Method of Information Security Risk Assessment Based on Improved Fuzzy Theory of Evidence

https://doi.org/10.3991/ijoe.v14i03.8422

Huang Xuepeng, Xu Wei(✉)
Hubei University of Police, P.R. China
42482852@qq.com

Abstract—A method based on improved fuzzy theory of evidence was presented to solve the problem that there exist all kinds of uncertainty in the process of information security risk assessment. The hierarchy model for the information systems risk assessment was established firstly, and then fuzzy sets were introduced into theory of evidence. The basic probability assignments were constructed using the membership function of fuzzy sets, and the basic probability assignments were determined. Moreover, weight coefficients were calculated using entropy weight and empirical factor, which combined the objective weights with the subjective ones, and improved the validity and reliability. An illustration example indicates that the method is feasible and effective, and provides reasonable data for constituting the risk control strategy of the information systems security.

Keywords—theory of evidence, fuzzy sets, entropy weight, information systems, risk assessment

1 Introduction

With the rapid development of computer networks and their applications, the Internet has gradually become the inevitable part of people's work and life. Modern society depends increasingly more on all kinds of information systems, with important information systems in all fields already being the key national infrastructures, including finance, electricity, telecommunications, government, business, etc., and their security is of great significance. For the security problem of information systems, performing thorough risk assessment for the systems in their life cycles has become the consensus [1]. Study on feasible security risk assessment methods for information systems has important theoretical and practical significance for exploiting the advantages of all kinds of information systems and ensuring the reliable and stable operation of modern society.

In the aspect of current study, researchers home and abroad have adopted fuzzy mathematics, neural networks, support vector machines (SVMs) and other methods [2-5] to establish security risk assessment models, and achieved many useful results. These results improved the theory of security risk assessment for information systems. However, the systematic assessment method has not been formed. Meanwhile,

for the determination of weight coefficients, traditional assignment methods have problems including difficulties in consistency check, differences of consistency understanding, and lack of scientific basis, etc [6]. As the extension of Bayesian inference, DS evidence theory can fuse evidences without knowing prior probability and conditional probability, which is in accordance with people's inference and decision process and has special advantages in dealing with uncertainties [7]. In this study, the information security risk assessment problem was solved by using an evidence theory based method. Focusing on the uncertain factors in assessment process, including assessment data lacking, incomplete knowledge, incomplete system modeling, inadequate risk identification, etc. the fuzzy relationship from index set to comment set was established, resulting in the basic support degree of each index toward comment set. The membership function of fuzzy set was used to construct the basic probability assignment (BPA) function in evidence theory, effectively solving the problem that the BPA function is difficult to determine. Moreover, combined weight assignment method integrating entropy weights and empirical factors was used for the determination of weight coefficients. The objective weights derived by the decision matrix were combined with the subjective expert weights, improving the scientificalness and effectiveness of weight determination.

2 Establishing index system of information security risk assessment

The risk assessment of information security is to use scientific methods and means to systematically analyze the threats faced by information and information systems, as well as the existing vulnerability; and to assess affected process the systems may suffer once security events happen [2]. In security risk assessment for information systems, the basis is to choose an appropriate index system, and objective assessment activities are impossible to be carried out without a scientific, feasible and credible index system.

The construction of index system is the hierarchical structure model based on the basic system features, which is established by analyzing system elements under the guidance of system goals. The basic hierarchy of it contains target layer, criterion layer and index layer. The factors in each layer have effect on the factors of the upper layer, and are affected by the factors of the lower layer. According to national standard ISO/IEC 15408 and GJB5095-2002, the security of information systems is divided into four aspects, namely physical environment security, network operation security, information security confidentiality, and security management. Delphi four round questionnaire method was used to establish the security risk assessment index system for information systems based on physical environment security, network operation security, information security confidentiality, and security management.

In the system, physical environment security is the basic premise of the safe operation of information systems. The continuity of information system services is guaranteed by the reliability of the information system environment, device security, and physical security. Safe operation is the necessary condition of successfully carrying

out all services of information systems, and the main function of it is to guarantee the reliability and stability of the safe operation of information systems. Information security confidentiality is the core of information system security, and it guarantees the safety of information systems and the information carried by these information systems by all kinds of technical means. Information system security depends not only on advanced technologies, but also on strict management. A safe and effective management system includes professional administrative organization, qualified management personnel, sound management system, complete technical facilities and strict password key management.

3 Mathematical model of fuzzy evidence theory

3.1 Basics of evidence theory

Evidence theory was based on the upper and lower probability, as well as its combination rules, proposed by A. P. Dempster in 1960s, and was established by G. Shafer in the publication "A mathematical theory of evidence" in 1976. Later, evidence theory gradually developed into a kind of important inference method dealing with uncertainty [8].

Let Θ be a finite discourse domain, usually called frame of discernment, which is composed of a set of mutually exclusive and exhaustive propositions to be studied. A subset of Θ, i.e., the element of 2^{Θ}, can be comprehended as a proposition.

Definition 1. Assume that Θ is a frame of discernment, and function $m: 2^{\Theta} \to [0,1]$ satisfies the following conditions:

$$m(\Phi) = 0 \tag{1}$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{2}$$

Then, $m(A)$ is called the BPA of A. $m(A)$ reflects the precise belief degree of proposition A, or the probability exactly assigned to proposition A. Condition (1) indicates that no belief is assigned to empty set, while condition(2) means that the sum of the BPA of all propositions equals to 1.

Definition 2. Assume that Θ is a frame of discernment, and function $m: 2^{\Theta} \to [0,1]$ is the BPA of Θ. The following function is defined.

$$\begin{aligned} &Bel: 2^{\Theta} \to [0,1] \\ &Bel(A) = \sum_{B \subseteq A} m(B) \ (\forall A \subset \Theta) \end{aligned} \tag{3}$$

Then function $Bel(A)$ is the belief function of Θ, which represents the sum of the possibility measure of all subsets of A, i.e., the total belief of A.

Definition 3. Assume that Θ is a frame of discernment, and function $Bel: 2^{\Theta} \to [0,1]$ is a belief function of discernment frame Θ. Define function:

$$\begin{aligned} &Pl: 2^{\Theta} \to [0,1] \\ &Pl(A) = 1 - Bel(\overline{A}) = \sum_{B \cap A \neq \Phi} m(B) \ (\forall A \subseteq \Theta) \end{aligned} \tag{4}$$

Then function $Pl(A)$ is called plausibility function, which represents the creditability of not denying A, and is the sum of BPA of all sets intersecting A. The credit interval of proposition A can be represented by $[Bel(A), Pl(A)]$.

Definition 4. Assume Bel_1 and Bel_2 are two belief functions of the same discernment frame Θ, and m_1 and m_2 are their corresponding BPA, respectively, with focal elements being $A_1, A_2, \cdots, A_k$ and $B_1, B_2, \cdots, B_n$, respectively. If

$$\sum_{A_i \cap B_j = \Phi} m_1(A_i) m_2(B_j) < 1 \tag{5}$$

then the following equation

$$m(A) = \begin{cases} 0 & A = \Phi \\ \dfrac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \Phi} m_1(A_i) m_2(B_j)} & \forall A \subseteq \Theta, A \neq \Phi \end{cases} \tag{6}$$

defines a function $m: 2^{\Theta} \to [0,1]$, and is called the direct sum of m_1 and m_2, denoted by $m_1 \oplus m_2$. Eq. (1) is called the evidential combination rule of m_1 and m_2, reflecting the combined supporting degree of the two pieces of evidence corresponding to m_1 and m_2.

3.2 Fuzzy evidence theory

The derivation of mass function has constantly been the research hotspot and difficulty of DS evidence theory. In the security risk assessment of information systems, the evidence in evidential space is usually described by fuzzy language. Therefore, in this study, evidence theory was extended to fuzzy sets, and the membership function of fuzzy sets was used to construct the BPA function in evidence theory, in order to

realize the effective integration of fuzzy set theory and evidence theory. The application process of fuzzy evidence theory is as follows.

Collection of evidence source data. The collection of evidence source data includes the establishment of assessment index system, the determination of assessment standard, and the determination of the basic support degree of each assessment index. The hierarchical structure model of the index system and its internal relation is consistent with the requirements of evidence combination rule [9].

In D-S evidence theory, assessment standard is usually composed of several assessment levels in ordinal scale. Generally, the following comment set is defined: $\Theta = \{h_1, h_2, \cdots, h_m\}$, where h_i represents a comment which is possibly given to an index. According to the fact that people's resolving power to the difference between the same attribute of different objects is within 7 ± 2 levels, the number of comment levels is set to 5, namely low, relatively low, ordinary, relatively high, and high.

The derivation of the basic support degree of assessment indices is to determine the membership of each index in the index system to comment set, in order to establish the fuzzy relation between the index set and assessment standard. When the basic support degree of indices is determined, in order to be more objective and reasonable, Delphi method and expert scoring can be used for determination, so that the impact of authoritative experts can be reduced and the results are more objective.

For index u_i, there exist v_{ij} pieces of h_j comments. Then the basic support degree of index u_i as a member of h_j is denoted as

$$\beta_{ij} = v_{ij} / \sum v_{ij} \tag{7}$$

Determination of index weights. In the determination of index weights, the weights derived by objective assignment are based on the real data in decision matrix. The data are refined and the weight assignment is more objective and effective. The limitation lies in that the weights are completely determined by data. Therefore, it is required that data should be universal and representative. The data quantity should be large, otherwise one-sidedness exists in the weight calculation, and conclusions contradictory to reality may be drawn.

Therefore, in the process of determining weights using the objective assignment method, fusing the prior knowledge of assessors should be emphasized. In other words, the objective index weights determined by real data and those determined by subjective prior knowledge should be integrated, finally generating the synthetic weights and realizing the unifying of subjective and objective.

$$w = a\ddot{o} + (1 - a)\tilde{a} \tag{8}$$

where a is called empirical factor with $0 \le a \le 1$, which represents the preference degree of assessors toward objective and subjective weights. The smaller a is, the more importance the assessors attach to experts' empirical knowledge; while the larger a is, the more importance the assessors attach to subjective weights.

The steps of weight determination for security risk assessment indices in information systems based on combined weight assignment method are as follows. (1) Invite experts to assess the affecting factors of information system security, and obtain the membership matrix of all factors. This will result in relative importance entropy, and the objective index weights can be derived by normalizing the entropy. (2) According to empirical knowledge, experts present the subjective weights of all indices in the index system. (3) Finally, the appropriate empirical factor *a* is selected according to the risk preference of decision maker, and the synthetic weights *w* of indices are calculated.

Determination of mass function. In D-S evidence theory, decision makers usually only have plausibility $\chi(0 \le \chi \le 1)$ for an evidence. Parameter χ is called preference coefficient, which means the trust degree of a decision maker to the basic support degree of indices, and it is usually set to $\chi = 0.9$.

Set the weights of each levels of indices as $w = \{w_1, w_2, \cdots, w_n\}$. The index with the maximum weight in the index set is key index, and others are non-key index. The basic support degree of key index u_k to assessment standard h_j is β_{kj}, then the trust degree of decision maker to key index is $m_{kj} = \chi\beta_{kj}$. For non-key index u_i in index set, the basic support degree of it to assessment standard h_j is β_{ij}, then the trust degree of decision maker to non-key index is:

$$m_{ij} = (w_i / w_k)\chi\beta_{ij} \tag{9}$$

where w_i is the weight of non-key index, and w_k is the weight of key index.

For an index at any level, the following equation holds.

$$m_i(h_j) = m_{ij}, \quad m_i(\Theta) = 1 - \sum_{j=1}^{m} m_{ij} \tag{10}$$

where m_{ij} represents the probability that the i–th index u_i supports the upper level index associated with comment h_j, and $m_i(\Theta)$ is unallocated probability which represents uncertain degree.

Therefore, the mass function of the indices at a certain level can be determined by m_{ij} of indices, and the mass function of the indices at a certain level forms a mass matrix M.

$$M(E_i) = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1m} & m_{1\Theta} \\ m_{21} & m_{22} & \cdots & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{r1} & m_{r2} & \cdots & m_{rm} & m_{r\Theta} \end{bmatrix} \tag{11}$$

Evidence combination and synthetic assessment. From the above, the mass function of indices at each level for the assessed object can be derived. The evidence is combined from top to bottom, resulting in the mass function of each first grade index. Combining them, we obtain the assessment matrix of first grade indices. Then, the key factors and non-key factors can be derived according to the weight of each index. According to the discount rate formula in (3), the discount rate and the assessment matrix of first grade indices can be used for calculating the mass matrix of assessed object. The mass function of assessed object can be obtained by combining evidence. Using the mass function of assessed object, the assessment result can be judged. Commonly used judgment criteria include maximum membership criterion and weighted average criterion [10]. Thus, the security risk of the assessed object can be determined.

4 Case analysis

We took an office automation information system as an example, and used the proposed method for security risk assessment. The steps are as follows.

Step 1. Establish the hierarchical structure model of security risk assessment for the office automation system.

In national standard ISO/IEC 15408 and "General requirements for information technology security", it is required that information security be divided into four aspects, namely physical security, operation security, information security and safe management. The hierarchical structure model of security risk assessment for information systems includes 4 first grade indices, namely physical environment security, network operation security, information security confidentiality and security management; and 22 second grade indices, namely environment security, device security, backup and recovery, etc. This assessment index system can reflect all aspects of the security of office automation system, without the need of cutting and expanding in the application of this case.

Step 2. Determine the assessment standard, and obtain the basic support degree of indices at each level according to the assessment standard.

In this case, the comments were classified into 5 grades, i.e., Θ = {low, relatively low, ordinary, relatively high, high}. Delphi method was used to determine the membership of each index to assessment standard, and the membership of indices at each level formed a membership matrix. Taking "network operation security" as an example, the basic support degree vector of "backup and recovery" is $a_{11} = \{0.10\ 0.30\ 0.40\ 0.15\ 0.05\}$. Similarly, the basic support degrees of other indices can be determined.

Step 3. Determine the weights of indices at each level.

The weight coefficients of each level in the index system were calculated by using weight assignment method combining entropy weight and empirical factors. Similarly, taking the indices of "network operation security", the objective weight vector of all indices is $\ddot{o} = \{0.121\ 0.189\ 0.218\ 0.131\ 0.152\ 0.189\}$. Meanwhile, expert opin-

ions were sought and the subjective weights of all indices were $\tilde{a} = \{0.25\ 0.1\ 0.15\ 0.15\ 0.15\ 0.25\}$. The empirical factor was set to $\acute{a} = 0.3$, which placing more emphasis on expert experience, resulting in synthetic weights $\grave{u} = \{0.211\ 0.1\ 27\ 0.170\ 0.144\ 0.151\ 0.253\}$. Similarly, the weights of the indices at other levels can be determined.

Step 4. Determine the mass matrix of indices at all levels.

Taking "network operation security" as an example again, according to (3) and (4), the mass matrix M_1 is as follows.

$$\begin{bmatrix} 0.050 & 0.151 & 0.201 & 0.075 & 0.025 & 0.498 \\ 0.035 & 0.175 & 0.210 & 0.210 & 0.070 & 0.300 \\ 0.042 & 0.253 & 0.168 & 0.253 & 0.126 & 0.158 \\ 0.135 & 0.360 & 0.180 & 0.180 & 0.045 & 0.100 \\ 0.027 & 0.136 & 0.163 & 0.163 & 0.054 & 0.457 \\ 0.066 & 0.264 & 0.264 & 0.033 & 0.033 & 0.340 \end{bmatrix}$$

Similarly, the mass matrix of indices at other levels can be derived.

Step 5. Determine the evidence combination and evaluation results.

According to evidence combination equation (1), the evidence of all mass matrices was combined, generating the mass function of the security risk level of the information system, i.e., $m(\theta_1) = 0.032$, $m(\theta_2) = 0.424$, $m(\theta_3) = 0.265$, $m(\theta_4) = 0.085$, $m(\theta_5) = 0.068$, $m(\Theta) = 0.126$. According to maximum membership criterion, the security risk level of this office automation system was relatively low, which accorded with the security status of the real operation of this system.

5 Conclusions

According to the security risk assessment problem for information systems, the hierarchical analysis model of security risk assessment of information systems was established, and a security risk assessment evaluation method for information systems was proposed based on improved fuzzy evidence theory. In this fuzzy evidence theory method, the fuzzy relationship between index set to comment set was established, resulting in the basic support degree of each index to comment set. By combining preference coefficient, the uncertainty in assessment process can be synthetically processed. The combined weight assignment method based on entropy weight and empirical factor integrates the objective weights derived by the data in decision matrices and the objective weights of experts, overcoming the one-sidedness of individual weight assignment methods, and improving the scientificalness of assessment results. This method can provide useful reference for the security risk assessment of information systems.

6 Acknowledgment

It is a project supported by Social Security Administration of Hubei Province Social Science key research bases(2017-19), General scientific research project of Hubei University of Police(2017).

7 References

[1] MA Chun-guang, WANG Cheng-hong, ZHANG Dong-hong, et al. A Dynamic Network Risk Assessment Model Based on Attacker's Inclination[J]. Journal of Computer Research and Development, 2015, 52(9): 2056-2068. (in Chinese)
[2] FU Yu, WU Xiao-ping, WANG Jia-sheng. An approach to information systems security risk assessment based on fuzzy-combinatorial neural network[J]. Journal of Naval University of Engineering, 2010, 22(1): 18-23. (in Chinese)
[3] LIU Jian, ZHAO Gang, ZHENG Yun-peng. Information security risk variety situation analysis model based on AHP and Bayesian network[J]. Journal of Beijing Information Science & Technology University, 2015, 30(3): 68-74. (in Chinese)
[4] MA Gang, DU Yu-ge, AN Bo, et al. Risk Evaluation of Complex Information System Based on Threat Propagation Sampling[J]. Journal of Computer Research and Development, 2015, 52(7): 1642-1659. (in Chinese)
[5] DANG De-peng, MENG Zhen. Assessment of information security risk by support vector machine[J]. Journal of Huazhong University of Science and Technology(Nature Science Edition), 2010, 38(3): 46-49. (in Chinese)
[6] WANG Jia-sheng, FU YU, WU Xiao-ping. Research on Security Risk Assessment of Information System Based on Improved Fuzzy AHP[J]. Fire Control & Command Control, 2011, 36(4): 33-36. (in Chinese)
[7] YE Qing, WU Xiao-ping, LI Mo-ci, et al. Abnormal evidence detection algorithm based on projection decomposition and k nearest neighbors distance[J]. Journal of Naval University of Engineering, 2015, 27(3): 9-13. (in Chinese)
[8] YE Qing, Dempster-Shafer evidence theory and its application in information fusion [D]. Journal of Naval University of Engineering,,2008.
[9] LIU Jian, ZHAO Gang, ZHENG Yun-peng. Information security risk variety situation analysis model based on AHP and Bayesian network[J]. Journal of Beijing Information Science & Technology University, 2015, 30(3): 68-74. (in Chinese)
[10] CHEN Yun-xiang, CAI Zhong-yi, ZHANG Zheng-min, et al. Method for group decision-making information integration based on evidence theory and intuitionistic fuzzy set[J]. Systems Engineering and Electronics, 2015, 37(3): 594-598. (in Chinese)

8 Authors

Xue-peng Huang, Male, born in Hanchuan City, Hubei Province, Lecturer, Research interest: Computer application and information security. He is with the Information & Network Center, Hubei University of Police, P.R. China.

Wei Xu, Male, born in Hanchuan City, Hubei Province, Lecturer at the Department of Information Technology, Hubei University of Police, P.R. China, Research interest: Computer application and information security. Address: No. 86 Jiefang Avenue, Hubei, Hubei University of Police, Wuhan, China ,Postcode: 430034.

Article submitted 11 February 2018. Final acceptance 05 March 2018. Final version published as submitted by the authors.