ICEB 2015 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-6-2015

An Assessment Model For Information System's Risk Based On Entropy Method And Grey Theory

Jinli Duan

Qishan Zhang

AN ASSESSMENT MODEL FOR INFORMATION SYSTEM'S RISK BASED ON ENTROPY METHOD AND GREY THEORY

Jinli Duan, School of pharmacy, Fujian University of Chinese Traditional Medicine, School of Economics and Management, Fuzhou University, Fuzhou City, Fujian Province, China, 78308776@qq.com
Qishan Zhang， School of Economics and Management, Fuzhou University, Fuzhou City, Fujian Province, China, zhangqs@fzu.edu.cn

ABSTRACT

In the process of risk assessment of information system, the risk assessment method and model are the key point. This paper analyzes the risk assessment methods of the information system, and points out the limitations of some methods. Considering the grey and dynamic characteristics of the evaluation index, the evaluation model based on Entropy Method and Grey Theory is presented, and the validity of the method is demonstrated by an example.

Keyword: Information system's risk; grey class; whitenization weight function ; risk level

INTRODUCTION

Methods for risk assessment of information system play an important role. From the analysis of the assets. threats,. vulnerabilities and security measures we can use some methods to evaluate the risk level of the information system and a scientific and effective risk assessment report is generated [3].

There are many methods of risk assessment, which can be divided into three categories: quantitative risk assessment method, qualitative risk assessment method, qualitative and quantitative assessment method. The qualitative method mainly focus on knowledge, experience, lessons of history and policy direction and special cases such as non- quantitative data [6]. It is based on in-depth interviews with the respondents to make a case record, and then through a theoretical derivation of the analytical framework of the interpretation of the data for encoding, at last the conclusion of the investigation about the level of information system risk is generated [2]. The typical method of qualitative analysis with factor analysis method, logical analysis, historical comparative method, Delphi method and so on [4]. Quantitative assessment is to assess the risk of information system using the quantitative indicators. The advantage of quantitative assessment method is to describe the results of the assessment using intuitive data, and the result is more clear and objective. Sometimes the quantitative analysis methods can make the research result more scientific, more rigorous and more profound [9]. For example a data is able to explain the problem that can not be explained clearly with a large section of the text. But it is very difficult to quantify risk assessment of information system. Firstly it is difficult to choose the grain size to quantitative assessment. Secondly attack source is widely distribute with all sorts of motives. It is impossible to predict the probability and frequency of security incidents. Information system is constantly updated and improved, the risk factors have continued to change, the previously acquired data is not adapted to the new situation, the data need to continue to change.

sometimes the complex things become too simple and fuzzy in order to quantify, and even too simple to represent things in themselves. Some quantified risk factors may be misunderstood and distorted. The advantage of qualitative assessment method is to avoid the shortcomings of the quantitative method, and it can dig out some deep thinking, which makes the evaluation more comprehensive and operational, but it is very subjective, so it needs the evaluators with highly professional knowledge and rich experience in the field of information system evaluation [7].

In the early days when research on information system risk assessment method, has just been put out, domestic and foreign scholars mainly study for risk assessment of information system from the qualitative point of view ,and mainly focuses some non system risk, such as personnel risk, management risk, environmental risk and so on.. . Then some scholars introduced the quantitative model to the information system risk assessment. At this time, the focus of the study is the systematic risk of the information system. There are two types of models. one is statistical model, such as mathematical programming. Bias model, clustering analysis and so on [8]. The greatest strength of the statistical model is that it has a clear explanation. But it has obvious defects which is too strict prerequisite. Attack source is widely distributed with all sorts of motives. It is impossible to identify and calculate the probability distribution of security incidents. The loss and potential impact of security incident is difficult to accurately estimate. In general statistical model needs large sample data. But the data of information system security is difficult to achieve. The other is artificial intelligence model, such as neural network, expert system, classification tree, etc.. The arrangement of every input weight is very important using the neural network technology. Information system is constantly updated and improved, the risk factors have continued to change, the previously acquired data is not adapted to the new situation, the data is real-time and dynamic. So the generalization ability of training sample is poor. Because of the complexity and dynamic of the information system, it is difficult to grasp the suitable assignment of initial weight. In addition, neural network technique is easy to form local optimum and can not get the whole optimum [1]. There is also a big limitation that it is difficult to find the training sample [5]. In view of the influence factors of information system, most of the indicators are dynamic and grey. The data of information system security is difficult to achieve The grey system theory , established by Julong Deng in 1982, is a new methodology that focuses on the study of problems involving small sample and poor information. It deals with uncertain system with partial known information through generating, excavating, and extracting useful information from what is available, so that systems, operational behaviors and their laws of evolution can be correctly described and effectively monitored. So the grey evaluation theory is suitable to evaluate the risk of information system. This model based on grey theory can overcome some defects of statistical model and neural network model. Firstly the Grey system theory can study on small sample and poor data. In the whole process of risk assessment ,at present we only achieve some objective and quantitative data by intrusion detection ,system audit, vulnerability scanning technology and so on. Secondly the grey system theory focus on the laws of evolution and can analysis the information real-time risk. It is important that the grey system can solve the problem of neural network that under the dynamic data the generalization ability of training sample of neural network is poor.

ASSESSMENT MODEL OF INFORMATION SYSTEMS RISK BASED ON ENTROPY METHOD AND GREY THEORY

Index System for Risk Assessment of Information System

It is the prerequisite to evaluate the risk of information system that the risk factors is correctly analyzed and the suitable assessment index system is established [10].Learning from the foreign BS7799 information security standards system and considering the actual situation of our country's information system security, index system for risk assessment of information system is generated as table 1.

Table 1. Evaluation Index System of Information System' Risk

Target layer	Criterion layer	Index layer
Information system' risk level (X)	Threat level, (X_1)	The risk of information's removing and stealing, (X_{11})
		The risk of network resource's destroying, (X_{12})
		The risk of information's abusing and tampering ,(X_{13})
		The risk of Service disruption and prohibition,(X_{14})
		The risk of Information leakage,(X_{15})
		The risk of fake access,(X_{16})
		The risk of bypass control,(X_{17})
		The risk of authorization violation,(X_{18})
	Vulnerability level, (X_2)	Management security (X_{21})
		Physical device security , (X_{22})
		Software security, (X_{23})
		Hardware security(X_{24})
		Personnel security,(X_{25})
		Environment security,(X_{26})
		Communication security,(X_{27})
	Protective measure, (X_3)	Recovery technique measures,(X_{31})
		Encryption measures, (X_{32})
		Anti hacking measures, (X_{33})
		Anti virus measures,(X_{34})
		Data backup measures,(X_{35})
	Consequence severity,(X_4)	The severity of environmental degradation,(X_{41})
		The severity of service degradation (X_{42})
		Information recovery cost, (X_{43})
		Service recovery cost (X_{44})

Measure the Weight of Each Indicator with Entropy Method

(1) According to the polarity of each indicator, standardize all index data of information systems risk. For the positive indicator data, use formula (1) to standardize; and for the negative index data, use formula (2) to standardize.

$$X'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \qquad (1)$$

$$X'_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \qquad (2)$$

(2) Calculate the index information entropy by the formula (3)

$$h_j = -k\sum_{i=1}^{m} \{(X'_{ij} \big/ \textstyle\sum_{i=1}^{m} X'_{ij}) * \ln(X'_{ij} \big/ \textstyle\sum_{i=1}^{m} X'_{ij})\} \qquad (3)$$

In formula (3), $k = \ln m$; and the index information entropy h_j =0, if the value of index data standardization X'_{ij}=0.

(3) Calculate the redundancy of each index's information entropy by the formula

$$d_j = 1 - h_j \qquad (4)$$

(4) Use the redundancy of information entropy to calculate the index weight through the formula

$$\eta_j = d_j \Big/ \sum_{j=1}^{n} d_j \qquad (5)$$

Evaluation of Grey Class

Due to the limitation of the expert level and the difference in the cognition angle, only a few of whiten weight of grey numbers are given. In order to truly reflect the level of a certain class, it is necessary to determine the evaluation of grey class.

In this setting, it is supposed that the whitenization weight function is the triangle whitenization weight function .and there are five levels in grey class, the grey grades h=1,2,3,4,5. Respectively, It is very low, low, medium, high,, very high. It is supposed that the grey number of every grey class is $\otimes_h \in [0,\, h,\, 2h]$ .The whitenization weight function is f_h

$$f_h(d_{ijk}) = \begin{cases} d_{ijk} \big/ h & d_{ijk} \in [0, h] \\ 2h - d_{ijk} \big/ 2h - h & d_{ijk} \in [h, 2h] \\ 0 & \\ & else \end{cases} \qquad (6)$$

Grey Evaluation Weight Vector and Weight Matrix

The grey class of belongingness of The index x_{ij} denote M_{ijk} ,and M_{ijk} is defined as $M_{ijk} = \sum_{k=1}^{m} f_k(d_{ijk})$,and $M_{ij} = \sum_{k=1}^{5} M_{ijk}$,then the belongingness of x_{ij} **in *the grey class*** h for m experts is q_{ijh} ,

$$q_{ijk} = M_{ijk} \big/ M_{ij} \qquad (7)$$

There are 5 grey classes , and h=1,2,3,4,5. the grey class of belongingness of x_{ij} in h is the grey evaluation weight vector q_{ij} ,and $q_{ij} = (q_{ij1}, q_{ij2}, q_{ij3}, qi_{j4}, q_{ij5})$ .Then a grey evaluation weight matrix is generated Q_i

$$Q_i = \left[q_{i1}, q_{i2}, q_{i3}, q_{i4}, q_{i5}\right]^T \qquad (8)$$

Comprehensive Evaluation

After the comprehensive evaluation of the second-level indexes, the grey evaluation weight vector B_i of X's comprehensive evaluation is obtained.

Where $B_i = w_i \times Q_i = (b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5})$ (9)

Then all the grades of belongingness of x_{ij} in ***the grey degree*** h produce the grey evaluation weight matrix Q ,Where $Q = \left[B_1, B_2, B_3, B_4, B_5\right]^T$

After the first-level indexes are evaluated comprehensively, the result of the evaluation produce the value of comprehensive assessment

$$B = w \times Q = (b_1, b_2, b_3, b_4, b_5) \qquad (10)$$

Assign values to the grey degree and produce the evaluation of grey level vector $C = (1, 2, 3, 4, 5)$

$$V_{Bi} = B_i \times C^T \qquad (11)$$

$$V_B = B \times C^T \qquad (12)$$

Here in the Formula (11) V_{Bi} denote the level of the impact factors, and in the Formula (12) V_B denote the level value of the risk value of the information system. So we can take appropriate measures to carry out risk control and risk aversion according to the risk state of the factors.

EMPIRICAL ANALYSIS

Five hospitals in Fujian province are selected —Fujian Medical University Union Hospital(Unit$_1$)， Fujian Provincial Hospital （Unit$_2$）， The First Affiliated Hospital of Fujian Medical University(Unit$_3$)， The Second Affiliated Hospital of Fujian Medical University(Unit$_4$) ， Nanjing General Hospital of Nanjing Military Command(Unit$_5$) to exemplify the validity of the Entropy Method and Grey Theory in the risk assessment. As the chart shows, there are four criterion layer indexes, and 24 index layer indicators.

Collect and Standardize Data

Quantitative indicators data in Table 1 are acquired from hospital information system statistical data, and qualitative data comes out from experts. After Standardization the index data of five units is as the following in Table 2.

Table2. Standardization of All Index Data

Index unit	Unit1	Unit2	Unit3	Unit4	*Unit5*
X_{11}	0.423	0.124	0.311	0.243	*0.215*
X_{12}	0.441	0.315	0.324	0.223	*0.207*
X_{13}	0.343	0.278	0.347	0.197	*0.223*
X_{14}	0,522	0.314	0.351	0.215	*0.264*
X_{15}	0.425	0.417	0.462	0.518	*0.178*
X_{16}	0.467	0.378	0.356	0.419	*0.224*
X_{17}	0.433	0.436	0.427	0.461	*0.215*
X_{18}	0.387	0.315	0.439	0.368	*0.308*
X_{21}	0.455	0.418	0.399	0.378	*0.267*
X_{22}	0.418	0.532	0.415	0.524	*0.312*
X_{23}	0.437	0.357	0.423	0.451	*0.311*

X_{24}	0.323	0.437	0.478	0.403	*0.204*
X_{25}	0.354	0.423	0.404	0.375	*0.211*
X_{26}	0.375	0.422	0.415	0.378	*0.279*
X_{27}	0.314	0.475	0.317	0.455	*0.246*
X_{31}	0.415	0.426	0.387	0.398	*0.231*
X_{32}	0.378	0.418	0.437	0.356	*0.319*
X_{33}	0.425	0.417	0.481	0.427	*0.306*
X_{34}	0.472	0.456	0.392	0.375	*0.278*
X_{35}	0.423	0.408	0.414	0.428	*0.139*
X_{41}	0.437	0.397	0.367	0.354	*0.231*
X_{42}	0.465	0.437	0.412	0.427	*0.271*
X_{43}	0.396	0.378	0.318	0.354	*0.197*
X_{44}	*0.456*	*0.471*	*0.428*	*0.369*	*0.234*

Calculate Weight of the Evaluation Index

Use entropy method to calculate weight of information's evaluation index, as is shown in Table 3.

Table3. Entropy and Weight of Evaluation Index

Criterion	Index	Entropy h_j	Redundancy d_j	Levels weight η_j	Combined weight η'_j
X_1	X_{11}	0.729	0.271	0.137	0.035
	X_{12}	0.676	0.324	0.103	0.026
	X_{13}	0.792	0.208	0.160	0.042
	X_{14}	0.767	0.233	0.132	0.032
	X_{15}	0.654	0.346	0.09	0.021
	X_{16}	0.763	0.237	0.113	0.026
	X_{17}	0.812	0.188	0.145	0.031
	X_{18}	0.673	0.337	0.12	0.023
X_2	X_{21}	0.791	0.209	0.129	0.023
	X_{22}	0.784	0.216	0.24	0.024
	X_{23}	0.79	0.21	0.131	0.036
	X_{24}	0.812	0.188	0.114	0.031
	X_{25}	0.736	0.264	0.107	0.022
	X_{26}	0.765	0.235	0.117	0.024
	X_{27}	0.853	0.147	0.162	0.027
X_3	X_{31}	0.713	0.287	0.163	0.035
	X_{32}	0.769	0.231	0.212	0.037
	X_{33}	0.661	0.339	0.311	0.039
	X_{34}	0.767	0.233	0.114	0.038
	X_{35}	0.734	0.266	0.20	0.022
X_4	X_{41}	0.459	0.541	0.216	0.311
	X_{42}	0.761	0.239	0.314	0.029
	X_{43}	0.729	0.271	0.268	0.025
	X_{44}	0.75	0.25	0.202	0.041

Calculate Whiten Weight Matrix

$$Q_1 = \begin{bmatrix} 0 & 0.293 & 0.894 & 0.576 & 0.329 \\ 0 & 0.544 & 0.677 & 0.783 & 0.219 \\ 0 & 0.634 & 0.522 & 0.617 & 0.322 \\ 0 & 0.542 & 0.466 & 0.723 & 0.355 \\ 0 & 0.615 & 0.712 & 0.521 & 0.423 \\ 0 & 0.772 & 0.624 & 0.822 & 0.543 \\ 0 & 0.612 & 0.542 & 0.754 & 0.566 \\ 0 & 0.237 & 0.576 & 0.433 & 0.321 \end{bmatrix} \qquad Q_4 = \begin{bmatrix} 0 & 0.233 & 0.415 & 0.673 & 0.534 \\ 0 & 0.231 & 0.553 & 0.462 & 0.567 \\ 0 & 0.677 & 0.715 & 0.644 & 0.516 \\ 0 & 0.435 & 0.617 & 0.733 & 0.563 \end{bmatrix}$$

Then use the formula $Q = \left(B_i\right) = \left(W_i ? Q_i\right)$ to calculate the comprehensive evaluation value of the five hospitals' information system risk,as the following Table 4

Table 4. The Comprehensive Evaluation Value of The Five Hospital Information System's Risk

Unit	B_i	$\left(B_i\right) = \left(W_i ? Q_i\right)$	V_{Bi}	V_B
Unit$_1$	B_1	$(0.214 \quad 0.673 \quad 0.552 \quad 0.417 \quad 0.369)$	0.325	
	B_2	$(0.324 \quad 0.433 \quad 0.615 \quad 0.756 \quad 0.519)$	0.416	0.453
	B_3	$(0.214 \quad 0.513 \quad 0.175 \quad 0.456 \quad 0.473)$	0.334	
	B_4	$(0.423 \quad 0.215 \quad 0.765 \quad 0.522 \quad 0.774)$	0.563	
Unit$_2$	B_1	$(0.334 \quad 0.231 \quad 0.167 \quad 0.655 \quad 0.528)$	0.342	
	B_2	$(0.224 \quad 0.261 \quad 0.187 \quad 0.644 \quad 0.788)$	0.516	0.447
	B_3	$(0.122 \quad 0.235 \quad 0.324 \quad 0.415 \quad 0.678)$	0.433	
	B_4	$(0.512 \quad 0.245 \quad 0.414 \quad 0.375 \quad 0.376)$	0.313	
Unit$_3$	B_1	$(0.612 \quad 0.345 \quad 0.517 \quad 0.355 \quad 0.361)$	0.417	
	B_2	$(0.432 \quad 0.645 \quad 0.557 \quad 0.335 \quad 0.343)$	0.438	0.412
	B_3	$(0.212 \quad 0.345 \quad 0.257 \quad 0.675 \quad 0.423)$	0.347	
	B_4	$(0.513 \quad 0.415 \quad 0.237 \quad 0.475 \quad 0.313)$	0.322	
Unit$_4$	B_1	$(0.323 \quad 0.455 \quad 0.264 \quad 0.325 \quad 0.412)$	0.357	
	B_2	$(0.533 \quad 0.511 \quad 0.474 \quad 0.625 \quad 0.318)$	0.461	0.437
	B_3	$(0.431 \quad 0.511 \quad 0.374 \quad 0.425 \quad 0.312)$	0.442	
	B_4	$(0.621 \quad 0.413 \quad 0.214 \quad 0.315 \quad 0.416)$	0.415	
Unit$_5$	B_1	$(0.301 \quad 0.313 \quad 0.374 \quad 0.285 \quad 0.316)$	0.356	
	B_2	$(0.431 \quad 0.323 \quad 0.404 \quad 0.385 \quad 0.226)$	0.378	0.305
	B_3	$(0.121 \quad 0.123 \quad 0.04 \quad 0.685 \quad 0.126)$	0.218	
	B_4	$(0.201 \quad 0.173 \quad 0.154 \quad 0.285 \quad 0.226)$	0.203	

From the Table , we can see that the five hospital information system are in the medium risk. the first-level indexes: threatening factors, vulnerability factors, protective measures and the consequences severity are in the state of a certain degree of risk. In the five hospital information system, we can see the index: protection measure is in higher risk than other three first-level indexes. Nanjing General Hospital of Nanjing Military Command(Unit$_5$)' s risk is lower than the other four hospital. This may be attributed to the attention of confidentiality for the military hospital. Fujian Medical University Union Hospital(Unit$_1$)' s risk is the highest .

especially in the consequences severity the score is 0.563. It maybe is a cause that there is the biggest infectious ward. The protection measures must be strengthened , or the information service security will become a hidden danger. the impact on Fujian Medical University Union Hospital is very large, leading to the deterioration of services, resulting in huge recovery costs. It is recommended that the hospital improve the installation of information security equipment and conduct regular checks and updates, such as firewall, access log, intrusion detection system, weak point scanner and so on.

CONCLUSIONS

The evaluation of information system risk is a systematic assessment problem with multi-index, multi-level and the index is grey and dynamic .To solve the problem, this paper presented a assessment method based Entropy Method and Grey Theory. The empirical study of five hospital indicated that it was effective to use the method to systematically evaluate information system risk and its evaluation result was comparatively objective and accurate.

REFERENCE

[1] A Der Kiureghian, PL Liu (2013) ‘Structural reliability under incomplete probability information’, *Journal of Engineering Mechanics,* Vol. 10, pp. 96-115.

[2] A. Saeed, R. De Lemos, & T. Anderson (1994) ‘An approach to the risk analysis of safety specifications’, *Proc. 9th Annu. Conf. Comput. Assurance,* pp.209 -221.

[3] Ernst G. Frankel (1988) ‘Analysis of maintained systems’, *Systems Reliability and Risk Analysis,* Springer Netherlands, New York.

[4] J.L Rouvroye & E.G van den Bliek (2010) ‘Comparing safety analysis techniques’, *Reliability Engineering & System Safety,* Vol. 75, No. 3, pp. 289–294.

[5] J Wang (2010) 'A subjective methodology for safety analysis of safety requirements specifications', *Fuzzy Systems,* Vol. 5 , No. 3, pp.33-46.

[6] J. Wang, J.B. Yang & P. Sen (1995) ‘Safety analysis and synthesis using fuzzy sets and evidential’, *reasoning, Reliability Engineering & System Safety*, Vol. 2, No. 1, pp. 53-65.

[7] Marco Bozzano & Adolfo Villafiorita (2013) *Improving system reliability via model checking: The FSAP/NuSMV-SA safety analysis platform,* Springer Berlin Heidelberg’ , Heidelberg, Germany.

[8] P. Fenelon, J. A. McDermid, M. Nicolson & D. J. Pumfrey (2011) ‘A simple method for the multi-item, single-level, information system safety analysis’, *Information safety Management Journal,* Vol. 38, pp. 39-47.

[9] Peter Fenelon & John A McDermid (2011) ‘An integrated tool set for software safety analysis’, *Journal of Systems and Software,* Vol. 21, No. 3, pp. 279–290.

[10] RL Borison (2012) ‘The role of cognition in the risk-benefit and safety analysis of antipsychotic medication', *Acta Psychiatrica Scandinavica*, Vol.13, pp.570-583.

AUTHORS

Jinli Duan, she is a doctoral candidate of School of Economics and Management in Fuzhou University. she has received his master's degree in technology innovation and management from School of Management in Fuzhou University in 2007. His current research interests include information management, system engineering and management, etc.

Qishan Zhang, he is a professor and doctoral supervisors of School of Economics and Management in Fuzhou University. His current research interests include information management ,data mining,etc