An Activity Pattern Based Wireless Intrusion Detection System

Nur Al Hasan Haldar
Comviva Technologies Pvt. Ltd.
Gurgaon, Haryana, India
Email: nurjamia@gmail.com

Muhammad Abulaish
Center of Excellence in Information Assurance
King Saud University, Riyadh, KSA
Email: mAbulaish@ksu.edu.sa

Syed Asim Pasha
IBM India Pvt. Ltd.
Noida, Uttar Pradesh, India
Email: asim.pasha2@gmail.com

***Abstract*—In this paper, we present an intrusion detection system which exploits pattern recognition techniques to model the usage patterns of authenticated users and uses it to detect intrusions in wireless networks. The key idea behind the proposed intrusion detection system is the identification of discriminative features from users activity data and use them to identify intrusions in wireless networks. The detection module uses PCA technique to accumulate interested statistical variables and compares them with the thresholds derived from users activities data. When the variables exceed the estimated thresholds, an alarm is raised to alert about a possible intrusion in the network. The novelty of the proposed system lies in its light-weight design which requires less processing and memory resources and it can be used in real-time environment.**

***Keywords*-intrusion detection; WLAN monitoring; feature extraction; activity pattern analysis;**

I. Introduction

Due to easy portability, mobility and enormous popularity of IEEE 802.11 standards, Wireless Local Area Network (WLAN) is increasingly being used by many companies, organizations and even individuals to install cost-effective networks in various locations such as offices, conference rooms, homes, and business areas. Although WLANs solve some problems that exist in traditional wired LANs, they also introduce new security issues [3]. There still exist certain vulnerabilities due to flaws in some IEEE 802.11 standard protocol which makes wireless network a highly desirable target for security breach. It is quite easy to spoof MAC addresses in WLAN using publicly available tools [1], making it possible to implement some wireless attack. Vulnerability exploits based remote attacks are one of the most destructive security issues faced by the security experts, as most of the high profile security threat of automatic dissemination attacks or worms are based on remote exploitation of vulnerabilities in compromised systems [4].

Though existing security techniques like, WEP, WPA or WPA2 can protect data frames, an attacker can still spoof control or a management frames to damage security. The open nature of wireless medium makes it easy for attacker to listen to network traffic. These attributes makes wireless network potentially vulnerable to several different types of attacks. Intrusion detection is generally used to secure any system in a network by comparing the set of baselines of the systems with their present behavior. A good number of research has been diverted towards developing signature-based as well as anomaly-based intrusion detection systems. Researchers have also developed some distributed protocols to detect router attack, by validating the traffic transmitted by one router and received by another router of the same transmission. If an existing network-based intrusion detection system can detect attacks that are exploiting vulnerabilities at the IP layer or above, it will detect the attack regardless of whether the packet travels any medium like wired or wireless. But, an existing network-based intrusion detection system cannot handle those attacks that exploits link layer protocol vulnerabilities [2]. Most of the attacks disrupt services by sending fraudulently management frames with spoofed source address or gained unfair channel access privilege by manipulating header field or inter-frame spacing. An evident solution of spoofing is to support per-frame source authentication for data frames and control frames.

In this paper, a new attempt is worked out for anomaly-based intrusion detection in wireless networks. We have proposed an activity pattern based wireless network intrusion detection system, which characterizes the packets to identify attacks using system behavior. Some of the features monitored by the proposed intrusion detection system are *ICMP packets sent*, *DNS query requests*, *ARP requests*, and *Internet Message Format (IMF)*. In order to increase the efficiency of the proposed system, to work in real-time environment, PCA (Principal Component Analysis) algorithm is applied on traffic data to reduce data dimensions and identify only relevant features.

II. Proposed Intrusion Detection System

The proposed intrusion detection system processes following four different types of data to learn the normal usage pattern of authorized users using the concept of principal Component Analysis (PCA) and generates a profile for them. The generated profiles are then used as a baseline to identify intrusions in the network.

ICMP packets sent – Internet Control Message Protocol (ICMP) is one of the core protocols of IP suite. It relays query message and also sends some error message indicating that a requested service is not available or a host or a router could not be reached. ICMP packet is selected as a parameter to calculate the number of packets sent by a particular MAC address.

DNS query requests – Domain Name System (DNS) query is used to translate domain names into the actual IP address of destination machine. So, the users' DNS requests are considered to model their domain of interests.

ARP requests – Address Resolution Protocol (ARP) is used to associate layer 3 (Network Layer) address (such as an IP address) with a layer 2 (Data Link Layer) address (MAC address). Based on the IP address, a machine can figure out if the destination IP is a local IP or not. So, ARP request is also an important feature to identify an intruder.

IMF – Internet Message Format (IMF) is a syntax for text messages that is sent between two computer users within the framework of electronic mail messages. Based on IMF usage pattern of already known users, IMF packets form an important feature to identify a probable intrusion in the system.

In line with the modeling and usage of classification methods like nave Bayes, Decision Tree, Support Vector Machine (SVM), Neural Network, etc., our intrusion detection method is implemented as a two-phase process. (i) *training phase* – this is also called profile generation phase. In this phase, activity data related to the authenticated users in the system are captured and their profiles are generated using PCA. Threshold value to determine the maximum profile deviation of a normal user from the profiles of the authenticated users is also determined during this phase. Further detail about the profile generation phase is presented in the following sub-section. (ii) *detection phase* – this is also called profile detection phase. In this phase, the learned profiles are used as a baseline to identify possible intruders in the system. If the Euclidean distance of a user profile with the learned profiles of the authentic users is greater than the pre-determined threshold then an alarm is generated and the user is classified as a possible intruder and consequently the packets are dropped. Otherwise, the user is classified as a normal user and its profile is added in the profile set of the authenticated users. In this way, the profile set and thereby the intrusion detection accuracy of the system increases with time.

A. Profile Generation and Threshold Determination

In this phase, we train the application to learn the usage patterns of the known users after analyzing their activity data. For this, we capture the activity data related to different type of features for each authenticated user, u_i, in the system and organize them into a $n \times m$ matrix, A, where n is the number of slots and m represents the number of days for which data are captured. Thereafter, we apply PCA on A to select $p \leq m$ eigenvectors corresponding to high eigenvalues, which forms a $n \times p$ matrix. We say this matrix a weight matrix and represent it using W. In order to get an aggregated pattern for the user under consideration over the considered time-period, we apply scalar product between each column vectors of W and A, i.e., $P = A^T W$. This scalar product gives a matrix P of order $m \times p$ in which columns correspond to the selected eigenvectors and rows correspond to the days considered for profile generation. Finally, each column of matrix P is averaged and used to generate a profile $\psi(u_i) =< w_1, w_2, \cdots, w_p >$ of user u_i as a p-dimensional vector, where the value of w_i is calculated using equation 1.

$$w_k = \frac{\sum_{j=1}^{m} col_{k_j}(P)}{m} \tag{1}$$

This process is repeated for each category of data to generate profiles of authenticated users. All profiles of a particular category are considered to form a single cluster and its centroid is calculated as an average of the corresponding elements in the profile vectors. Dissimilarity between a pair of two profile vectors is calculated as an Euclidean distance between them, and the maximum of these distances over all user profiles is decided as threshold value, which is used to classify new users profiles during detection phase of the system. For a new user, her profile is generated using the above-discussed method and its distance from the generated centroid is calculated. If the distance value is within the threshold then the profile and thereby the data packet is considered as benign, otherwise it is considered as malicious and an alarm is raised.

III. Conclusion

In this paper, a new attempt to model authenticated users' activities to identify intrusions in wireless networks is presented. The proposed intrusion detection system characterizes different types of data packets and uses them as a baseline to identify intrusions in the wireless networks. In order to increase the efficiency of the proposed system and to work in real-time environment, PCA algorithm is applied on traffic data to identify relevant features through mapping high-dimensional data into lower dimensional space.

References

[1] J. Wright, *Detecting wireless LAN MAC address spoofing*, Technical report, Jan 2003, URL: http://www.net-security.org/article.php?id=364.

[2] F. Guo and T. Chiueh, *Sequence Number-Based MAC Address Spoof Detection*, in Proceedings of the 8th International Symposium on Recent Advances in Intrusion Detection (RAID), pp. 309–329, 2006.

[3] J. S. Park and D. Dicoi, *WLAN Security: Current and Future*, IEEE Internet Computing, pp. 60–65, Sep-Oct 2003.

[4] H. J. Wang, C. Guo, D. R. Simon, and A. Zugenmaier, *Shield: Vulnerability-driven network filters for preventing known vulnerability exploits*, in Proceedings of the ACM SIGCOMM, conference on Applications, technologies, architectures, and protocols for computer communications, ACM, pp. 193–204, 2004.