Journal of Software Engineering

ISSN 1819-4311

Academic Journals Inc.

www.academicjournals.com

Journal of Software Engineering 9 (4): 818-827, 2015
ISSN 1819-4311 / DOI: 10.3923/jse.2015.818.827

Study on Security Risk Assessment for Information System Based on Fuzzy Set and Entropy Theory

Sha Fu, Zhongli Liu, Guang Sun, Hangjun Zhou and Wenbin Liu
Department of Information Management, Hunan University of Finance and Economics, 410205, People's Republic of China

Corresponding Author: Sha Fu, Department of Information Management, Hunan University of Finance and Economics, 410205, People's Republic of China

ABSTRACT

In allusion to accuracy of security risk assessment for information systems in campus, put forth a method to evaluate security risks of the information system combining fuzzy set with the entropy theory. By analyzing risk factors involved in the information system with fuzzy set, this method constituted the matrix on dependence degrees of evaluation sets corresponding to various factors and then determined weights of risk factors with the entropy coefficient method to reduce subjective bias. During the evaluating process, security risk values of different factors have been integrated with the system integration method, so that levels on security risks of information system in campus can be determined. Case analysis proved that this method can be effectively applied in security risk assessment for information system in campus.

Key words: Fuzzy set, entropy weight, information system, risk assessment, database server

INTRODUCTION

In today's information age when cyber attacks are increasingly expanding the scale, security risks confronted to information system in campus are getting serious day by day, whose security concerns fundamental interests of the country and information systems users. As a result of its special nature, campus has a lot of important and confidential information resources. Unfortunately, defense capability of the campus on information security is relatively weak compared to other industries. In order to protect their operation in safe and normal states, we must find out the serious flaws that could lead to standstill and one of the effective ways to solve this problem is to conduct security risk assessment for information systems of the campus. In view of literatures at home and abroad, a number of scholars have established evaluation models for security risks of the information system by applying fuzzy mathematics theory (Fu *et al*., 2010) grey system theory (Zhang *et al*., 2012), rough set theory, neural network, Bayesian network (Fu *et al*., 2006) and others, who have made some achievements in evaluating process and method (Dzazali and Zolait, 2012). Gehani (2003) proposed a set of host-based models for risk management and decision-making in his doctoral thesis. Basic idea on decision-making methods relating to information security risks proposed by Bilar (2003) in his doctoral thesis goes as: A variety of harmful consequences are owing to vulnerability of software in the network host, if overall risk of those software exceeds the threshold value set by the system administrator, security decision-making shall be made, i.e., replace software to reduce its overall risk. Zhao *et al*. (2007) proposed the model of fuzzy risk assessment for information system and got the final comprehensive

evaluation with the weight sum. Although, the method is simple, there are still relatively great subjective preferences on systematic characteristics and accuracy. Zhang *et al.* (2010) have evaluated the overall risk by constructing risk matrix, but this method is not applicable for analysis on models in multi-level structure with many risk factors.

Evaluation on security risk is an important part of safety engineering for information system, acting as the basis and prerequisite for building up the security system. Security risk of information system refers to the combination of potential possibility of impact and potential impact onto asset or asset group in the system produced by threats which make use of existing vulnerability (Liu *et al.*, 2010). Under evaluation method proposed by this study for security risks of the information system, it was analyzed from the perspective of system engineering, constituted matrix about dependence degrees of evaluation sets with fuzzy sets on the basis of overall analysis for various risk factors involved in information system, determined weight vectors of those factors with the entropy coefficient method (Ahmed *et al.*, 2012) and thus obtained the security risk value for the system. This method is simple and practical with strong sense of pertinence, providing comprehensive evaluation on security risks of information system with a new thought in scientific and feasible manners, which serves as the directional guidance for designing the supportive system to evaluate security risks of the information system.

PRELIMINARIES

Security risk assessment for information system: According to BS 7799 standard (BS 7799-2, 1999), risk is defined as the possibility of loss or damage caused by main threat taking advantage of vulnerability of the assets. Security risk of the information system-R is seen as a function among asset, threat and vulnerability, namely:

$$R = g(a, t, v)$$

where, A stands for the impact onto assets; t refers to frequency of the threat towards system, v is the severity of vulnerability (Zhang *et al.*, 2008). By analyzing the concept and physical logic about security of the information system, the relational graph relating to basic elements of risk assessment was obtained as shown in Fig. 1. GB/T 20984-2007 standard defined asset impact, threat frequency and severity of vulnerability all in five grades, specifically stated as: Very high, high, medium, low, very low.

Fuzzy mathematics and fuzzy set: In practice, fuzzy mathematics is extensively and successfully applied in different areas of the national economy, especially in science and technology,

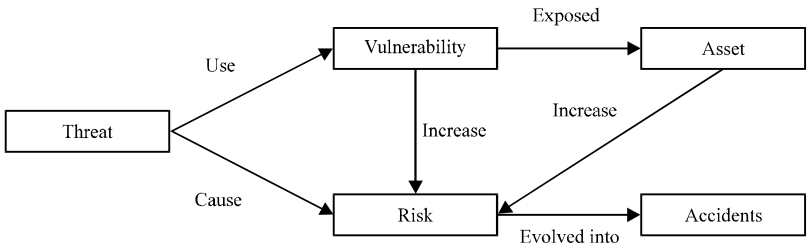

Fig. 1: Relation among basic elements of risk assessment

economic management and social sciences. Fuzzy mathematics has been founded by L.A. Zadeh in 1921 an American cyberneticist. In 1965, he published the paper entitled as "Fuzzy Sets", which declared the birth of fuzzy mathematics.

Definition: Set U as the discourse domain, called as mapping:

$$\mu_A: U \rightarrow \{0, 1\}, x \mid \rightarrow X_A(x) \in [0, 1]$$

Which determines the fuzzy subset A on U. Mapping μ_A is called as A's membership function, $\mu_A(x)$ is called as x's degree of membership corresponding to A. The above definition showed that fuzzy subset A is solely determined by the membership function μ_A. Afterwards, fuzzy subset A is always considered as equal to membership function μ_A. The collection of fuzzy subsets on U is called as U's fuzzy power set and recorded as P(A). For simplicity, $\mu_A(x)$ can be replaced by A(x). Fuzzy subset is simply called as fuzzy set and degree of membership is called as membership degree. Membership degree and membership function are basic ideas of fuzzy mathematics.

Information entropy theory: The concept of entropy was first proposed in 1865 by the German physicist Rudolf Clausius. In 1948, father of the information theory-Claude Elwood Shannon introduced the concept of entropy into the field of information, which used "Information entropy" as a measure to estimate disorder degree of the information (Liang and Qian, 2008). Shannon systematically presented the measuring method:

$$H(X) = H(p_1, p_2, \ldots, p_n) = -k\sum_{i=1}^{n} P_i \log P_i$$

for information going as used entropy to measure uncertainty or information quantity of a random event with probability statistics, which laid the scientific theoretical basis for modern information theory and extended the quantified application of entropy into studies on uncertainty and random quantification of the system (Ding *et al*., 2010).

In different systems, entropy can be used to measure state confusion or disorder, uncertainty or lack of information, inhomogeneities or richness and etc. Therefore, appraising order degree of the system architecture with extensive application of the theory on information entropy became a new approach for today's system architecture evaluation.

METHOD ON SECURITY RISK ASSESSMENT FOR INFORMATION SYSTEM BASED ON FUZZY SET AND ENTROPY THEORY

To determine entropy coefficients of risk factors: Assume that the system may be in n different states: $\{S_1, S_2, \ldots, S_n\}$, P_i presents the probability of the system in the state of S_i, where:

$$i = 1, 2, \ldots, n; 0 \le P_i \le 1, \sum_{i=1}^{n} P_i = 1$$

then entropy of the system X is:

$$H = -\sum_{i=1}^{n} P_i \ln P_i \quad (1)$$

Extreme values of the entropy showed that entropy value will get larger as P_i closer to equal, where uncertainty of the security risk factors for evaluation on system risks will be greater. Thus, weight of indicators with information entropy according to supportive degree P_{ij} of various security risk factors for indicators in the judgment set can be calculated. Relative importance of the risk factor A_i can be measured by entropy:

$$H_i = -\sum_{j=1}^{m} p_{ij} \ln p_{ij} \tag{2}$$

Entropy has the extreme value. When the system state is in equal probability, i.e., $P_i = 1/n$, entropy reaches to the maximum value: $H_{max} = \ln m$. Conducting normalization for Eq. 2 with H_{max}, we can get the entropy value of relative importance for measuring the security risk factor A_i:

$$e_i = -\frac{1}{\ln m} \sum_{j=1}^{m} p_{ij} \ln p_{ij} \tag{3}$$

When p_{ij} (j = 1, 2,... , m) are at equal values, entropy e_i reaches to the maximum 1. Namely, as $0 \le e_i \le 1$ when entropy reaches to the maximum, we can use $(1-e_i)$ to measure weight of the security risk factor A_i. After the normalization, we can obtain weight ϕ_i of the risk factor A_i as:

$$\phi_i = \frac{1}{n - E}(1 - e_i) \tag{4}$$

Where:

$$E = \sum_{i=1}^{n} e_i$$

ϕ_i met with:

$$0 \le \phi_i \le 1, \sum_{i=1}^{n} \phi_i = 1$$

Fuzzy set and membership matrix: Among factors involved in security risk assessment for information system, those with more obvious fuzzy features include asset impact, threat frequency and severity of vulnerability. An analysis can be conducted by applying fuzzy set, where steps for constructing security risk factor set and judgment set are as follows: (1) Constitute the security risk factor set, set $U = \{u_1, u_2, \ldots, u_n\}$, where n is the number of factors. (2) Constitute the judgment set. Relatively with different criteria on the previous level, experts make comments on risk degree to each risk factor and divide them into m levels to measure performance on each indicator and relevant risk degrees therefrom. Different judgment sets can be established for different criteria. For example, set judgment sets of $V = \{v_1, v_2, \ldots, v_m\}$ for asset, threat and vulnerability, where m is the number of elements in the judgment set. (3) Invite experts to evaluate factors in the risk factor set U by referring to judgment set V and constitute fuzzy mapping f: $U \rightarrow F(V)$, where F(V) refers to all those in the fuzzy set on V, $u_i \mid \rightarrow f(u_i) = (p_{i1}, p_{i2}, \ldots, p_{im}) \in F(V)$, fuzzy mapping f indicates

supportive degree of risk factor u_i for comments in the judgment set. f can induce a fuzzy relation $R_f \in F(U \times V)$, namely, $R_f(u_i, v_j) = f(u_i)(v_j) = p_{ij}$, which can obtain the membership matrix P [10] as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{pmatrix}$$

Different membership matrix will be made for asset impact, threat frequency and severity of vulnerability according to impacts onto assets imposed by factors relating to security risks of the information system, respectively as P_a, P_t and P_v, where corresponding weight vectors is $\phi = (\phi_1, \phi_2, \ldots, \phi_n)$. As calculating the asset impact, endow indicators in the judgment set with appropriate weight to obtain the indicator weight vector $U = (u_1, u_2, \ldots, u_{n1})$, where n_1 is the number of elements in the judgment set of asset impact going as:

$$R_a = \phi \bullet P_a \bullet U'$$

Similarly, we obtain the indicator weight vector of judgment set for threat frequency going as $V = (v_1, v_2, \ldots, v_{n2})$, where n_2 is the number of elements in the judgment set of threat frequency and the threat frequency is $R_i = \phi \bullet P_a \bullet U'$; Then, we obtain the indicator weight vector of judgment set for severity of vulnerability going as $W = (w_1, w_2, \ldots, w_{n3})$, where n_3 is the number of elements in the judgment set of vulnerability severity and vulnerability severity is $R_v = \phi \bullet P_v \bullet U'$.

Security risk rating: We can obtain values of asset impact, threat frequency and severity degree (R_a, R_t, R_v) in the way of calculating described above. By integrating security risk value of each element with system integration method, we can obtain system risk values as follows:

$$R = g(a,t,v) = R_a \oplus R_t \oplus R_v = k_1 R_a + k_2 R_t + k_3 R_v$$

where, k_1, k_2 and k_3 represent the relative importance of the three elements and $k_1+k_2+k_3 = 1$. Determining risk rating with Table 1, resolving action priority in accordance with risk rating and developing appropriate measures, we can ultimately make a reasonable and feasible plan for risk disposal.

RESULTS AND DISCUSSION

Taking database server of the information system in a university at the central region as an example, calculates security risks in comprehensive manners with the evaluation model we have

Table 1: Risk level description and quantitative expression

Risk level	Quantitative expression of risk	Risk level	Measure to take
1	$0.0<\phi\leq0.2$	Low risk	Accept risk
2	$0.2<\phi\leq0.4$	Plain risk	Notice and prevention
3	$0.4<\phi\leq0.6$	Medium risk	Enhance prevention and control
4	$0.6<\phi\leq0.8$	High risk	Inform related department of purposeful control
5	$0.8<\phi\leq1.0$	Very high risk	Take control measure to reduce risk or change system

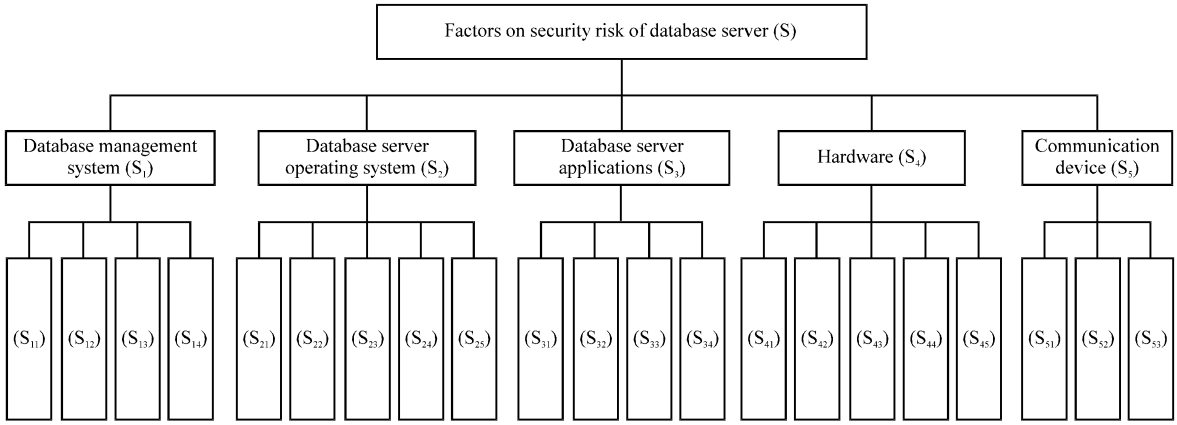

Fig. 2: Factors on security risk of database server

built. Risk factors S influencing security of the database server include $S = \{S_1, S_2, S_3, S_4, S_5\}$ = {database management system, database server operating system, database server applications, hardware, communication device}, as shown in Fig. 2.

These five groups of risk factors are composed by the following risk factors (Gong *et al.*, 2011):

- $S_1 = \{S_{11}, S_{12}, S_{13}, S_{14}\}$ = {data file destroyed, errors occur during query or restoration, errors occur during data modification, deletion error}
- $S_2 = \{S_{21}, S_{22}, S_{23}, S_{24}, S_{25}\}$ = {buffer overflow, register destruction, file and directory structure damaged, rewriting of user account and priority, conflict among different applications}
- $S_3 = \{S_{31}, S_{32}, S_{33}, S_{34}\}$ = {function error, can not access the required resources, illegal data input, the conflict with the operating system version}
- $S_4 = \{S_{41}, S_{42}, S_{43}, S_{44}, S_{45}\}$ = {memory failure, CPU failure, hard drive failure, power failure, bus malfunction}
- $S_5 = \{S_{51}, S_{52}, S_{53}\}$ = {network hardware or interface failure, communication protocol failure, router failure}

According to the expertise and experience of experts and aiming at five groups of risk factors mentioned above, respectively determine weight coefficient for indicators relating to security risks confronted with database server and calculate the value on comprehensive security risk.

Step 1: Constitute factor set and judgment set on security risk: Taking "hardware (S_4)" as an example for analysis, the security risk factor set $U = \{u_1, u_2, u_3, u_4, u_5\}$, where u_i (i = 1, 2, 3, 4, 5), respectively representing judgment sets U for risk factor sets of ""memory failure", "CPU failure," "hard drive failure," "power failure" and "bus malfunction" including $V_a = \{v_{a1}, v_{a2}, v_{a3}, v_{a4}, v_{a5}\}$, $V_t = \{v_{t1}, v_{t2}, v_{t3}, v_{t4}, v_{t5}\}$, $V_v = \{v_{v1}, v_{v2}, v_{v3}, v_{v4}, v_{v5}\}$.

Invite a number of experts to rate the risk factor set U. Calculate the probability of risk factors as member of various indicators and then we can obtain the membership matrix P_a, P_t and P_v as shown in Table 2, 3 and 4.

Step 2: Calculate entropy coefficient for risk factors: Referring to Table 2, use Eq. 3 to calculate the e_i vector for "asset" membership matrix.

Table 2: "Asset" membership matrix P_a

Parameters	V_{a1}	V_{a2}	V_{a3}	V_{a4}	V_{a5}
u_1	0.4	0.3	0.2	0.1	0.0
u_2	0.0	0.4	0.4	0.1	0.1
u_3	0.1	0.3	0.4	0.2	0.0
u_4	0.3	0.5	0.2	0.0	0.0
u_5	0.1	0.2	0.5	0.1	0.1

Table 3: "Threat" membership matrix P_t

Parameters	V_{t1}	V_{t2}	V_{t3}	V_{t4}	V_{t5}
u_1	0.1	0.3	0.4	0.1	0.1
u_2	0.1	0.5	0.1	0.2	0.1
u_3	0.3	0.4	0.1	0.2	0.0
u_4	0.2	0.3	0.3	0.1	0.1
u_5	0.5	0.2	0.2	0.1	0.0

Table 4: "Vulnerability" membership matrix P_v

Parameters	V_{v1}	V_{v2}	V_{v3}	V_{v4}	V_{v5}
u_1	0.4	0.4	0.1	0.1	0.0
u_2	0.5	0.3	0.2	0.0	0.0
u_3	0.2	0.2	0.3	0.2	0.1
u_4	0.1	0.2	0.3	0.3	0.1
u_5	0.5	0.2	0.2	0.1	0.0

$$e_1 = -\frac{1}{\ln m}\sum_{j=1}^{m} r_{1j} \ln r_{1j}$$
$$= -\frac{1}{1.6094}\left[0.4\times(-0.916) + 0.3\times(-1.204) + 0.2\times(-1.609) + 0.1\times(-2.303) + 0\right]$$
$$= 0.7952$$

Similarly, to obtain e_2, e_3, ..., e_5. Thus, e_i vector goes as (0.7952, 0.7416, 0.7952, 0.6398, 0.8445). And then, calculate weight vector ϕ_i of each risk factor according to Eq. 4:

$$E = \sum_{i=1}^{n} e_i = 3.8163$$

$$\phi_1 = \frac{1}{n-E}(1-e_1) = \frac{1}{5-3.8163}\times(1-0.7952) = 0.1730$$

Similarly, to obtain ϕ_2, ϕ_3, ..., ϕ_5, then the weight vector ϕ_i goes as (0.1730, 0.2183, 0.1730, 0.3043, 0.1313).

Table 3, use Eq. 3 to calculate the e_i vector for "threat" membership matrix going as (0.8814, 0.8445, 0.7952, 0.9350, 0.7584). Calculate weight vector ϕ_i of each risk factor with Eq. 4 going as (0.1510, 0.1979, 0.2607, 0.0828, 0.3076).

Table 4, use Eq. 3 to calculate the e_i vector for "vulnerability" membership matrix going as (0.7416, 0.6398, 0.9675, 0.9350, 0.7584). Calculate weight vector ϕ_i of each risk factor with Eq. 4 going as (0.2698, 0.3761, 0.0339, 0.0679, 0.2522).

To sum up:

$$\phi_a = (0.1909, 0.3696, 0.3177, 0.0869, 0.0350)$$

$$\phi_t = (0.2834, 0.3349, 0.1926, 0.1459, 0.0432)$$

$$\phi_v = (0.4357, 0.2916, 0.1832, 0.0794, 0.0102)$$

Step 3: Determine weight of indicators in the evaluation set: On the "asset" membership matrix, determine weight of standard w_1, w_2, ..., w_5, respectively as 1/15, 2/15, 3/15, 4/15, 5/15. Thus, weight of indicators in this evaluation set goes as:

$$W_a = (1/15,\ 2/15, 3/15, 4/15,\ 5/15)$$

Similarly, we can get:

$$W_t = (1/15, 2/15, 3/15,\ 4/15,\ 5/15)$$

$$W_v = (1/15, 2/15, 3/15, 4/15, 5/15)$$

Step 4: Calculate security risk values for "hardware (S_4)" and other groups: To obtain security risk values for elements of "hardware (S_4)" by calculating:

$$R_a = \phi_a \bullet P_a \bullet W_a^T = (0.1909, 0.3696, 0.3177, 0.0869, 0.0350) \bullet \begin{bmatrix} 1/15 \\ 2/15 \\ 3/15 \\ 4/15 \\ 5/15 \end{bmatrix} = 0.1604$$

Similarly, we can get:

$$R_t = \phi_t \bullet P_t \bullet W_t^T = 0.1554,\ R_v = \phi_v \bullet P_v \bullet W_v^T = 0.1291$$

As $k_1 = k_2 = k_3 = 1/3$, then:

$$R_4 = k_1 R_a + k_2 R_t + k_3 R_v = 0.1483$$

Similarly, we can get security risk values of "database management system (S_1)", "database server operating system (S_2)", "database server application (S_3)" and "communication device (S_5)", respectively going as $R_1 = 0.2317$, $R_2 = 0.1618$, $R_3 = 0.3154$, $R_5 = 0.1967$.

Step 5: Calculate values of the database server on comprehensive security risk: Under the thought of comprehensive evaluation for the system, giving full consideration to relative

importance of each module in the database server, calculate security risk values of the database server with weighed average method. Assume each module has same extent of significance to the entire database server, that is:

$$g_1 = g_2 = g_3 = g_4 = g_5 = 1/5$$

Then, the comprehensive security risk value goes as:

$$R = g_1R_1 + g_2R_2 + g_3R_3 + g_4R_4 + g_5R_5 = 0.2108$$

Table 1 showed that security risk of this database server was at a low risk level and the system was relatively safe and reliable, it consistent with their security situation in actual operation. The case verified the proposed method can better make up the lack of certainty in indicators, it will help further improve the scientific of decision-making. Compared with the risk assessment values obtained in the references, the proposed method can better meet the needs of practical applications and have good maneuverability.

CONCLUSION

Aiming at security risk assessment for information system in campus and combining risk factors under fuzzy set involved in database server, this study conducted analysis from asset impact, threat frequency and severity of vulnerability and made corresponding descriptions on rating, constituted membership matrix of each factor matched along with judgment set and determined weight of factors with entropy coefficient to reduce subjective bias resulted from traditional method for weight determination. By calculating risks of a database server in a campus information system, we know that method proposed by this study has strong sense of pertinence, feasibility and effectiveness, which provides an effective approach to realize intellectual evaluation for security risks of the information system in campus.

ACKNOWLEDGMENTS

This study was supported by the Scientific Research Fund of Hunan Provincial Education Department (No. 14C0184) and also supported by the Hunan Province Philosophy and Social Science Foundation (No. 14YBA065). Supported by the construct program of the key discipline in Hunan province.

REFERENCES

Ahmed, M., L. Sharif, A. Issa-Salwe and A. Alharby, 2012. Information security: Securing a network device with passwords to protect information. Trends Inform. Manage., 6: 62-67.

BS 7799-2, 1999. Information security management: Specification for information security management systems. British Standard Institute (BSI), UK., Revised May 15, 1999. http://shop.bsigroup.com/ProductDetail/?pid=000000000030039679.

Bilar, D., 2003. Quantitative risk analysis of computer networks. Ph.D. Thesis, Dartmouth College, Hanover, NH., USA.

Ding, W.P., J.D. Wang, Z.J. Guan and H. Zhu, 2010. Algorithm of attribute reduction based on extension entropy of variable precision thresholding in incomplete information system. J. Chin. Comput. Syst., 31: 2372-2376.

Dzazali, S. and A.H. Zolait, 2012. Assessment of information security maturity: an exploration study of Malaysian public service organizations. J. Syst. Inform. Technol., 14: 23-57.

Fu, Y., X.P. Wu and C. Yan, 2006. The method of information security risk assessment using Bayesian networks. Wuhan Univ. (Nat. Sci. Edn.), 52: 631-634.

Fu, Y., X.P. Wu, Q. Ye and X. Peng, 2010. An approach for information systems security risk assessment on fuzzy set and entropy-weight. Acta Electronica Sinica, 38: 1489-1494.

Gehani, A., 2003. Support for automated passive host-based intrusion response. Ph.D. Thesis, Duke University, Durham, NC., USA.

Gong, J., J. Zhang, X. Wu and S. Liu, 2011. Applying information system security risk assessment in campus network. Jisuanji Yingyong yu Ruanjian, 28: 285-288.

Liang, J.Y. and Y.H. Qian, 2008. Information granule and entropy theory in information system. Sci. China (Ser. E: Inform. Sci.), 38: 2048-2065.

Liu, Y., Q. Lin and K. Meng, 2010. Research on quantitive security risk assessment method of an enterprise information system based on information entropy. Comput. Sci., 37: 45-48,56.

Zhang, L., J. Peng, Y. Du and Q. Wang, 2012. Information security risk assessment survey. J. Tsinghua Univ. Sci. Technol., 52: 1364-1369.

Zhang, T., D.J. Mu, S. Ren and L. Yao, 2010. Risk assessment model of information security based on risk matrix. Comput. Eng. Applic., 5: 93-95.

Zhang, W.H., B.C. Du and Y.Y. Yang, 2008. Application of fuzzy analytic hierarchy process to TV and radio information assurance evaluation indicator systems. Acta Electronica Sinica, 10: 2060-2064.

Zhao, D.M., J.F. Ma and Y.S. Wang, 2007. Model of fuzzy risk assessment of the information system. J. Commun., 28: 51-64.