Information exposure rating based on hierarchical model for big data

Jianbo Yao[1,*], and *Chaoqiong* Yang[2]

[1]School of Mathematics and Computer, Hezhou University, Hezhou Guangxi 542899, China
[2]Audit Division, Zunyi Normal College, Zunyi Guizhou 563006, China

Keywords: Information, Exposure information, Exposure quantification, Hierarchical.

Abstract. Big data improves the function of information, but also increases the exposure of information leakage. From the perspective of information leakage, Information entropy is used to quantify information, and the quantitative index of security exposure are identified as mutual information, and a hierarchical exposure assessment model is established. With the help of local priority principle, classify the types of information leaks and attack methods, the mutual information is determined as quantitative indicators for the exposure of information disclosure. The security exposure of information leakage is evaluated by fuzzy comprehensive evaluation

1 Introduction

The concept and application of big data have gradually penetrated into the accounting field, and the big data has improved the function of information and the efficiency of accounting work. The information of the big data era, the data sources is extensive, storage is relatively concentrated, contains a detailed record of accounting behavior, such as enterprise management data, customer information and privacy, big data had increased the exposure of information disclosure [1-3].

Big data and cloud computing the depth of the fusion, to promote the rapid development of the information system, but new information system, such as cloud accounting, is faced with some exposures, cloud accounting network information system is a special kind of information processing system, in addition to general information processing system of security features, but also has its own characteristics for some security, in the process of information system development, the stability of the platform system, identity authentication and management of the loopholes and defects such as data encryption system security problems appear, is likely to be pregnant and exposure factor in the production of information[4].

For complex and diversified information exposure in big data environment, this paper establishes the quantitative exposure assessment of a hierarchical model, effectively

* Corresponding author: 2055164364@qq.com

quantitative evaluation caused the security exposures by the information disclosure.

The remainder of this paper is organized as follows: In Section 2, basic concepts of exposure assessment. In Section 3, common exposure assessment methods. In Section 4, calculate the security exposure of information leakage. In Section 5, hierarchical model of exposure assessment. In Section 6, conclusions and future work.

2 Basic concepts of exposure assessment

The premise of studying the security exposure assessment of information system is to clarify the formation mechanism of security exposure of information system. In general, exposure is the uncertainty of the outcome of actions or events. no matter the result is positive opportunity or negative threat, people can only through the possibility of these uncertainties, and the actual after the effects and consequences to evaluate exposure. The security exposure of information system is to lead the security problems of information system, the possibility or the actual negative threat. The composition of security exposure of information system is divided into five aspects: origin, mode, pathway, receptor and consequence. The origin is the sponsor of the threat, called the threat source; The means by which the threat is threatened is called threat behavior; The path is the weak link used by the threat source, called vulnerability or vulnerability; The receptor is the recipient of the threat, the asset; The consequence is the loss caused by the threat to the source, called the impact., the relationship between them can be expressed as the origin of one or more of the exposure, using one or a variety of ways, through one or more channels, violation of one or more receptor, system adverse consequences[5]..

3 Common exposure assessment methods

In the process of exposure assessment, the most important point is how to calculate the security exposure. Figure1 shows the process of exposure calculation:

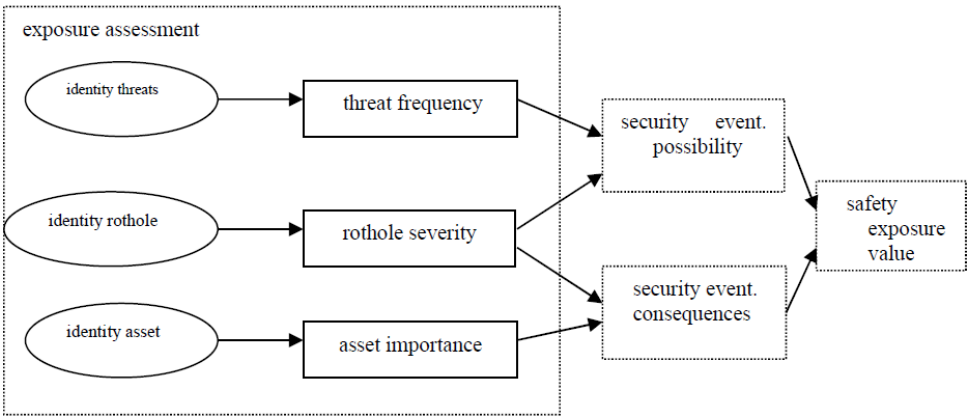

Fig. 1. Safety exposure calculation process.

Exposure assessment is generally divided into the following steps:

(1) identify information assets and assign values to information assets.

(2) identify possible threats and assign values to them.

(3) find vulnerabilities and assign values to them.

(4) the probability and threat of security incidents and the combination of vulnerability and vulnerability.

(5) the exposure value of information assets is calculated based on the importance of assets and the possibility of occurrence of security incidents.

The above gives the general steps and procedures for the calculation of security exposures. but in the actual exposure assessment process, due to the differences each system, the choice of assessment methods will directly affect the outcome of exposure assessment. Therefore, it is necessary to choose the appropriate exposure assessment method according to the specific situation of the system.

4 Calculate the security exposure of information leakage

The safety exposure index method in the literature [6] is used to calculate the security exposure of information leakage, which R is the exposure index:

$$R = f(P_s, C_s) = P_s + C_s - P_s C_s \tag{1}$$

P_s is the probability function of the exposure event, C_s is the consequence function, estimating P_s use the fuzzy theory, measuring C_s use the mutual information $I(leak_i, Key)$.

The process of information leakage can be seen as the process of information entropy reduction, and the secret information entropy is reduced by the attack of the attacker's statistical analysis. If the information entropy $leak_i$ is $H(leak_i)$, the conditional entropy of the secret key Key is $H(Key \mid leak_i)$ in $leak_i$, the mutual information of $leak_i$ and Key is $I(leak_i, Key)$.

Maximum acquisition $I(leak_i, Key)$ is the target of information disclosure:

$$I(Key, leak_i) = H(Key) - H(Key \mid leak_i) \tag{2}$$

Mutual information quantification is the quantitative calculation parameter in (2).

5 Hierarchical model of exposure assessment

In view of the complex and diversified information disclosure, the classification of information disclosure and attack methods, Based on the local priority principle, the exposure quantification is carried out from local to global level, and the multi-factor problem is solved to solve the single factor problem and reduce the complexity of the problem Because of the uncertainty and variability of threats, different attacks use fuzzy security threat weight factors to distinguish them. The evaluation steps in the hierarchical model of information disclosure exposure assessment are as follows [7-8].:

(1) Establishment of the collection of security exposure factors for information disclosure.

Assume different types of information disclosure collections is LS :

$$LS = \{leak_1, leak_2, ..., leak_i, ..., leak_p\} \tag{3}$$

The disclosure of $leak_i$ has the information attacks set AS_i consisting of all possible attack methods:

$$AS_i = \{a_{i1}, a_{i2}, ...a_{ij}, ..., a_{iq}\} \tag{4}$$

a_{ij} is the j attack method for the i .Classification of leakage and attack, the collection of multi-level leakage exposure, including single leakage of local exposure and the multiple disclosure of global exposures.

(2) Definition of exposure indicators.

Disclosure can affect the security of information systems, if the information system has s sub-system, the security exposure index R is defined as:

$$R = \{r_1, r_2, ...r_k, ..., r_s\} \quad , \quad r_k = I_k(leak_i, key) \tag{5}$$

r_k is the exposure value of the $k(1 \le k \le s)$ subsystem, r_k is quantized by $I_k(leak_i, key)$.

(3) Determine the weight coefficient of safety factors.

The threats of different attacks are different and fuzzy. Assumption $U=\{u_1,...,u_t\}$ is influencing factors set, in order to determine the weight coefficients at_{ij} of security threats in set U , its calculating use fuzzy hierarchical analysis method:

To comprehensively consider the various factors of collection U , using the scale of $0.1 \sim 0.9$ to compare each factor of two attack sets AS_i , the greater the threat level, the larger the scale value;

Establishing of weight fuzzy complementary judgment matrix $A = (a_{ij})_{q*q}$ .The weight coefficient of fuzzy complementary matrix can be obtained by literature [9]:

$$w_i = \frac{\sum_{j=1}^{q} a_{ij} + \frac{q}{2} - 1}{q(q-1)}, \qquad 1 \le i, j \le q \tag{6}$$

For AS_i , the determination of fuzzy weight coefficient for different attack methods at_{ij} , it can be obtained according to formula (6):

$$W_i = [w_{i1}, w_{i2}, ..., w_{iq}] \tag{7}$$

This indicates that different attack methods have different attack efficiency.

(4) Local exposures and global exposures.

Introduce local exposure and global exposure assessment information single leakage and multiple leakage security exposures. The proportion of a single attack a_{ij} in a exposure set R :

$$P_{ij} = [p_{ij1}, ..., p_{ijk}, ..., p_{ijs}]$$

$$p_{ij1} + ... + p_{ijk} + ... + p_{ijs} = 1 \tag{8}$$

The local exposures S_i is determined by a single leak $leak_i$:

$$S_i = T_i \cdot R^T$$

$$T_i = W_i \cdot P_i = [t_{i1}, ..., t_{ik}, ..., t_{is}] \quad , \quad t_{ik} = \sum_{j=1}^{q} w_{ij} p_{ijk} \ , \ 1 \le i \le p, 1 \le k \le s \tag{9}$$

The global exposure S is determined by multiple leaks:

$$S = T \cdot R^T$$

$$T = [T_1, T_2, ..., T_p] = [W_1 \cdot P_1, W_2 \cdot P_2, ..., W_p \cdot P_p] \tag{10}$$

T_i, T is the exposure assessment matrix, R^T is the exposure quantization matrix.

6 Conclusions and future work

In this paper, a hierarchical model of information exposure assessment is established. The local first, after the overall analysis, step by step, which effectively divides the types of information leakage and attack methods. The quantitative assessment of the information disclosure from the local exposures to the global exposure level, thus the complicated environment of information system security analysis simplified, established the definition of an effective assessment of the exposures of the information disclosure.

As future work, we will further refine the model and apply the model to the actual situation.

This research was supported by the Doctor Fund project of Hezhou University under grant HZUBS201809.

References

1. Zhang Yi-Li. Analysis of information security protection in the era of big data. National business sentiment (economic theory research), 2016 (3), 107-108.
2. Li Heng. Research on information security in the age of big data. Tax payment, 2017 (36), 51-51.
3. Huang Li. Security research of information system in the era of big data. New economy, 2015 (11), 108.
4. Peng Chao-Ran. Exposure factors and preventive measures for accounting informatization in the era of big data. Public Finance Research, 2014 (4), 73-76.
5. Chen Liang. Research on security exposure assessment model of information system. Journal of Chinese Peopleπ s Public Security University (Science and Technology). 2007 (4), 50-53.
6. Qian Gang. Information system security management. Nanjing: southeast university press, 2004.
7. Zhang Tao. Key technology research on the bypass attack on the password-oriented chip: [doctoral thesis]. Chengdu: university of electronic science and technology 2008.
8. Yao Jian-Bo, Zhang Tao. Quantitative evaluation model of side channel attack exposure. Computer engineering and application, 201248 (26), 84-87.
9. Ji Dong Dynasty, Song Pen-Feng, Yu Tian-Xiang. Fuzzy analytic hierarchy process and its application in design optimization. System engineering and electronic technology, 2006,28 (11) : 1692-1694.