Computer Communications 31 (2008) 4212–4219

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Traffic flooding attack detection with SNMP MIB using SVM ☆

Jaehak Yu, Hansung Lee, Myung-Sup Kim *, Daihee Park

Department of Computer and Information Science, Korea University, Yeongi-Gun, Republic of Korea

ARTICLE INFO

Article history:
Received 18 March 2008
Received in revised form 10 September 2008
Accepted 10 September 2008
Available online 19 September 2008

Keywords:
Intrusion detection
SNMP
MIB
DoS/DDoS
Support vector machine

ABSTRACT

Recently, as network flooding attacks such as DoS/DDoS and Internet Worm have posed devastating threats to network services, rapid detection and proper response mechanisms are the major concern for secure and reliable network services. However, most of the current Intrusion Detection Systems (IDSs) focus on detail analysis of packet data, which results in late detection and a high system burden to cope with high-speed network traffic. Little or no integration exists between IDS and SNMP-based network management, in spite of the extensive monitoring and statistical information provided by SNMP agents implemented on network devices and systems. In this paper we propose a lightweight and fast detection mechanism for traffic flooding attacks. Firstly, we use SNMP MIB statistical data gathered from SNMP agents, instead of raw packet data from network links. The involved SNMP MIB variables are selected by an effective feature selection mechanism and gathered effectively by the MIB update time prediction mechanism. Secondly, we use a machine learning approach based on a Support Vector Machine (SVM) for attack classification. Using MIB and SVM, we achieved fast detection with high accuracy, the minimization of the system burden, and extendibility for system deployment. The proposed mechanism is constructed in a hierarchical structure, which first distinguishes attack traffic from normal traffic and then determines the type of attacks in detail. Using MIB datasets collected from real experiments involving a DDoS attack, we validate the possibility of our approaches. It is shown that network attacks are detected with high efficiency, and classified with low false alarms.

1. Introduction

With the ever more rapid development of the Internet, the Internet is currently an infrastructure for all kinds of network service. For example, the Web and VoIP service became the most popular and general services for the Internet. Many other new services such as IPTV service are emerging in the Internet. The significant increase of our dependency on Internet-based services in everyday life has intensified the survivability of networks. Because of extensive public availability, the Internet has become the main target of malicious attacks. Both the systems connected to the Internet and the network devices comprising the Internet, can all be severely compromised by intrusions. Recently, network flooding attacks such as DoS/DDoS and Internet Worm have posed devastating threats to network services. Moore et al. [1] reported that the DoS/DDoS attack is the main threat to the entire Internet, and the majority of them (90–94%) are deployed by using TCP. As a result, rapid detection and fast response mechanisms are the major concern for secure and reliable network services [2–6].

Intrusion Detection attempts to detect network and system attacks by examining various data records obtained from target systems and network. The network and system attacks are categorized into two types: host-based attacks [HAs] and network-based attacks [NAs] [7]. HAs target a machine and try to gain access to privileged services or resources on that machine. The detection systems for HAs usually obtain and analyze the system call data from an audit-process which tracks all system calls made on behalf of each user. On the other hand, NAs make it difficult for legitimate users to access various network services, by intentionally occupying or sabotaging network resources and services. This can be done by sending large amounts of network traffic, after exploiting well-known faults in networking services, overloading network or system, etc. The detection systems for NAs use network traffic data (packet data, flow data, MIB data, etc.) to examine traffic addressed to the machines being monitored. In this paper we focus on the NAs.

There are two general approaches to Intrusion Detection: Misuse Intrusion Detection (MID) and Anomaly Intrusion Detection (AID) [7–10]. Similar to virus detection, MID is based on pattern matching to identify intrusions. These known patterns are referred to as signatures, which are extracted from the known attacks. MID

☆ This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2007-331-D00387) and IT National Scholarship Program of MIC, Korea.

* Corresponding author. Tel.: +82 41 860 1347.
E-mail addresses: dbzzang@korea.ac.kr (J. Yu), mohan@korea.ac.kr (H. Lee), tmskim@korea.ac.kr (M.-S. Kim), dhpark@korea.ac.kr (D. Park).

0140-3664/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2008.09.018

can detect attacks with high accuracy for pre-learned attacks. However, MID has the intrinsic disadvantage of low flexibility, because the signature must be manually updated for new types of attack. However, AID constructs a normal usage behavior profile, named a historical or long-term behavior profile. And, the analysis model examines deviations of the short-term behavior profile from the norm. The deviations can be treated as the baselines for distinguishing attack activities from normal behaviors. The basic assumption of the AID is that an intruder's behavior will be noticeably different from that of legitimate users. This approach is useful to detect new types of attack, but has the unavoidable limitation that it is difficult to take proper action against the attack, due to the lack of detail information. The main concern of current research in this area is to use the advantages of MIDs and AIDs by overcoming their limitations. In this paper we use a support vector machine (SVM) based AID mechanism for attack detection and in-detail attack type classification.

Most of the current Intrusion Detection Systems (IDSs) investigate packet data to evaluate the security status of the network and system, which results in a significant processing burden and eventually, late detection time. Little or no integration exists between IDS and SNMP-based Network Management Systems (NMS) in spite of the extensive monitoring and statistical information offered by SNMP agents running on network elements. For example, SNMP MIB-II, an IETF standard MIB supported by all the SNMP agents, provides a large number of traffic performance information on different layers and protocols: IP, TCP, UDP, ICMP, etc. SNMP agents are already implemented in most current network elements, thus, MIB statistical data are available to easily collect for security analysis. This can be extended to collect additional data pertinent to network activities and is independent of the operating systems. By enlisting the MIB data, IDS promises a lower processing overhead for analysis, and high flexibility of deployment.

Recently, it has been reported that SVM is one of the most successful classification algorithms in the data mining area and provides good performance for anomaly network intrusion detection [9,11,12]. However, most of the SVM based and other AID approaches assume raw packet data inspection for the input training and testing dataset of the system, because their performance was evaluated by testing with the KDD-Cup99 dataset [13], which is extracted from raw tcpdump packet data. It is worthwhile to apply these approaches to the SNMP MIB dataset for the analysis of network and system security.

Some studies [4,5,14–18] used SNMP MIB data for a intrusion detection. Li et al. [4] developed a system named as MAID which uses SNMP MIB-II data for anomaly detection. They periodically collected 27 MIB variables from 4 MIB-II groups [19] (Interface, IP, TCP, and UDP), and converted them into a probability density function (PDF) to calculate statistical similarity metrics which is the input data of the attack classifier. For the detection mechanism, they used a neural network classifier, a typical backpropagation (BP) network, other than SVM. Cabrera et al. [6,14] also used SNMP MIB for symptom analysis of the DDoS attack. Puttini et al. [15] applied the associated Bayesian classification to the SNMP MIB variables to detect anomalous network traffic behavior in Mobile Ad Hoc Networks (MANET). Ramah et al. [16] developed an anomaly detection system using periodic SNMP data collection which is derived from a PCA (Principle Component Analysis) based unsupervised anomaly detection scheme proposed by Shyu et al. [17]. According to our literature review, these studies focused on the detection of intrusion from normal traffic, but most of them did not consider the determination of attack types, such as TCP-SYN Flooding, UDP flooding, ICMP flooding, etc. These classification mechanisms have their own structural limitations such that BP mechanism should reconstruct a training data from the beginning when a new attack occurs and Bayesian mechanism has difficulty in dealing with the features having continuous values. Furthermore, SVM-based classification mechanisms are not utilized even though other machine learning approaches (BP, Bayesian, etc.) are selectively used.

In this paper, we propose a lightweight and fast detection mechanism for the traffic flooding attacks such as DoS/DDoS and Internet Worm. Firstly, we use the SNMP MIB statistical data gathered from SNMP agents, instead of the raw packet data from network links. The involved SNMP MIB variables are selected by an effective feature selection mechanism named correlation feature selection (CFS) [27]. The SNMP MIB data are effectively retrieved from the target system as soon as the MIB variables are updated at the target system. Here our MIB update time prediction mechanism is applied for the fast detection. Secondly, we use a machine learning approach based on a Support Vector Machine (SVM) for attack classification. Our overall objective is to integrate IDSs with NMSs by attempting to construct SVM-based IDSs working on SNMP MIB data. Using the MIB data and SVM, we achieved fast detection with high accuracy, minimization of the processing burden, and flexibility of system deployment. The proposed mechanism is constructed as a hierarchical two-level structure. At the first level, an one-class SVM distinguishes attack traffic from normal traffic. At the second level, a multi-class SVM identifies the type of attacks in detail: TCP-SYN flooding, UDP flooding, ICMP flooding, etc. Using MIB datasets collected every 15 seconds during 10 days from real experiments involving DDoS attack, we tested the possibility of our approaches. Our experimental results showed that the detection accuracy of the proposed mechanism approaches 99.40%, with a false positive rate (FPR) and false negative rate (FNR) of 1.8% and 0.6%, respectively. It is shown that network attacks are detected with high efficiency, and classified with relatively low false alarms.

The paper is organized as follows. Section 2 describes three considering points for the SNMP-based attack detection process. In Section 3 we present the proposed SVM-based hierarchical two-level structure for traffic flooding attack detection. Section 4 describes the experiments and results. Section 5 closes this paper with our conclusion and possible future work.

2. Considering points for SNMB-based traffic flooding attack detection

SNMP provides a universal method of exchanging data for purposes of monitoring systems that reside on a network. The use of SNMP is most dominant in the modern industry. But, to utilize SNMP for traffic flooding attack detection, we need to consider the following three points in the use of the SNMP MIB variables which affects the performance and accuracy of the detection system: (1) Proper selection of SNMP MIB variables for attack detection, (2) Determination of the detection timing about when and how often, (3) Algorithm for attack detection using the selected MIB variables.

The first, we need to select proper SNMP MIB variables for attack detection. This selection should be done to meet that the number of SNMP MIB variables involved is minimized and the range of attack types covered is maximized. Yoo et al. [30] utilized tcpInErrs, udpNoPorts, and icmpOutEchoReps SNMP MIB variables for the purpose of detecting DoS/DDoS attacks. However, most of modern attack tools get matured and generate error-free packets to a valid port in a victim system. Attackers scan vulnerabilities of a victim host first before sending a large flood of packets targeting the victim host, which sabotage both systems resources and network resources. Therefore, the MIB variables used in [30] may not contribute much as before on detecting modern attacks. We solve this MIB selection problem by the correlation based feature selection algorithm (CFS) [27].