UDC 004.413.4

Assessing Security Risks Using the Apparatus of Fuzzy Logic Theory

V. I. Chernish, K. I. Ivanov, A. A. Zamula
Kharkiv National University of Radioelectronics, Ukraine

This work is dedicated to the use of the apparatus of fuzzy logic for assessing information security risks. In this work we consider the urgency of evaluating the risks. Also the risk estimation technique with the use of the apparatus of the theory of fuzzy logic is shown.

***Key words**: risk, fuzzy set, approximate reasoning, term.*

Данная работа посвящена вопросам использования аппарата нечеткой логики для оценивания рисков информационной безопасности. В работе рассматриваются актуальность проблемы оценивания рисков, а также приведена методика оценивания рисков с использованием аппарата теории нечеткой логики.

***Ключевые слова:** риск, нечеткое множество, приближенное рассуждение, терм.*

Ця робота присвячена питанням використання апарату теорії нечіткої логіки для оцінювання ризиків інформаційної безпеки. У роботі розглядаються актуальність проблеми оцінювання ризиків, а також детально наведена методика оцінювання ризиків з використанням апарату теорії нечіткої множини.

***Ключові слова:** ризик, нечітка множина, наближені розсуди, терм.*

1. Relevance of the work

The widespread introduction of information technologies (IT) in the structure of modern organizations has become reality. At all managerial levels there is a need to expand information and communications capabilities through the introduction of modern IT. As a result, the information structures of organizations grow: local network of organizations become connected to the Internet; networks are combined into multi-tiered structures of the office and expand to the level of distributed corporate networks. But the introduction of IT is accompanied by the possibility breaches of information security (IS).

We have not any questions about evaluation value of information risk (IR) before certain moment, and the problem is compounded by the fact that security issues usually are being solved by IT professionals who are unable to determine the real implications for the organization of IS threats. The company's managers, who would assess the real impact of threats, often have low skills in the IT field to estimate IR.

Thus, organizations often lack the expert who could assess the real risks associated with the use of IT, although risk assessment is needed not only in terms of security of already existing infrastructure, but also for a correct assessment of prospects of development and security of IT. Moreover, the timely assessment of risk should be viewed not only as a function of IR management, but also as protection for the entire organization and its ability to perform functions. Consequently, the risk management process should not be regarded as a technical function that can be instructed the technician, but as the basic control function of organization.

2. The risk of IS and its evaluation

Standard of IS [1] defines the evaluation of information risk as a comprehensive assessment of the two indicators, the so-called two-factor model:

- potential damage to the company in violation of information security;
- the probability occurrence of such violation.

Both measures may not always be accurately determined numerically, so a possible solution is to carry out their assessments on fuzzy data. Such data are opinions of experts. Further assessments passed specialist of IS, which decides which risks should be reduced, and which does not. In the known method of estimating the risks at the level of risk assigned to the product of the magnitude of possible damage and probability of risk. Then choose from a variety of risks that the level of which is the greatest, and he is eliminated. Behind him is eliminated second highest risk, then the third, etc. Risk assessment of this model can be interpreted as the average expected loss of risk for the billing period. However, if the risk is going to happen, the damage from the risk is received completely, not partially.

A major shortcoming of this approach is the lack of value that determines the cost of risk reduction when deciding the fate of the risk. In fact, the highest level of risk may require as many costs as four following risks.

It is seen that decreasing of the highest level risk is less profitable than reduction of the next four risks. You should also remember that there are dependent on each other risks. Dependence can be of two kinds: appearance of first increases the probability of appearance of second and appearance of first reduces that probability. The dependence can also occur while minimizing the risks: by eliminating one of the risks, we would reduce the probability of occurrence of another. The calculation of such relationships is quite complex and should be done by an expert.

The introduction of a new variable, the value of risk reduction to an acceptable level, of course, will complicate the decision-making system. In addition, the frequent situation is when a particular risk is not fully statistical. If you expand the above-described two-factor model with the new options, you will need to obtain a precise formula for assessing risk and deciding on risk. In the case where there are no accurate data there is a problem in the reliability of risk assessment the transition from quantitative to qualitative assessment of the variables and fuzzy inference rules such as "if A then B", is possible problem. For these purposes, a theory of fuzzy sets is acceptable, which simplifies the decision-making in complex and formalized problems.

3. The basics of using the theory of fuzzy sets and linguistic variables in evaluating the risk IS

Frequently the risk analysis is a task without a clearly defined measures of success, because it basically implies the best guesses and intuitive judgments, resulting in the fuzzy (vague) data [1].To express the risk of fuzzy numbers you should find its main parameters (initial and central moments) that characterise the membership function.

The concept of IS risk assessment with the use of the theory of fuzzy logic is the logical-linguistic model, which is based on the theory of fuzzy sets and linguistic variables. Fuzzy set in a non-empty space U is the set of ordered pairs [].

$$\{x_i / \mu_A(x_i)\},$$

where $\mu_A(x)$: $U \to [0, 1]$ - function of belonging of x to A, that assigns to each element $x \in U$ the degree of its belonging to the fuzzy subset A. Function $\mu_A(x)$ takes its values in a totally ordered set M = [0, 1], which is called the set of accessories [3, 4].

Linguistic variable is characterized by a set of *(L, T (L), U, G, M)*, where *L* - the variable name; *T (L)* - term-set of the variable *L*; *U* - universal set of basic values (the area in which the values of linguistic variable are determined); *G* - syntactic rule; *M* - semantic rule [3].

Term-set *T (L)* is a set of terms that names the linguistic values of the variable *L*. Each term corresponds to a fuzzy subset of *U*, which defines the linguistic meaning of the term.

In other words, the meaning of the term is characterised by a function $\mu: U \to [0, 1]$ that assigns to every element $u \in U$ the value of this element. Syntactic rule G generates terms. Term that consists of a single word or several words, that are always appear together, is called an atomic term.

Term, consisting of one or more atomic terms, is called a composite term. The semantic rule M assigns to each atomic term a meaning in the form of a fuzzy set. In addition, the semantic rule M binds the accessories of atomic terms in a composite linguistic value to belonging to the composite value.

4. Fuzzy logic and approximate reasoning

The interpretation of truth as a linguistic variable leads to fuzzy logic. Fuzzy logic is the basis of approximate reasoning that is kind of reasoning, in which the values of truth and inference rules are unclear. The basis of fuzzy logic consists of operations of negation, conjunction, disjunction and implication, extended for the case of statements that are not numeric and linguistic truth values [3].

It should be noted that the logical operations of negation, conjunction and disjunction can be applied to terms which characterise the truth, and the value of a true statement. The result of the operation of the first type is an additional term which describes the truth, and the result of the operation of the second type is the truth value of compound statement. Logical implication is based on the compositional rule of inference in approximate form. Compositional rule of inference in general has the form

$$A\square$$

$$\frac{A \to B}{B'},$$

where A - the premise, B – result of the implications, $A \to B$ - a rule of inference that specifies a causal relation between premise and conclusion. $A\square$ is somewhat close to A and $B\square$ is close to B. A rule of inference in the general case is a binary fuzzy relation defined in $U \times V$, where U - the universal set, in which the parcel A is defined, and V - the universal set, in which the conclusion B is defined. Thus, the conclusion of the implication B can be written as

$$B' = A' \bullet (A \to B),$$

where «•» - convolution (compositional rule of fuzzy inference) [5].

5. Features of the application of the theory of fuzzy sets in evaluating the IS risks

Lets describe the existing two-factor model of risk assessment in terms of fuzzy sets with the help of standard NIST 800-30 [6]. To describe the probability of appearance we take a linguistic variable «P» with three terms: «high», «middle» and «low». To describe the value of damage caused by risk we take the linguistic variable «D» from the three terms: «high», «middle» and «low». To describe the level of risk we take the linguistic variable «RiskLevel» with three terms: «high», «middle», «low».

According to the standard NIST 800-30 risk level is defined as the smallest of the levels of damage and probability of risk [6]. Thus, we obtain the following set of rules [6]:

1) If «P» = «low» and «D» = «low», then «RiskLevel» = «low».
2) If «P» = «low» and «D» = «middle», then «RiskLevel» = «low».
3) If «P» = «low» and «D» = «high», then «RiskLevel» = «low».
4) If «P» = «middle» and «D» = «low», then «RiskLevel» = «low».
5) If «P» = «middle» and «D» = «middle», then «RiskLevel» = «middle».
6) If «P» = «middle» and «D» = «high», then «RiskLevel» = «middle».
7) If «P» = «high» and «D» = «low», then «RiskLevel» = «low».
8) If «P» = «high» and «D» = «middle», then «RiskLevel» = «middle».
9) If «P» = «high» and «D» = «high», then «RiskLevel» = «high».

Using the software environment FuzzyTECH 5.54d we construct a table output of the theory of fuzzy logic (Fig. 1).

IF		THEN	
P	U	DoS	RiskLevel
low	low	1.00	low
low	medium	1.00	low
low	high	1.00	low
medium	low	1.00	low
medium	medium	1.00	medium
medium	high	1.00	medium
high	low	1.00	low
high	medium	1.00	medium
high	high	1.00	high

Figure 1. Fuzzy inference rules that are specified in the table

To manually apply the above rules in practice, we do the following steps. First, we "fuzzycate" risk indicators that mean that we calculate the degree of belonging of each of them to each term of the corresponding term-set. We can do this in two ways. Either group of experts assesses specific risk indicators to certain level, that gives us a ready degree of belonging, Or we set model in terms of numerical indicators (i.e., for each term - the membership function), and the group of experts assess the exact value of the risk indicators. Next, we convert these estimates to the degree of belonging.

The second step considers each of the above rules. In this example, there are nine, but there may be more or less. For each rule, we have to do the following. First, assess the validity of each of equations. Truth value will be equal to the belonging function of the variable to a term listed right on the mark "equality". Then we need to evaluate the truth of the rule. It will be equal to the minimum value of the truth of each of the equations from the rule. The result of each rule is a fuzzy set, which repeats the term, specified as result, but the degree of belonging of this set is the truth of the rules in those cases where the former above the latter.

After a cycle we get the fuzzy sets in the number of fuzzy inference rules. The end result is a fuzzy set that is equal to the union of all obtained fuzzy sets.

To avoid doing all these procedures by ourselves, we can use the FuzzyTECH. FuzzyTECH, that is developed and constantly updated by INFORM GmbH (Inform Software Corporation, Germany), is designed to solve different problems of fuzzy modelling. Unlike the MATLAB, the program FuzzyTECH is a specialized tool that allows us to develop and explore a variety of fuzzy models in graphical mode, and convert them into code for a programming language with the possibility of implementing programmable microcontrollers. This program allows us to operate with linguistic variables and create them to product the rules. In interactive mode, you can observe the values of each of the variables, as well as the level of validity of each of the rules.

6. Using the cost calculation of the IS risk reduction

A model based on two risk factors, namely the probability of appearance and the amount of damage is incomplete. In some cases, you need to know the cost of risk reduction by reducing the approximate damage caused by risk. If you do not take into account this value, then more resources can be spent on eliminating the risk than expected loss from risk.

In ISO / IEC 17799-2000 [1] it is recommended for risk assessment to take into account the ratio of the cost of risk reduction to the approximate damage caused by the risk. If this value exceeds unity, then the risk is recommended for adoption. If this ratio is less than unity, then it makes sense to eliminate this risk.

Note that the final decision based not only on this ratio, but the magnitude of the risk level.

To add this variable to a risk assessment using fuzzy sets we need to enter the linguistic variable that indicates the cost of risk reduction. We call it «V», it has few terms: «high», «middle», «low». We also introduce an output variable «doing», which shows expediency of risk reduction. In this variable there are five terms, which are called «very_positive», «positive», «zero», «negative», «very_negative». Then we introduce the fuzzy inference rules for transforming the input variables «D» and «V» into the output variable «doing». Values of the terms for the «D» and «V» should be defined in the same way, that means, if some level of damage corresponds to the term «high», then the same value of the cost of risk reduction will correspond to the term «high». It is a prerequisite for comparing the cost of the risk reduction and the amount of damage.

We define the fuzzy inference rules as follows:

1) If «V» = «high» and «D» = «high», then «doing» = «zero».

2) If «V» = «high» and «D» = «middle», then «doing» = «negative».

3) If «V» = «high» and «D» = «low», then «doing» = «very_negative».
4) If «V» = «middle» and «D» = «high», then «doing» = «positive».
5) If «V» = «middle» and «D» = «middle», then «doing» = «zero».
6) If «V» = «middle» and «D» = «low», then «doing» = «negative».
7) If «V» = «low» and «D» = «high», then «doing» = «very_positive».
8) If «V» = «low» and «D» = «middle», then «doing» = «positive».
9) If «V» = «low» and «D» = «low», then «doing» = «zero».

IF		THEN	
U	Z	DoS	doing
low	low	1.00	zero
low	medium	1.00	negative
low	high	1.00	very_negative
medium	low	1.00	positive
medium	medium	1.00	zero
medium	high	1.00	negative
high	low	1.00	very_positive
high	medium	1.00	positive
high	high	1.00	zero

Figure 2. Fuzzy inference rules that are specified in the table

With FuzzyTECH we can get from the initial values of variables «V» and «D» the value of «doing». If the variable takes the value of «positive» and «very_positive», it is expedient to eliminate this risk, because risk reduction will cost less than would be lost in the implementation of risk. If the value of «doing» takes the value «negative» and «very_negative», it is inexpedient to reduce the level of risk. The presence of consoles «very» indicates an extreme value, when the variables of cost of risk reduction and the variable of possible harm take opposite values. In the case where the variable takes the value of «zero» the expediency of risk reduction is difficult to judge, since the cost of risk reduction and potential damage from it takes alinguistic value.

Thus, two variables, «V» and «D» characterise the level of risk and expediency of risk reduction. When you want to reduce the overall level of risk, first of all most dangerous risks are eliminated. After that, the less dangerous risks are eliminated, etc. It is noteworthy that two obtained variables are simple to understand and use.

7. Conclusions

As shown above, the use of fuzzy set theory is an effective way of evaluating risk when the risk indicators are uncertain, as well as during the selection from the group of risks several risks that meet certain criteria. Modalities for the functioning of informational system (funding, critical information, the possibility of threats, etc.) specify the actions for risk reduction (eliminate, the adoption of). The disadvantages of the use of fuzzy sets are the subjectivity of assessment of risk in fuzzy terms and the subjectivity of inference rules.

ЛИТЕРАТУРА

1. International Standard ISO/IEC 17799:2000 Code of Practice for Information Security Management

2. Черныш В.И. Методы оценивания информационных рисков компании / В.И.Черныш // Материалы XV Международного юбилейного молодёжного форума «Радиоэлектроника и молодежь в XXI веке»: Сб. тезисов, 18–20 апреля 2011 р., Т.5. - Харьков: ХНУРЭ. 2011. – С. 195.
3. Замула О.А. Аналіз міжнародних стандартів в галузі оцінювання ризиків інформаційної безпеки / О.А.Замула, В.І. Черниш // Системи обробки інформації . – Харків: ХУ ПС, 2011. – Вип.2(92). – С.53-56
4. Замула А.А. Оценивание рисков информационной безопасности в современных информационных системах / А.А. Замула, В.И. Черныш, К.И. Иванов // XIV Международная научно-практическая конференция «Безопасность информации в информационно – телекоммуникационных системах», тезисы докладов. – К.: ЧП «ЕКМО», НИЦ «ТЕЗИС» НТУУ «КПИ», 2011. - С. 31.
5. Zadeh L. A. Fuzzy sets. Information and Control. – 1965. – Vol. 8, № 3. – Pp. 338-353.
6. Замула А.А. Математические методы оценивания информационных рисков компании / А.А. Замула, В.И. Черныш, Ю.В. Землянко // Прикладна радіоелектроніка: наук.-техн. Журнал. – 2011. Том 9. №1.- С.123-127.
7. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечёткая логика и искусственные нейронные сети. Учеб. пособие. – М.: Издательство Физико-математической литературы, 2001. – С. 224.
8. Леоненков А. В. Нечёткое моделирование в среде MATLAB и FuzzyTech. – СПб.: БХВ - Петербург, 2005. – С. 739.
9. Сидоров А. О., Торшенко Ю. А., Павлютенков А. А., Осовецкий Л. Г. Разработка методики структурированной оценки риска // Научно-технический вестник СПбГУ ИТМО № 55. Системы: управление, моделирование, безопасность 2008 г., Санкт-Петербург, с. 108-110.
10. Сергей Петренко, Сергей Симонов «Методики и технологии управления информационными рисками» - IT Manager, 2004.
11. Замула А.А. Математические методы оценивания информационных рисков компании / А.А. Замула, В.И. Черныш, Ю.В. Землянко // Прикладна радіоелектроніка: наук.-техн. Журнал. – 2011. Том 9. №1.- С.123-127.

Надійшла – 31.08.2011.