The Key Techniques of the Network Anomaly Detection Based on Data Mining

HE Xiaobo[1, a]

[1]Chongqing water resources and electric engineering college, Chongqing 402160,China

[a]hexiaobo@126.com

Keywords: Network Anomaly Detection, Data Mining, Intrusion Detection

Abstract. Network realized the information sharing, including individuals, businesses and the government and the whole society. With the expanding of network application scope, for all kinds of network attack and destruction is growing. Computer network security is an international problem, worldwide each year because of the computer network security system was damaged caused economic losses amounted to tens of billions of dollars. Network security is a very complicated problem, it involves many aspects of network engineering, such as network technology, network protocol, intrusion detection system is built and encryption technology, etc., this thesis mainly studies the network intrusion detection system based on data mining is the subject.

Introduction

With the continuous development of network technology, network attack means also emerge in endlessly, people put forward higher request to network security. Intrusion detection is a kind of active safety technology, as an important component of information security architecture, research on intrusion detection technology has caused more and more people's attention [1]. Traditional intrusion detection system to network packets of fetching pattern matching with the rules in the library rules, with the continuous improvement of network bandwidth, there is a huge challenge on test efficiency. And the rules in the library according to the knowledge of the "experts" manual coding, can only detect known attacks, while the unknown attack or variations of known attacks powerless, adaptive differential [2].

Network security is the cornerstone of the development of the network information. How to effectively prevent and detect the attack on network has become a key technology of computer network. Data mining technology can find various intrusion behavior from the massive audit data and the normal behavior pattern, the introduction of the technique of data mining to intrusion detection, will effectively improve the efficiency of intrusion detection system detection and adaptive.

Intrusion detection technology for the network

Intrusion detection is found for invading behavior. It collects the key points of the computer network or system information and analysis, to determine whether the network or system in violation of the security policy or be attacked. Intrusion detection and the difference between its traditional security technology can advance warning and afterwards discovered, so the larger extent, improved the security of the network system [3]. Intrusion detection is mainly composed of three steps: data acquisition, data analysis and response. Data acquisition is responsible for the collection of multiple point in the network system of information, mainly including system log, network packets and application data. Data analysis is to point to analyze the collected information and try to find the intrusion behavior, the main methods are statistics, pattern matching, data mining, such as protocol analysis. Response is when detect intrusion or attack inform user and take appropriate response measures, such as screen display, voice alarm, mail, system logging, etc., can even to active timely blocking illegal connection [4].

US Internet project working group on intrusion detection (IDWG) from architecture, communication mechanism, language format, the respect such as API specification for the standard

of the IDs and launched a series of draft proposal, put forward the common intrusion detection framework. This model will be a basic intrusion detection system is divided into event generator, event generators), event analyzer (event analyzers), event database (event databases) and response unit (the response units) four parts, as shown in figure 1.

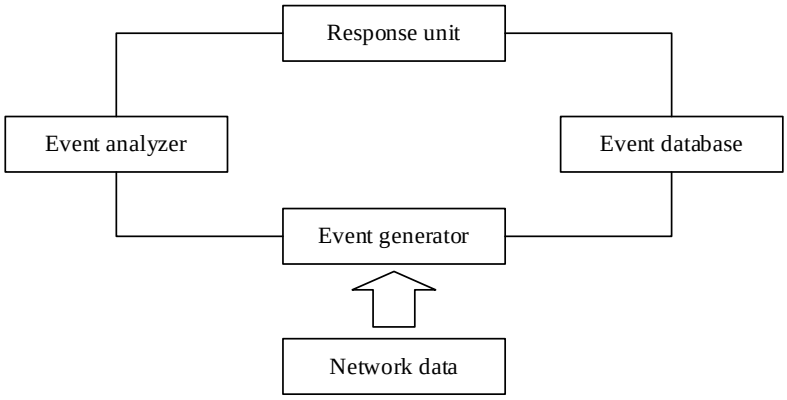

Figure 1.The model of intrusion detection system

Application of data mining in the network anomaly detection

Data mining is a complex process, can be roughly divided into three stages: data preparation, data mining and discovery mode of interpretation and evaluation [5]. Figure 2 describes the basic process of data mining.

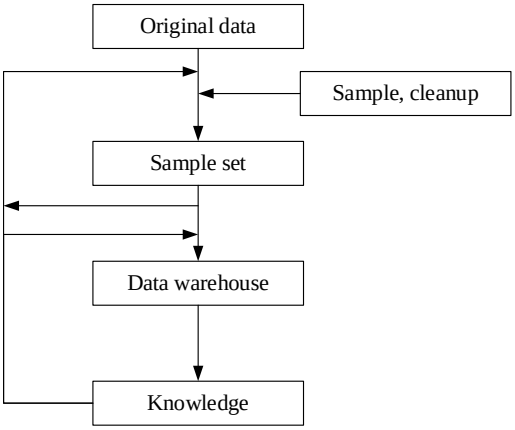

Figure 2.The basic process of data mining

Data preparation. The data preparation phase is divided into three sub steps: data selection, data preprocessing and data transformation. Data selection is to determine the purpose of discovery task operation object, namely target data. Data preprocessing might include eliminate noise data, calculate the default value is derived, filter duplicate records, the conversion of data types and other operations. Data transformation is to eliminate data dimension or dimension reduction, that is, from the initial features find useful to analysis the real, to reduce the digging characteristics or the number of variables to consider.

Data mining. Data mining stage must first determine the mining tasks or mining purposes, such as the discovery of rules of classification, clustering and pattern, and so on. Determine the need to select a suitable mining algorithm after mining tasks. Select mining algorithm is mainly from two factors: one is the data characteristic; The second is the user's requirements and system environment. To determine the right after the mining algorithms can be the operation of data mining.

Find the pattern evaluation and interpretation. Excavated pattern data mining stage, after the user or machine evaluation, there may be redundant or irrelevant pattern, it must be eliminated;

May also be excavated mode cannot meet the needs of users, must return to knowledge discovery. Knowledge discovery process should pay attention to the following three points: 1) data mining is only one step in the KDD process. Data mining quality depends on the validity of the data mining algorithm and used for data mining the data quality and quantity of the two factors.2) the whole process of data mining is a process of constant feedback.3) visualization at all stages of the data mining has a very important role.

According to the analysis of the basic steps of data mining, intrusion detection based on data mining process is divided into data preparation phase and detection phase, the process of intrusion detection based on data mining is shown in figure 3.

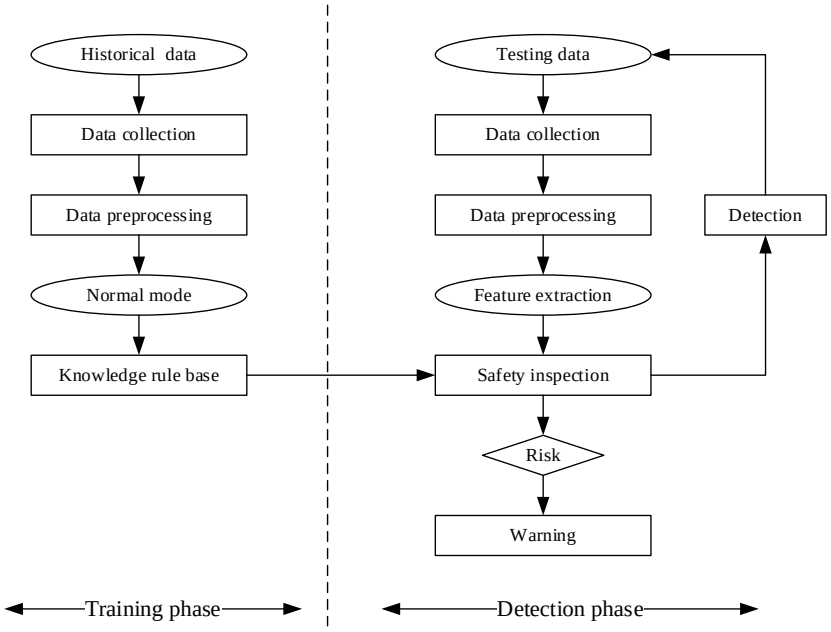

Figure 3.The process design of intrusion detection based on data mining

The network anomaly detection system structure design based on data mining

For network anomaly detection of intrusion detection systems including event generator module, the clustering analysis module, event analyzer module, database module, response unit module and association rules analysis module six big modules, as shown in figure 4.

1) Event generator module. Capture data packets from the network and will capture the packets from the bottom up analysis along the protocol stack decoding and form the corresponding packet class, for subsequent processing module.

2) Clustering analysis module. It is based on the improved algorithm to build network normal behavior patterns.

3) Event analyzer module. It is the core module of the whole system, and analysis the center of the packet. The anomaly detection and misuse detection engine of two modules. The anomaly detection engine use to build a good network normal behavior model to test the packet after pretreatment, normal packet filtering.

4) Events database. The module to maintain a rule base, with a certain grammatical form save intrusion detection rules, and provide a basis for pattern matching analysis of misuse detection engine.

5) Response unit module. Triggered when the judgment in intrusion behavior, carry on the alarm and logging or database records.

6) Association rules analysis module. To cook according to the package of association rule mining, find intrusion patterns of behavior and translated into intrusion rules added to the repository.

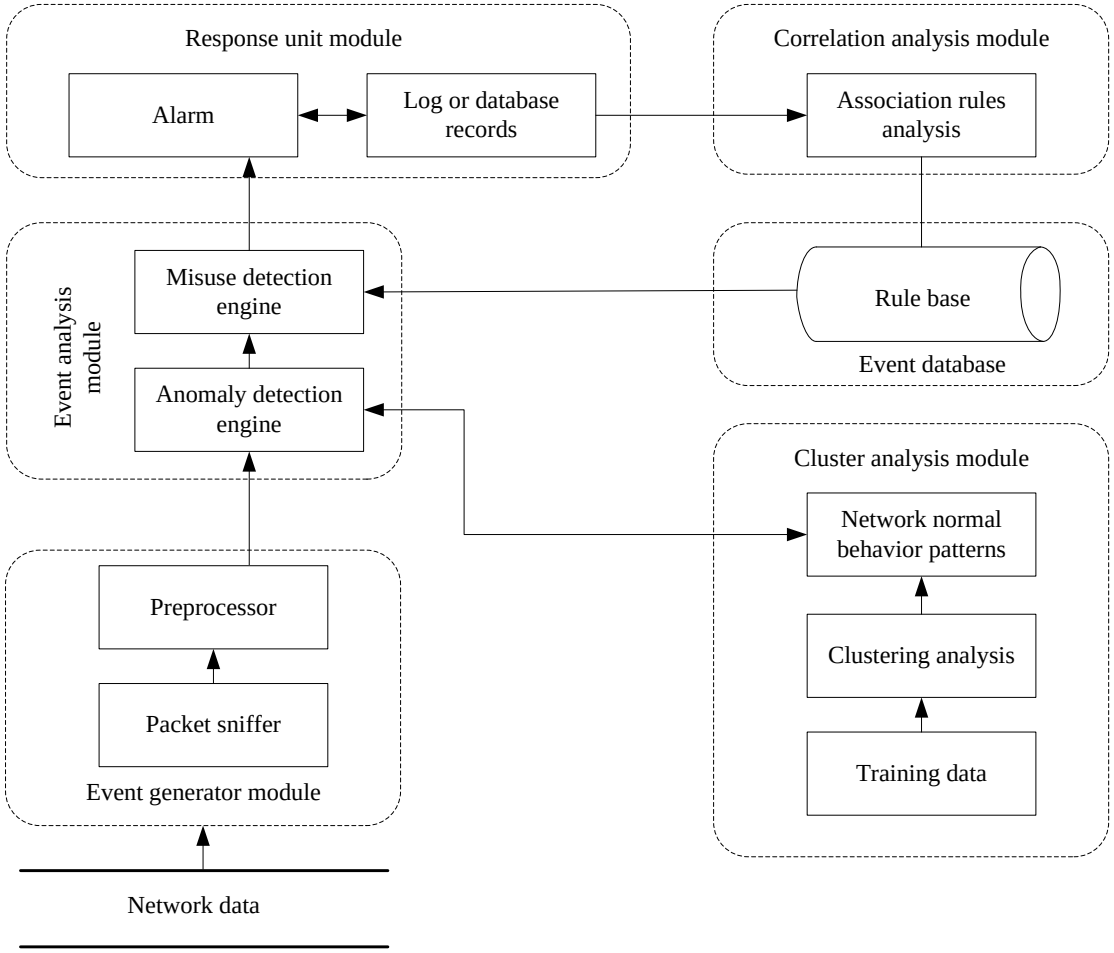

Figure 3. The network anomaly detection system based on data mining

Conclusion

With the continuous development of computer network technology, network security problem is increasingly outstanding. The traditional network security protection technology is difficult to cope with new problems in the field of network security, intrusion detection technology as a kind of active safety technology, can quickly find the attack behavior, tracking the invasion, the implementation of alarm response. Intrusion detection system can effectively ensure the security of the network, in computer network security system is becoming more and more important.

References

[1] Garcia-Teodoro P, Diaz-Verdejo J, Maciá-Fernández G: computers & security, 2009, 28(1): 18-28.

[2] Aydın M A, Zaim A H, Ceylan K G: Computers & Electrical Engineering, 2009, 35(3): 517-526.

[3] Muniyandi A P, Rajeswari R, Rajaram R: Procedia Engineering, 2012, 30: 174-182.

[4] Corchado E, Herrero Á: Applied Soft Computing, 2011, 11(2): 2042-2056.

[5] M.D. Dikaiakos, D. Katsaros, and P. Mehra: Internet Computing, Vol. 13(2009) No.5, p. 10