Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/080245

An Information Security Threat Assessment Model based on Bayesian Network and OWA Operator

Kehe Wu[1,*] *and Shichao Ye*[2,*]

[1] Beijing Engineering Research Center of Electric Information Technology, North China Electric Power University, Changping 102206, Beijing, China
[2] Department of Control and Computer Engineering School, North China Electric Power University, Changping 102206, Beijing, China

Received: 17 Apr. 2013, Revised: 18 Aug. 2013, Accepted: 20 Aug. 2013
Published online: 1 Mar. 2014

Abstract: Information security threat assessment involves two aspects, namely, technology and management. A great amount of uncertainties exist in the assessment, which cannot be strictly quantized. Thus, the completely objective information security risk assessment is hard to realize. To this end, this research proposed an information security threat assessment model based on Bayesian Network (BN) and OWA operator. Firstly, with the integration of expert knowledge, the conditional probability matrix of reasoning rules in BN was clarified, as a basis of the establishment of information security threat assessment model. Then, with the group-decision method of OWA operator, the subjective judging information of experts on the threat level of target information system was integrated, which was taken as the prior information of the threat level of target information system. Meanwhile, with the observation nodes of objective assessment information, subjective and objective security threat level was integrated, which realized the continuity and accumulation of the security assessment. Finally, the rationality and effectiveness of this model were verified through the simulation example.

Keywords: Information security, Bayesian Network, OWA operator, quantitative assessment

1 INTRODUCTION

With the development of computer technology and the Internet, new attacks with pitfalls of system security have been extensively used by illegal intruders and hackers. Moreover, security risks and threats faced by the security of information system have been gradually severed. The security of information system has been a focus among people.

The information security risk assessment is one of the effective methods of addressing the security issues of information system, including fault tree analysis, analytic hierarchy process (AHP) and fuzzy comprehensive evaluation. Such methods have been used by security assessment personnel. Yet, the impact of human factors and administrative management measures on the information system was insufficiently considered so far. Meanwhile, information security threat assessment involves two aspects, namely, technology and management. A great amount of uncertainties exist in the assessment, which cannot be strictly quantized. Thus, the completely objective information security risk assessment is hard to realize.

In this research, the subjective and objective security assessment information was integrated, and the information security threat assessment model based on Bayesian Network (BN) and OWA operator was established. First of all, the group-decision method based on OWA operator sufficiently uses the experiences and knowledge of each decision makers to assess the target information system. This, to a great extent, makes up the one-sidedness of individual judgment of decision makers; secondly, similar to the neural network, BN can fully depict the reasoning process of human beings. BN-based security assessment can not only quantitatively interpret the process of security assessment, but also reflect the continuity and accumulation of the security assessment. Hence, information security threat assessment model based on BN and OWA operator can sufficiently consider the subjective judging information of each decision-maker but also demonstrate the continuity and

[*] Corresponding author e-mail: epuwkh@126.com, ye_shichao@126.com

accumulation of security assessment. Besides, it improves the confidence level of BN prior information.

2 SUBJECTIVE THREAT GROUP DECISION BASED ON OWA OPERATOR

2.1 OWA Operator and Its Weight Endowment Method

Definition 1: Given $F : R_n \to R$, there is a n-dimensional weight vector correlated to $F, w_i \in [0,1], 1 \leq i \leq n$, and $\sum_{i=1}^{n} w_i = 1$, to make:

$$F\left(a_1, a_2, \ldots, a_n\right) = \sum_{i=1}^{n} w_i b_i \tag{1}$$

Where b_i is the ith maximum factor of the array $(a_1, a_2, \ldots, a_n)$. Then F is called the n-dimension OWA operator.

OWA operator is a kind of operator lying between the maximum operator and the minimum operator.

When $w = (1, 0, 0, \ldots, 0)$:

$$F(a_1, a_2, \ldots, a_n) = \max(a_1, a_2, \ldots, a_n) = b_1 \tag{2}$$

OWA operator is equivalent to the ”or” operator in fuzzy operation:

When $w = (0, 0, 0, \ldots, 1)$:

$$F(a_1, a_2, \ldots, a_n) = \min(a_1, a_2, \ldots, a_n) = b_n \tag{3}$$

OWA operator is equivalent to the “and” operator in fuzzy operator:

When $w = (1/n, 1/n, 1/n, \ldots, 1/n)$:

$$F\left(a_1, a_2, \ldots, a_n\right) = \frac{1}{n} \sum_{i=1}^{n} a_i \tag{4}$$

OWA operator is equivalent to the arithmetic average operator.

The identification of weight vector of OWA operator is directly related to the size of the data set. In order to ensure the fairness and reasonableness of the decision results, this research discretized the Gaussian distribution to clarify the weight vector of position. In this method, the discretion value was ranked in a position with a relatively small weighted value, which efficiently eliminated the adverse effects of emotional factors on the decision-making process.

Set μ is the mathematical expectation of $(1, 2, \ldots, n)$ endowed with the weight vector $w = (1/n, 1/n, \ldots, 1/n)$; σ is the standard deviation of $(1, 2, \ldots, n)$ in μ and weight vector w, thus we have:

$$\begin{aligned}
\mu_n &= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2} \\
\sigma_n &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (i - \mu_n)^2} \\
\omega' &= \frac{1}{\sqrt{2\pi\sigma_n}} e^{-\frac{(i-\mu_n)^2}{2\sigma_n^2}} \\
\omega &= \frac{\omega'}{\sum_{i=1}^{n} \omega'}
\end{aligned} \tag{5}$$

2.2 Group decision making subjective judgments threat based on OWA operator

For the evaluation of n information systems, The set of decision-making groups is $D = (d_1, d_2, \ldots, d_n)$, where $d_k (k = 1, 2, \ldots, m)$ represents the K-th decision makers, subjective judgment given in the form of decision are: the Utility value is $u^{(k)} = \left(u_1^k, u_2^k, \ldots, u_n^k\right)^T$, the fuzzy language evaluation value is $S = (s_0, s_1, \ldots, s_T)$,the fuzzy complementary judgment matrix is $p_{(k)} = \left(p_{ij}^{(k)}\right)_{n*n}$, Therefore, the information consistent with these judgments into utility value is:

1) fuzzy complementary judgment matrix into utility values:

$$u_i^{(k)} = \left(\sum_{j=1}^{n} p_{ij}^{(k)} + \frac{n}{2} - 1\right) \Big/ n(n-1), i = 1, 2, \ldots, n \tag{6}$$

2) evaluation value into fuzzy linguistic approach utility value:

We can make the fuzzy language evaluation value c as described in natural language corresponding to a utility value. For example,

$$\begin{aligned}
S = \{&s_0 = level_1 = 0, s_1 = level_2 = 0.1, s_2 = level_3 = 0.3, \\
&s_3 = level_4 = 0.5, s_4 = level_5 = 0.7, s_5 = level_6 = 0.9, \\
&s_6 = level_7 = 1\}
\end{aligned}$$

Thus,the utility value is converted to the formula is:

$$u_i^{(k)} = S_i^{(k)} \Big/ \sum_{i=1}^{n} S_i^{(k)}, i = 1, 2, \ldots, n \tag{7}$$

Assembled using OWA operator making a threat on the target population levels of subjective judgment of the i-th information is:

$$u_i = OWA_w\left(u_i(1), u_i(2), \ldots, u_i^{(n)}\right), i = 1, 2, \ldots, n \tag{8}$$

Decision-making groups $u = (u_1, u_2, \ldots, u_n)^T$ which is the subjective judgment of the information is:

$$u_i = u_i \Big/ \sum_{i=1}^{n}, i = 1, 2, \ldots, n \tag{9}$$

3 BN and reasoning algorithm

BN is also known as Belief Network, comprising of a series of combinations expressing causal rules. Most communications reasoning algorithm was proposed by Pearl, as a reasoning algorithm that is appropriate for simply connected space BN. In the algorithm of multi-tree communication method, assume at a node X, then there are m child nodes $(Y_1, Y_2, \ldots, Y_n)$ and n father nodes $(Z_1, Z_2, \ldots, Z_n)$. Assume *Bel* as a posterior probability distribution, then λ is the information of evidence acquired from child nodes π and is the information of evidence acquired from father nodes. $M_{X|Z} = P(X = x|Z = z)$ shows the probability of Event x in the child node X for a father node Z in a situation z. As X has discreteness, $\lambda(x)$ and $\pi(x)$ are actually vectors. Its element is related with each discrete value of :

$$\begin{aligned} \lambda(x) &= [\lambda(X = x), \lambda(X = x_2), \ldots, \lambda(X = x_l)] \\ \pi(x) &= [\pi(X = x), \pi(X = x_2), \ldots, \pi(X = x_l)] \end{aligned} \tag{10}$$

The reasoning algorithm of BN centers on single node. λ can be obtained from child node and π from father node. After that, *Bel*, λ and π at this node were calculated, triggering the updates of adjacent nodes. The renewal process is as follows:

Step 1: renewal of its own posterior probability: $Bel(x) = \alpha\lambda(x)\pi(x)$ Where α was the normalizing factor, so we have:

$$\sum Bel(x) = 1, \lambda(x) = \prod \lambda_{y_j}(x), \pi(x) = \prod \pi_{z_i} M_{X|Z} \tag{11}$$

Step 2: bottom-up renewal:

$$\lambda_x(z) = \lambda(x) M_{X|Z} \tag{12}$$

Step 3: top-down renewal:

$$\pi_y(x) = \alpha\pi(x) \prod_{k \neq j} \lambda_{y_j}(x) \tag{13}$$

4 BN–based information security threat assessment model

4.1 Analysis of assessment factors that influence information security threat level

Information security incidents originated from external causes (threats) and internal factors (fragility). Through the assessment of threats and fragility of information, the possibility of incidents can be acquired. Meanwhile, the impact of information security incidents is correlated with capital. Thus, the impact can be acquired through the assessment of capital.

Information security risks can be viewed as an influence on capital. To simplify the model, the following factors will only be considered: influence on capital , frequency of threats on capital as well as the fragility f of capital. The threat level was TL. On this basis, the BN-based information security threat assessment model was established.

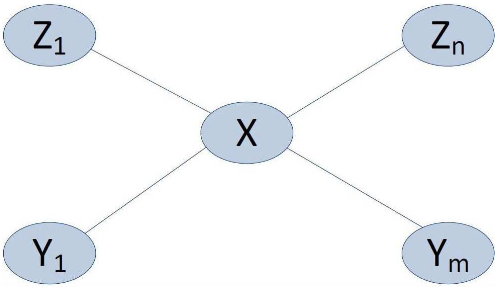

Fig. 1: The structure of Bayesian Network

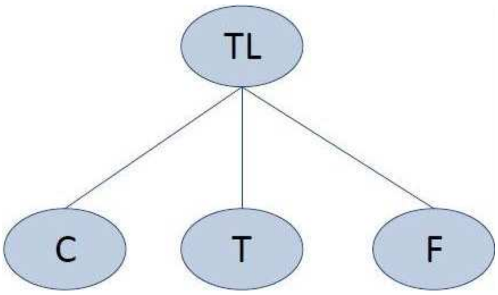

Fig. 2: An Information Security Threat Assessment Model based on Bayesian Network

States of variables in the model are gathered as follows:

$$\begin{aligned} TL &= \{high, medium, low\} \\ C &= \{big, middle, small\} \\ T &= \{high, medium, low\} \end{aligned} \tag{14}$$

4.2 Establishment of conditional probability matrix of reasoning rules

Conditional probability matrix reflects experts' opinions on causal relationship among the association node in the network, which form the expert knowledge. For example, if TL is high, the possibility of small, medium and large loss of capital is 10%, 30% and 60% respectively; if TL is medium, possibility of small, medium and large loss of capital is 40%, 40% and 20%; if TL is low, possibility of small, medium and large loss of capital is 60%, 30% and 10%. The interpretation of t and f is similar with the above descriptions, as shown in the following Table.

It should be noted that conditional probability matrix is an expert knowledge, thus showing certain subjectivity. Thus, repeated testing of sample data can be used to properly adjust the matrix so that the credibility of the assessment results can be improved.

Table 1: Inference rules conditional probability matrix.

Threat level	$P(C\|TL)$	$P(t\|TL)$	$P(f\|TL)$
	small middle big	low medium high	not serious ordinary serious
high	0.1 0.3 0.6	0.1 0.5 0.4	0.8 0.1 0.1
medium	0.4 0.4 0.2	0.4 0.3 0.3	0.6 0.3 0.1
low	0.6 0.3 0.1	0.6 0.2 0.2	0.1 0.45 0.45

5 Analysis of examples

The decision-making group comprised of four experts assessed the TL of a target. Assume the target TL was respectively high, medium and low. Threat judging information given by four decision makers is:

$U_1 = (0.2, 0.6, 0.2)$

$U_2 = (0.3, 0.3, 0.4)$

$U_3 = (0.07, 0.33, 0.6)$

$U_4 = (0.37, 0.3, 0.33)$

According to the equation, it can be obtained that the OWA operator weight vector is:

$$w = (0.155, 0.345, 0.155, 0.345) \tag{15}$$

Then TL assessment value of the decision-making group was:

$$U = (0.247, 0.367, 0.387) \tag{16}$$

After the BN initialization with prior information and conditional probability, the assessment system was fully prepared and put into the waiting state. When the system obtained new assessment information, leaf node of the network was renewed, and triggered the network reasoning. After the renewal of probability distribution of node state of the entire network, the condition of probability distribution of root node state was obtained, and the TL assessment was completed. Assume the probability of the following influential factors was logged in:

$$\lambda_c = [0\ 0\ 1]\ \lambda_t = [0\ 1\ 0]\ \lambda_f = [0\ 1\ 0] \tag{17}$$

Example 1: Assume there is no prior subjective assessment information of threat from the decision-making group, we set the prior information of TL in an information system as $\pi(TL)$, which reflected the insufficient possibility assessment from information starvation. Thus, it can be considered that each condition was closer. Thus, the assessment results were as shown in the figure. Bel1 indicates the maximal possibility of low threat.

Example 2 Assume that TL is generated from the group-decision method based on OWA operator, and then BN prior information is Π(TL). The assessment results were shown in the figure. Bel2 indicates the increase of probability of medium TL. And the probability of the other two levels was decreased. It can be seen that the subjective TL judging information of the decision-making group obviously influenced the assessment results.

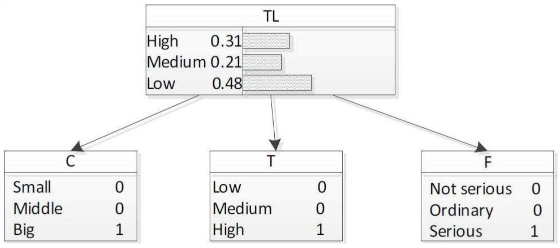

Fig. 3: Example 1 assessment results

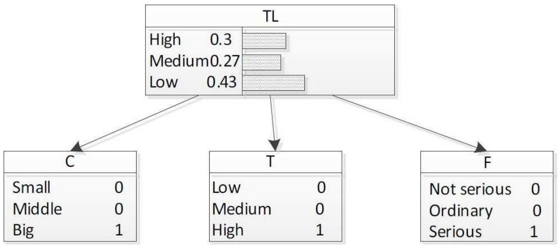

Fig. 4: Example 2 assessment results

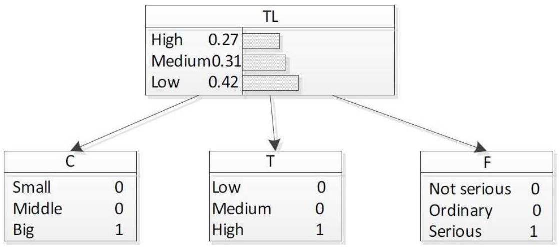

Fig. 5: Example 3 assessment results

Example 3 After a certain period, this information system was assessed again. In this round of assessment, the results of the previous round were taken as the prior information for calculation. At this time, prior information became Π(TL). Assume the probability of each influencing factor remain unchanged. Then the results were shown in the figure. In other words, the probability of medium TL continued to increase while the rest kept decreasing.

To sum up, due to different prior information, the results were differed. Common prior information includes two aspects, namely, the prior information that should be set in the initiation of the algorithm, and the prior information as the results of the previous period in the operation period of algorithm. Results of the previous simulation examples demonstrated that to take the subjective judgment of threat level of target information system from the decision-level group as BN prior information can more efficiently reflect the real TL of the target.

6 Conclusion

The traditional information security threat assessment model does not take into account subjective threat judgment information given by decision-makers based on their professional experiences. For the overall assessment model, it was a kind of information loss. In this research, on the basis of systematic analysis of information security threat elements, subjective TL judging information and objective situation information were combined so as to establish the information security threat assessment model based on BN and OWA operator. This model is verified to be more consistent with the actual process of information security assessment, which can relatively reflect the real TL accurately. The algorithmic examples proved the effectiveness of the method, which could provide a new perspective for the assessment of information security threat.

Acknowledgements

This work was supported by the National Basic Research Program of China (973 Program). The first author acknowledges the financial support by the National Basic Research Program of China (973 Program) from China's Ministry of Science and Technology.The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] Yager R R and Filev D P., Induced ordered weighted averaging operators. IEEE Transactions on Systems, Man and Cybernetics, **29**, 141-150 (1999).

[2] Yager R R, Families of OWA operators. Fuzzy Sets and Systems, **59**, 125-148 (1993).

[3] Merigó J M and Casanovas M, The fuzzy generalized OWA operator and its application in strategic decision making. Cybernetics and Systems, **41**, 359-370 (2010).

[4] Fu Y, Wu X, Yan C, The method of information security risk assessment using Bayesian networks. Journal-Wuhan University Natural Sciences Edition, **52**, 631 (2006).

[5] Maglogiannis I, Zafiropoulos E, Platis A, et al, Risk analysis of a patient monitoring system using Bayesian Network modeling. Journal of Biomedical Informatics, **39**, 637-647 (2006).

[6] Yager R R, On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transactions on Systems,Man and Cybernetics, **18**, 183-190 (1988).

[7] Li X, Ji Q, Active affective state detection and user assistance with dynamic Bayesian networks. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, **35**, 93-105 (2005).

[8] Yang Z, Wright R N, Privacy-preserving computation of Bayesian networks on vertically partitioned data. Knowledge and Data Engineering, IEEE Transactions on, **18**, 1253-1264 (2006).

[9] Yager R R. Quantifier guided aggregation using OWA operators[J]. International Journal of Intelligent Systems, **11**, 49-73 (1996).

[10] Jensen F V. An introduction to Bayesian networks[M]. London: UCL press, (1996).

[11] Yager R R. Induced aggregation operators[J]. Fuzzy Sets and Systems, **137**, 59-69 (2003).

[12] Neapolitan R E. Learning bayesian networks[M]. Upper Saddle River: Pearson Prentice Hall, (2004).

[13] Yager R R. A new approach to the summarization of data[J]. Information Sciences, **28**, 69-86 (1982).

[14] Rigaux C, Ancelet S, Carlin F, et al. Inferring an Augmented Bayesian Network to Confront a Complex Quantitative Microbial Risk Assessment Model with Durability Studies: Application to Bacillus Cereus on a Courgette Puré Production Chain[J]. Risk Analysis, (2012).

[15] Wang Y, Vassileva J. Bayesian network-based trust model[C]//Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on. IEEE, 372-378 (2003).

[16] van Steensel B, Braunschweig U, Filion G J, et al. Bayesian network analysis of targeting interactions in chromatin[J]. Genome research, **20**, 190-200 (2010).

[17] Yager R R, Filev D. On the issue of defuzzification and selection based on a fuzzy set[J]. Fuzzy sets and Systems, **55**, 255-271 (1993).

[18] Correa E, Goodacre R. A genetic algorithm-Bayesian network approach for the analysis of metabolomics and spectroscopic data: application to the rapid identification of Bacillus spores and classification of Bacillus species[J]. BMC bioinformatics, **12**, 33 (2011).

[19] Johnson B, Grossklags J, Christin N, et al. Are security experts useful? Bayesian Nash equilibria for network security games with limited information[M]//Computer Security–ESORICS 2010. Springer Berlin Heidelberg, 588-606 (2010).

[20] Feng N, Xie J. A Bayesian networks-based security risk analysis model for information systems integrating the observed cases with expert experience[J]. Scientific Research and Essays, **7**, 1103-1112 (2012).

[21] Zhang S, Song S. A Novel Attack Graph Posterior Inference Model Based on Bayesian Network[J]. J. Information Security, **2**, 8-27 (2011).

[22] Manshaei M, Zhu Q, Alpcan T, et al. Game theory meets network security and privacy[J]. ACM transaction on Computational Logic, 5 (2011).

[23] Bohacik J, Davis D N. Estimation of cardiovascular patient risk with a Bayesian network[C]//Proc. Ninth European conference of young research and scientific workers (TRANSCOM 2011), 7-40 (2011).

[24] Sharma S K, Pandey P, Tiwari S K, et al. An improved network intrusion detection technique based on k-means clustering via Naïve bayes classification[C]//Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on. IEEE, 417-422 (2012).

[25] Guo X, Hu R. The effectiveness evaluation for security system based on risk entropy model and Bayesian network theory[C]//Security Technology (ICCST), 2010 IEEE International Carnahan Conference on. IEEE, 57-65 (2010).

Kehe Wu born in September 1962, engaged in computer application technology and network information security research, is China earlier in the IM (Intelligent Management), electricity ERP technology, electric power information system security protection technology research and development scholars . He led the “Network Software Research” specializes in power intelligent software technology and network information security technology research; achievements include: electric power information security protection system, power generation business intelligence management systems, network real-time data management and integrated data mining application system.

Shichao Ye born in 1986.2, doctoral students in reading. Main research directions: information security. The study includes: Power Information Security Protection evaluation; PKI technologies. Research Information System for power (including management information systems and production control system) security protection features, focusing on network security assessment and trusted computing environment to build theory and related technology, the network environment, reliability, stability, real-time , business continuity, operational privacy and non-repudiation, etc. related theoretical research and technology development. Work was supported by the National Natural Science Foundation of China: Energy thermal power process control and optimization of basic research (ID: 51036002) and national 973 project: coordination of multiple distributed power generation and intelligent scheduling (ID: 2012CB215203).