk, information security, linguistic variable, membership function, neuro-fuzzy network, adaptive neuro-fuzzy inference system, neuro-fuzzy designer.

Citation: Glushenko S.A. (2017) An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security. *Business Informatics*, no. 1 (39), pp. 68–77.
DOI: 10.17323/1998-0663.2017.1.68.77.

[1] This research has been carried out with financial support of RFBR within the framework of scientific project No. 16-31-00285 "Fuzzy logic methods and models in risk management decision support systems"

Introduction

The introduction of information technologies and computing facilities in the production and management of modern enterprises provides an effective tool to increase labor productivity. However, the enterprise IT-infrastructure often takes an unstructured form, which leads to an uncontrolled growth of information security (IS) vulnerabilities and risks to the enterprise as a whole.

Information security is an "information and supporting infrastructure protection against accidental or intentional effects of a natural or artificial character, which can cause unacceptable harm" [1].

Paper [2] analyzes the most common methods of IS risk assessment – NIST [3] and CRAMM [4], describes their disadvantages and proposes to use fuzzy logics for these purposes. The proposed fuzzy production model (FPM) includes seven input linguistic variables (*Table 1*) describing risk factors, four output linguistic variables (*Table 2*), characterizing risks of various areas of information security, as well as four rules bases (*Table 3*) [2, 5].

When forming the input linguistic variables, the following term sets can be used, which determine the levels of factors [6]:

- T2 = {Low (L); High (H)};
- T3 = {Low (L); Medium (M); High (H)};
- T4 = {Very Low (VrL); Low (L); Medium (M); High (H)};
- T5 = {Very Low (VrL); Low (L); Medium (M); High (H); Very High (VrH)}.

When developing output linguistic variables, the following term sets can be used which determine risk factors [6]:

Table 1.

Risk factors of the information security of organization

Designation	Name of linguistic variable	Type of term set and interpretation of levels of factors
x_1	Software/hardware protection	T3. L – satisfactory to ensure the initial security level; M – sufficient for basic information security; H – completely complies with the information confidentiality level
x_2	Organizational protection	T3. L – deficient planning and lack of monitoring of vulnerabilities; M – planning and monitoring of vulnerabilities are conducted irregularly; H – timely planning and monitoring of vulnerabilities
x_3	Legal protection	T3. L – fragmentary and incomplete documentation; M – documentation is available, but short on details; H – documentation is complete and synchronized
x_4	Motivation of a threats source (TS)	T5. VrL – none; L – rare manifestation of interest; M – may well provoke interest; H – most likely will be interested in it; VrH – necessarily will take interest in it
x_5	Possibilities of a threats source (TS)	T5. VrL – none; L –insignificant level of TS infrastructure; M – medium level of infrastructure; H – rather high level of infrastructure; VrH – TS has significant possibilities
x_6	Market value of an information resource (IR)	T5. VrL – clear information; L – IR is of little value; M – IR is commercially confidential; H – highly confidential data; VrH – catastrophic value for the organization (strategic planning)
x_7	Volume of information resource data (IR) of organization	T5. VrL – very small part; L – minor part; M – half of IR; H – major part; VrH – full volume of IR

Table 2.

Risk factors of information security of an organization

Designation	Name of linguistic variable	Remark
y_1	Risk of effective protection reduction	Characterizes the capability to reduce/increase the effective protection in relation to the required effectiveness for a particular enterprise
y_2	Risk of potential threats	Characterizes the possibility of potential threats to an enterprise
y_3	Risk of material damage	Characterizes the possibility of material damage to an enterprise when the organization's information security parameters are violated
y_4	Risk to an organization's IS	Integrated risk describing the enterprise's assurance level of information security

Table 3.

Fuzzy production rules of a model (fragment)

Rule designation	Antecedent	Consequent
	Rules base R1	
R1.1	$(x_1 = \text{L} \wedge x_2 = \text{L} \wedge x_3 = \text{L}) \vee (x_1 = \text{M} \wedge x_2 = \text{L} \wedge x_3 = \text{L}) \vee (x_1 = \text{L} \wedge x_2 = \text{M} \wedge x_3 = \text{L})$	$y_1 = \text{VrHER}$
R1.2	$(x_1 = \text{H} \wedge x_2 = \text{L} \wedge x_3 = \text{L}) \vee (x_1 = \text{M} \wedge x_2 = \text{M} \wedge x_3 = \text{L}) \vee (x_1 = \text{L} \wedge x_2 = \text{H} \wedge x_3 = \text{L}) \vee (x_1 = \text{M} \wedge x_2 = \text{H} \wedge x_3 = \text{L}) \vee (x_1 = \text{L} \wedge x_2 = \text{L} \wedge x_3 = \text{M}) \vee (x_1 = \text{L} \wedge x_2 = \text{M} \wedge x_3 = \text{M}) \vee (x_1 = \text{L} \wedge x_2 = \text{H} \wedge x_3 = \text{M}) \vee (x_1 = \text{L} \wedge x_2 = \text{L} \wedge x_3 = \text{H})$	$y_1 = \text{HER}$
R1.3	$(x_1 = \text{H} \wedge x_2 = \text{M} \wedge x_3 = \text{L}) \vee (x_1 = \text{H} \wedge x_2 = \text{H} \wedge x_3 = \text{L}) \vee (x_1 = \text{M} \wedge x_2 = \text{L} \wedge x_3 = \text{M}) \vee (x_1 = \text{H} \wedge x_2 = \text{L} \wedge x_3 = \text{M}) \vee (x_1 = \text{M} \wedge x_2 = \text{M} \wedge x_3 = \text{M}) \vee (x_1 = \text{M} \wedge x_2 = \text{H} \wedge x_3 = \text{M}) \vee (x_1 = \text{M} \wedge x_2 = \text{L} \wedge x_3 = \text{H}) \vee (x_1 = \text{L} \wedge x_2 = \text{M} \wedge x_3 = \text{H}) \vee (x_1 = \text{M} \wedge x_2 = \text{M} \wedge x_3 = \text{H}) \vee (x_1 = \text{L} \wedge x_2 = \text{H} \wedge x_3 = \text{H})$	$y_1 = \text{MER}$
R1.4	$(x_1 = \text{H} \wedge x_2 = \text{M} \wedge x_3 = \text{M}) \vee (x_1 = \text{H} \wedge x_2 = \text{H} \wedge x_3 = \text{M}) \vee (x_1 = \text{H} \wedge x_2 = \text{L} \wedge x_3 = \text{H}) \vee (x_1 = \text{H} \wedge x_2 = \text{M} \wedge x_3 = \text{H}) \vee (x_1 = \text{M} \wedge x_2 = \text{H} \wedge x_3 = \text{H})$	$y_1 = \text{LER}$
R1.5	$x_1 = \text{H} \wedge x_2 = \text{H} \wedge x_3 = \text{H}$	$y_1 = \text{VrLER}$

– T1 = {Low evidence of risk (LER); Medium evidence of risk (MER); High evidence of risk (HER)};

– T2 = {Very low evidence of risk (VrLER); Low evidence of risk (LER); Medium evidence of risk (MER); High evidence of risk (HER); Very high evidence of risk (VrHER)}.

In order to build a fuzzy model, it is necessary to determine all its elements: rules base, number and type of membership functions for each variable model, parameters of membership functions, logical operators, etc.

The structure of the fuzzy production model for assessing risk to an organization's information security is provided in *Figure 1.*

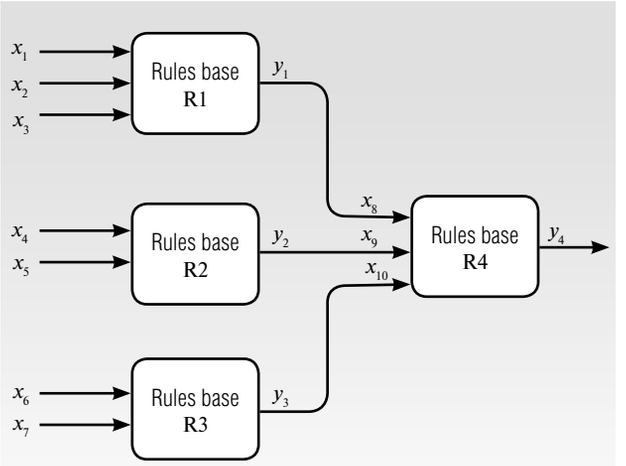

Fig. 1. Structure of fuzzy productive model

1. Statement of the problem

The developed model is based on expert knowledge about the modeled information security system (ISS) [2, 5, 7]. The system information was obtained with the involvement of an expert in the domain area, with the information obtained converted in a fuzzy model. This method is efficient if the expert has a perfect knowledge of the ISS. In practical work, the expert's knowledge is often insufficiently complete and accurate, and sometimes even contains contradictions. Therefore, the model should be based on objective system information, which can be presented by data from measuring the system input and output values [8].

These circumstances predetermine the relevance of development of the fuzzy self-tuning model for ISS risk assessment. The fuzzy model setting should primarily mean a process of determining parameters of the membership functions of input and output linguistic variables which minimize errors of the model outputs relative to the observed prototype system.

For model setting, namely, the optimization of its parameters, the following methods are most often used [9]:

- ✦ methods based on using neuro-fuzzy networks;
- ✦ searching methods;
- ✦ clustering-based methods.

The first group methods are associated with the conversion of a fuzzy model in a neuro-fuzzy network (NFN) and application of network training methods based on measurements of input and output system data for settings of the model network.

The second group methods are direct search methods for optimal parameters of the fuzzy model. The search process can be both ordered and unordered (trial-and-error method). The most commonly used ordered search method is a method based on the use of genetic algorithms.

Clustering-based methods combine model parameters setting and its structuring. They are used for building fuzzy self-organizing models which have control over their essential input parameters, define an optimal number of fuzzy sets for input and output linguistic variables, and establish the form and number of rules.

Currently, the most studied methods are the first group methods. They make it possible to [10]:

- ✧ optimize (set) parameters of the membership functions of linguistic parameters based on measurements of input and output dependences of the actual system;
- ✧ correct fuzzy models which are not developed adequately enough by experts;
- ✧ extend the expert-developed fuzzy models to the area of the studied model, where expert knowledge is limited.

The listed peculiarities explain the applicability of methods based on the use of neuro-fuzzy networks for setting a fuzzy model for assessing risk to an organization's information security.

2. Conversion of fuzzy model to neuro-fuzzy network

A conversion of fuzzification block elements is provided in *Figure 2*, which depicts a conversion of the piecewise linear membership functions to a fragment of the neural network.

To set parameters a_i of the membership functions in the course of network training, it is necessary to calculate derivatives of output values of the fuzzification block using appropriate parameters.

Block fuzzification results in calculated values of a degree of membership of input values by fuzzy set A_{ij}, each of which represents a linguistic range of definition.

The conversion of block elements of the rules base predicts presentation of the rule condition in the form of a neural network fragment. In this case, operations "AND" and "OR" can be performed using T- and S-norms and standards, or through other operators.

The input parameters of the defuzzification block are activation degrees $\mu_{B_l}(y)$ of fuzzy sets B_l at the model output. The center-of-gravity method is used for their conversion to a distinct number.

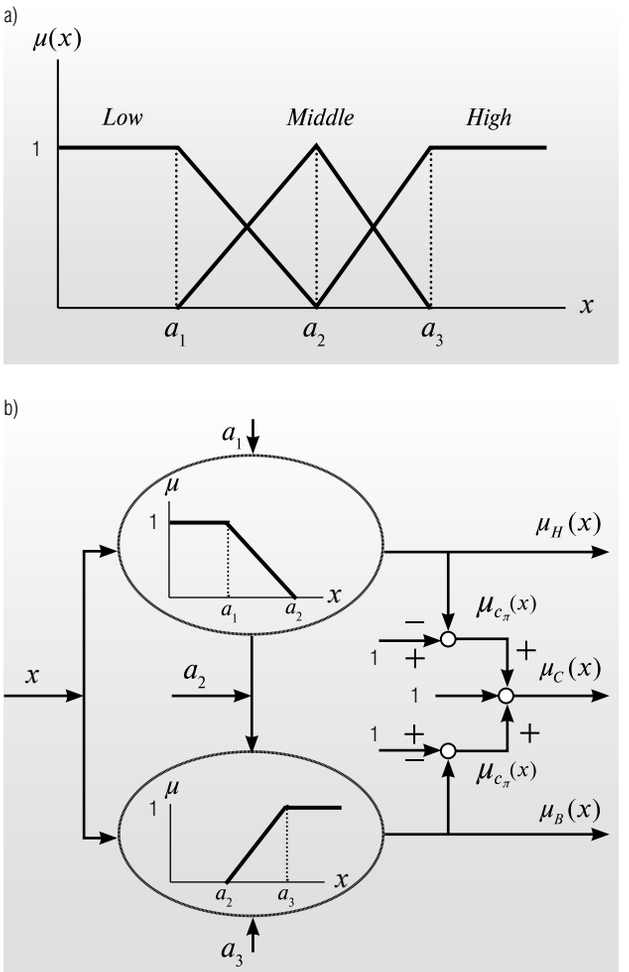

Fig. 2. Conversion of piecewise linear membership functions (a) to neural network fragments (b)

As a result, the neuro-fuzzy network corresponding to the fuzzy model in *Table 3* will have the structure given in *Figure 3*.

3. Application of Neuro-Fuzzy Designer package for building the NFN

The developed NFN is built based on the adaptive neuro-fuzzy inference system (ANFIS) [8, 9] through the Neuro-Fuzzy Designer package of MATLAB software [11]. ANFIS is a neural network with several inputs and one output, which in their turn are fuzzy linguistic variables. In this case, the terms of input and output linguistic variables are described by membership functions that are coherent with the developed fuzzy self-tuning model of the IS risk assessment.

In the fuzzification phase, the triangular membership functions (*Figure 4*) for term sets of input (x_1, x_2, x_3) and output (y_1) linguistic variables (LV) were specified:

x_1 – LV "Software and hardware protection" (*SwHwPrt*);

x_2 – LV "Organizational protection" (*OrgPrt*);

x_3 – LV "Legal protection" (*LegPrt*);

y_1 – LV "Risk of effective protection reduction" (*RiskPrt*).

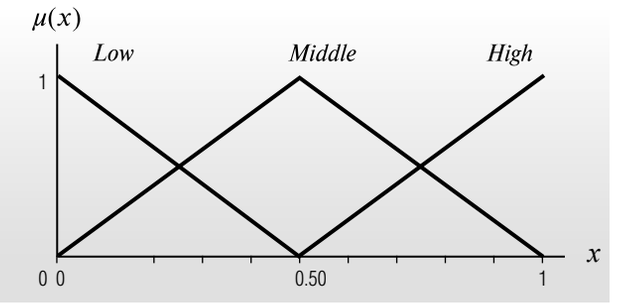

Fig. 4. Membership function for input variable *SwHwPrt*

The generated fuzzy inference system, which contains 27 rules of fuzzy products, is depicted in *Figure 5*.

NFN training was performed based on the training sample, which contained 200 sets representing a vector of values of factor levels having an impact on the risk, (input LV) and values of IS risk level (output LV). The data was obtained by generalizing the domain expert opinions using the Delphi method within the approach proposed in paper [12]. To generate training sets, the data received from the intrusion detection systems, antivirus programs, firewalls and other systems included in the ISS can be also used.

The Neuro-Fuzzy Designer package enables you to train using the method of backpropagation, the main purpose of which is to set up all multilayer structure layers by changing weights of intermediate layers, and the hybrid method, which is a combination of the method of least squares and the method of backpropagation. The results of applying NFN training methods for risk assessment of information security are provided in *Table 4*.

Table 4.

Application of Training Methods for Neutron-Fuzzy Network

Training method	Error value	Number of stages[1]
Method of backpropagation	0.0271	200
Hybrid method	0.0108	28

[2] Number of stages required for achieving a stated error value

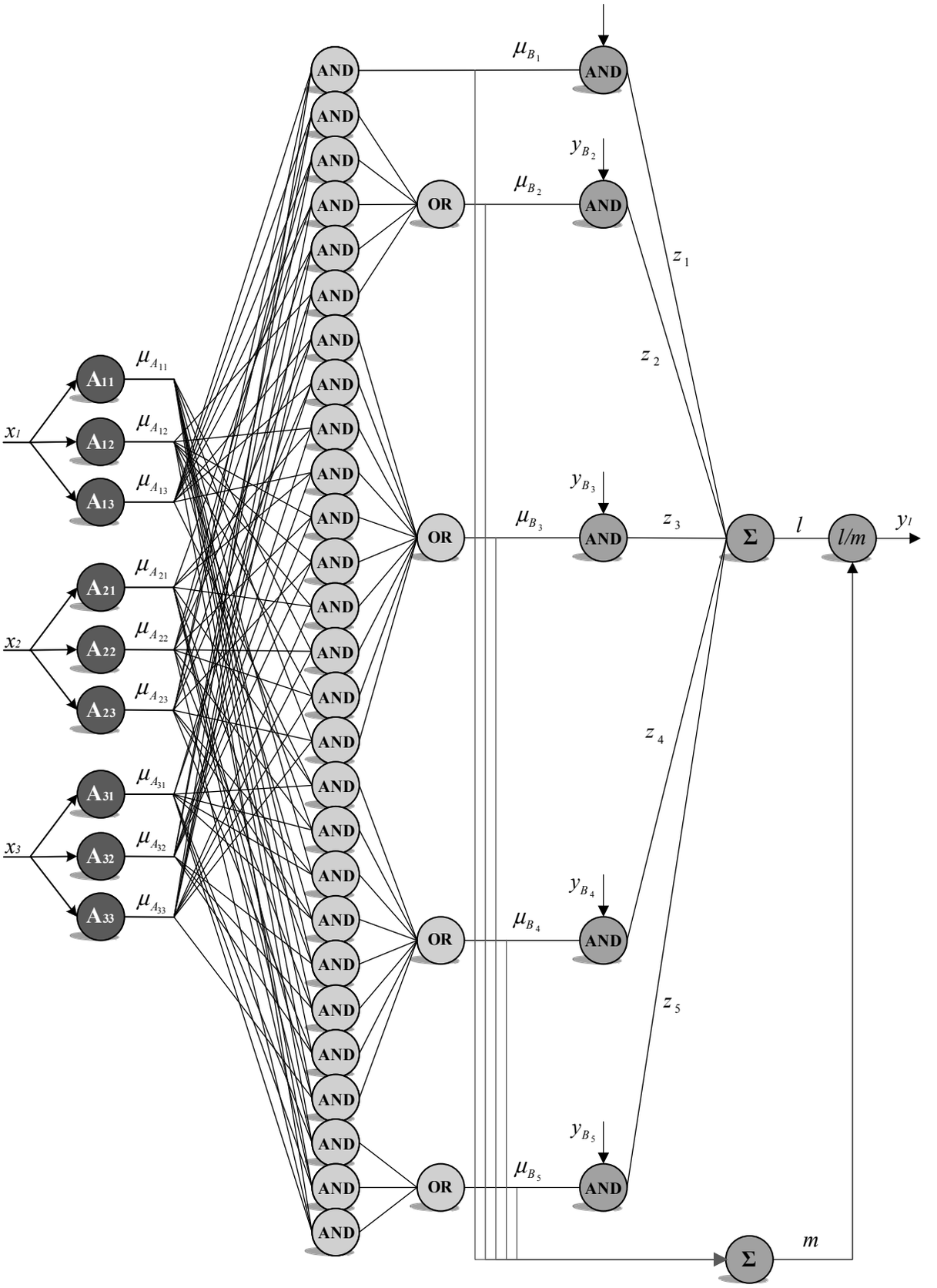

Fig. 3. Neuro-fuzzy network (fragment)

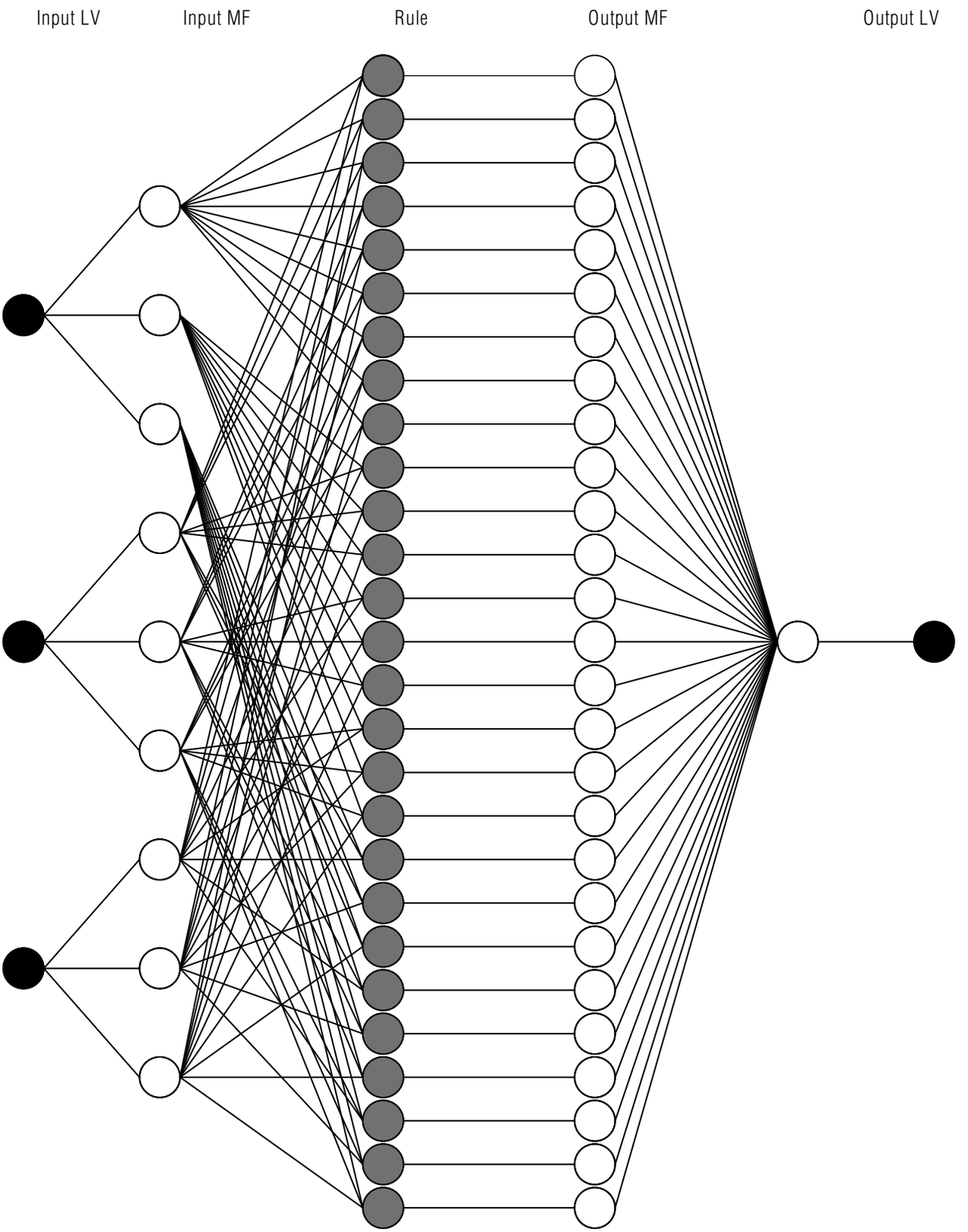

Fig. 5. Structure of fuzzy inference system

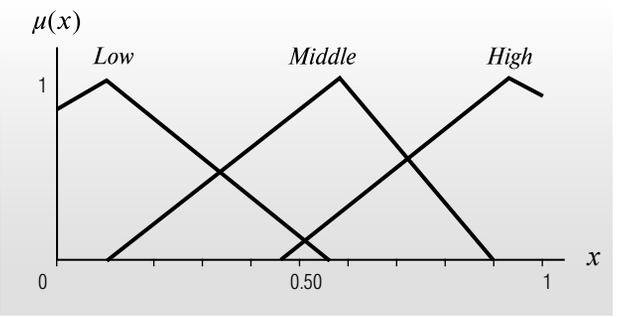

Fig. 6. Optimized membership functions

As the Table shows, the hybrid training method enables us to get better results of network errors value during a fewer number of stages. With this in mind, for configuring the membership function parameters the choice was made in favor of the hybrid method. *Figure 6* depicts a result of optimization (setting) of membership function parameters of the linguistic variable *SwHwPrt*.

Figure 7 depicts a surface of the trained fuzzy model, which shows how output LV depends on two input LV; meanwhile, the third variable value is fixed.

The graphical view of dependence of output LV (*RiscPrt* – "Risk of effective protection reduction") on input LV (*SwHwPrt* – "Software and hardware protection" and *OrgPrt* – "Organizational protection") shows an expected increase of the value of risk of effective protection reduction of an organization with the decrease of the hardware-software protection and organizational protection [2].

Therefore, a smooth and monotone dependence diagram of the reduced "output surface" implies a good "quality" of the output mechanism and adequacy and consistency of the used inference rules.

The risk assessment mechanism based on NFN has broad capabilities. In particular, it can be adapted to the existing risk management models, and can be also modified taking into account the actual conditions of the organization's information security policy [7].

Conclusion

The fuzzy production model described in the introduction was the first approximation for the subject domain concerned and was to be set up. The developed fuzzy self-tuning model of the information security risk assessment enabled us to adjust parameters of the membership functions for linguistic variables for the information security systems under study and obtain a more adequate fuzzy production model.

The fuzzy self-tuning model we implemented enables us to perform continuous analysis of the information security risk, and the information obtained as a result of fuzzy modeling enables IT managers to identify risk priorities (from "very high" to "very low") and to develop effective action plans to reduce the impact of the most hazardous threats.

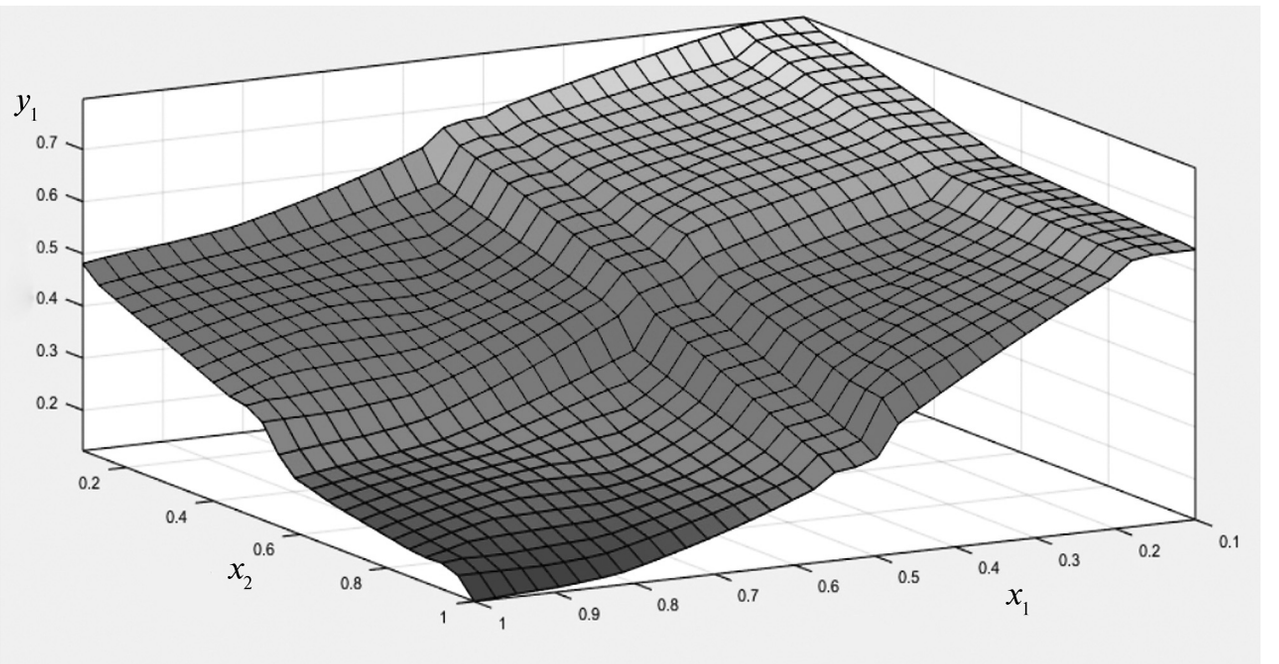

Fig. 7. Surface of the fuzzy model system

References

1. State Technical Commission (2005) *Informatsionnye tekhnologii. Osnovnye terminy i opredeleniya v oblasti tekhnicheskoy zashchity informatsii. Rekomendatsii po standartizatsii (R 50.1.053-2005)* [Information technologies. Basic terms and definitions in scope of technical protection of information. Recommendations on standardization (P 50.1.053-2005)]. Moscow, State Technical Commission (in Russian).
2. Glushenko S.A. (2013) Primenenie sistemy MATLAB dlya otsenki riskov informatsionnoy bezopasnosti organizatsii [Risk assessment information security systems organization with MATLAB system]. *Business Informatics*, no. 4 (26), pp. 35–42 (in Russian).
3. National Institute of Standards and Technology (2002) *Risk management guide for information technology systems. Special publication 800-30.* Gaithersburg, MD: NIST.
4. Simonov S.V. (1999) Analiz riskov, upravlenie riskami [Risk analysis, risk management]. *Jet Info*, no. 1 (68), pp. 2–28 (in Russian).
5. Glushenko S.A. (2013) Nechetkaya produktsionnaya model' otsenki riskov informatsionnoy bezopasnosti organizatsii [Fuzzy production model for assessment of an organization's information security risks]. *Economics and Law Issues*, issue 11. Rostov-on-Don, RSUE (RINE), p. 147 (in Russian).
6. Dolzhenko A.I. (2009) Model' analiza riska potrebitel'skogo kachestva proektov ekonomicheskikh informatsionnykh sistem [Risk analysis model of consumer quality of economic information systems projects]. *Bulletin of North-Caucasus State Technical University*, vol. 18, no. 1, pp. 129–134 (in Russian).
7. Glushenko S.A., Dolzhenko A.I. (2015) Sistema podderzhki prinyatiya resheniy nechetkogo modelirovaniya riskov informatsionnoy bezopasnosti organizatsii [A decision support system for fuzzy modeling of an organization's information security risks]. *Information Technologies*, no. 1, pp. 68–74 (n Russian).
8. Borisov V.V., Kruglov A.S., Fedulov A.S. (2012) *Nechetkie modeli i seti* [Fuzzy models and networks]. Moscow: Hotline – Telecom (in Russian).
9. Piegat A. (2013) *Nechetkoe modelirovanie i upravlenie* [Fuzzy modeling and control]. Moscow: BINOM. Knowledge Laboratory (in Russian).
10. Rutkowska D., Pilinski M., Rutcowski L. (2006) *Neyronnye seti, geneticheskie algoritmy i nechetkie sistemy* [Neural networks, genetic algorithms and fuzzy systems]. Moscow: Hotline – Telecom (in Russian).
11. Leonenkov A.V. (2005) *Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH* [Fuzzy modeling using MATLAB and fuzzyTECH]. Saint Petersburg: BHV-Petersburg (in Russian).
12. Khubaev G.N. (2011) Poluchenie gruppovoy ekspertnoy otsenki znacheniy pokazateley: poshagovaya protsedura i programmnoe obespechenie [Obtaining a group expert estimate of metrics values: a stepwise procedure and software]. *Software and Systems*, no. 2, pp. 13–16 (in Russian).

Адаптивная нейро-нечеткая система оценки рисков информационной безопасности организации[3]

С.А. Глушенко

кандидат экономических наук, старший преподаватель кафедры информационных систем и прикладной информатики
Ростовский государственный экономический университет (РИНХ)
Адрес: 344002, г. Ростов-на-Дону, ул. Большая Садовая, д. 69
E-mail: gs-gears@yandex.ru

Аннотация

В статье обосновывается важность применения оценки рисков при реализации системы обеспечения информационной безопасности. Рассматриваются наиболее распространенные методики оценки риска и предлагается использовать для этих целей теорию нечеткой логики. Описывается предложенная нечеткая продукционная модель (НПМ), в которой определены семь входных лингвистических переменных, характеризующих факторы риска, четыре выходных лингвистических переменных, характеризующих риски различных областей информационной безопасности, а также четыре базы правил.

[3] Исследование выполнено при финансовой поддержке РФФИ, в рамках научного проекта № 16-31-00285 «Методы и модели нечеткой логики в системах принятия решений управления рисками»

Отмечается, что НПМ является первым приближением для рассматриваемой предметной области и требует оптимизации с целью минимизации ошибки выходов модели. Рассматриваются наиболее распространенные методы оптимизации параметров нечетких моделей и обосновываются преимущества применения методов, основанных на использовании нейро-нечетких сетей (ННС).

Описывается процесс преобразования элементов нечеткой модели, таких как блок фаззификации, блок базы правил и блок дефаззификации во фрагменты нейронной сети. Результатом данного процесса является нейро-нечеткая сеть, соответствующая нечеткой модели.

Построение разработанной ННС осуществляется на основе системы нейро-нечеткого вывода (adaptive neuro-fuzzy inference system, ANFIS) посредством применения специализированного пакета Neuro-Fuzzy Designer программного средства MATLAB. Обучение модели было выполнено гибридным методом, который представляет собой комбинацию методов наименьших квадратов и обратного распространения ошибки. Результатом данного процесса является оптимизация (настройка) параметров функций принадлежности входных лингвистических переменных.

Использованный подход нейро-нечеткого моделирования позволил получить более адекватную нечеткую продукционную модель, которая позволяет проводить лингвистический анализ рисков информационной безопасности организации. Полученные с ее помощью сведения позволяют ИТ-менеджерам определять приоритеты рисков и разрабатывать эффективные планы мероприятий по снижению влияния наиболее опасных угроз.

Ключевые слова: риск, информационная безопасность, лингвистическая переменная, функция принадлежности, нейро-нечеткая сеть, система нейро-нечеткого вывода, дизайнер нейро-нечетких сетей.

Цитирование: Glushenko S.A. An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security // Business Informatics. 2017. No. 1 (39). P. 68–77. DOI: 10.17323/1998-0663.2017.1.68.77.

Литература

1. Информационные технологии. Основные термины и определения в области технической защиты информации. Рекомендации по стандартизации (Р 50.1.053-2005). М., 2005.
2. Глушенко С.А. Применение системы MATLAB для оценки рисков информационной безопасности организации // Бизнес-информатика. 2013. № 4 (26). С. 35–42.
3. Risk management guide for information technology systems. Special publication 800-30. Gaithersburg, MD: NIST, 2002.
4. Симонов С.В. Анализ рисков, управление рисками // Jet Info. 1999. № 1 (68). С. 2–28.
5. Глушенко С.А. Нечеткая продукционная модель оценки рисков информационной безопасности организации // Вопросы экономики и права: Сборник статей аспирантов и соискателей ученой степени кандидата наук. Выпуск 11. Ростов-на-Дону: РГЭУ (РИНХ), 2013. С. 147.
6. Долженко А.И. Модель анализа риска потребительского качества проектов экономических информационных систем // Вестник Северо–Кавказского государственного технического университета. 2009. Т. 18. № 1. С. 129–134.
7. Глушенко С.А., Долженко А.И. Система поддержки принятия решений нечеткого моделирования рисков информационной безопасности организации // Информационные технологии. 2015. № 1. С. 68–74.
8. Борисов В.В., Круглов А.С., Федулов А.С. Нечеткие модели и сети. 2-е изд. М.: Горячая линия – Телеком, 2012. 284 с.
9. Пегат А. Нечеткое моделирование и управление. 2-е изд. М.: БИНОМ. Лаборатория знаний, 2013. 798 с.
10. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Горячая линия – Телеком, 2006. 452 с.
11. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб: БХВ-Петербург, 2005. 736 с.
12. Хубаев Г.Н. Получение групповой экспертной оценки значений показателей: пошаговая процедура и программное обеспечение // Программные продукты и системы. 2011. № 2. С. 13–16.