A Survey of Anomaly Detection Methods in Networks

Weiyu Zhang, Qingbo Yang, Yushui Geng
Modern Educational Technology Center
Shandong Institute of Light Industry
Jinan, China
zwy@sdili.edu.cn

***Abstract*--Despite the advances reached along the last 20 years, anomaly detection in networks is still an immature technology, Nevertheless, the benefits which could be obtained from a better understanding of the problem itself as well as the improvement of these methods. Therefore, in this paper we present a survey on anomaly detection in networks. In order to distinguish between the different approaches used for anomaly detection in networks in a structured way, we have classified those methods into four categories: statistical anomaly detection, classifier based anomaly detection, anomaly detection using machine learning and finite state machine anomaly detection. We describe each method in details and give examples for its applications in networks.**

Keywords-Anomaly detection; Machine learning; Intrusion detection; Network security

I. INTRODUCTION

Communication networks make physical distances meaningless. While we are enjoying the ease of being connected, it is also recognized that an intrusion of malicious users from one place can cause severe damages to wide areas. Computer network's security becomes a critical issue and it is important to develop mechanisms to defense against the intrusions.

The existing intrusion detection methods fall in two major categories: signature recognition and anomaly detection [1]. For signature recognition techniques, signatures of known attacks are stored and monitored events are matched against the signatures. The techniques signal an intrusion when there is a match. An obvious limitation of these techniques is that they cannot detect new attacks whose signatures are unknown. Anomaly detection, on the other hand, builds models of normal data and detects any deviation from the normal model in the observed data. Given a set of normal data to train from, and given a new piece of test data, the goal is to determine whether the test data belong to "normal" or to an anomalous behavior. The anomaly detection techniques have the advantage that they can detect new types of intrusions as deviations from normal usage [2]. However, their weakness is the high false alarm rate.

The remainder of this article is organized as follows. In Section 2, we introduce the problem through the presentation of theoretical considerations. In Section 3, we provide detailed discussions on the various techniques used in anomaly detection. Finally, in section 4 we conclude this paper.

II. PROBLEM STATEMENT

To pose the problem of anomaly detection in any system implies the existence of a subjacent concept of normality. The notion of 'normal' is usually provided by a formal model that expresses relations between the fundamental variables involved in the system dynamics. Consequently, an event is catalogued as anomalous because its degree of deviation in relation to the profile of characteristic behavior of the system, specified by the model of normality, is high enough.

Formally, an anomaly detection system S can be defined as a pair S = (M,D), where M is the model of normal behavior of the system and D is a similarity measure that allows obtaining, given an activity record, the degree of deviation that such activities have with regard to the model M. Therefore, the core of the system is constituted by two main modules: the modeling subsystem and the detection subsystem. The first of them works during a training stage, and performs an event processing in order to obtain the model M of the normal behavior of the system. The obtained model is subsequently used by the detection engine to evaluate new events. As was stated before, this evaluation is a measurement of the degree of deviation that such events present in relation to the model of the system. These two modes of operation are usually carried out separately. However, it is important to note that systems evolve and, therefore, the model should be reconstructed periodically in order to provide a way of adaptation to the new environment.

III. ANOMALY DETECTION METHODS

A. Anomaly detection using statistics

In statistical methods for anomaly detection, the system observes the activity of subjects and generates profiles to represent their behavior. Typically, two profiles are maintained for each subject: the current profile and the stored profile. As the network events are processed, the system updates the current profile and periodically calculates an anomaly score by comparing the current profile with the stored profile using a function of abnormality of all measures within the profile. If the anomaly score is higher than a certain threshold, the system generates an alert.

978-1-4244-5273-6/09/$26.00 ©2009 IEEE

Statistical anomaly detection has a number of advantages. Firstly, these systems do not require prior knowledge of security flaws and/or the attacks themselves. In addition, statistical approaches can provide accurate notification of malicious activities that typically occur over extended periods of time. However, statistical anomaly detection schemes also have drawbacks. Firstly, it can be difficult to determine thresholds that balance the likelihood of false positives with the likelihood of false negatives. In addition, statistical methods need accurate statistical distributions, but, not all behaviors can be modeled using purely statistical methods.

Haystack [3] is one of the earliest examples of a statistical anomaly-based intrusion detection system. It used both user and group-based anomaly detection strategies, and modeled system parameters as independent, Gaussian random variables. Haystack defined a range of values that were considered normal for each feature. If during a session, a feature fell outside the normal range, the score for the subject was raised. It was designed to detect six types of intrusions. But, one drawback of Haystack was that it was designed to work offline.

Statistical Packet Anomaly Detection Engine (SPADE) [4] is a statistical anomaly detection system. SPADE was one of the first papers that proposed using the concept of an anomaly score to detect port scans, instead of using the traditional approach of looking at p attempts over q seconds. In [4], the authors used a simple frequency based approach, to calculate the 'anomaly score' of a packet. The fewer times a given packet was seen, the higher was its anomaly score. Once the anomaly score crossed a threshold, the packets were forwarded to a correlation engine that was designed to detect port scans. However, the one major drawback for SPADE is that it has a very high false alarm rate. This is due to the fact that SPADE classifies all unseen packets as attacks regardless of whether they are actually intrusions or not.

B. Anomaly detection using a classifier

In this section we focus on the anomaly detection using a classifier. Anomaly detection depends on the idea that normal characteristics behavior can be distinguished from abnormal behavior. A classifier can be used to predict the normal incoming event given the current event. If during the monitoring phase the next event is not the one predicted by the classifier, it is considered as an anomaly. The classification process typically involves the following steps: 1. Identify class attributes and classes from training data. 2. Identify attributes for classification. 3. Learn a model using the training data. 4. Use the learned model to classify the unknown data samples. A variety of classification techniques have been proposed in the literature. These include inductive rule generation techniques, fuzzy logic and genetic algorithms-based techniques.

Inductive rule generation algorithms typically involve the application of a set of association rules and frequent episode patterns to classify the audit data. The advantage of using rules is that they tend to be simple and intuitive, unstructured and less rigid. As the drawbacks they are difficult to maintain, and in some cases, are inadequate to represent many types of information. A number of inductive rule generation algorithms have been proposed in literature. Some of them first construct a decision tree and then extract a set of classification rules from the decision tree. Other algorithms directly induce rules from the data by employing a divide-and-conquer approach.

Fuzzy logic techniques have been in use in the area of network security since the late 1990's [5]. Dickerson et al. [6] developed the Fuzzy Intrusion Recognition Engine (FIRE) using fuzzy sets and fuzzy rules. FIRE uses simple data mining techniques to process the network input data and generate fuzzy sets for every observed feature. The fuzzy sets are then used to define fuzzy rules to detect individual attacks. FIRE does not establish any sort of model representing the current state of the system, but instead relies on attack specific rules for detection.

Genetic algorithms, a search technique used to find approximate solutions to optimization and search problems, have also been extensively employed in the domain of intrusion detection to differentiate normal network traffic from anomalous connections. The major advantage of genetic algorithms is their flexibility and robustness as a global search method. The earliest attempt to apply genetic algorithms to the problem of intrusion detection was done by Crosbie and Spafford [7] in 1995, when they applied multiple agent technology to detect network based anomalies.

C. Anomaly detection using machine learning

Machine learning aims to answer many of the same questions as statistics. However, unlike statistical approaches which tend to focus on understanding the process that generated the data, machine learning techniques focus on building a system that improves its performance based on previous results. In other words systems that are based on the machine learning paradigm have the ability to change their execution strategy on the basis of newly acquired information.

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest. When used in conjunction with statistical techniques, Bayesian networks have several advantages for data analysis [8]. Several researchers have adapted ideas from Bayesian statistics to create models for anomaly detection. Valdes et al. [9] developed an anomaly detection system that employed naive Bayesian networks to perform intrusion detection on traffic bursts.

Bayesian techniques also have been frequently used in classification and suppression of false alarms areas. Kruegel et al. [10] proposed a multi-senor fusion approach where the outputs of different IDS sensors were aggregated to produce a single alarm. This approach is based on the assumption that any anomaly detection technique cannot classify a set of events as an intrusion with sufficient confidence. Although using Bayesian networks for intrusion detection can be effective in certain applications, their limitations should be considered in the actual implementation. Since the accuracy of this method is dependent on certain assumptions that are typically based on the behavioral model of the target system, therefore, selecting an accurate model is the most important things towards solving the problem. Unfortunately selecting an accurate behavioral model is a difficult task as typical networks are complex.

Typical datasets for intrusion detection are very large and multidimensional. To tackle the problem of high dimensional datasets, researchers have developed a dimensionality reduction technique known as principal component analysis (PCA). PCA is a technique where n correlated random variables are transformed into d<n uncorrelated variables. The uncorrelated variables are linear combinations of the original variables and can be used to express the data in a reduced form. Shyu et al. [11] proposed an anomaly detection scheme, where PCA was used as an outlier detection scheme and was applied to reduce the dimensionality of the audit data and arrive at a classifier that is a function of the principal components.

Mahoney et al. [12–14] presented several methods that address the problem of detecting anomalies in the usage of network protocols by inspecting packet headers. The common denominator of all of them is the systematic application of learning techniques to automatically obtain profiles of normal behavior for protocols at different layers. Packet Header Anomaly Detector (PHAD) [12], LEarning Rules for Anomaly Detection (LERAD) [13] and Application Layer Anomaly Detector (ALAD) [14] use time-based models in which the probability of an event depends on the time. For each attribute, they collect a set of allowed values and flag novel values as anomalous. PHAD, ALAD, and LERAD differ in the attributes that they monitor. PHAD monitors 33 attributes from the Ethernet, IP and transport layer packet headers. ALAD models incoming server TCP requests: source and destination IP addresses and ports, opening and closing TCP flags, and the list of commands in the application payload. Depending on the attribute, it builds separate models for each target host, port number (service), or host/port combination. LERAD also models TCP connections. The authors break down the multivariate problem into a set of univariate problems and sum the weighted results from range matching along each dimension. The advantage of this approach is that it makes the technique more computationally efficient and effective at detecting network intrusions.

D. Anomaly detection using finite state machines

A finite state machine (FSM) is a model of behavior composed of states, transitions and actions. In this model, a state stores information about the past, a transition indicates a state change and is described by a condition that would need to be fulfilled to enable the transition. An action is a description of an activity that is to be performed at a given moment.

The finite state machine has been used to detect attacks on the DSR protocol in [15]. First, an algorithm for monitor selection for distributed monitoring all nodes in networks was proposed and then the correct behaviors of the nodes according to DSR were manually abstracted. Using this method has the advantage of detecting intrusions without the need of trained data or signatures, also unknown intrusions can be detected with few false alarms. As a result, a distributed network monitor architecture which traces data flow on each node by means of finite state machine was proposed. In Ref. [16], Sekar et al. present a specification-based model as well as a prototype with excellent detection performance. The model proposed by authors consists of developing protocol specifications by using Extended Finite State Automata (EFSA).

IV. CONCLUSIONS

Networks are becoming increasingly complex at the same time that security concerns do not cease to grow and require more and more attention. Hence, there is a strong need for anomaly detection as a frontline security research area for network security. In order to give a clear vision about the use of this technique, we present in this paper a classified survey of the methods that are used for anomaly detection in networks. We believe that a deeper knowledge is required until this technology achieves a solid maturity.

REFERENCES

[1] H. S. Javitz and A. Valdes, The SRI Statistical Anomaly Detector, Proceedings of the 1991 IEEE Symposium on Research in Security and Privacy, May 1991.

[2] D. E. Denning, An Intrusion Detection Model, IEEE Transactions on Software Engineering, SE-13, pp. 222-232, 19517.

[3] S.E. Smaha, Haystack: An intrusion detection system, in: Proceedings of the IEEE Fourth Aerospace Computer Security Applications Conference, Orlando, FL, 1988, pp. 37–44.

[4] S. Staniford, J.A. Hoagland, J.M. McAlerney, Practica automated detection of stealthy portscans, Journal of Computer Security 10, 2002, pp. 105–136.

[5] H.H. Hosmer, Security is fuzzy!: applying the fuzzy logic paradigm to the multipolicy paradigm, in: Proceedings of the 1992-1993 Workshop on New Security Paradigms Little Compton, RI, United States, 1993.

[6] J.E. Dickerson, J.A. Dickerson, Fuzzy network profiling for intrusion detection, in: Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society (NAFIPS), Atlanta, GA, 2000,pp. 301–306.

[7] M. Crosbie, G. Spafford, Applying genetic programming to intrusion detection, in: Working Notes for the AAAI Symposium on Genetic Programming, Cambridge, MA, 1995, pp. 1–8.

[8] D. Heckerman, A Tutorial on Learning With Bayesian Networks, Microsoft Research, Technical Report MSRTR-95-06, March 1995.

[9] A. Valdes, K. Skinner, Adaptive model-based monitoring for cyber attack detection, in: Recent Advances in Intrusion Detection Toulouse, France, 2000, pp. 80–92.

[10] C. Kruegel, D. Mutz, W. Robertson, F. Valeur, Bayesian event classification for intrusion detection, in: Proceedings of the 19th Annual Computer Security Applications Conference, Las Vegas, NV, 2003.

[11] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, L. Chang, A novel anomaly detection scheme based on principal component classifier, in: Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, Melbourne, FL, USA, 2003, pp. 172–179.

[12] M.V. Mahoney, P.K. Chan, PHAD: Packet Header Anomaly Detection for Identifying Hostile Network Traffic Department of Computer Sciences, Florida Institute of Technology, Melbourne, FL, USA, Technical Report CS-2001-4, April 2001.

[13] M.V. Mahoney, P.K. Chan, Learning Models of Network Traffic for Detecting Novel Attacks Computer Science Department, Florida Institute of Technology CS-2002-8, August 2002.

[14] M.V. Mahoney, P.K. Chan, Learning nonstationary models of normal network traffic for detecting novel attacks, in:Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Canada, 2002, pp. 376–385.

[15] P. Yi, Y. Jiang, Y. Zhong, and S. Zhang, Distributed Intrusion Detection for Mobile Ad hoc Networks, Proceedings of the 2005 Symposium on Applications and the Internet Workshops (SAINTW'05),pp. 94-97.

[16] R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, S. Zhou, Specification-based anomaly detection: a new approach for detecting network intrusions, Proceedings of the Ninth ACM Conference on Computer and Communications Security, Washington, DC, USA, November 18–22, 2002, pp. 265–274.