Advances in Computer Science Research (ACSR), volume 61
7th International Conference on Education, Management, Computer and Society (EMCS 2017)

The Research and Application of Multiple Detectors of Bayesian Classifier Fusion Technology

Xin Sui

College of Humanities and Sciences of Northeast Normal University, Chang Chun, 130117, China

Keywords: Network security; Intrusion detection; Dynamic; Bayesian algorithm; Fusion technology

Abstract. Nowadays, people more and more get the attention of the network security problem, and Intrusion Detection technology is a kind of security mechanism, which is a kind of dynamic monitoring, preventing and resisting intrusion behavior. As one of the core technologies of network security, Intrusion Detection system has become an essential infrastructure for most organizations. The paper combing of fusion technology for the lack of Naive Bayesian algorithm, based on the analysis of the collected information and the result of the fusion, give the result of the detection system, and then to determine whether the system has been invaded, and effectively improve the rate of detection of intrusion detection system.

Introduction

As of January 2014, the number of Internet users is close to 2.5 billion people worldwide, including 590 million of Chinese users. At present, web-based services has been extended to all fields of society, the network of the confidential information of the number of substantial growth, this makes the security of information transmission in the network is very important, network security has become especially important. We should generate a variety of network security technology, firewall and intrusion detection, security scanning technology, etc. [1]. Focus in the study of intrusion detection is to capture and record network that exist in a large number of dynamic information, effectively distinguish normal behavior and malicious attacks, to extract the effective detection rules, and found the regularity of intrusion behavior, and thus enhanced the accuracy of intrusion detection alarm [2]. Intrusion detection technology has the characteristics of active defense, has become the research hot spot.

The Analysis Method of the Intrusion Detection Technology

There are two main types of analysis method of intrusion detection technology: misuse detection and anomaly detection.

Misuse detection is based on knowledge. It contains known in advance to create a database, all kinds of intrusion attack, with each kind of specific way to represent each known intrusion attack mode, and then through the collection, analysis, network packet judgment whether against a database of attack pattern matching, and then determine whether attack behavior. Common misuse detection techniques mainly include pattern matching method, the expert system method, state transition analysis method.

Misuse detection technology of high accuracy, low rate of false positives, deficiency is cannot detect the unknown intrusion behavior, thus the non-response rates higher.

Anomaly detection technology is based on the behavior of the intrusion detection technology. The characteristics of the technology first used normal user behavior to establish a model of a normal behavior, then the user's current behavior compared with normal behavior model, the degree of difference are obtained. With a range, or to measure the degree of the difference threshold, the degree of difference is beyond a range or threshold. Divided roughly such kinds anomaly detection technology, quantitative analysis, statistical analysis and detection on the basis of data mining technology, based on artificial neural network intrusion detection technology, detection technology based on genetic algorithm and detecting technology based on naive Bayesian algorithm.

Anomaly detection technology can effectively detect the unknown intrusion attack, but the

performance of the system request is higher.

Bayesian Algorithm and Bayesian Classifier

Intrusion detection system detection in computer network or system is the purpose of illegal intrusion behavior occurs. Research data is collected, judgment and classification, clustering, behavior as well as to the intrusion alarm and response, etc. [3].

Intrusion detection technology based on naive Bayes algorithm is simple and practical, has mature mathematical theory foundation, and the calculation efficiency, greatly improving the accuracy and completeness of the intrusion detection, application in intrusion detection more, also become hot topics in the study of intrusion detection pattern classification.

Bayes' theorem is according to the new information from the prior probability of posterior probability is obtained a kind of method; the key is to use probability according to various forms of uncertainty [4]. Bayesian classification learning algorithm is simple, able to handle large training data set, has quite a good classification prediction effect. When working in the classifier, does not need to be calculated very accurately every case is belong to which kind of posterior probability. Thus effectively simplifies the classifier in the posterior probability calculation problem.

The Bayesian Classifier in the Application of Intrusion Detection System

Classification learning method based on Bayes' theorem can be applied to the intrusion detection field. Using Bayes classification algorithm to represent the process behavior of the system call sequences in classification, need in sequential patterns add a statistic about the model number. Using sliding window handler execution path, automatic statistics out every normal or abnormal patterns in the training data of frequency, the sum of these statistical frequency values determine the number of training data model. On the basis of the frequency values can be concluded that normal behavior subsequence and abnormal behavior subsequence probability and conditional probability in the training data set. Get the probability, after using the Bayesian theorem sequence belongs to category of probability are obtained. The method based on statistical method from the training data set in normal and abnormal sequences corresponding to the sequence of location selection probability distribution, which can more accurately response model of the overall information, makes the forecast of category information more accurate and reasonable.

Naive Bayes classifier is mainly based on a small amount of training data set can estimate the necessary parameters, less sensitive to missing data. Naive Bayesian classification model is simple, easy to implement, classified the advantage of fast speed, high accuracy, is currently one of the most widely used classification model.

Multiple Detectors Based on Bayesian Classifier Fusion Technology

Data fusion technology was first used in military field, 1973 research institutions in the United States funded by the department of defense to carry out the study of sonar signal interpretation system. Since then, especially since the gulf war, dedicated to the data fusion study the dramatic increase in the number of papers and works greatly promote the development of the technology of data fusion. At present, in the industrial control, robotics, air traffic control, marine monitoring and management, and other fields and development in the direction of multiple sensor fusion. More than 30 years, target data fusion technology for military application in industry, agriculture, space, target tracking and inertial navigation etc. is widely attention and application. Technology as a kind of data integration and data fusion processing technology, is of many traditional disciplines and the integration and application of new technology, from a general perspective of data fusion, including communications, computer science, artificial intelligence, neural network pattern recognition, decision theory, signal processing, estimation theory, uncertainty theory and optimization technology, etc. Fusion is the basic function of related, estimation and recognition, with a focus on the estimation and recognition.

Using different detection technology and the model of intrusion detection module, have their respective advantages and disadvantages on the function [5-6]. In order to more effectively identify malicious intrusion attempts in network system and network intrusion detection system can integrate a variety of technologies of intrusion detection, the ability to monitoring and control system of audit information into and analyses system line [7-8]. In the intrusion detection system, use a variety of analysis and detection mechanism, aiming at different full information in system is analyzed, and the result of them to make decision fusion, will effectively improve the detection rate of the whole system, effectively reduce the rate of false positives. Detector data fusion technique, is introduced to research and test results summary or recognition result of the decision problem [9-10]. For multiple detectors monitoring results, the most intuitive judgment method is using the k/N vote method, when N in the detector to determine k intrusion behavior, that the system of intrusion attack (11-12). Multiple detectors based on Bayesian classifier fusion result decision scheme considering the influence of different probe detection performance, expect to get maximum detection rate.

Detection subsystem A1, A2, A3... An of the same by the monitoring system for testing, just because of its test result depends on detection mechanism and analyze the data of the detectors, and test results of independence. The result of the output is to 1, detected intrusion; when the output is 0, detector did not detect intrusion. When N detection system works at the same time, the testing results of the composition of a one-dimensional is vector of length N. By collecting N detection subsystem test results of the vector, and according to the actual behavior is normal for N test results of two kinds of data sets, and use it as training data to create a Bayesian classifier. By using the vector classifier for each test system of test result for decision fusion based on the results of the Bayesian inference classification, simply by comparing various detection subsystem of the test results of vector belongs to the "normal" and "abnormal" the size of the a posteriori probability, you can judge system is really the invasion attack.

The Experiment

Experimental data from the international knowledge discovery and data mining (KDD CUP 1999 tool race of data as the training and testing data, the data contains a variety of simulation, simulation attack (including normal request) is mainly divided into five categories.

By the experiment result of multiple detectors based on Bayesian classifier fusion algorithm than the general Bayesian algorithm has higher accuracy of the classification of the progress and, in a larger extent, effectively improve the accuracy and completeness of the intrusion detection.

Naive Bayes classification algorithm is presented in this paper, on the basis of the multiple detector based on Bayesian classifier fusion algorithm is summarized. By comparison with the result of the experiment proves that the multiple detectors based on Bayesian classifier fusion algorithm on the accuracy and recall rate have obvious increase, the improved algorithm has certain practical value and the value of further research.

References

[1] Wang H F,Gao G Y. The research progress of network intrusion detection[J]. Computer Security, 2014-12: 58-61.

[2] Zhang X,Tao J,Zhang J D. The research of intrusion detection based on the variable precision and rough[J]. Journal of Qingdao University of Science and Technology (Natural Science Edition) , 2014,35(2): 196-199.

[3] Cao D Y. Intrusion Detection Technology[M]. POSTt and TELECOM PRESS, 2011-9: 197.

[4] MICHALSIK R S,CARBONELL J, MICHE L T. Machine Learning: An artificial intelligence approach[M]. San Mateo:Morgan Kaufmann,1983: 463-482.

[5] Guo W J,Sun J Y,Ren J. A distributed intrusion detection system based on data fusion[J]. The Computer Technology and Development, 2006-2.

[6] He X H. Alarm fusion in the distributed intrusion detection system[D]. Beijing Jiaotong University, 2007-3: 100.

[7] Li Y. The design of distributed intrusion detection system architecture based on data fusion and data mining technology[J]. Computer Knowledge and Technology, 2004, 8.

[8] Chen Y. The research of network abnormal behavior detection based on classifier fusion [D]. JiangNan University, 2013-6-1.

[9] Liu Q. Intrusion detection in some applications of neural network and data fusion method [D]. Nanjing University of Science and Technology, 2003-12-1.

[10] Jiang J G. Distributed intrusion detection system and the research and practice of information fusion technology [D]. SiChuan University2003-9-10.

[11] Su L. Data fusion technology in the application of distributed intrusion detection research [D]. GuangDong University of Technology, 2008-4-1.

[12] Cao W W. Intrusion detection system alerts fusion research [D]. Beijing Jiao Tong University, 2014-3-1.