International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-2 Issue-4, April 2013

A Research Paper on Hybrid Intrusion Detection System

Amit Kumar, Harish Chandra Maurya, Rahul Misra

Absrtact-_An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a Management Station. Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, and reporting attempts. In addition, organizations use IDPSes for other purposes, such as identifying problems with security policies, documenting existing threats and deterring individuals from violating security policies. IDPSes have become a necessary addition to the security infrastructure of nearly every organization. Different methods can be used to detect intrusions which make a number of assumptions that are specific only to the particular method. Hence, in addition to the definition of the security policy and the access patterns which are used in the learning phase of the detector, the attack detection capability of an intrusion detection system also depends upon the assumptions made by individual methods for intrusion detection. The purpose of an intrusion detection system is to detect attacks. However, it is equally important to detect attacks at an early stage in order to minimize their impact. I have used Dataset and Classifier to refine Intruders in Networks.

Keywords- (IDS), (IDPS), IDPSes.

I. INTRODUCTION

In the 21st century the development of telecommunications networks has taken giant leaps from circuit and packet switched networks towards all-IP based networks. This development has created a unified environment where communication of applications and services (data and voice) are being transferred on top of the IP-protocol.

Although the development of communication networks has been towards a better sustainability of technologies it has also raised new unwanted possibilities. Threats that were applicable only in the fixed networks are now feasible in the radio access networks. When taken into account that threats are becoming more and more sophisticated it also means that the security systems have to become more intelligent. The basic security measurements such as firewalls and antivirus scanners are in their limits to cope with the overgrowing number of intelligent attacks from the Internet. A solution to enhance the overall security of the networks is to increase the security layers with intrusion detection systems. To understand what role intrusion detection has in telecommunications networks it can be thought through a simple example.

Manuscript published on 30 April 2013.

***** Correspondence Author (s)
Amit Kumar, M.Tech 2nd Year, Bhagwant University,Ajmer, India.
Harish Chandra Maurya, Asst.Professor, Bhagwant University,Ajmer, India.
Rahul Misra, M.Tech 2nd Year, Bhagwant University,Ajmer, India.

Think of intrusion detection as a security guard that is guarding the front gate of a factory premises.

The premises of the factory represent the network of a mobile operator and the fence surrounding the factory is the operator's firewall. Employees of the factory represent the traffic in the operator's network. It is know that factories are well protected and they do not want to let people inside the premises that do not have the required clearances. The fence or firewall in this case, is in charge to keep all unwanted visitors outside the factory premises. Just like in a firewall, a fence has holes (gates) in it to let employees move in and out of the factory premises. These holes in the fence though leave the factory vulnerable to the unwanted visitors and this is why the factory has a security guard guarding the gate. Depending on the role that the security guard is in, while he is monitoring the people going in and out of the factory premises, he either notifies the head of security when he detects a suspicious looking person walking through the gate. Or he steps in and prevents this person from entering the factory premises. The basic functionality of an intrusion detection system is the first example of the security guard. IDS generate an alarm when it detects something suspicious and then the security personnel of the network operator further investigate the cause of the alarm. An intrusion detection system (IDS) is a device, typically a designated computer system, which monitors activity to identify malicious or suspicious alerts. It is placed inside an organisation to monitor what occurs within the network of the organisation. The goal of an intrusion detection system is to accurately detect computer security incidents, and notify network administrators. A distinction is made between alerts and incidents by an intrusion detection system. Alerts are defined as all the observable actions on the computer network that are picked up by the sensors of an intrusion detection system. Incidents are malicious or suspicious alerts that have a high enough value to be considered a security-relevant system event in which the system's security policy is disobeyed or otherwise breached. An IDS consists of four components, according to the Common Intrusion Detection Framework (CIDF); event generators, analysers, event databases and response units. In the research of this thesis, Dataset is used to provide attacks and normal data to analyzer. An effort will be made to choose a machine learning method that can be used as an analyser, which improves the detection rate alerts from incidents. An event database will be used to train the analyser, and to evaluate its predictions. The response units will not be within the scope of this thesis, but can be controlled by the decisions of the analyser.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

II. INTRUSION DETECTION AND INTRUSION DETECTION SYSTEM

The intrusion detection systems are a critical component in the network security arsenal.

2.1 Principles and Assumptions in Intrusion Detection

Denning defines the principle for characterizing a system under attack. The principle states that for a system which is not under attack, the following three conditions hold true:

1. Actions of users conform to statistically predictable patterns.
2. Actions of users do not include sequences which violate the security policy.
3. Actions of every process correspond to a set of specifications which describe what the process is allowed to do.

Systems under attack do not meet at least one of the three conditions. Further, intrusion detection is based upon some assumptions which are true regardless of the approach adopted by the intrusion detection system. These assumptions are:

1. There exists a security policy which defines the normal and (or) the abnormal usage of every resource.
2. The patterns generated during the abnormal system usage are different from the patterns generated during the normal usage of the system; i.e., the abnormal and normal usage of a system results in different system behavior. This difference in behavior can be used to detect intrusions.

As we shall discuss later, different methods can be used to detect intrusions which make a number of assumptions that are specific only to the particular method. Hence, in addition to the definition of the security policy and the access patterns which are used in the learning phase of the detector, the attack detection capability of an intrusion detection system also depends upon the assumptions made by individual methods for intrusion detection.

2.2 Components of Intrusion Detection Systems

An intrusion detection system typically consists of three sub systems or components:

1. **Data Preprocessor** – Data preprocessor is responsible for collecting and providing the audit data (in a specified form) that will be used by the next component (analyzer) to make a decision. Data preprocessor is, thus, concerned with collecting the data from the desired source and converting it into a format that is comprehensible by the analyzer.Data used for detecting intrusions range from user access patterns (for example, the sequence of commands issued at the terminal and the resources requested) to network packet level features (such as the source and destination IP addresses, type of packets and rate of occurrence of packets) to application and system level behavior (such as the sequence of system calls generated by a process.) We refer to this data as the audit patterns.
2. **Analyzer (Intrusion Detector)** – The analyzer or the intrusion detector is the core component which analyzes the audit patterns to detect attacks. This is a critical component and one of the most researched. Various pattern matching, machine learning, data mining and statistical techniques can be used as intrusion detectors. The capability of the analyzer to detect an attack often determines the strength of the overall system.
3. **Response Engine** – The response engine controls the reaction mechanism and determines how to respond when the analyzer detects an attack. The system may decide either to raise an alert without taking any action against the source or may decide to block the source for a predefined period of time. Such an action depends upon the predefined security policy of the network

The authors define the Common Intrusion Detection Framework (CIDF) which recognizes a common architecture for intrusion detection systems. The CIDF defines four components that are common to any intrusion detection system. The four components are; Event generators (E-boxes), event Analyzers (A-boxes), event Databases (D-boxes) and the Response units (R-boxes). The additional component, called the D-boxes, is optional and can be used for later analysis.

III. PROPOSED WORK

We use two classification techniques for our proposed architecture, in a combined manner. Consequently, an increasing number of approaches have been developed for accomplishing such purpose, including k-nearest-neighbor (KNN) classification, Naïve Bayes classification, support vector machines (SVM), decision tree (DT), neural network (NN), and maximum entropy. Our choice among all available classification techniques is depends upon our studies about all classifier. We put our motivations for these classifiers in below topic at a glance.

3.1 Bayes' Theorem

Let $\boldsymbol{X}$ be a data tuple. In Bayesian terms, $\boldsymbol{X}$ is considered "evidence." As usual, it is described by measurements made on a set of n attributes. Let H be some hypothesis, such as that the data tuple $\boldsymbol{X}$ belongs to a specified class C. For classification problems, we want to determine $P(H|\boldsymbol{X})$, the probability that the hypothesis H holds given the "evidence" or observed data tuple $\boldsymbol{X}$. In other words, we are looking for the probability that tuple $\boldsymbol{X}$ belongs to class C, given that we know the attribute description of $\boldsymbol{X}$. $P(H|\boldsymbol{X})$ is the posterior probability, or *a posteriori probability*, of H conditioned on $\boldsymbol{X}$.

In this way, Bayes' theorem adjusts the probabilities as new information on evidences appears.

According to its classical formulation, given two events A and B, the conditional probability

- $P(A|B)$ that A occurs if B occurs can be obtained if we know
- P(A), the probability that A occurs
- $P(B)$, the probability that B occurs,
- $P(B|A)$ the conditional probability of B given A,

(As shown in equation):

$$P(A\,|B) = \frac{P(B\,|A)P(A)}{P(B)}$$

3.1.2 Naïve Bayes Classifier for intrusion Detection

In Bayesian classification, we have a hypothesis that the given data belongs to a particular class. We then calculate the probability for the hypothesis to be true. This is among the most practical approaches for certain types of problems.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

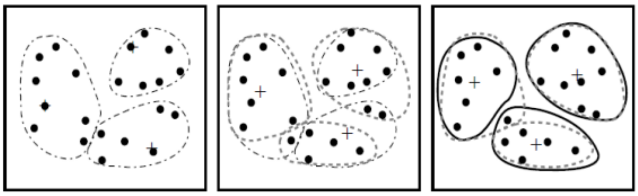
The approach requires only one scan of the whole data. Also, if at some stage there are additional training data, then each training example can incrementally increase/decrease the probability that a hypothesis is correct. Thus, a Bayesian network is used to model a domain containing uncertainty .

3.2 K-means Clustering

The *k*-means algorithm takes the input parameter, *k*, and partitions a set of *n* objects into *k* clusters so that the resulting intra-cluster similarity is high but the inter-cluster similarity is low. Cluster similarity is measured in regard to the *mean* value of the objects in a cluster, which can be viewed as the cluster's *centroid* or *center of gravity*.

Network intrusion class labels are divided into four main classes, which are DoS, Probe, U2R, and R2L. Fig. 1(a) to Fig. 1(c) shows the steps involved in K-Means clustering process. Fig.2 will later show the final overall result with application of the classification approach. The main goal to utilize K-Means clustering approach is to split and to group data into normal and attack instances.K-Means clustering methods partition the input dataset into k- clusters according to an initial value known as the seed-points into each cluster's centroids or cluster centers. The mean value of numerical data contained within each cluster is called centroids. In our case, we choose k = 3 in order to cluster the data into three clusters (C1, C2, C3). Since *U2R* and R2L attack patterns are naturally quite similar with normal instances, one extra cluster is used to group U2R and R2L attacks.

Back to Fig. 1(b), each input will be assigned to the closest Centroid by squared distances between the input data points and the centroids. New centroids will then be generated for each cluster by calculating the mean values of the input set assigned to each cluster as shown in Fig. 1(c).



(a) (b) (c)

Algorithm: *k*-means. The *k*-means algorithm for partitioning, where each cluster's center is represented by the mean value of the objects in the cluster.

Input:

k: the number of clusters,

D: a data set containing *n* objects.

Output: A set of *k* clusters.

Method:

(1) Arbitrarily choose *k* objects from *D* as the initial cluster centers;
(2) Repeat
(3) (Re) assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster;
(4) Update the cluster means, i.e., calculate the mean value of the objects for each cluster;
(5) Until no change;

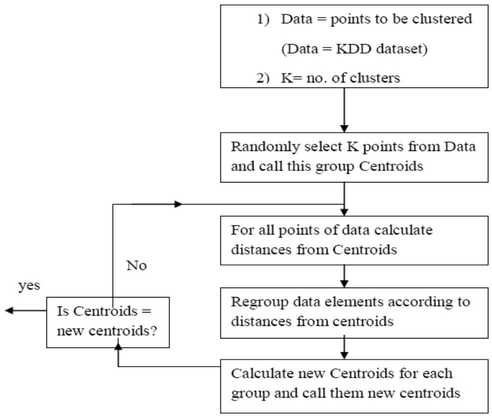


Figure 3.2: block diagram for K-means clustering

IV. FUTURE SCOPE

In the future:

- We recommend considering the Hybrid Intrusion Detection System which is better at detecting R2L and U2R attacks.
- The misuse detection approach better at detecting R2L and U2R attacks more efficiently as well as anomaly detection approach.
- Work for approach which is better at detecting attacks at the absence of match signatures as provided in the misuse rule files.

The critical nature of the task of detecting intrusions in networks and applications leaves no margin for errors. The effective cost of a successful intrusion overshadows the cost of developing intrusion detection systems and hence, it becomes critical to identify the best possible approach for developing better intrusion detection systems.

Every network and application is custom designed and it becomes extremely difficult to develop a single solution which can work for every network and application. In this thesis, we proposed novel frameworks and developed methods which perform better. However, in order to improve the overall performance of our system we used the domain knowledge for selecting better features for training our models. This is justified because of the critical nature of the task of intrusion detection. Using domain knowledge to develop better systems is not a significant disadvantage; however, developing completely automatic systems presents an interesting direction for future research.

The field of intrusion detection has been around since 1980's and a lot of advancement has been made in the same. However, to keep pace with the rapid and ever changing networks and applications, the research in intrusion detection must synchronize with the present networks. Present networks increasingly support wireless technologies, removable and mobile devices. Intrusion detection systems must integrate with such networks and devices and provide support for advances in a comprehensive manner.

REFERENCES

[1] Stefan Axelsson. Research in Intrusion-Detection Systems: A Survey. Technical Report 98-17, Department of Computer Engineering, Chalmers University of Technology, 1998.

[2] SANS Institute - Intrusion Detection FAQ. Last accessed: August 30, 2012. http://www.sans.org/resources/idfaq/.
[3] Kotagiri Ramamohanarao, Kapil Kumar Gupta, Tao Peng, and Christopher Leckie. The Curse of Ease of Access to the Internet. In Proceedings of the 3 rd International Conference on Information Systems Security (ICISS), pages 234–249. Lecture Notes in Computer Science, Springer Verlag, Vol (4812), 2008.
[4] Overview of Attack Trends, 2002. Last accessed: November 30, 2008. http://www. cert.org/archive/pdf/attack_trends.pdf.
[5] Kapil Kumar Gupta, Baikunth Nath, Kotagiri Ramamohanarao, and Ashraf Kazi. Attacking Confidentiality: An Agent Based Approach. In Proceedings of IEEE International Conference on Intelligence and Security Informatics, pages 285–296. Lecture Notes in Computer Science, Springer Verlag, Vol (3975), 2006.
[6] The ISC Domain Survey. Last accessed: Novmeber 30, 2008. https://www.isc. org/solutions/survey/.
[7] Peter Lyman, Hal R. Varian, Peter Charles, Nathan Good, Laheem Lamar Jordan, Joyojeet Pal, and Kirsten Swearingen. How much Information. Last accessed: Novmeber 30, 2008. http://www2.sims.berkeley.edu/research/projects/how-much-info-2003.
[8] Animesh Patcha and Jung-Min Park. An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends. Computer Networks, 51(12):3448–3470, 2007.
[9] CERT/CC Statistics. Last accessed: Novmeber 30, 2008. http://www.cert.org/stats/.
[10] Thomas A. Longstaff, James T. Ellis, Shawn V. Hernan, Howard F. Lipson, Robert D. Mcmillan, Linda Hutz Pesante, and Derek Simmel. Security of the Internet. Technical Report The Froehlich/Kent Encyclopedia of Telecommunications Vol (15), CERT Coordination Center 1997. Last accessed: Novmeber 30, 2008. http://www.cert.org/encyc_article/tocencyc.html.
[11] KDD Cup 1999 Intrusion Detection Data. Last accessed: Novmeber 30, 2008. http: //kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.
[12] Kapil Kumar Gupta, Baikunth Nath, and Kotagiri Ramamohanarao. Applicatio base Intrusion Detection Dataset. Last accessed: Novmeber 30, 2008. http://www.csseunimelb.edu.au/˜kgupta.
[13] Stefan Axelsson. Intrusion Detection Systems: A Taxomomy and Survey. Technical Report 99-15, Department of Computer Engineering, Chalmers University of Technology, 2000.
[14] Anita K. Jones and Robert S. Sielken. Computer System Intrusion Detection: A Survey. Technical report, Department of Computer Science, University of Virginia, 1999. Last accessed:Novmeber 30, 2008. http://www.cs.virginia.edu/˜jones/IDS-research/Documents/jones-sielken-survey-v11.pdf.
[15] Peyman Kabiri and Ali A. Ghorbani. Research on Intrusion Detection and Response: A Survey. *International Journal of Network Security*, 1(2):84–102, 2005.
[16] Joseph S. Sherif and Tommy G. Dearmond. Intrusion Detection: Systems and Models. In *Proceedings of the Eleventh IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises. WET ICE*, pages 115–133. IEEE, 2002.
[17] Mikko T. Siponen and Harri Oinas-Kukkonen. A Review of Information Security Issues and Respective Research Contributions. *SIGMIS Database*, 38(1):60–80, 2007. ACM.
[18] Teresa F. Lunt. A survey of intrusion detection techniques. *Computers and Security*, 12(4):405–418, 1993. Elsevier Advanced Technology Publications.
[19] Emilie Lundin and Erland Jonsson. Survey of Intrusion Detection Research. Technical Report 02-04, Department of Computer Engineering, Chalmers University of Technology, 2002.
[20] James P. Anderson. Computer Security Threat Monitoring and Surveillance, 1980. Last accessed: Novmeber 30, 2011. http://csrc.nist.gov/publications/ history/ande80.pdf.
[21] Dorothy E. Denning. An Intrusion-Detection Model. *IEEE Transactions on Software Engineering*, 13(2):222–232, 1987. IEEE.
[22] H. S. Javitz and A. Valdes. The SRI IDES Statistical Anomaly Detector. In *Proceedings of*
[23] *the IEEE Symposium on Security and Privacy*, pages 316–326. IEEE, 1991.
[24] S.E.Smaha. Haystack: An Intrusion Detection System. In *Proceedings of the 4th Aerospace Computer Security Applications Conference*, pages 37–44. IEEE, 1988.
[25] Paul Innella. The Evolution of Intrusion Detection Systems, 2001. Last accessed: Novmeber 30, 2008. http://www.securityfocus.com/infocus/1514.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)