한국산학기술학회논문지
Vol. 10, No. 4, pp. 811-817, 2009

Design of NePID using Anomaly Traffic Analysis and Fuzzy Cognitive Maps

Hyeock-Jin Kim[1], Sang-Ryul Ryu[1] and Se-Yul Lee[1*]
[1]**Dept. of Computer Science, Chungwoon University**

비정상 트래픽 분석과 퍼지인식도를 이용한 NePID 설계

김혁진[1], 류상률[1], 이세열[1*]
[1]청운대학교 컴퓨터학과

Abstract The rapid growth of network based IT systems has resulted in continuous research of security issues. Probe intrusion detection is an area of increasing concerns in the Internet community. Recently, a number of probe intrusion detection schemes have been proposed based on various technologies. However, the techniques, which have been applied in many systems, are useful only for the existing patterns of probe intrusion. They can not detect new patterns of probe intrusion. Therefore, it is necessary to develop a new Probe Intrusion Detection technology that can find new patterns of probe intrusion. In this paper, we proposed a new network based probe intrusion detector(NePID) using anomaly traffic analysis and fuzzy cognitive maps that can detect intrusion by the denial of services attack detection method utilizing the packet analyses. The probe intrusion detection using fuzzy cognitive maps capture and analyze the packet information to detect syn flooding attack. Using the result of the analysis of decision module, which adopts the fuzzy cognitive maps, the decision module measures the degree of risk of denial of service attack and trains the response module to deal with attacks. For the performance evaluation, the "IDS Evaluation Data Set" created by MIT was used. From the simulation we obtained the max-average true positive rate of 97.094% and the max-average false negative rate of 2.936%. The true positive error rate of the NePID is similar to that of Bernhard's true positive error rate.

요 약 IT 시스템 기반의 네트워크 환경의 급속한 발전은 지속적인 연구방향의 중요한 이슈의 결과이다. 침입시도 탐지는 관심분야의 하나인 것이다. 최근에 다양한 기술을 기반으로 하는 침입시도탐지들이 제안되고 있으나 이러한 기술은 여러 형태의 침입시도의 패턴 중에 한가지 형태 및 시스템에 적용이 가능한 것이다. 또한 새로운 형태 침입시도를 탐지하지 못하고 있다. 그러므로 새로운 형태를 인식하는 침입탐지 관련 기술이 요구되어 지고 있다. 본 연구에서는 퍼지인식도와 비정상 트래픽 분석을 이용한 네트워크 기반의 침입탐지기법(NePID)을 제안한다. 이 제안은 패킷분석을 통하여 서비스거부공격과 유사한 침입시도를 탐지하는 것이다. 서비스거부공격은 침입시도의 형태를 나타내며 대표적인 공격으로는 syn flooding 공격이 있다. 제안한 기법은 syn flooding을 탐지하기 위하여 패킷정보를 수집 및 분석한다. 또한 퍼지인식도와 비정상 트래픽 분석을 적용하여 판단모듈의 분석 결과를 토대로 기존의 서비스 거부 공격의 탐지 툴과의 비교분석을 하였으며 실험데이터로는 MIT Lincoln 연구실의 IDS 평가데이터 (KDD'99)를 이용하였다. 시뮬레이션 결과 최대평균 positive rate는 97.094% 탐지율과 negative rate는 2.936%을 얻었으며 이 결과치는 KDD'99의 우승자인 Bernhard의 결과치와 유사한 수준의 값을 나타내었다.

Key Words : Probe Intrusion Detection, False Errors, Anomaly Traffic, Patterns Analysis, Fuzzy Cognitive Maps

1. Introduction

The rapid growth of network in information systems has resulted in the continuous research of security issues. One of the research areas is probe intrusion detection that many companies have adopted protect their information

*Corresponding Author : Se-Yul Lee(pirate@chungwoon.ac.kr)
Received March 03, 2009 Revised April 17, 2009 Accepted April 22, 2009

assets for several years. In order to address the security problems, many automated intrusion detection have been developed. However, between 2005 and 2008, more than 100 new attack techniques were created and published which exploited Microsoft's Internet Information Server, one of the most widely used web servers. Recently, severally Intrusion Detection has been proposed based on various technologies. A "false positive error" is and error that IDS sensor misinterprets on or more normal packets or activities as an attack. IDS operators spend too much time on distinguishing events. On the other hand, a "false negative error" is an error resulting from attacker is misclassified as a normal user. It is quite difficult to distinguish intruders from normal users. It is also hard to predict all possible false negative errors and false positive errors due to the enormous varieties and complexities of today's networks. IDS operators rely on their experience to identify and resolve unexpected false error issues.

Recently, according to the CERT-CC (Computer Emergency Response Team Coordination Center), hacking is increasing about 300% each year. A variety of hacking techniques are known : DoS, DDoS(Distributed DoS), PDoS(Permanent DoS), Probe Attack, Vulnerability Scan Attack and others. Among them, PDoS, Vulnerability Scan Attack and Probe Attack are the three most frequently used methods. Port scan or vulnerability of network as abnormality intrusion of network is based on anomaly probe detection algorithms such as scanlogd[1], RTSD (Real Time Scan Detector)[2], and Snort[3]. Such open source programs have some problems in invasion probe detection. That is scanlogd and RTSD can not detect slow scan, while Snort does not provide open port scan. Therefore, a new algorithm that can provide slow scan and open port scan is required.

The main objective of the paper is to improve the accuracy of intrusion detection by reducing false alarm rate and minimize the rate of false negative by detecting unexpected attacks. In an open network environment, intrusion detection rate is rapidly improved by reducing false negative errors rather than false positive errors. We propose a network based probe detection model using the fuzzy cognitive maps that can detect intrusion by the DoS attack detection method. A DoS attack appears in the form of the probe and syn flooding attack, which is a typical example. The syn flooding attack takes advantage of the vulnerable 3-way handshake between the end-points of TCP[4-7]. The proposed NePID[8] captures and analyzes the packet information to detect syn flooding attack. Using the results of detection module, which utilizes the fuzzy cognitive maps, the detection module measures the degree of risk of the DoS and trains the response module to deal with attacks[6,7].

The rest of this paper is organized as follows. The background and related work is summarized in Section 2. Section 3 describes the proposed new NePID model. Section 4 illustrates the performance evaluation of the proposed probe intrusion detection model. Conclusions and future work are presented in Section 5.

2. Related work

Previous studies of DoS attack detection can be divided into three categories : attack prevention, attack source trace-back and identification, and attack detection and filtering. Attack prevention obviously provides avoidance of DoS attacks. With this method, server system may be securely protected from malicious packet flooding attack. There are indeed known scanning procedures to detect them based on real experience[9,10]. Attack source trace-back and identification is to identify the actual source of packet sent across network without replying to the source in the packets[11]. Attack detection and filtering are responsible for identifying DoS attacks and filtering by classifying packets and dropping them[12]. The performance of most of DoS detection is evaluated based on false positive error and false negative error. The detection procedure utilizes the victim's identities such as IP address and port number. Packet filtering usually drops attack packets as well as normal packets since both packets have the same features. Effectiveness of this scheme can be measured by the rate of the normal packet which is survived in the packet filtering. Among these schemes, attack prevention has to recognize how DoS attack is performed and detect attack pattern using predefined features[13]. Therefore, when a new attack detection tools are developed, new features that detect the pattern of attack needs to be defined. Current IP trace-back solutions are not always able to trace the source of the packets. Moreover, even though

the attack sources are successfully traced, stopping them from sending attack packets is another very difficult task.

A DoS attacked traffic is quite difficult to distinguish from legitimate traffic since packet rates from individual flood source are usually too low to catch warning by local administrator. Thus, it is efficient to use inductive learning scheme utilizing the Quinlan's C4.5 algorithm approach to detect DoS attack[14]. Inductive learning systems have been successfully applied to the intrusion detection. Induction is formalized by inductive learning using decision tree algorithm which provides a mechanism for detecting intrusion. The key idea of this approach is to reduce the rate of false errors. The false error rates of the known intrusion detection schemes are summarized in Table 1.

As shown in Table 1, FSTC(False Scan Tool and Clustering) provides the largest false negative error, while the Fuzzy ART scheme provides the smallest false negative errors and the largest false positive error. In the meantime, Inductive Learning System provides moderate false negative and false positive error on the average. From the above results, it is highly recommended to develop a new DoS detection scheme based on fuzzy cognition.

[Table 1] False errors of Intrusion Detection Methodology[2]

Methodology	False Negative Error	False Positive Error
FSTC	22.65%	20.48%
Inductive Learning System	9.79%	9.10%
K-Means (Average Value)	9.37%	20.45%
Fuzzy ART (ρ = 0.9)	6.03%	38.73%

3. NePID Model

3.1 NePID Algorithm

The NePID(Network Probe Intrusion Detector) model is a network-based detection scheme that utilizes network data to analyze packet information. Based on the analysis of each packet, probe detection is performed. In order to determine intrusion detection, various features of packet is utilized including source IP address, source port number, destination IP address, destination port number, flags, data size, timestamp, and session pattern as given by (1).

Packet X = (src_ip, src_port, dst_ip, dst_port, flag, data, timestamp, pattern,···) (1)

Now it is needed to quantize each feature parameter based on comparison criterion to determine attack detection. The procedure to assign effect values can be summarized as follows.

[state 1] *Feature Equality*

$$FE(x) = \begin{cases} 0(x \neq a) \\ 1(x = a) \end{cases}$$

a:standard, x:comparison

[state 2] *Feature Proximity*

$$FP(x) = \frac{k}{|x-a|}$$

a:standard, x:comparison, k:constant

[state 3] *Feature Separation*

$$FS(x) = k|x-a|$$

a:standard, x:comparison, k:constant

[state 4] *Feature Covariance*

$$FC(x,y) = |\mathrm{cov}(x(t), y(t))|$$

x, y:comparison, t:time, cov():degree of dispersion

[state 5] *FeatureFrequency*

$$FF(x) = \log_2 \frac{1}{\Pr(x)}$$

Pr(x):x's probability

Using the above state variables, the total degree of abnormality for a packet can be calculated as in (2).

$$A_{total}(x) = \omega_1 A_1 + \omega_2 A_2 + \ldots + \omega_n A_n = \sum_{i=1}^{n} \omega_i A_i \quad (2)$$

$A_{total}(x)$: Abnormality per packet

ω_i : Weight value of packet

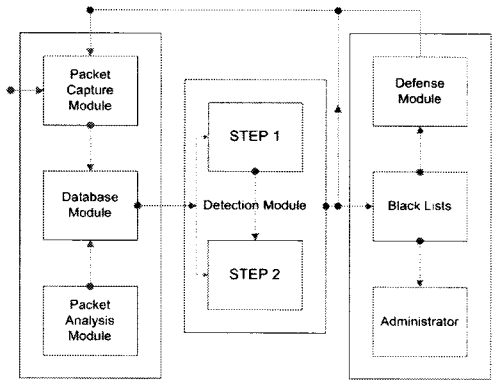
A_i : Abnormality of packet

n : Total feature number of abnormality

If the total degree of abnormality for a packet is greater than the threshold of attack attempt, the associated packet is classified as abnormal.

3.2 NePID Architecture

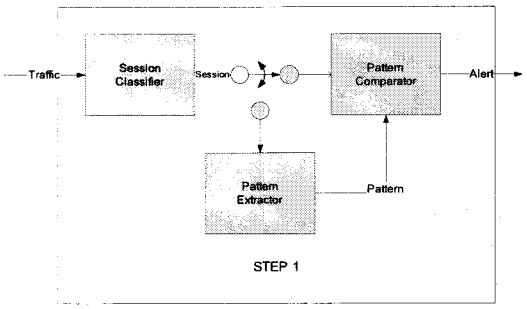
The NePID architecture consists of network-based intrusion detection system and monitoring tool as shown in Fig. 1[5,9]. As monitoring tool(Z-monitor 1.0), a protocol analyzer(Wireshark 1.0.5) is used, whereas the detection system(Intel core2duo E4500, 2G Memory)is directly connected to the router(Cisco 1750), which interconnects 100Mbps LANs. The NePID algorithm is obviously implemented on the detection system.



[Fig. 1] NePID architecture

The NePID adopts the problem solving methodology which uses previous problem solving situations to solve new problems. The model does preprocessing by packet analysis module and packet capture module. The packet capture module captures and controls packet. The packet capture module does realtime capturing and packet filtering b using the monitoring tool of Detector4win ver. 1.2. In the packet filtering processing, packets are stored according to the features which distinguish between normal packets and abnormal packets. The packet analysis module stores data and analyzes half-open state. After storing packets, the packets, which are extracted by audit record rules in the packet analysis module, are sent to the detection module.

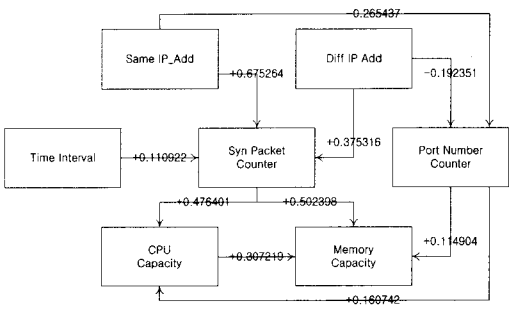
The session classifier takes packet of the traffic and checks whether or not the source is the same as the destination. There is a buffer for the specific session to be stored. And, if the next packet is arrived, it is stored in the correspond buffer. If all packets of the corresponding buffer are collected, all packets of the corresponding buffer are output as on session. The output session becomes an input to the pattern extractor or pattern comparator according to action mode. The action mode consists of learning mode and pre-detection mode. The output session from the session classifier is sent to the pattern extractor in the learning mode and to the pattern comparator in the pre-detection mode. Fig. 2 is the block diagram of the STEP 1.



[Fig. 2] A Block Diagram of STEP 1

As the variable events dependent on the detection module, we can set the identity of IP address, the time interval of half-open state, the rate of CPU usability, the rate of memory, and syn packet. For example, the weight between the two nodes is bigger than 0 since the rate of CPU usability increases in proportion to the size of syn packet.

In Fig. 3, each rectangular box represents feature event, while each number denotes effect value in FCM.



[Fig. 3] Weight value of Path using FCM in STEP2

4. Performance Evaluation

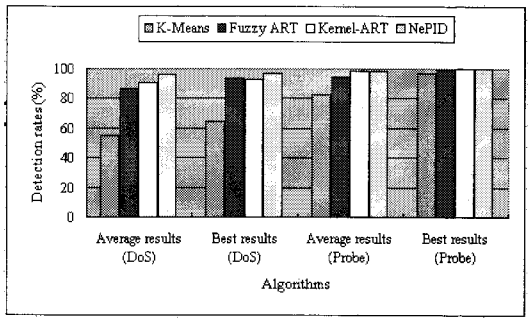
For the performance evaluation of the proposed NePID model, we have used the KDD data set (Knowledge Discovery Contest Data) by MIT Lincoln Lab, which consists of labeled data (training data having syn flooding and normal data) and non-labeled data (test data). Since the TCP syn flooding attacks come from abnormal packets, detection of abnormal packets is similar to detection of syn flooding attacks in TCP networks.

The best detection and false error rates are summarized in Table 2. The simulation results for the connection records of DoS attacks are collected for 10 days. The average rate of true positive is measured of 97.064%. According to the KDD'99 competition results, the best rate of the Bernhard's true positive is known as 97.1%[15].

[Table 2] Best detection and error rates

Day	True Positive	False Positive	True Negative	False Negative
Day1	95.623%	0.000%	100.000%	4.377%
Day2	87.861%	0.000%	100.000%	12.139%
Day3	96.098%	0.001%	99.999%	3.902%
Day4	99.569%	0.000%	100.000%	0.431%
Day5	100.000%	0.000%	100.000%	0.000%
Day6	98.930%	0.000%	100.000%	1.070%
Day7	100.000%	0.001%	99.999%	0.000%
Day8	87.701%	0.000%	100.000%	12.299%
Day9	100.000%	0.000%	100.000%	0.000%
Day10	97.917%	0.000%	100.000%	2.083%
Average	97.064%	0.000%	99.999%	2.936%

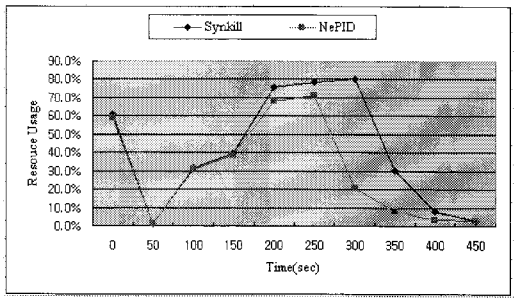
Comparing Bernhard's true positive rate with that of NePID, we realized that the result of NePID is as good as Bernard's. In addition, the false negative rate of the proposed scheme, 2.936%, is considerably smaller than that of the Bernhard's, 3.91%.



[Fig. 4] Detection rates of DoS vs Probe

Fig. 4 illustrates the performance of four different detection algorithms for both DoS and probing. The key difference between NePID and the others is that the former is resource based probe detection algorithm, whereas the latter's are basically rule-based detection algorithms. Thus, the proposed algorithm is able to detect probe regardless of input patterns and the number of features. The key advantage of the NePID over the other algorithms is the ability of real-time update of effect values in FCM. Therefore, as shown in Fig. 4, the proposed NePID algorithm outperforms the other algorithms in both DoS and probe.

In order to evaluate the performance from the viewpoint of resource usage, system resource usage of the NePID is compared to that of Synkill, which is a well-known syn flood attack detection tool developed by Purdue University[16].



[Fig. 5] Comparison of system resource usage

Fig. 5 shows the system resource usage of both Synkill and NePID when DoS attack is applied at 100 seconds and the two detection tools are activated at 200 seconds. Both NePID and Synkill take care of the attack from 200 seconds to 350 seconds. In Fig 5, we can see that

resource usage of NePID drops drastically at about 250 seconds, while resource usage of Synkill drops rapidly at around 300 seconds. This results from the fact that the attack detection tools detect the attack and discard abnormal packets. Also, Fig. 5 illustrates that the proposed NePID outperforms Synkill using less system resources. The main reason that the NePID performs better than Synkill is that NePID is basically a probe detection scheme which is activated in advance for false errors, whereas Synkill is in operation after the attack, which results in longer time delay.

5. Conclusions

In this paper, we proposed a network based intrusion detection model using fuzzy cognitive maps which can detect intrusion by DoS attack. A DoS attack appears in the form of the intrusion attempt. The syn flooding attack takes advantage of the weak point of three way handshake between the end points of TCP connections. The NePID model captures and analyzes the packet information to detect SYN flooding attack. Using the results of the FCM detection module, the detection module measures the degree of risk of the DoS and trains the response module to deal with attacks.

For the performance evaluation of the proposed model, the average rates of the true positive and false negative errors are measured. The true positive error rate of the NePID is similar to that of Bernhard's true positive error rate. However, the false negative rate of the proposed scheme is considerably smaller than that of the Bernhard's.

In addition, system resource usage of the NePID is compared to that of Synkill, which is a well-known syn flood attack detection. The proposed NePID outperforms Synkill in system resource usage and time delay. The better performance results from the fact that the NePID is basically a probe detection scheme which is activated in advance for false errors. For further research, the NePID detection method needs to be extended to general purpose intrusion detection system.

References

[1] Solar, "Designing and Attacking Port Scan Detection Tools", Phrack Magazine, Vol. 8, Issue 53, pp. 13-15, 1998.

[2] http://www.krcert.or.kr

[3] http://silicondefense.com

[4] R. Axelrod, "Structure of Decision: The Cognitive Maps of Political Elites," Princeton, NJ: Princeton University Press, 1976.

[5] J. Cannady, "Applying Neural Networks to Misuse Detection," In Proceedings of the 21st National Information System Security Conference, 1998.

[6] STRC, Intrusion Detection System and Detection Rates Report, KISA, 2007.

[7] L. Feinstein, D. Schnackenberg, R. Balupari, D. Kindred, "Statistical Approaches to DDoS Attack Detection and Response", DARPA Information Survivability Conference and Exposition, 2003.

[8] S. Y. Lee, "An Adaptive Probe Detection Model using Fuzzy Cognitive Maps", Ph. D. Dissertation, Daejeon University, 2003.

[9] S. Gibson, "The Strange Tale of the Denial of Service Attacks Agent GRC.COM", http://grc.com/dos/grcdos.htm

[10] S. A. Hofmeyr, S. Forrest, and A. Somayaji, "Intrusion detection using sequences of system calls," Journal of Computer Security, Vol. 6, pp.151-180, 1998.

[11] http://wireshark.com

[12] S. Savage., D. Wetherall, A. Karlin., "Practical Network Support for IP Trace back," In Proceedings of ACM SIG COMM, 2000.

[13] P. Ferguson, D. Sene, "Network Igress Filtering: Defeating Denial of Service Attacks Which Employ IP Source Address Spoofing," RFC 2827, 2000.

[14] K.C Chang, "Defending against Flooding-Based Distributed Denial of Service A Tutorial," IEEE Communications Magazine, 2002.

[15] W. Lee, S. J. Stolfo., "A Framework for Constructing Features and Models for Intrusion Detection Systems," In Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2000.

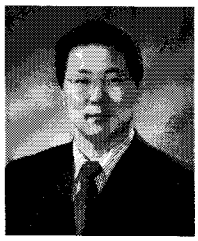
[16] C. L. Schuba, I. V. krsul, M. G. Kuhn., "Analysis of a Denial of Service Attack on TCP, "Proceedings of IEEE Symposium on Security and Privacy, pp.208-223, 1997.

Hyeock-Jin Kim [Regular Member]

- Aug, 1999 : Dept. of Computer Engineering, AJou University (Ph. D.)
- Mar, 1992 : Assistant Professor at Gimcheon College
- Mar, 1997 ~ current : Currently Associate Professor in Dept. of Computer Science at Chungwoon University

<Research Area>
CG, CAGD and web technology.

Sang-Ryul Ryu [Regular Member]



- Feb, 1990 : Dept. of Computer Engineering, Kyungpook National University(M.S.)
- Feb, 1997 : Dept. of Computer Engineering, Kyungpook National University(Ph. D.)
- Mar. 1998 ~ current : Currently Associate Professor in Dept. of Computer Science at Chungwoon University

<Research Area>
Algorithm constructing and analysis, real-time image processing and Networks.

Se-Yul Lee [Regular Member]

- Aug, 2003 : Dept. of Computer Engineering, Daejeon University (Ph. D.)
- Jan, 1999 : Researcher of ETRI and Insopack Ltd
- Mar. 2004 ~ current : Currently Assistant Professor in Dept. of Computer Science at Chungwoon University

<Research Area>
Network Security, Computer Network, Grid Middleware, and Fuzzy Neural Networks.