International Journal of Computer Applications in Engineering Sciences
[VOL I, ISSUE IV, DECEMBER 2011] [ISSN: 2231-4946]

Hybrid Intrusion Detection with Weighted Signature Generation

Sahana Devi K. J., Bharathi M.

Dept of CSE, SJCIT,Chickballapur

sahanadevi5@gmail.com, bharathi_m_p@yahoo.com

***Abstract*- An intrusion detection system (IDS) inspects all inbound and outbound network activity and identifies suspicious patterns that may indicate a network or system attack from someone attempting to break into or compromise a system IDS. Since IDS only works by matching the incoming transaction record with its predefined attack patterns stored in the database, it is necessary to develop a system which can automatically detect any new attack and record it in the database. Hence, we propose the Anomaly Detection System (ADS) as an enhancement in mending IDS loopholes by using a technique called Signature-based generation which analyzes normal profile against anomaly profile and automatically build its signature to be later on stored in the database. The paper reports the design principles and evaluation results of a new experimental hybrid intrusion detection system(HIDS) by combining the advantages of low false-positive rate of signature-based intrusion detection system (IDS) and the ability of anomaly detection system (ADS) to detect novel unknown attacks. A weighted signature generation scheme is developed to integrate ADS with SNORT by extracting signatures from anomalies detected and adds signature generated into the SNORT signature database for fast and accurate intrusion detection.**

***Keywords*—Network security, intrusion detection systems, anomaly detection, signature generation, SNORT and Bro systems,**

I. INTRODUCTION

Intrusions and anomalies are two different kinds of abnormal traffic events in an open network environment. An intrusion takes place when an unauthorized access of a host computer system is attempted. An anomaly is observed at the network connection level. Both attack types may compromise valuable hosts, disclose sensitive data ,deny services to legitimate users, and pull down network base computing resources The intrusion detection system (IDS) offers intelligent protection of networked computers or distributed resources much better than using fixed-rule firewalls. Existing IDSs are built with either signature-based or anomaly-based systems. Signature matching is based on a misuse model, whereas anomaly detection is based on a normal use model. The design philosophies of these two models are quite different, and they were rarely mixed up in existing IDS products from the security industry. A signature-based IDS like SNORT employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts analyzing previous attacks. The collected signatures are used to match with incoming traffic to detect intrusions. These are conventional systems that detect known attacks with low false alarms. However, the signature-based IDS cannot detect unknown attacks without any precollected signatures or lack of attack classifiers. Furthermore, signature matching performs well only for single-connection attacks. With the sophistication of attackers, more attacks involve multiple connections. This limits the detection range by signature matching.

On the other hand, an anomaly-based system uses a different philosophy. It treats any network connection violating the normal profile as an anomaly .A network anomaly is revealed if the incoming traffic pattern deviates from the normal profiles significantly. Through a data mining approach, anomaly detection discovers temporal characteristics of network traffic. This system can detect unknown attacks and handles multiconnection attacks well. However, anomaly detection may result in higher false alarms. The newly proposed HIDS is designed to solve these problems with much enhanced performance. In this project a new hybrid intrusion detection system (HIDS) is presented. This system combines the positive features of both intrusion detection models to achieve higher detection accuracy, lower false alarms, and, thus, a raised level of cybertrust. Our HIDS is network-based, which should not be confused with the host-based IDS with the same abbreviation by other authors. An adaptive base support threshold is applied on selected axis attributes in mining the Internet episode rules. The episode rules are used to build the HIDS, which detects not only known intrusive attacks but also anomalous connection sequences.

II. SYSTEM ANALYSIS

In the world of globalization, the data is being the backbone of IT industries. Every IT companies competing among them with Existing technologies to save their data transmission. Proposed System has been developed to overcome the above problem. A weighted signature generation scheme is
developed to integrate ADS with SNORT by extracting signatures from anomalies detected. HIDS extracts

signatures from the output of ADS and adds them into the SNORT signature database for fast and accurate intrusion detection.

A. Analysis on Existing Networks

A signature-based IDS like SNORT employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts analyzing previous attacks. The collected signatures are used to match with incoming traffic to detect intrusions. These are conventional systems that detect known attacks with low false alarms. However, the signature-based IDS cannot detect unknown attacks without any precollected signatures or lack of attack classifiers . Furthermore, signature matching performs well only for single-connection attacks. With the sophistication of attackers, more attacks involve multiple connections. This limits the detection range by signature matching.

B. Idea on proposed network

An anomaly-based system uses a different philosophy .A network anomaly is detected if the incoming traffic pattern deviates from the normal profiles significantly. This system combines the positive features of both intrusion detection models to achieve higher detection accuracy, lower false alarms, and, thus, a raised level of cyber trust. This system combines the positive features of both intrusion detection models to achieve higher detection accuracy, lowerfalse alarms, and, thus, a raised level of cyber trust. Our HIDS is network-based, which should not be confused with the host-based IDS with the same abbreviation by other authors. An adaptive base support threshold is applied on selected axis attributes in mining the Internet episode rules. The episode rules are used to build the HIDS, which detects not only known intrusive attacks but also anomalous connection sequences.

III. PROBLEM ANALYSIS

A. Function Problem & Solving technology

This is a challenging question because today's Internet is unique in thefollowing respects. First, topologies and traffic demands of the Internet are not arbitrary but have certain structures. The worst-case results may not be applicable to realistic topologies and traffic demands. A general open question is whether selfish routing results in poor performance in Internet-like environments (i.e., under realistic network topologies and traffic demands). Second, users in overlay networks do not have full flexibility in specifying their end-to-end paths. Due to limited availability of source routing support in the routers, the path between any two network nodes is dictated by the Internet routing protocols, such as OSPF, MPLS, or BGP. While overlay networks provide another mechanism to enable users to control their routes by relaying through overlay nodes, the route between two overlay nodes is still governed by the underlying routing protocol. A natural question is how to model such selfish overlay routing and whether selfish overlay routing results in poor performance. Third, even if selfish overlays (i.e., overlays consisting of selfish traffic) yield good performance, they can be deployed only incrementally. As a result, traffic and overlay traffic will interact with each other. Such interactions are called as horizontal interactions. An important question is how such selfish traffic affects the remaining traffic routed using the traditional routing protocols. A related question is whether multiple overlays result in poor performance. Fourth, the way in which selfish users choose their routes can interact with traffic engineering. Such interactions are called as vertical interactions, which can be viewed as the following iterative process. First, Internet Service Providers (ISPs) adjust network-level routing according to traffic demands, using schemes in, to minimize network cost. Then selfish users adapt to changes in the underlying default routes by choosing different overlay paths to optimize their end-to-end performance. Such adaptation changes traffic demands and triggers traffic engineering to readjust the default routes, which in turn makes selfish users adapt to new routes. Given the mismatch between the objectives of selfish routing and traffic engineering, an interesting question is whether selfish routing interacts poorly with traffic engineering. In this project, the above questions are answered through extensive simulations. A game-theoretic approach is taken to compute the traffic equilibrium of various routing schemes and then evaluate background their performance and focus on intra-domain network environments because recent advances in topology mapping and traffic estimation allow us to use realistic network topologies and traffic demands for such scenarios. Understanding selfish routing in inter-domain environments is also of great interest but is more challenging. First, no realistic models are present for inter-domain traffic demands. Second, despite some recent progress towards understanding autonomous System relationships more research efforts are needed to develop realistic models for inter-domain routing policies. Finally, the large size of inter-domain topologies makes it computationally prohibitive to derive traffic equilibrium.

B. Performance Problem & Solving Technologies

An anomaly is observed at the network connection level. Both attack types may compromise valuable hosts, disclose sensitive data, deny services to legitimate users, and pull down network based computing resources. The intrusion detection system (IDS) offers intelligent

protection of networked computers or distributed resources much better than using fixed-rule firewalls. Existing IDSs are built with either signature-based or anomaly-based systems. Signature matching is based on a misuse model, whereas anomaly detection is based on a normal use model.The design philosophies of these two models are quite different, and they were rarely mixed up in existing IDS products from the security industry. A signature-based IDS like SNORT employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts analyzing previous attacks. The collected signatures are used to match with incoming traffic to detect intrusions. . However, the signature-based IDS cannot detect unknown attacks without any precollected signatures or lack of attack classifiers. Furthermore, signature matching performs well only for single-connection attacks. With the sophistication of attackers, more attacks involve multiple connections. This limits the detection range by signature matching. On the other hand, an anomaly-based system uses a different philosophy. It treats any network connection violating the normal profile as an anomaly. A network anomaly is revealed if the incoming traffic pattern deviates from the normal profiles significantly. Through a data mining approach, anomaly detection discovers temporal characteristics of network traffic. This system can detect unknown attacks and handles multiconnection attacks well. However, anomaly detection may result in higher false alarms. The newly proposed HIDS as in Fig 1 is designed to solve these problems with much enhanced performance.

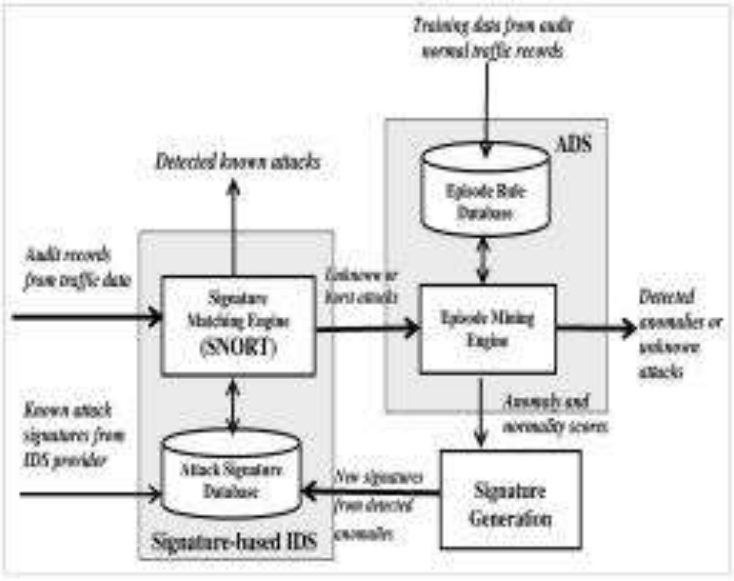

Fig.1: A hybrid intrusion detection system built with a SNORT and an anomaly detection subsystem (ADS) through automated signature generation from Internet episodes.

IV. IMPLEMENTATION DETAILS

The implementation can be preceded through Socket in java but it will be considered as peer to peer communication. For proactive routing we need dynamic routing. So java will be more suitable for platform independence and networking concepts. For maintaining route information we go for MY-SQL as database back end.

A. Module Description

1) Application Development:

File upload and download, Online shopping : In this first module, upload and download the file and here, add the new event online shopping.

4) Anomaly detection :

2) Feature Extraction (Analyzing connection services using log report)

i. Source IP
ii. Destination IP
iii. Source sent Bytes
iv. Dest sent Bytes
v. No. of bytes sent

3) Generation of FER

The frequent item sets generated are added to the database for further reference.

Compare with db and if it is deviate from knowledge base or exceeding the threshold value or if

events exceeds the threshold value of already fixed anomaly is detected.

5) Weighted signature generation

These are the events which represents no intruder
Definition: If there is any anomaly detected it will generate the signature based on anomaly detection.

B. Algorithm: Base-Support Traffic Data Mining Algorithm(Lee's algorithm,)

1: INPUT: Base-support threshold f0, all axis attributes and the set T of all network connections
2: OUTPUT: New FERs to add into existing rule set L
3: for each axis item set X in T, do
4: calculate support(X);
5: end for
6: scan T and compute L = {itemsetY|f(Y) >= f0};
7: repeat
8: generate new epsiode rule sets E = {e0; e1; . . . ; en}, where support(e0; e1; . . . en) >=f0 * min{base_sup(ei)};
9: if E is not empty, then
10: generate FERs from E with confidence e above minimum confidence eo;
11: add the generated FERs into rule set L;

V. Experimental Results

The HIDS achieved a low 47 percent detection rate at 1 percent false alarms. However, the detection rate can be raised to 60 percent if the false alarms can be tolerated up to 30 percent. SNORT has almost a constant 30 percent detection rate with almost zero false alarms. ADS can reach a 30 percent detection rate if we can tolerate 30 percent false alarms. The enhanced HIDS performance is obviously driven by the ADS performance. As a comparison, Bro has a 22 percent detection rate with almost zero false alarms. These results confirm the claimed advantages of HIDS. The balancing point is up to the designer's choice between detection accuracy and tolerance of false alarms.

VI. Conclusion And Future Plan

A new base-support data mining scheme for generating frequent episode rules. Our HIDS results in a detection rate of 60 percent, which doubles the 30 percent in using SNORT and almost triples the 22 percent in using Bro alone. To achieve an even higher detection rate, the false alarms must be maintained below 3 percent. Alerts from intrusions and anomalies detected can be correlated to result in an even smaller overhead in the detection process. For further research, we suggest the two following issues for continued research and development effort. Both issues demand prototyping and benchmark experiments.

VII. Acknowledgement

I would like to thank all the people who have helped in completion of dissertation work. To name a few my project guide Bharathi.M,and Principal of EWIT Dr.K.Channakeshavalu for their constant support and guidance.

References

[1] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, "A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data," Applications of Data Mining in Computer Security, Kluwer Academic Publishers, 2002.

[2] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Proc. Second Int'l Conf. Knowledge Discovery and Data Mining, 1996.

[3] W. Fan, M. Miller, S. Stolfo, W. Lee, and P. Chan, "Using Artificial Anomalies to Detect Unknown and Known Network Intrusions," Proc. First IEEE Int'l Conf. Data Mining, Nov. 2001.

[4] U.M. Fayyad and K.B. Irani, "Multi-Interval Discretization of Continuous-Valued Attributes from Classification Learning," Proc. Int'l Joint Conf. Artificial Intelligence (IJCAI '93), pp. 1022- 1027, 1993.

[5] S. Floyd and V. Paxson, "Difficulties in Simulating the Internet," IEEE/ACM Trans. Networking, vol. 9, no. 4, pp. 392-403, Aug. 2001.

[6] K. Hwang, Y. Chen, and H. Liu, "Defending Distributed Computing Systems from Malicious Intrusions and Network Anomalies," Proc. IEEE Workshop Security in Systems and Networks (SSN '05) held with the IEEE Int'l Parallel & Distributed Processing Symp., 2005.

[7] K. Hwang, Y. Kwok, S. Song, M. Cai, Y. Chen, and Y. Chen, "DHT-Based Security Infrastructure for Trusted Internet and Grid Computing," Int'l J. Critical Infrastructures, vol. 2, no. 4, pp. 412- 433, Dec. 2006.

[8] Kaleton Internet, "Combination of Misuse and Anomaly Intrusion Detection Systems," http://www.kaleton.com, Ma