Research on the Data Mining Method based on Information Security Vulnerability Ontology

ZhangHan[1], LvYali[2]

[1]Peking university, Beijing 100871,China;

[2]Henan University of Traditional Chinese Medicine, Zhengzhou 450008,China

Xuexi123@163.com

Keywords: Information security, Data mining, Vulnerability ontology, Related information.

Abstract. Information security has experienced a long time from the concept to deepen. Nowadays, with the development and in-depth of computer science and technology, the problem of information security grow with each passing day out. In June 2014, the freely available WiFi has trap vulnerability event, it specifically steals important information in the user's mobile phone. Information security and people's lives are closely linked, causing everyone to worry about the safety of confidence. Internationally, the United States, Germany, Britain, Japan and Russia and some other developed countries are relatively early for the study of information security, the larger of their human and financial investment has obtained a more fruitful results, at the same time the use of security information system is also more common. In recent years, China has begun to research and information security related to the basic theory and core technology issues, but because the specialized information security senior personnel is relatively small, the development of domestic information security is slow. This paper studies and explore the data mining method based on the information security vulnerabilities, the method can dig out the hidden association information behind the data through the establishment of security information system, but also judge the authenticity of the vulnerability ontology according to the relevant data information, and the feasibility and the application of the research direction are determined by experiments.

Introduction

In recent years, with the popularization of computer network technology and the rapid development of the information technology industry, the number of personal PC and unit host server is increasing faster, especially the national administrative agencies, military security and other departments on the computer's large-scale application, pociety's dependence on computers is becoming more and more significant, in case the application of the computer system or server security are destroyed, it will not only make the social life in turmoil, but also give the national, units or individuals to cause irreparable losses [1-3]. In the world, the network crime rate is relatively high, and the economic loss caused by the network crime every country in the world is amazing, which greatly exceeds the sum of the loss of common economic crimes [4,5]. Today, network information security has become a common concern of the international community. From a large amount of data, the algorithm can solve the problem of information security in the computer field, it also has important significance to optimize the information security vulnerabilities.

Data Mining Principles

The concept of data mining. Data mining is a new interdisciplinary science in many field, it includes many theoretical knowledge, and it is influenced by many subjects, including data system, computer control system, machine learning, visualization and electronic information science and so on [6-8]. Data mining is the non trivial process of acquiring effective, novel, potentially useful and ultimately understandable patterns from the massive database. Under the premise of no clear assumption, data mining begins to dig the information and discover knowledge, it is not a tool for the software package, it can be simple to buy and run into the business intelligence control environment,

but it can not automatically start to produce effective business rules [9-11]. Data must be real, massive and noisy in the database, it is mainly used to tap the information that is interested by the users, and the information is acceptable, understandable and can be used. The information in data mining is an important basis, and it must be correct and useful.

Data mining technology. Data mining technology is the research and development result of the database technology after a long time, which can access the database through the database to obtain the effective contact potential between data, as so to promote the transmission of information [12,13]. The technology can be cleaned, transformed and mining on data, which mainly includes information gathering, data integration, data reduction, data cleaning, data mining, data mining, data pattern evaluation, data knowledge representation and so on. Data mining techniques are shown in Figure 1.

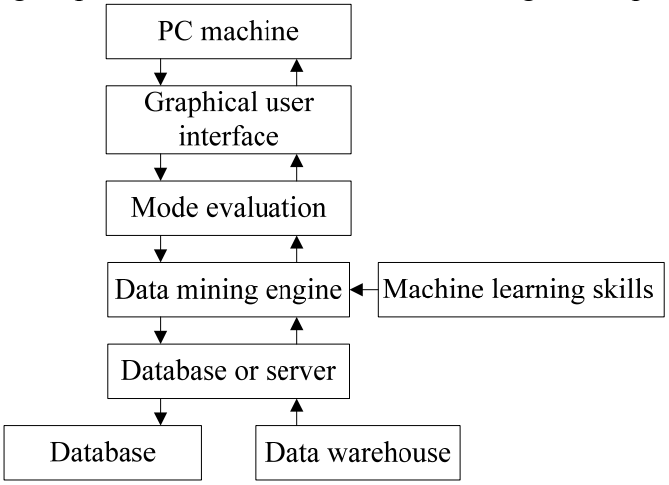

Fig. 1 Data mining technology system

As shown in Figure 1, data mining is a process of continuous learning and continuous cycle, if each process is not to achieve the desired goal, it needs to return to the last step to continue to implement the steps and to obtain useful results. Data mining technology is to find the specific rules and the characteristics of interest through the analysis of a large amount of time series data, including mining sequence patterns, periodicity, trend and deviation.

The Basic Framework of Information Security Vulnerability Ontology Security System

The basic framework of information security system. Information security involves information confidentiality, integrity and availability. Information security system includes human, technical means and management system, which can constitute the dynamic information and network security system framework WPDRR model, to achieve the system's security [14-16]. Information security is a complex, large and interconnected. Information security system is not a single technology, including many aspects of knowledge content. In the whole system, the use of firewall, the system needs to use the information encryption and other protective measures, and the use of intrusion detection and other technical tools understand and evaluate the security of the entire information security system [17-19]. According to the results of the detection, the system can be real-time adjustment security level that is the highest level of security or the state highest of risk level, and using the backup method will recover the system when the system is destroyed, finally through real-time monitoring system, the system carries out track for the intrusion behavior. The basic framework of information security vulnerability ontology system is shown in Figure 2.

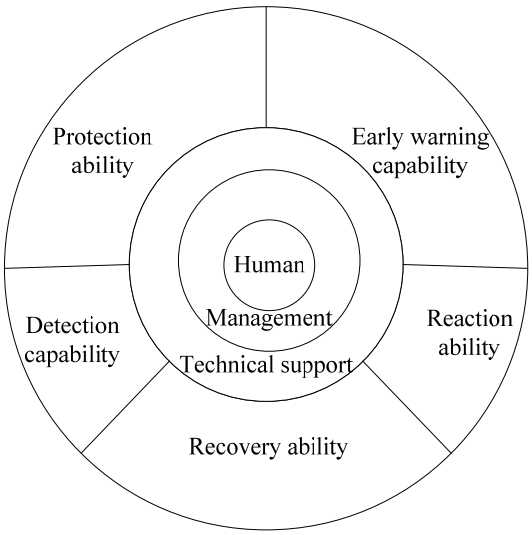

Fig. 2 The construction of Information security vulnerabilities ontology framework

The construction of information security vulnerability ontology security system. The construction of information security vulnerabilities ontology security system is divided into the following three parts [20-21]:

*The technical model of information security vulnerabilities ontology security system.*No matter what kind of single protocol technology, the demand for any security information will not be able to handle good, and we should choose appropriate information security vulnerability ontology security system, including security service, protocol layer and system safety. The technical model of information security vulnerability ontology security system is shown in Figure 3.

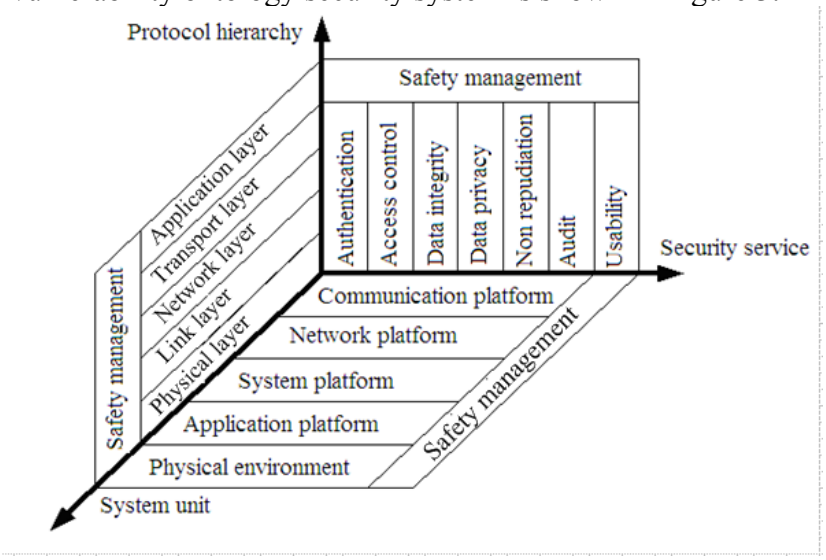

Fig.3 The technical model of information security vulnerability ontology security system

Determining the nature of information security vulnerability. Vulnerability is actually a software system or hardware design flaws. For some specific intrusion or threat sensitivity, the system can be divided into three types that are respectively environment, code and operation. The definition of information security vulnerability type is analyzed by using PROTEGE software.

Storing the information security vulnerability ontology security system. In this paper, the construction of information security vulnerability ontology security system is stored in relational database, and the relationship table of detailed structure is stored in the relation table.

Research on Data Mining Method based on Information Security Vulnerability Ontology

To mine the information security vulnerability data, we need to build a unique information vulnerability ontology security system, and then mining model is established by the association rule algorithm, finally the use of specific algorithm mines the database [22]. After a specific point of the experiment, the sysrem can get the mutual information between the vulnerability and the system software, and the model will give a warning of the vulnerability in different software, to complete the forecast of the number of vulnerabilities.

Vulnerability ontology data mining method. Data mining generally obtains no grain size distinction, no hierarchy data from some vulnerabilities database, and then it uses an iterative approach to carry on the rules of mining through the algorithm. As the probability of each type of vulnerability, it is relatively low, so it needs a limited number of supporting, and then to dig out a certain number of security error vulnerabilities. This paper proposes a data mining method based on information security vulnerability ontology, we need to carry out association rules mining data. According to the system framework, we can get more practical relevance information. The process of vulnerability ontology data mining method is shown in Figure 4.

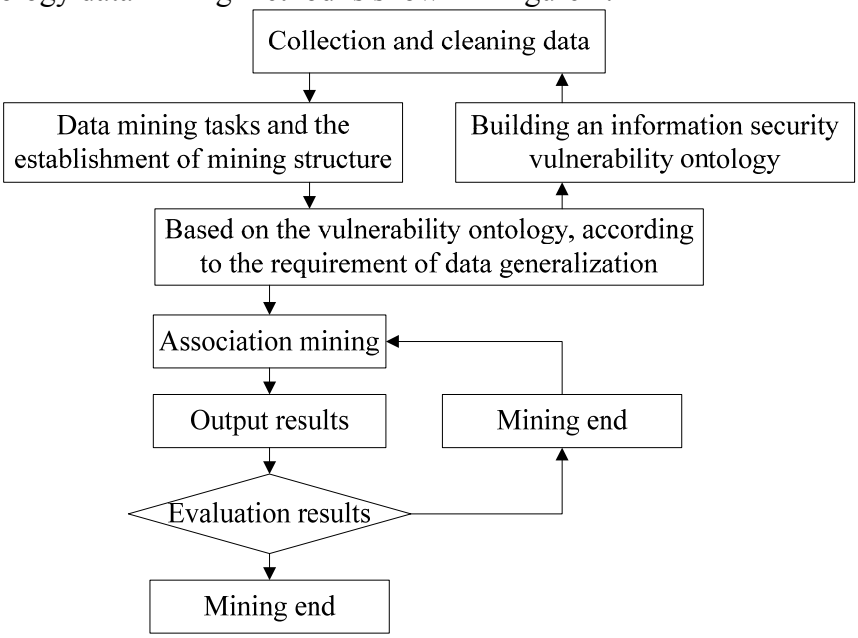

Fig. 4 The flow chart of vulnerability ontology data mining method

The association rules of vulnerability ontology system. Different mining techniques can obtain different data mining algorithms, and then to obtain more data association information, at the same time to find out the association rules that are beneficial to the research. The association rules of vulnerability ontology system plays a very important role in data mining, it carries out the process stage of selecting the results, and finally vulnerability ontology and mining algorithms can be linked, which can get the best evaluation results.

For the different users and research management staff, the perspective is far due to the focus of different perspectives. Therefore, according to the different vulnerability ontology, using the different association rules carries out data mining. In this paper, the object of the research is research management personnel, the use of the association rules is ADARF filter, which is mainly divided into the following sections:

(1) In accordance with the association rules, the system carries out structure formation to extract all rules algorithm, each type of vulnerability ontology type encloses the convenient distinction label;

(2) Design of the system hierarchy, the hierarchy is L_0 to L_4 from top to bottom, finding the difference between the level of each node. The difference is $L_d=L_i-L_0$ (I is 0,1,2,3,4);

(3) According to the research management object, this paper develops the limits of information attention, to determine whether the L_d meets the rule algorithm; if satisfied, it is an effective algorithm, the system will automatically save.

Experiment

In order to verify the feasibility of this system, this experiment uses the system analysis method to compare the vulnerabilities created in the experiment. According to the existing data rules, the system can mine the correlation degree of different vulnerabilities in the same system. On the five levels, the system is the implementation of data parallel computing method, and according to the degree of vulnerability definition, the system sets up the mining model and data analysis. In this experiment, the minimum confidence level is set to 0.5. Experimental results are shown in Table 1.

Table 1. The experimental results of vulnerability ontology system association rules

No.	Category	Credibility (before)	Credibility (after)
1	Verification encryption degree	3	10
2	Verify that the configuration error is encrypted	1	6
3	Verify permission	2	8
4	Verification access restrictions	1	3
5	Verification resource management	5	11

From the above table, the credibility of the project is obviously improved, after using vulnerability ontology system, and the system mining association rules very high.

Summary

The computer is the large-scale popularization, the network attack method is more diverse and complex, so the system software security has become the focus of world people's attention. According to the data mining technology, this paper uses the ADARF association rule filter to establish the information security vulnerability ontology, it mines a lot of association rules and finds a lot of effective information to meet the needs of users. Experiments show that the system can find and reduce the potential information security risk in the system, and the system is more efficient and feasible, which has higher practical significance.

References

[1] J.Y. Xu, X.Q. Zhang, C.H. Gu. Analysis of information security vulnerabilities based on ontology. Journal of East China University of Science and Technology, 2014(1):125-131.

[2] J. Bai. The application of data mining technology in the electronics store. HeFei University of Technology, 2014: 1-13.

[3] Z.G. Pan. Research on Web news text automatic classification based on the content, 2013: 7-9.

[4] D.G. Li. Study of network security risk assessment Based on agent and ontology. North China Electric Power University, 2013: 1-9.

[5] Y. Deng. Research on data warehouse and data mining technology theory as well as its application. Gansu Industrial University, 2013: 1-12.

[6] X. Chen. Research on mining association rule algorithm. Electronic Science and Technology University, 2013: 3-15.

[7] C.S. Cheng. Web mining technology and its application. Shandong University, 2014: 2-11.

[8] Z.G. Yuan. The massive scientific data mining based on Bayesian theory. Electronic Science and Technology University, 2014: 1-6.

[9] Y.B. Ouyang. Clustering method and its application for spatial data mining. Electronic Science & Technology University, 2014: 2-13.

[10] M. Zhang. Application of data mining in statistical work. Heilongjiang University, 2013: 2-9.

[11] H.Y. Yin. Research on computer virus prevention under the network environment. Net world, 2013(4): 22-24.

[12] S.H. Zheng. Research on the distributed sequential pattern mining based on massive traffic information flow. Hangzhou Electronic Science & Technology University, 2014: 12-16.

[13] J.Y. Xu. Research on data mining technology based on information security vulnerability ontology. East China University of Science and Technology, 2013: 11-13.

[14] M. Li, Y. Deng. The study of Web data mining techniques and its tools. Computer engineering and applications, 2013(20): 92-94.

[15] J.X. Bi, Q.S. Zhang. Overview of the association rule mining algorithm. China Engineering Science, 2014(4): 88-94.

[16] L.C. Cai, H.Y. Zhao, X.J. Li. Apriori algorithm overview of association rule mining. Journal of Sichuan University of Science and Engineering, 2014(1): 66-70.

[17] W.J. Luo. Research and implementation of visualization technology in data mining. Electronic Science and Technology University, 2013: 1-12.

[18] X. Zhong. Research and application of spectral clustering in outlier data minin. Chongqing University, 2013: 27-32.

[19] C. Liu. Research on vulnerability mining technology based on protocol analysis. Beijing University of Posts and Telecommunications, 2013: 7-12.

[20] H.L. Feng. Research on some key algorithms in data mining. Xi'an University of Science and Technology, 2013: 3-9.

[21] P. Chen. Research on some key technologies in data mining grid. Beijing University of Posts and Telecommunications, 2013: 10-18.

[22] Y.X. Yu. Research on personalized service based on web log mining. East China Normal University, 2012: 1-6.