Journal of Theoretical and Applied Information Technology
15th June 2018. Vol.96. No 11
© 2005 – ongoing JATIT & LLS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ASSESSING INFORMATION SECURITY RISK WITH THE FUZZY SET THEORY

[1]MURATKHAN RAIKHAN, [2]KHABDOLDA BOLAT, [3]ZHUMABEKOV MEIRAM, [4]OMAROVA ALTYNAY

[1]Senior Lecturer, Master of Applied Mathematics, Department of Applied Mathematics and Informatics Sciences, Karaganda State University named after academician E.A. Buketov, Kazakhstan

[2]Teacher, Master of Natural Science, Department of Medical Biophysics and Informatics, Karaganda State Medical University, Kazakhstan

[3]Candidate of Philological Sciences, Head of Department of Journalism, Karaganda State University named after academician E.A. Buketov, Kazakhstan

[4]Senior Lecturer, Master of Technical Sciences, Department of General Education, Academia Bolashak, Kazakhstan

E-mail: [1]rmuratkhan@yahoo.com, [2]berserbol@mail.ru, [3]merik.75@mail.ru, [4]altynaiom@mail.ru

ABSTRACT

There are two important issues when it comes to information system development – business processes modeling and business project risk assessment. However, in practice, risk is assessed at the later stages of project development (design and implementation). This enhances the importance of the question whether performed business processes are safe. Traditionally, information system risk is defined as the combination of probable negative events and possible consequences. However, information security risk of a modern organization is a multidimensional complex concept that includes a set of interrelated variables. Values of risk factors often cannot be accurately calculated. Therefore, information security risk assessment can be considered as an unclear problem. This article describes methods of information security risk assessment with the fuzzy set theory. An example of organization's information security risk assessment is considered according to the requirements of international standards and preferences of the owner of information resources.

Keywords: *Risk, Fuzzy Set Theory, Risk Assessment Method*

1. INTRODUCTION

Modeling, based on creating and studying models describing how the automated system (AS) functions, is a common approach to assessing AS operability. Such models allow analyzing and optimizing information collecting, storing and processing, as well as choosing data protection technologies [1]. Mathematical model of the system allows assessing various characteristics of the AS adequately by reflecting physical essence of its processes. However, classical modeling methods require clear model input values.

Methods of business processes risk analysis are powerful tools to help people manage uncertainty. A detailed analysis, risk assessment and assessment results can provide valuable support for decision-making. There are many risk analysis methods that calculate and assess risks. These methods can be either qualitative or quantitative, depending on the available information and detail level [2]. Quantitative methods are based mostly on statistical approaches, including the Monte Carlo simulation [2], fault and event tree analysis [2-3], sensitivity analysis [3], annual expected losses [4], Risk Impact [5], failure type and analysis of effects [3], etc. On the other hand, qualitative methods are based mainly on judgment than on statistical calculations, such as scenario analysis [4], fuzzy set theory (FST) [4], etc. Quantitative and qualitative methods both have advantages and disadvantages. FST method seems to be appropriate for risk analysis and assessment, since such an analysis is highly subjective and associated with inaccurate and vague information.

Since FST was introduced by Zadeh [6] to deal with uncertain problems, linguistic meanings were widely used to bring the reason to the truth. Fuzzy set theory provides an approximate model for information risk assessment through a linguistic approach.

AS security analysis is different, as fuzzy values (namely, expert estimates) are often used as input data while assessing information security (IS) risks. This necessitates the use of fuzzy models, which major advantage is associated with the possibility of using much smaller volumes of information about the simulated system in comparison with traditional mathematical models. At the same time, information can be crude and fuzzy. This is effective when it comes to such a complex and ambiguous process as risk assessment in automated systems [6-8].

The purpose of this paper is to create a model that will allow assessing information security risks with incomplete and ambiguous data on their components. We will consider the business process risk management based on asset analysis. Based on this approach, business process analysts can trace the origin of risk (namely, vulnerable assets).

2. METHODS

If we want to develop a fuzzy model suitable for assessing information security risks, we have to analyze the capabilities of existing models based on the FST.

A fuzzy set A in X is defined as the set $A = \{(x, \mu_A(x))|x \in X\}$, where X is a range of values, and $\mu_A(x)$ is the membership function characterizing the membership degree of x element to a fuzzy set A. Thus, there are three cases:

1. $\mu_A(x)$= 1 – complete membership of the x element in a fuzzy set A – A∈x;
2. $\mu_A(x)$= 0 – x element is not in a fuzzy set A – A∉x;
3. 0<$\mu_A(x)$<1 – partial membership of the x element in a fuzzy set A.

As a rule, fuzzy models are developed for fuzzy control systems. Therefore, their typical structure involves 4 blocks [9]:

1) linguistic variable formalization;
2) fuzzification block (calculating the membership degree of clear input parameters of the model to input fuzzy sets);
3) output block (rule base – a set of logical rules defining the cause-effect relationship between the input and output values – is the major element);
4) defuzzification block (calculating clear output value based on the resulting membership function, calculated by the output mechanism in the output block).

Different types of fuzzy models differ in the way the blocks are implemented.

Currently, Mamdani model is the most frequently used type of fuzzy model [10]. In the framework of Mamdani's method, simulated system is considered as a "black box", characterized by insufficient information about the physical phenomena occurring inside it. The model performs such a mapping of input space (vector X) to an output space Y that provides the most accurate approximation of the real system (for example, in the sense of a mean absolute error). This type of mapping will be possible if there is a geometric surface (mapping surface) in space given defined by the Cartesian product X×Y. The Mamdani model is a set of rules:

IF (x ∈ A) THEN (y ∈ B),

Where: A, B – fuzzy sets.

Each rule sets a certain fuzzy point in the specified space. Based on a set of fuzzy points, fuzzy graph is formed. At the same time, mechanism of interpolation between points depends on the used fuzzy logic apparatus.

There have been developed other types of fuzzy models, including Takagi-Sugeno-Kang models (TSK models) that are the most important ones. Mamdani and Takagi-Sugeno-Kang models differ in the form of rules [11]. In the case of TSK model, the rules are as follows:

IF (x ∈ A) THEN (y=f(x)),

Where: fuzzy sets are replaced with the function f(x), which can be nonlinear. Although, linear functions y=ax+b are usually used.

Since conclusion made on the basis of a TSK model has a more complex mathematical representation and is less visible than the conclusion made on the basis of a Mamdani model, Mamdani model is more suitable for assessing information security risks, as visibility of the overall situation is more important, than the value accuracy.

2.1 Linguistic variable formalization

If we want to use the Mamdani model to assess information security risks, we have to determine the input data. It follows from the definition of information security risk that risk magnitude R is a function of potential damage (value of information, resource or asset) AV, information security threat probability P(T) and asset vulnerability to threats V:

$$R=V*P(T)*AV \quad (1)$$

Thus, input factors will involve expert estimates for three fuzzy variables ("threat probability ", "asset value", " asset vulnerability to threats ") described by linguistic term set: {very low, low, medium, high, very high} (Table 1).

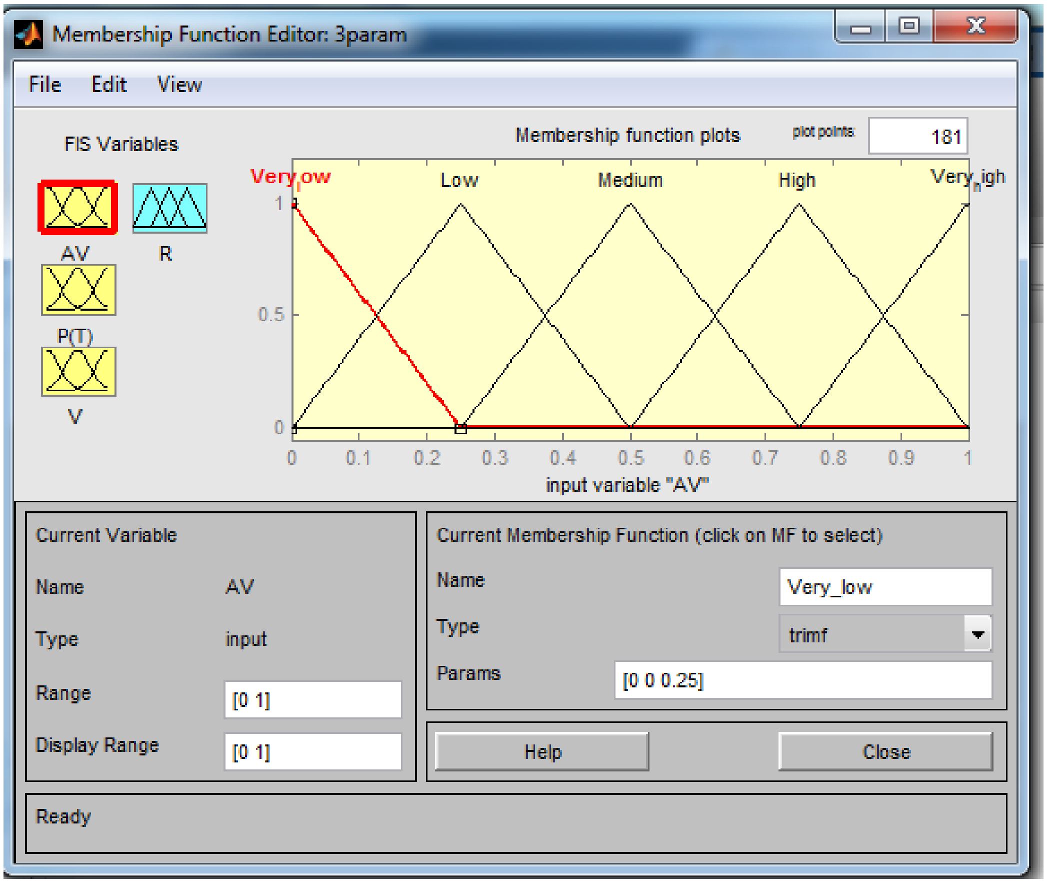
Table 1: Assessment Scale Levels in regards to Threats, Damage and Vulnerabilities.

As a result, we will get an information security risk assessment that can be described by an extended linguistic term set: {negligible, very low, low, below average, moderate, above average, high, very high, critical} (Table 2).

Table 2: IS Risk Assessment Scale Levels.

2.2 Fuzzification

Let's apply the Mamdani model to assess information security risks. We will use the MATLAB Fuzzy Logic Toolbox package to computerize the process of obtaining clear values of the "information security risk" with the Mamdani fuzzy inference algorithm. In this study, membership functions of linguistic terms are characterized by triangular fuzzy numbers, as they are very often used in applications, such as fuzzy controllers. They are also often used in decision-making, business and finance, and social sciences, etc. [12]. Membership functions of four fuzzy sets (threat probability, asset value, asset vulnerability to threats and IS risk) are shown in Figures 1, 2, 3 and 4, respectively.



Membership Function Editor: 3param

Current Variable

Name AV

Type input

Range [0 1]

Display Range [0 1]

Current Membership Function (click on MF to select)

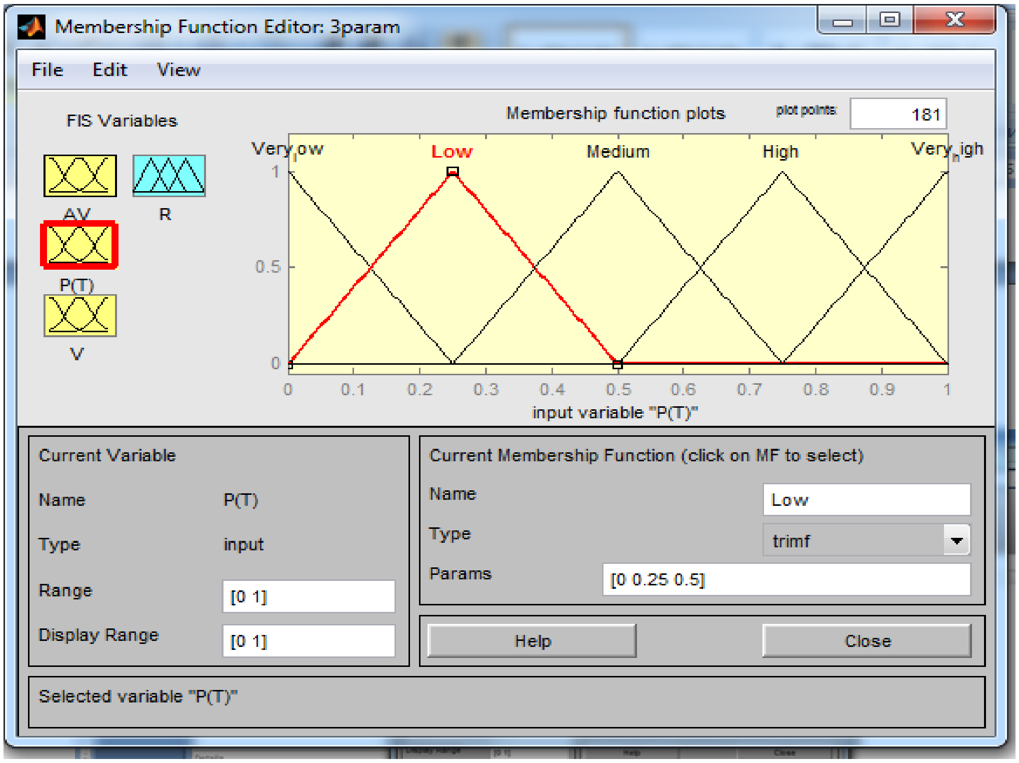
Name Very_low

Type trimf

Params [0 0 0.25]

Ready

Figure 1: Membership function of a linguistic variable "AV"



Membership Function Editor: 3param

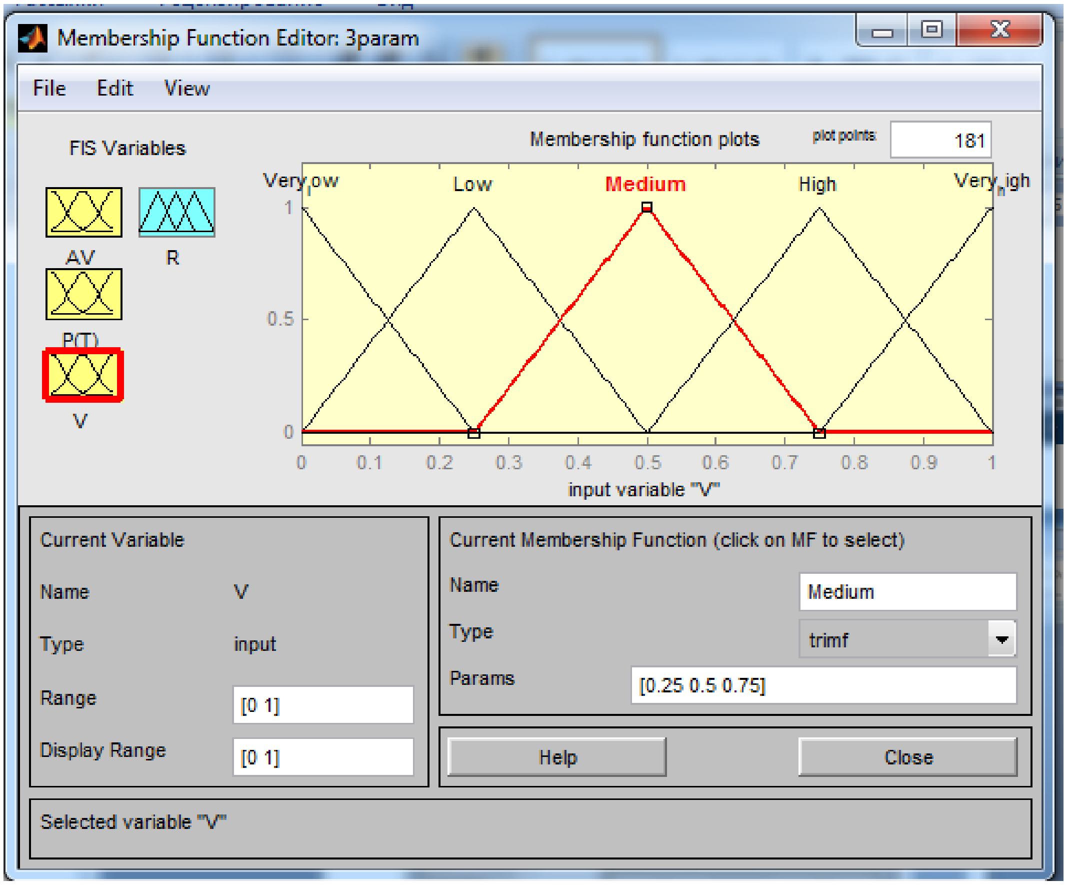
Membership function plots plot points: 181

Current Variable
Name P(T)
Type input
Range [0 1]
Display Range [0 1]

Current Membership Function (click on MF to select)
Name Low
Type trimf
Params [0 0.25 0.5]

Selected variable "P(T)"

Figure 2: Membership function of a linguistic variable "P(T)"



Membership Function Editor: 3param

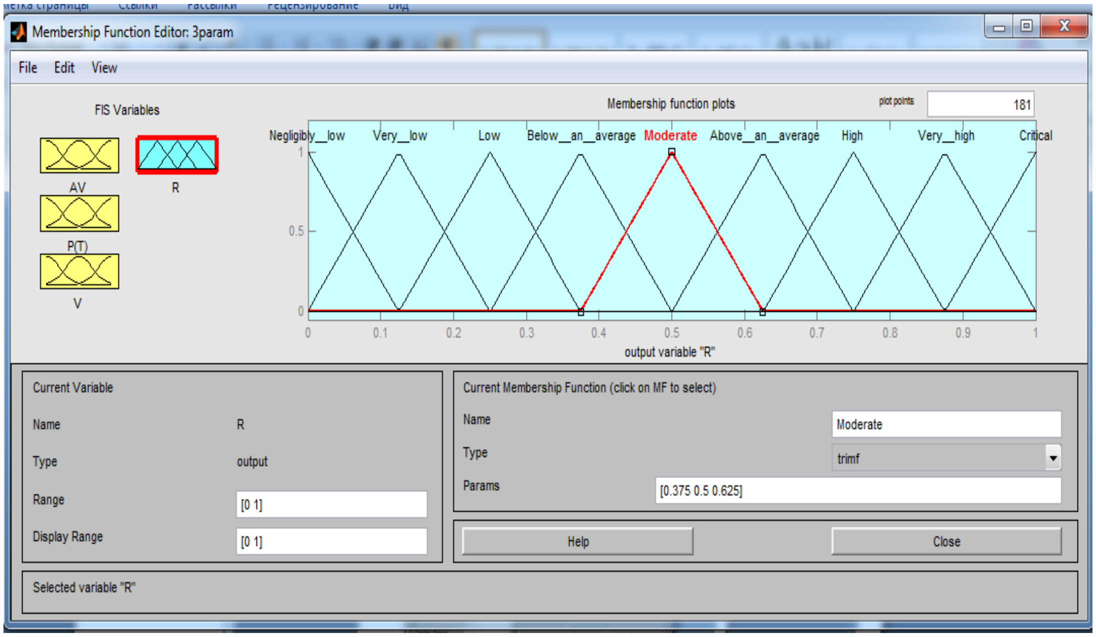
Membership function plots plot points: 181

Current Variable
Name V
Type input
Range [0 1]
Display Range [0 1]

Current Membership Function (click on MF to select)
Name Medium
Type trimf
Params [0.25 0.5 0.75]

Selected variable "V"

Figure 3: Membership function of a linguistic variable "V"



plot points 181

Current Variable

Name R

Type output

Range [0 1]

Display Range [0 1]

Current Membership Function (click on MF to select)

Name Moderate

Type trimf

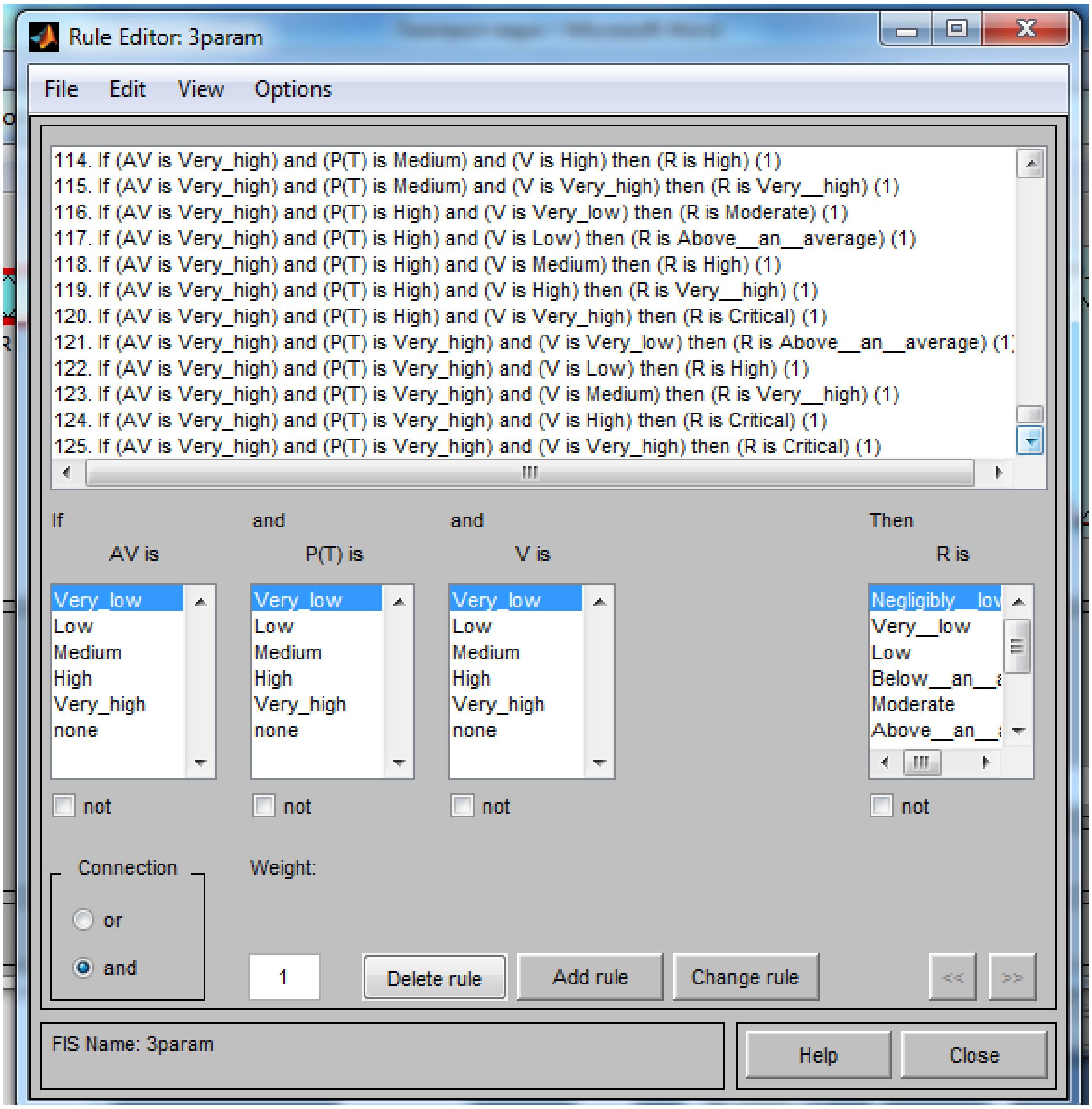
Params [0.375 0.5 0.625]

Selected variable "R"

Figure 4: Membership function of a linguistic variable "R"

2.3 Output block

Risk assessment mechanism is an expert system, which knowledge base is composed of rules reflecting the logic of relationships between input (AV, P(T), V) and output values (R). In the simplest case, this is a "table" logic, in general – a more complex logic reflecting real relationships formalized with production rules like "If ..., then". We have used the following production rules (Figure 5).



114. If (AV is Very_high) and (P(T) is Medium) and (V is High) then (R is High) (1)
115. If (AV is Very_high) and (P(T) is Medium) and (V is Very_high) then (R is Very__high) (1)
116. If (AV is Very_high) and (P(T) is High) and (V is Very_low) then (R is Moderate) (1)
117. If (AV is Very_high) and (P(T) is High) and (V is Low) then (R is Above__an__average) (1)
118. If (AV is Very_high) and (P(T) is High) and (V is Medium) then (R is High) (1)
119. If (AV is Very_high) and (P(T) is High) and (V is High) then (R is Very__high) (1)
120. If (AV is Very_high) and (P(T) is High) and (V is Very_high) then (R is Critical) (1)
121. If (AV is Very_high) and (P(T) is Very_high) and (V is Very_low) then (R is Above__an__average) (1
122. If (AV is Very_high) and (P(T) is Very_high) and (V is Low) then (R is High) (1)
123. If (AV is Very_high) and (P(T) is Very_high) and (V is Medium) then (R is Very__high) (1)
124. If (AV is Very_high) and (P(T) is Very_high) and (V is High) then (R is Critical) (1)
125. If (AV is Very_high) and (P(T) is Very_high) and (V is Very_high) then (R is Critical) (1)

FIS Name: 3param

Figure 5: Specified production rules

2.4 Defuzzification

Defuzzification is a process of converting a fuzzy set into a clear number.

In the fuzzy set theory, defuzzification procedure is similar to the process of finding position characteristics (mathematical expectation, modes, medians) of random variables described by the probability theory. The simplest way to perform defuzzification is to select a clear number corresponding to the maximum membership function value.

Defuzzification of a fuzzy set $\widetilde{A} = \int_{[\underline{u},\overline{u}]} \mu_A(u)/u$ by the center of gravity method is performed according to the formula:

$$a = \frac{\int_{\underline{u}}^{\overline{u}} u \cdot \mu_A(u)\, du}{\int_{\underline{u}}^{\overline{u}} \mu_A(u)\, du}$$

Defuzzification of a fuzzy set $\tilde{A} = \int_{[\underline{u},\overline{u}]} \mu_A(u)/u$ by the median method consists in finding such a number a, that

$$\int_{\underline{u}}^{a} \mu_A(u)\ du = \int_{a}^{\overline{u}} \mu_A(u)\ du$$

Defuzzification of a fuzzy set $\tilde{A} = \int_{[\underline{u},\overline{u}]} \mu_A(u)/u$ by the Center of Maximum method is performed according to the formula:

$$a = \frac{\int_G u\,du}{\int_G du}$$

Where: G – set of all elements belonging to the interval $[\underline{u},\ \overline{u}]$, which degree of membership in Ã is the maximum possible [13].

3. DISCUSSION

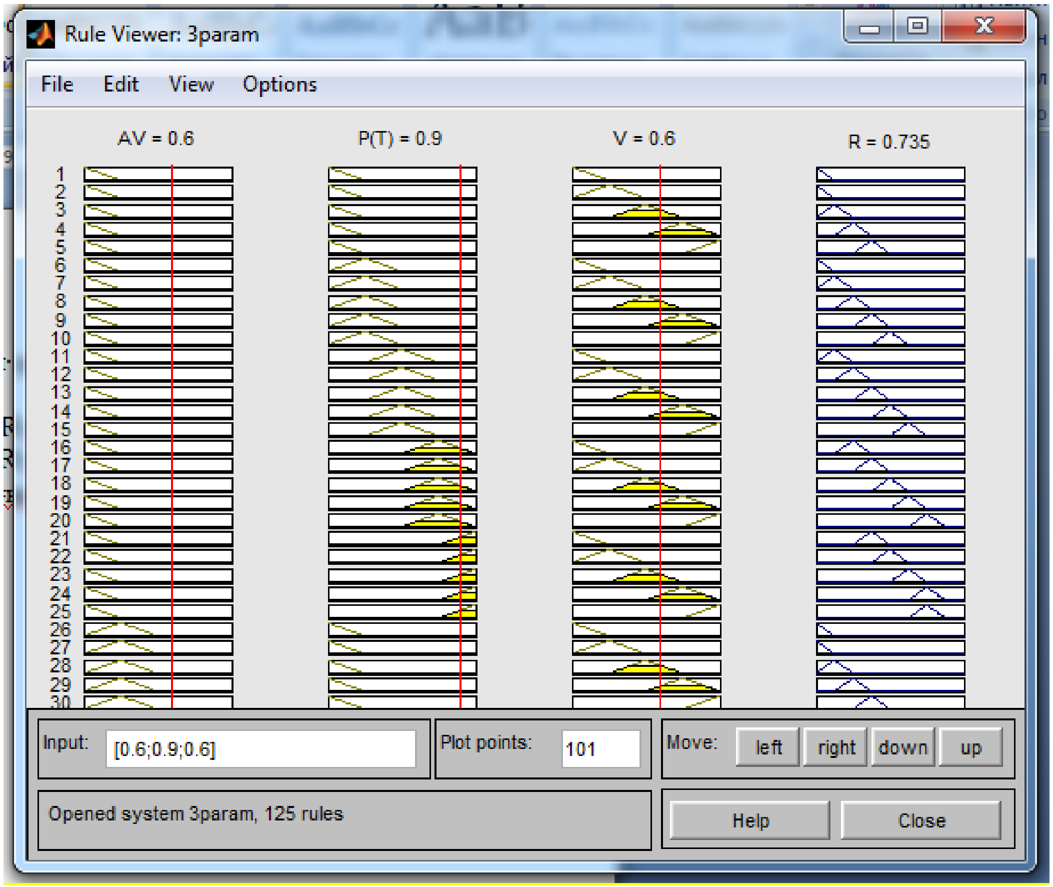
In [14], there are considered examples with three input data (AV, P (T), V); IS risk levels (R) are assessed according to the Microsoft methodology for a single asset, but with different vulnerabilities (Table 3).

Table 3: IS Risk Assessment according to the Microsoft Methodology [10].

Risk assessment results for these examples will be presented below, but with a fuzzy model.

Figure 6 shows the graphical interpretation of a Mamdani fuzzy inference algorithm with triangular membership functions for the first example (AV=0.6, P(T)=0.9, V=0.6) and obtained R (0.735). The latter corresponds to the linguistic variable – high risk.

Figure 7 shows the graphical interpretation of a Mamdani fuzzy inference algorithm with trapezoid membership functions for the considered threat example (AV=0.6, P(T)=0.9, V=0.6) and obtained R (0.71). The latter corresponds to the linguistic variable – high risk.

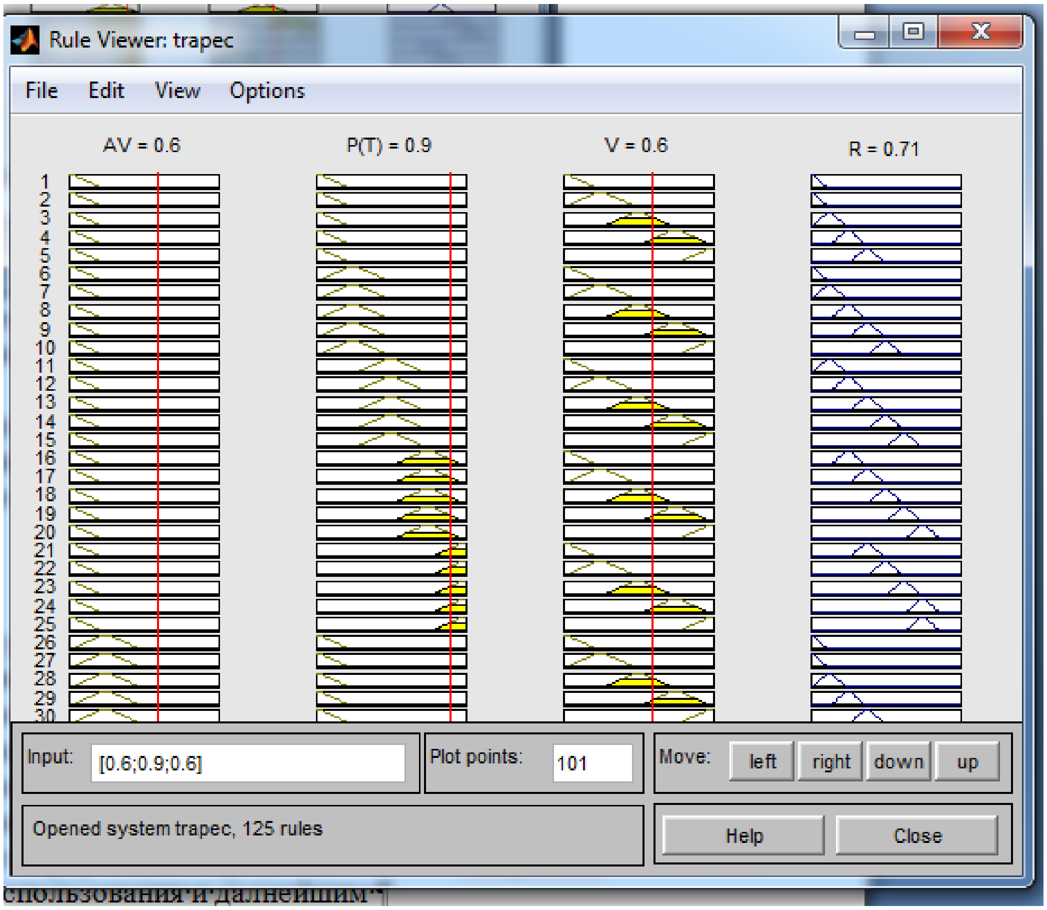


Input: [0.6;0.9;0.6]

Plot points: 101

Opened system 3param, 125 rules

Figure 6: Graphical interpretation of the Mamdani fuzzy inference algorithm with triangular membership functions



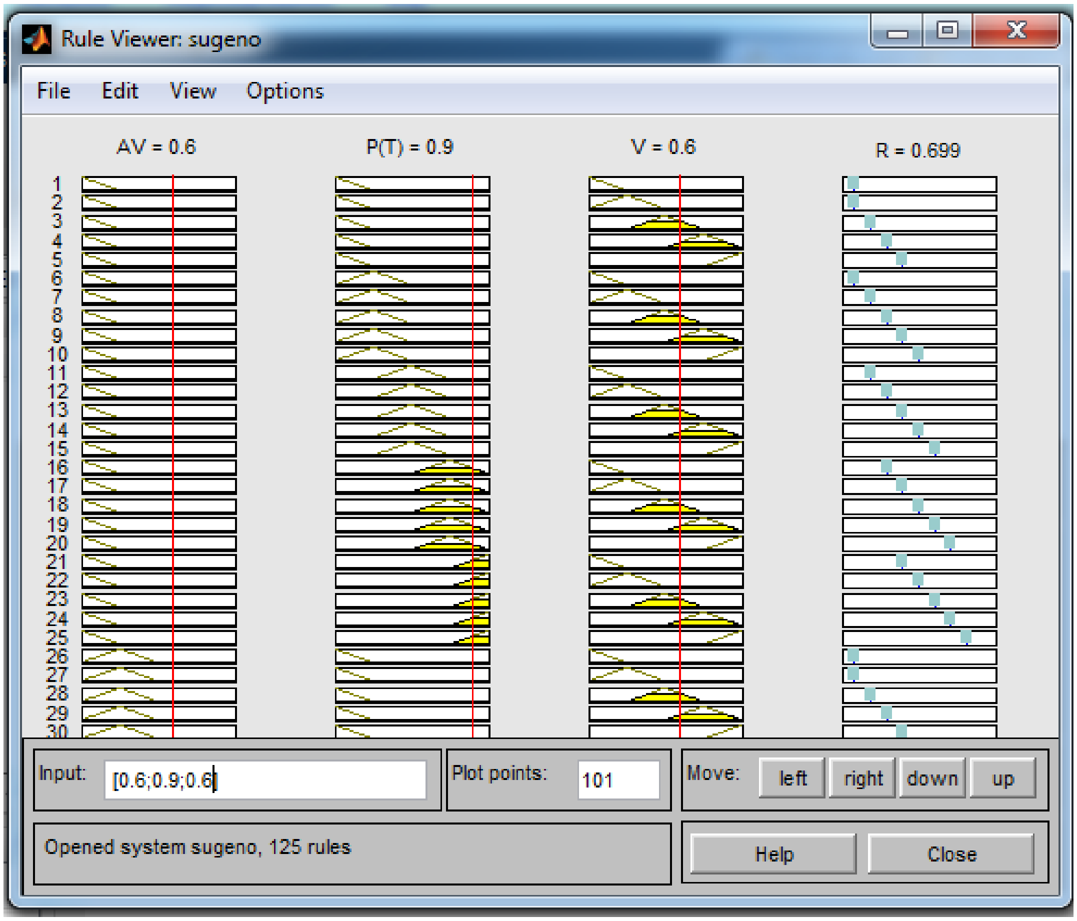
Input: [0.6;0.9;0.6]

Plot points: 101

Opened system trapec, 125 rules

Figure 7: Graphical interpretation of the Mamdani fuzzy inference algorithm with trapezoid membership functions

Figure 8 shows a graphical interpretation of the Sugeno fuzzy inference algorithm for the considered threat example (AV=0.6, P(T)=0.9, V=0.6) and obtained R (0.699). The latter corresponds to the linguistic variable – high risk.



AV = 0.6 P(T) = 0.9 V = 0.6 R = 0.699

Input: [0.6;0.9;0.6]

Plot points: 101

Opened system sugeno, 125 rules

Figure 8: Graphical interpretation of the Sugeno fuzzy inference algorithm

Similarly, we have calculated data for other cases presented in Table 3; results are presented in Table 4. Information security risk levels calculated with a fuzzy sets apparatus and fuzzy logic correspond to IS risk levels calculated according to the generally accepted Microsoft methodology (Table 4). This proves that this article proposes an adequate fuzzy model of IS risk assessment. Existing qualitative methods of information security risk assessment do not provide sufficiently accurate results, and quantitative assessment is reduced to probabilistic methods that do not provide reliable results without incident statistics. Models, based on the fuzzy set theory and fuzzy logic, do not have the aforementioned shortcomings and can be used to process expert estimates.

Table 4: Comparative Analysis Of Fuzzy Logic Based Information Security Risk Assessment Methods.

№	R [14]	Mamdani fuzzy inference		Sugeno fuzzy inference
		Triangular membership functions	Trapezoid membership functions	
1	High	0.735	0.71	0.699
2	High	0.777	0.81	0.816
3	Low	0.271	0.28	0.261

4. CONCLUSIONS

Fuzzy logic based mechanism of risk assessment allows considering the quality of input data and reliability of information sources. It has wide capabilities allowing it to adapt to a variety of application profiles and build in its own risk management system.

Methodology for information risk assessment that we propose was developed for practical use. We have proposed an approach to risk assessment that involves business project development based on a fuzzy set theory.

Many well-known methods and software products used for information risk assessment provide us with quantitative or qualitative assessment. Our methods will help to get both quantitative and qualitative assessments.

In this paper, less attention is paid to risk management planning, resolution and control that come with business project development. Further research should be conducted in regards to risk management planning. Besides, risk must be controlled regularly in order to track the status of identified risks.

REFRENCES:

[1] T.I. Buldakova, and A.Sh. Dzalolov, "Analysis of Data Processes and Choices of Data-processing and Security Technologies in Situation Centers", *Scientific and Technical Information Processing*, Vol. 39, No 2, 2012, pp. 127-132. DOI: 10.3103/S0147688212020116

[2] J.C. Bennett, G.A. Bohoris, E.M. Aspinwall, and R.C. Hall, "Risk Analysis Techniques and their Application to Software Development", *European Journal of Operational Research*, Vol. 95, 1996, pp. 467–475.

[3] D. White, "Application of Systems Thinking to Risk Management: a Review of the Literature", *Management Decision*, Vol. 3, No. 10, 1995, pp. 35–45.

[4] R.K.J.R. Rainer, C.A. Snyder, and H.H. Carr, "Risk Analysis for Information Technology', *Journal of Management Information Systems*, Vol. 8, No. 1, 1991, pp. 129–147.

[5] B.W. Boehm, "Software Risk Management", *IEEE Computer, Society Press* (Washington, DC), 1989.

[6] L.A. Zadeh, "Fuzzy Sets", *Information and Control*, Vol. 8, 1965, pp. 338–353.

[7] D. Satybaldina, R. Muratkhan, and D. Kabenov, "Ontology and Fuzzy Measures Based System for Information Security Risk Assessment", *WOSIS - 9th International Workshop on Security in Information Systems* (Wroclaw, Poland), 28 june – 1 july, 2012, pp. 77-85.

[8] P.A. Balashov, R.I. Kislov, and V.P. Bezguzikov, "Fuzzy Logic Based Information Security Risk Assessment. Computer Systems Security", Confident: *Information and Methodological Journal*, Vol. 6, 2003, pp. 60-65.

[9] N.G. Yarushkina, "Fundamentals of the Theory of Fuzzy and Hybrid Systems", *Finance and Statistics Publishing House* (Moscow), 2004, 320 p.

[10] E.H. Mamdani, and S. Assilian, "An Experiment in Linguistic Synthesis with Fuzzy Logic Controller", *International Journal of Man-Machine Studies,* Vol. 7, No. 1, 1975, pp. 1-13.

[11] T. Takagi, and M. Sugeno, "Fuzzy Identification of Systems and its Applications to Modeling and Control. *IEEE Transactions on Systems, Man and Cybernetics,* Vol. SMC-15, No 1, 1985, pp. 116-132. DOI: 10.1109/TSMC.1985.6313399

[12] G. Bojadziev, and M. Bojadziev, "Fuzzy Logic for Business, Finance, and Management", *World Scientific*, Singapore, 1997.

[13] N.B. Khaptakhaeva, S.V. Dambaeva, and N.N. Ayusheeva, "Introduction to the Fuzzy Set Theory: Textbook", Part I. Ulan-Ude: *East-Siberian State University of Technology and Management*, 2004, pp. 68.

[14] D. Baranov, and I. Koneev, "Issues of Transition from Qualitative to Quantitative Risk Analysis", *Depositarium Journal*, Vol. 9, No. 67, 2008, pp. 26-31.

Table 1: Assessment Scale Levels in regards to Threats, Damage and Vulnerabilities.

Scale Level	Threat Probability (*P(T)*)	Asset Value (*AV*)	Asset Vulnerability to Threats (*V*)	Value
Very Low	The event almost never happens	Insignificant losses of tangible assets and/or resources that can be quickly replenished, or insignificant reputational exposure	Ignorable vulnerability	(0; 0; 0.25)
Low	The events are rare	More noticeable losses of tangible assets, more significant reputational exposure or infringement of interests	Insignificant vulnerability that can be easily eliminated	(0; 0.25; 0.5)
Medium	The event can happen under certain circumstances	Ample losses of tangible assets and/or resources, ample damage to reputation and interests	Moderate vulnerability	(0.25; 0.5; 0.75)
High	The event is most likely to happen under attacks	Significant damage to reputation and interests that may be a threat to further business	Serious vulnerability that can be eliminated with significant costs	(0.5; 0.75; 1)
Very High	The event will probably happen if there are attacks	Destructive consequences and inability to continue business activity	Critical vulnerability that casts some doubt on whether there is a possibility to eliminate it	(0.75;1; 1)

Table 2: Is Risk Assessment Scale Levels.

Scale Levels	Risk Description	Value
Negligible	Risk can be neglected	(0; 0; 0.125)
Very Low	It is necessary to determine whether there is a need in corrective actions, or whether there is an opportunity to take this risk	(0; 0.125; 0.25)
Low	Risk level allows working, but there are prerequisites for a malfunction	(0.125; 0.25; 0.375)
Below Average	Corrective action plan should be developed and applied within an acceptable period of time	(0.25; 0.375; 0.5)
Moderate	Risk does not allow working stably; there is an urgent need in corrective actions changing the operating mode towards risk reduction	(0.375; 0.5; 0.625)
Above Average	The system can continue its operation, but the corrective action plan must be applied as quickly as possible	(0.5; 0.625; 0.75)
High	At this risk level, business processes are unstable	(0.625; 0.75; 0.875)
Very High	Measures on risk reduction must be immediately taken	(0.75; 0.875; 1)
Critical	Risk level is very high and unacceptable for the organization; system operation has to be stopped; radical measures must be taken to reduce the risk	(0.875; 1; 1)

Table 3: IS Risk Assessment According To The Microsoft Methodology [10].

№	Asset name	AV	Risk Description	P(T)	Vulnerability Description	V	R
1	Data on client’s investments	average	Unauthorized access to client data through the stolen financial adviser's account	high	Local account theft due to untimely updated anti-virus software, network configuration or security systems	medium	high
2	Data on client’s investments	average	Unauthorized access to client data through the stolen financial adviser's account	high	Remote client account theft due to untimely updated anti-virus software, network configuration or security systems	high	high
3	Data on client’s investments	low	Unauthorized access to client data through the stolen financial adviser's account	medium	Account is stolen by a good employee abusing his official position	low	low