The research of Intrusion Detection based on Support Vector Machine

Li Bo
Network Information Center
Chongqing University of Technology
Chongqing, 400050, China
libo@cqut.edu.cn

Chen Yuan Yuan
Network Information Center
Chongqing University of Technology
Chongqing, 400050, China
cyy@cqut.edu.cn

***Abstract*—Intrusion detection is developed quickly because which has important position in network security. The method of SVM based on statistics learning theory is used in the intrusion detection system, which classifies detecting data efficiently, and achieves the aim that SVM can accurately predict the abnormal state of system. By the use of this method, the limitation of traditional machine learning method is avoided and ensures the stronger extension ability which makes intrusion detection system to have the better detecting performance.**

Keywords-computer network; intrusion detection; normal action; abnormal action; distort rate; miss probability

I. INTRODUCTION

With the development of computer and network, addition of network intrusion affair, people find it's not enough to build security system only from a defense point of view, intrusion detection has been a new generation network security guarantee technology after some traditional security guarantee measure such as "Firewall", "Data Encrypt" etc. Intrusion detection system can be defined the system responds to vicious purpose and action aiming at computers or network resource. In many times it's cost much to build an IDS, career man need to use characters that are refined from analysis of known attacking methods and known system's leak into Misuse Detection, or adopt many statistic analytical methods to do Anomaly Detection, but the expansibility and adaptability aren't strong. For all these reasons, we import SVM into the intrusion detection system, then build more complete character repository to detect intrusion through analyzing physical data to refine normal act characters and summing rules of intrusive action.

II. INTRUSION DETECTION SUMMARY

Intrusion detection is a new network security technology emerging in recent 20 years which can initiatively help computer protects itself from hacker's attack. Moreover, it's an important part of P2DR (Policy Protection Detection Response) dynamic security model. It can find the security policy-offending actions as well as the signs [1] of being attacked through collecting and analyzing information at the key points in either computer network or system. The combination of software and hardware applied in intrusion detection is IDS (Intrusion Detection System). As an active and initiative security preventive technology, intrusion detection can provide real-time protection from inner attack, outer attack and misoperation.etc. besides, it can intercepts and responds to intrusion before network system getting hurt.

III. INTRUSION DETECTION METHOD ANALYSIS [2]

Intrusion detection methods are the point of intrusion detection research. Intrusion detection methods analyze various data resources to judge whether intrusion happens. Dealing with intrusion detection problems from different aspects, these methods use all kinds of technologies to build intrusion detection's normal model or attacking model. Each of them has speciality and scarcity, as the whole we classify them into three: Misuse Detection, Anomaly Detection, and Combination Detection.

A. Misuse Detection

Misuse Detection builds intrusive characteristic repository through describing characters of known system's leak and known attacking model, then filtrates the affair data got from practical detection done to know whether it includes the sign of intrusive action. If found a meeting requirement's match, then it's indicated an attacking action happened once. Misuse Detection's work mode is similar to commercial antivirus software's. It's possible to cause underreport, because Misuse Detection can detect known attacking action in effect, but can do nothing to unknown new attacking action. As the emergence of attack type against new leaks or new attack mode aims as old leaks, to be detected attack needs field experts in that field or other machine learning systems to refine new attack character mode and add it into attack character repository. So the key to the research of Misuse Detection is character repository updating automatically to ensure completeness of system detection capability. Typical Misuse Detection methods include Pattern March, State Transition Analysis, Consequence Model—Based Approach.

B. Anomaly Detection

Anomaly Detection builds normal action mode character repository, and compares normal action character figure with user's practical action, then achieves the aim intrusion can be detected by judging the degree of deviation between user's action and normal action. If the degree of deviation exceeds a given threshold, then it will indicate an illegal attack action.

Anomaly Detection is irrelevant to the system, so it has good detection adaptability, and possesses the ability of detecting unknown attack mode, however it exists problems about high distort rate and bad accuracy. The causes are as follows: ① it's hard to describe system's normal action mode correctly, which leads to a mass of misinformation. ② It's difficult to identify normal action mode's normal deviation (normal mutative action) and

978-0-7695-3906-5/09 $26.00 © 2009 IEEE
DOI 10.1109/ICCCS.2009.43

abnormal deviation (intrusive action), even if system captures some new attack types, it's the same difficult to detect attack action from normal action. So how to keep low distort rate and possess detecting unknown attack mode ability is the problem Anomaly Detection method must face with. Typical Anomaly Detection methods include Statistic Profile—based Approach, Pattern Predicting based, Machine Learning Based.

C. Mixed Detection

Mixed Detection synthesizes the advantages of Anomaly Detection and Misuse Detection. For the reason Anomaly Detection and Misuse Detection are complementary, the integration of the two can learn from others' strong points to offset one's weakness, which improves the whole detecting ability efficiently. The methods include Genetic Algorithm, Neural Network, and Biology Immunity.

IV. SUPPORT VECTOR MACHINE[3]

SVM (Support Vector Machine) is a machine learning method proposed by Vapnik in 1990s, whose core thought is based on Mercer core expansion theorem, which maps sample space to a higher dimensional even an infinite dimensional character space by the use of nonlinear mapping function ϕ, and applies linear learning machine method to resolve problems about high nonlinear classifying and regression. It's easy to see the essence of SVM classifying is the question solving optimal hyperplane.

To suppose a training sample set $\{(x_i,y_i),i=1,2,\ldots,L\}$ with size of L is composed of two classes, if $x_i \square RN$ belongs to the first class, then y_i will be marked plus(y_i=1), if it belongs to the second class , then y_i will be marked minus (y_i=-1). The purpose of learning is to construct a discriminate function, and classify test data accurately as far as possible.

When training sample set is nonlinear, by using a nonlinear mapping function ϕ we match training sample set data x to a higher linear character space in which the dimension may be infinite dimensional, construct optimal classifying hyperplane, and get classifier's discriminate function. So in the case of nonlinearity, classifying hyperplane is

$$w \bullet \varphi(x_i) + b = 0 \ ; \qquad (1)$$

Discriminate function is

$$y(x) = sign[w \bullet \phi(x) + b] \ ; \qquad (2)$$

Optimal classifying hyperplane question is described as

$$\left.\begin{aligned} &\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{l}\xi_i \\ &s.t. \quad y_i(w \bullet \phi(x) + b) \geq 1 - \xi_i \\ &\xi_i \geq 0, i = 1,2,...,l \end{aligned}\right\} \qquad (3)$$

C is penalty parameter, Bigger is C denotes penalty to wrong classifying is bigger. Use Lagrangian multiplier method to solve the linear constraint quadratic programming question, as

$$\left.\begin{aligned} &\max_{\alpha}\{L_D = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j \phi(x_i) \bullet \\ &\phi(x_j) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j K(x_i,x_j)\} \\ &s.t. \quad 0 \leq \alpha_i \leq C \\ &\sum_{i=1}^{l}\alpha_i y_i = 0 \end{aligned}\right\}$$

(4)

In the formula: $K(x_i, x_j) = \phi(x_i) \bullet \phi(x_j)$ is called core function, discriminate function

$$y(x) = sign(\sum_{x_i \in SV}\alpha_i \, y_i K(x_i, x) + b \ ) \qquad (5)$$

is thresh old value b is

$$b = \frac{1}{N_{NSV}}\sum_{x_i \in SV}(y_i - \alpha_i \, y_i K(x_j, x_i)) \qquad (6)$$

$K(x_i, x) = \tanh(kx_i \bullet x + \theta)$ 。

As we know from (4) ~ (6), by the use of nonlinear mapping function sample data can be matched to higher dimensional even infinite dimensional character space in which we can construct optimal classifying hyperplane, but it needn't to compute nonlinear function at the time solving optimal question and computing discriminate function, only compute core function, then avoid character space dimension disaster. Core function's election must meet the requirement [4] of Merce. Familiar core functions include linear function $K(x_i, x) = x_i \bullet x$, multinomial function $K(x_i, x) = (x_i \bullet x + 1)^d$, radial basis function $K(x_i, x) = \exp(-\|x - x_i\|^2 / \delta^2)$, Multilayer preceptor function $K(x_i, x) = \tanh(kx_i \bullet x + \theta)$.

V. INTRUSION DETECTION SYSTEM BASED ON SUPPORT VECTOR MACHINE

A. SVM and Anomaly Detection

In fact, Intrusion Detection can be vested in classifying problem that is to divide abnormal state data from normal state data during detecting. But as data relationships are complex in intrusion detection and intrusive sample gained is limit, moreover the favorable characters of SVM make it a better anomaly detection technology in the field of Intrusion Detection. Even under the lack of enough prior background knowledge, it still can ensure higher accurate rate.

Similar to other abnormal detection methods, the work that adopts SVM to detect intrusion is classified two steps: 1.training; 2.detecting. In the phase of training, if appropriate SVM core function and performance parameter are elected, under some SVM training

arithmetic using sample data to train classifier, we can gain SVM decision function. In the phase of detecting, firstly normalizing various data and transforming to the input vector format that can be accepted by SVM, then using decision function got in step backward to classify input vector, the result we get is detection conclusion.

B. SVM system model[5]

Intrusion Detection System based on SVM via SVM classifying technology identifies system calls and parameters as two classes which are normal and abnormal. System frame as shown in Figure1, in which it includes five parts: data collection, data preprocessing, data-base, SVM classifier and system decision.

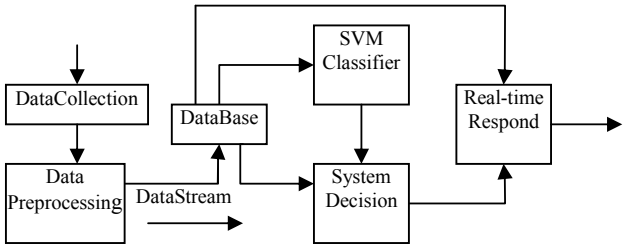

Figure 1. SVM system model

Data collection module collects system calls and parameters, sends these sequences into data preprocessing module which distills short sequences from system calls and parameters and makes them digital vectorization, forms character vector can be solved by SVM, then inputs standard data after preprocessing into data-base. Data-base conserves training data after preprocessing, real-time data and detection result. Training process of SVM is using standard data of data-base to train SVM classifier, which includes determination of SVM model and determination of core function's parameter in SVM etc. Via training it gets SVM classifier used to decide. In the final, it uses the deciding process of SVM classifier to identify a new connection, saves the result into data-base and transports to response module.

C. Training sample

For SVM's parameters gained by training, so it needs two kinds of training samples that are normal short sequence sample and abnormal short sequence. Using sliding window with size of L to scan known normal system call execution trace, it can get normal system call short sequence sample. For only a little part of program occupied by intrusion illegal action, so abnormal short sequence makes up only a small section of abnormal executing trace. When using sliding window with size of L to scan abnormal execution trace, it can get system call short sequences list including not only normal short sequences but also abnormal short sequences.

Compared this set of short sequences with acquired normal short sequences, the system call short sequences different from normal short sequences constitute abnormal short sequence samples. If we get a sample, mark all the short sequence samples as the form of (x_1, y_1) , (x_2, y_2) , □, $(x_1, y_1) \in Rk \times \{\pm 1\}$. Among them normal sample y_i is +1, abnormal sample y_i is -1, gained normal and abnormal system call short sequence training data samples as shown in table 1:

TABLE I. NORMAL AND ABNORMAL SYSTEM CALL SHORT SEQUENCE TRAINING DATA SAMPLE

short sequences： ($x_i \in R6$)						**style(y_i)**
21	***3***	***7***	***10***	***12***	***7***	***abnormal： -1***
4	***2***	***62***	***69***	***5***	***141***	***normal：+1***
3	***66***	***66***	***8***	***140***	***66***	***normal： +1***
4	***5***	***11***	***9***	***6***	***6***	***abnormal： -1***
…	…	…	…	…	…	…

VI. CONCLUSIONS

The method of SVM based on statistics learning theory is used in the intrusion detection system, which classifies detecting data efficiently, and improves the efficiency of distilling system's rule, thereby greatly reduces system's distort rate and miss probability, provides good basis for intrusion detection's real-time detection. So SVM is very fitful for the real-time detection in which action characters change frequently and attack action emerges endlessly, which will be applied more and more widely in network security.

REFERENCES

[1] Zhao Xiao Lin, Peng Zhu Lin, Wang Ya Bin. Network Security Technology tutorial [M]. Beijing: National defense industry Press. 2002.1,245—245.

[2] Ma Zhan Fei, Zheng Xue Feng, Research of Intrusion Detection System Based on Computer Network [J]. Microcomputer Information(integration of management and control), 2006.12-3

[3] Du Shu Xin, Wu Tie Jun, Support Vector Machines for pattern recognition [J]. Journal of Zhejiang University (Engineering Science), 2003.37(5)

[4] VAPNIK V N. The nature of statistical learning[M]. Berlin; Springer, 1995.

[5] Bi Xiao Dong, Research of Intrusion Detection Model Based on 1 class Support Vector Machine [J]. Journal of Shandong Normal University (Natural Science) 2006.21(4).