ISSN 2229-5518

Monitoring and Identifying Abnormal Activities on Computer System

Prof.Sharmishta Desai, Yogita Bhutambare, Mandodari Wakade

Abstract— An intrusion detection system (IDS) is monitoring system which is used to identify abnormal activities in a computer system. Intrusion detection system reports alarms to system operator when it detects any abnormal condition. IDS is working in dynamically changing environment .Traditionally working of IDS depends on security experts which requires manual tuning.Basically an IDS consists of prediction engine which analyses data and outputs the prediction on the data. By seeing predictions system operator is able to know that data record is normal or is affected by any attack. Therefore prediction engine is the heart intrusion detection system.

In this paper we present three types of intrusion detection based on the source of detection – host based, network based and hybrid intrusion detection and also focuses on intrusion detection techniques i.e. misuse detection and anomaly detection techniques, supervised and unsupervised based learning based on the different approaches.

Index Terms—*attack, anomaly detection, IDS (Intrusion Detection System), intrusion, Intrusion Detection, misuse detection, signature based etc.*

1 INTRODUCTION

Protection of any system is an important aspect of any computing system. Protection encompasses the integrity, confidentiality and availability of the resources provided by a computing system. Three aspects of network systems make these systems more vulnerable to attack than independent machines-

- Networks typically provide more resources than independent machines
- Network systems are typically configured to facilitate resource sharing
- Global protection policies which are applied to all of the machines in a network are rare.

Any set of actions that attempt to compromise the integrity, confidentiality or availability of resources is called as intrusion. An intruder is the individual or group of individuals who initiates the actions in the intrusion. Also intruder can be from inside system that is someone with permission to use the computer with normal user privileges, who uses a hole in some operating system to escalate their privilege level, or it can be from outside system that is someone on another network or perhaps even in another country who exploits a vulnerability in an unprotected network service on the computer to gain unauthorized entry and control. Intrusion detection system is based on the fact that an intrusion will be reflected by a change in the 'normal' patterns of resources.

Intrusion detection is a methodology by which undesirable or abnormal activity can be detected. An intrusion detection system is a monitoring system which reports alarms to the system operator whenever it infers from its detection model. Intrusion Detection System (IDS) is software, hardware or combination of both used to detect intruder activity.

Intrusion detection is the process of identifying and responding to malicious activities targeted at computing and networking resources [4].

2. BASIC ARCHITECTURE OF IDS

One approach to designing a network security is to define network behaviour patterns that indicate or improper use of the network and look for the occurrence of those patterns. While such an approach may be capable of detecting known varieties of intrusive behaviour, it would allow new or undocumented types of attack to go undetected. As a result, our decision was to build a system which monitors and learns normal network behaviour and then detects deviations from it.

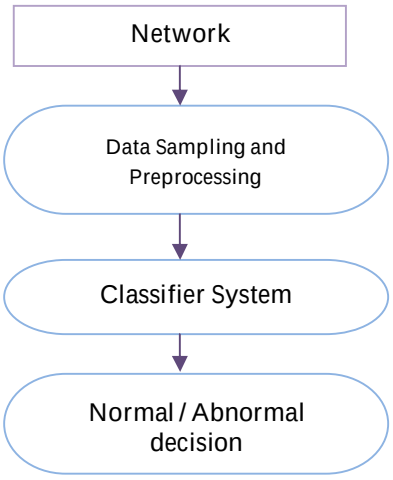

Fig 1: Basic architecture of IDS

2.1 Data Sampling:

The first step in collecting data is to determine exactly what type of data should be collected. Since the goal of this project is directed toward intrusion detection at the network level a natural choice of data is the network transmission packet. The network provides two types of information to study, transport information user information. But for this only transport information is selected. Transport information consists of source destination ordered pair and some type of checksum on which integrity of packet is determined. Transport information is added to packet as part of network transmission protocol. Transport information which cannot be made deceptive by fraudulent user therefore we call this information is unbiased data .and user information contains information which is going to be transformed from one machine to another. This can be easily modified by fraudulent user so we call it as biased data.

The second step in collecting data is to develop some mechanism for monitoring network packets. Since detecting an intrusion is not depending on the specific method used to monitor packets, any mechanism capable of obtaining a valid data sampling is satisfactory.

The final step in collecting is to process it in such a way that it is transformed into a format acceptable to the classifier system.

2.2 Data Pre-processing:

There are some values which are important to classifier .these are packet size value, timestamp value and Ethernet source-destination ordered pair.

There are two reasons to pre-processing data:

1) In the case of source and destination address and packet sizes, the raw data can be compressed without loss of relevant information. This results in data which is easier for classifier system to manipulate .also which requires less disk storage space.
2) In the case of time stamp information, the basic second count provided is augmented. To include contextual information of hour of day and day of weak. This allows the building of network behaviour that is based on human temporal patterns.

2.3 Classifier system:

The classifier system is a parallel, message-passing, rule based system. All rules are of the condition -action form. The condition is receipt of messages and action is the sending of messages when the rule is satisfied. All messages contain a tag specifying their origin and an information field.

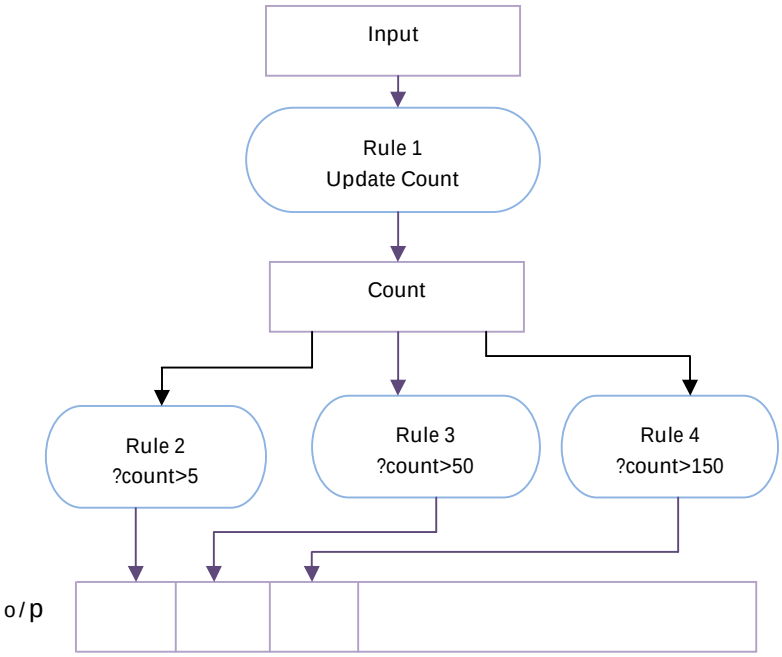

Fig 2: A classifier system

A classifier system then consists of four parts:

1) An input interface:
   In this case an input interface is message that contains information from a 4-tuple describing an individual packet information.
2) The classifiers:
   These are the rules which define the ways in which the system consumes and creates messages
3) The message list:
   A list of all messages yet to be considered by the classifier rules. The messages may be from input interface or from satisfied rules.
4) The output interface:
   An output interface is message indicating whether current network behaviour is believed to be normal or abnormal.

Why should I use Intrusion Detection Systems?

Intrusion detection allows organizations to protect their systems from the threats that come with increasing network connectivity and reliance on information systems. Given the level and nature of modem network security threats, the question for security professionals should not be *whether* to use intrusion detection, but *which* intrusion detection features and capabilities to use.

IDSs have gained acceptance as a necessary addition to every organization's security infrastructure. Despite the documented contributions intrusion detection technologies make to system security, in many organizations one must still justify the ac-

quisition of IDSs. There are several compelling reasons to acquire and use IDSs:

1. To prevent problem behaviors by increasing the perceived risk of discovery and punishment for those who would attack or otherwise abuse the system,

2. To detect attacks and other security violations that are not prevented by other security measures,

3.To detect and deal with the preambles to attacks (commonly experienced as network probes and other "doorknob rattling" activities),

4. To document the existing threat to an organization

5. To act as quality control for security design and administration, especially of large and complex enterprises

6. To provide useful information about intrusions that do take place, allowing improved diagnosis, recovery, and correction of causative factors.

System performed functions:

Intrusion detection systems perform a variety of functions [5]:
• Monitoring and analysis of user and system activity
• Auditing of system configurations and vulnerabilities
• Assessing the integrity of critical system and data files
• Recognition of activity patterns reflecting known attacks
• Statistical analysis for abnormal activity patterns
• Operating system audit trail management, with recognition of user activity reflecting policy violations
the information for symptoms of security, the goal of IDS are:

- Detection of attacks
- Prevention of attacks
- Detection of policy violations
- Enforcement of use policies
- Enforcement of connection policies
- Collection of evidence

3. REQUIREMENT OF IDS

When building IDS, a certain number of requirements must be fulfilled in order for the system to be efficient. Before listing the requirements of IDS, it is necessary to introduce four important terms.

- False Positive (FP): represents the number of instances that are classified by the IDS as being anomalous when in fact they are legitimate.
- True Positive (TP): represents the number of instances that are classified by the IDS as being anomalous and that really are anomalous.
- False Negative (FN): represents the number of instances that are classified by the IDS as being legitimate when in fact, they are anomalous.
- True Negative (TN): represents the number of instances that are classified by the IDS as being legitimate and that really are legitimate.

An intrusion detection system has to full the following requirements [1].

- **Accuracy:** Also referred as soundness, this property ensures that the IDS does not classify legitimate instances as anomalous. As mentioned above, the problem of false positives limits the use of IDS using anomaly detection in real-world applications.
- **Performance:** IDS must be able to classify the traffic without adding a noticeable overload to the network.
- **Completeness:** This property is the core of the IDS. It states that IDS should be able to detect all intrusion attempts leading to a false negative rate equal to 0. In practice, this property is very hard to achieve because the IDS must be able to detect known attack as well as unseen ones.
- **Fault Tolerance:** IDS must itself be resistant to attacks.
- **Scalability:** An IDS must be able to process the traffic of the network in real-time without dropping any packets because of a higher bandwidth than what the IDS can handle.

The IDS must be designed in order to be robust in the worst case scenario. For example, in a 10 Gb/s Ethernet network, the largest number of packets that can go through the wire at one moment is 14,880,960. An IDS performing on this type of network should be able to handle that many packets per second. Furthermore, if the IDS uses an audit trail extracted from each host of the network, it must be able to cope with an increase in the number of hosts.

4. TYPES OF IDSS

Intrusion detection systems can be classified into three groups according to their location: host based IDS, network-based IDS and Hybrid-based IDS. These three types of IDSs are briefly described below.

3.2.1 Host-Based

Host-based IDSs normally utilize information sources of two types, operating system audit trails, and system logs. Operating system audit trails are usually generated at the innermost (kernel) level of the operating system, and are therefore more detailed and better protected than system logs. However, system logs are much less obtuse and much smaller than audit trails, and are furthermore far easier to comprehend. Some host-based IDSs are designed to support a centralized IDS management and reporting infrastructure that can allow a single management console to track many hosts. Others generate messages in formats that are compatible with network

management systems.
The data from a single host is used to detect signs of intrusion as the packets Enters or exits the host. Host-based systems are becoming more and more popular due to their effectiveness at handling insider misuse. This is mainly due to the IDS gathering data (log files) from each critical machine inside the network, whereas network-based systems can only view the data that passes by a particular network node.
Host-based systems excel at stopping the following:

- Data Access/Modification: The makeup of mission critical data is different for every organization, but includes things like the Web site, customer or member databases, proposal information, and personnel records. By keeping an eye on the access of this data and taking note of changes, host-based IDS's are good at knowing when something changed that shouldn't have.
- Abuse of Privilege: This is probably one of the most serious problems in most organizations, and an area where host-based IDS's excel. By keeping track of changes to permissions, the host-based system can notify security personnel when the doors are swinging too wide. In addition, most host-based systems allow security administrators to get a quick view of the privileges that exist across their organization, and can ensure that people like former employees are removed from all systems.

3.2.2 Network-Based

The majority of commercial intrusion detection systems are network-based. These IDSs detect attacks by capturing and analyzing network packets. Listening on a network segment or switch, one network-based IDS can monitor the network traffic affecting multiple hosts that are connected to the network segment, thereby protecting those hosts.

Network-based IDSs often consist of a set of single-purpose sensors or hosts placed at various points in a network. These units monitor network traffic, performing local analysis of that traffic and reporting attacks to a central management console. As the sensors are limited to running the IDS, they can be more easily secured against attack. Many of these sensors are designed to run in "stealth" mode, in order to make it more difficult for an attacker to determine their presence and location.

Network-based systems excel at outsider attacks, and focus on catching people before they are authenticated. Areas where they shine include stopping the following:

- DOS & Packet Manipulation: A denial of service (DOS) attack is when someone sends an overload of network packets to a single resource, causing it to either crash or become so slow as to be unresponsive. A more advanced version is the Distributed Denial of Service attack, in which multiple computers all attack the resource simultaneously. Many network attacks involve sending network packets that are of incorrect size or configuration, which often causes the targeted resource to crash. Network-based IDS's, because they can process huge amounts of network traffic and sit in an optimal location, are excellent for blocking such attacks. However, note that they can also be a prime target for these attacks.
- Unauthorized Use: This is the most common attack type that people think of when they hear about IT security. Network-based IDS's are ideal for tracking unauthorized access, meaning intruders that are attempting to login to a machine without the proper credentials, compromise a machine to create a jump-off point, and those that are looking to grab passwords or data.

3.2.3 Hybrid-based

Hybrid systems that mix features of both host-based and network-based systems are becoming the norm, but most IDS's still are stronger in one area or the other. Many organizations find success by using a mixture of tools and systems to make up an overall intrusion detection strategy. A host-based system complemented by a handful of inexpensive network monitoring tools can make for a complete strategy.

5. Intrusion Detection Techniques

The techniques for the intrusion detection can be divided into two categories:

- Anomaly Intrusion Detection
- Misuse Intrusion Detection

These techniques are categorized based approaches like Statistics, Data mining, Neural Network Based and Self Organizing Maps Based approaches etc.

1. Anomaly Intrusion Detection

Anomaly detection works by establishing a baseline for normal network behaviour. A system trained for anomaly detection analyzes known "normal" network behaviour to establish this baseline and, after successful training, is used to detect any network behaviour outside of the established baseline. This method is quite successful in detecting any type of "non-normal" behaviour, but has trouble defining specific intrusion types, and has a very high false-positive rate. The high false positive rate is due to the system's lack of experience with the overall set of normal network behaviour. In essence, the system flags any unknown traffic as an intrusion [2].

Anomaly Detection Techniques includes Statistical, Neural Network, Immune System, file checking and Data Mining based approaches for the detection of attacks [3].

Statistical based methods:
Statistical methods monitor the user/network behavior by measuring certain variables statistics over time].

Distance based methods:
These methods try to overcome limitations of statistical outlier detection approach when the data are difficult to estimate in the multidimensional distributions.

Profile based methods:
This method is similar to rule based method but in this profile of normal behavior is built for different types of network traffics, users, and all devices and deviance from these profiles means intrusion.

Model based methods:
Other approaches based on deviance normal and abnormal behavior is modeling them but without creating several profiles for them .In model based methods, researchers attempt to model the normal and/or abnormal behaviors and deviation from this model means intrusion.

Signature based:
Matching available signatures in its database with collected data from activities for identifying intrusions.

Rule based:
Rule based system uses a set of "if-then" implication rules to characterize computer attacks. State transition: in this approach IDSs try to indentify intrusion by using a finite state machine that deduced from network. IDS states correspond to different states of the network and an event make transit in this finite state machine. An activity identifies intrusion if state transitions in the finite state machine of network reflect to sequel state.

Neural Network Based:
This Neural Network model solved normal attack patterns and the type of the attack. When given data was presented to the model.

1. Misuse Intrusion Detection

Misuse detection is the most common approach used in the commercial IDS [3].Misuse detectors analyze system activity, looking for events or sets of events that match a predefined pattern of events that describe a known attack. As the patterns corresponding to known attacks are called signatures, misuse detection is sometimes called "signature-based detection." The most common form of misuse detection used in commercial products specifies each pattern of events corresponding to an attack as a separate signature. However, there are more sophisticated approaches to doing misuse detection (called "state-based" analysis techniques) that can leverage a single signature to detect groups of attacks.

Misuse Intrusion Detection uses the pattern of known attacks or weak spots of the system to match and identify the attacks. So there are some ways to represent the attack in the form of pattern or an attack signature so that even variations of same attack can be detected. The main object of misuse detection focuses to use an expert system to identify intrusions based on a predetermined knowledge base [17].This approach detects all the known attacks and tries to recognize known bad behavior. Misuse Detection Techniques includes genetic algorithm, expert system, pattern matching, state transition analysis and keystroke monitoring based approaches for the detection of attacks.

Genetic Algorithm Based Detection:
There are many researchers who used GAs in IDS to detect malicious intrusion from normal use. The Genetic Algorithm provides the necessary population breeding, randomizing, and statistics gathering functions.

Expert System Based Detection:
Expert System is a system of software or combined software and hardware capable of competently executing a specific task usually performed by a human expert. Expert system is highly specialized computer systems capable of simulating a human specialist's knowledge and reasoning into Knowledge-base and is characterized by a set of facts and heuristic rules.

State transition based:
In this approach IDSs try to indentify intrusion by using a finite state machine that deduced from network. IDS states correspond to different states of the network and an event make transit in this finite state machine. An activity identifies intrusion if state transitions in the finite state machine of network reflect to sequel state. The main problem in this technique is to find out known signatures that include all the possible variations of pertinent attack, and which do not match non intrusive activity.

COMPARISON OF IDS TECHNIQUES

Table I. Comparison of intrusion detection techniques

Sr. no.	Detection Technique	Approach	Author	Detection of known Attack	Detection of Unknown Attack
1.	Misuse Based Detection	Genetic Algorithm	28,29,30, 31,32	Yes	No
2.		Expert system	33,34,35	Yes	No
3.		State Transition	36	Yes	No
4.	Anomaly Based Detection	Data Mining	37,38,39	Yes	Yes
5.		Rule Based	40,41	Yes	Yes
6.		Decision Tree	42,43,44	Yes	Yes
7.		Statistical	45,46,47	Yes	Yes
8.		Signature	48,49,50	Yes	Yes
9.		Neural network	51,52	Yes	Yes

Table 1 shows the comparison of intrusion detection techniques with different approaches and their strength and weakness

RELATED WORK

- Mohammad Sazzadul Hoque [10] explained, an implementation of IDS using GENETIC ALGORITHM. (IJNSA), Vol.4, No.2, March 2012.
- Multi-stage IDS[11] considers every stage used by recent intrusions and applies them to the HiddenMarkov Model algorithm to determine which intrusion is used in the audit data.(Do-hyeon Lee, Doo-young Kim, Jae-il Jung)(2008 IEEE).
- Charles Elkan [12] explaining Results of the KDD'99 Classifier Learning (ACM SIGKDD, January 2000.)
- Jie-Fang Liu and Fang-Min Dong [14] explaining A Dynamic Adaptive Load Balance Algorithm (novel algorithm) in Parallel Intrusion Detection System.. The experimental results shows that the algorithm can dispatch data packets reasonably and utilize all the sensors' sources effectively.(2008 IEEE)
- Design of Intrusion detection system based on pattern matching algorithm by ZHANG Hu [13]. (2009 IEEE).
- Research and optimization of Pattern Matching Algorithm Based on Instrusion Detection System explained by QIN Hai-sheng, LI Xin-hua, WEI Hai-lan and LI Jun-hui [15](2011 IEEE).
- Machine Learning-based Intrusion Detection Algorithms, which aims to study the efficiency of the method based on machine learning in intrusion detection, including artificial neural networks and support vector machine explained by Hua TANG [16] . (Journal of Computational Information Systems5 December 2009).
- Mr.C.Saravanan,Mr.M.V.Shivsankar, Prof.P.Tamije Selvy and ,Mr.S.Anto giving demonstration that high attack detection accuracy can be achieved by using Memetic algorithm for feature selection with Layered conditional Random Fields.(ACM SIGKDD, January 2000)
- Maheshkumar Sabhnani and Gursel Serpen [6] explain the evaluation performance of a comprehensive set of pattern recognition and machine learning algorithms on four attack categories as found in the KDD 1999 Cup intrusion detection dataset.

6. CONCLUSIONS

In this survey paper, we have described the generic architectural model of IDS with their types. Specifically we have focused on two important techniques of intrusion detection system: Misuse and Anomaly based detection based on number of different approaches with their strengths and weakness. Researchers proposed several intrusion detection approaches and each detection approach is suitable only for detecting a particular type of attack. Because of limited attack coverage of each approach, there is an urgent need to arrive of a generic detection approach that handles almost all types of attacks.

For that it is required to understand and analyze the techniques that are already investigated by several researchers. We hope this study will be useful for researchers to carry forward research on system security for designs of a IDS that not only will have identified strengths but also overcome the drawbacks.

REFERENCES

[1] Alexandre Balon-Perin, "Ensemble-based methods for intrusion detection" Master thesis realized at the department of Computer and Information Science of the Norwegian University of Science and Technology (*NTNU*)(2011-12).

[2] Ryan Wilson, Charlie Obimbo," Improvements on Self-Organizing Feature Maps for User-to-Root and Remote-to-Local Network Intrusion Detection on the 1999 KDD Cup Dataset" (*IJISR*), Volume 1, Issue 3, September 2011

[3] Sandip Sonawane, Shailendra Pardeshi and Ganesh Prasad," A survey on intrusion detection techniques", *(NCETIT-2012).*

[4] Prof.D.P.Gaikwad, Pooja Pabshettiwar, Priyanka Musale , Pooja Paranjape , Ashwini S. Pawar," A Proposal for Implementation of Signature Based Intrusion Detection System Using Multi-

threading Technique", *(ijceronline.com) Vol. 2 Issue. 7, November-2012.*

[5] "An Introduction to Intrusion Detection and Assessment", Prepared by Rebecca Bace from Infidel, Inc. for ICSA, Inc.

[6] Maheshkumar Sabhnani, Gursel Serpen, Gursel Serpen "Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context"

[7] An Automatically Tuning Intrusion Detection System "Zhenwei Yu, Jeffrey J. P. Tsai, Fellow and Thomas Weigert, *IEEE(2007)*

[8] Hua TANG," Machine Learning-based Intrusion Detection Algorithms *",(Journal of Computational Information Systems5 December 2009).*

[9] Weiming Hu, Wei Hu, and Steve Maybank, "AdaBoost-Based Algorithm for Network Intrusion Detection", *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART B: CYBERNETICS, VOL. 38, NO. 2, APRIL 2008.*

[10] Mohammad Sazzadul Hoque," an implementation of IDS using GENETIC ALGORITHM". (IJNSA), Vol.4, No.2, March 2012.

[11] Do-hyeon Lee, Doo-young Kim, Jae-il Jung" Multi-stage IDS considers every stage used by recent intrusions and applies them to the Hidden Markov Model algorithm". *(2008 IEEE).*

[12] Charles Elkan, "Results of the KDD'99 Classifier Learning", (*ACM SIGKDD, January 2000*.)

[13] ZHANG Hu, "Design of Intrusion detection system based on pattern matching algorithm", (2009 *IEEE*).

[14] Jie-Fang Liu, Fang-Min Dong "A Dynamic Adaptive Load Balance Algorithm in Parallel Intrusion Detection System". *(2008 IEEE)*

[15] QIN Hai-sheng, LI Xin-hua, WEI Hai-lan and LI Jun-hui," Research and optimization of Pattern Matching Algorithm Based on Instrusion Detection System", *(2011 IEEE).*

[16] Hua TANG, "Machine Learning-based Intrusion Detection Algorithms", (Journal of Computational Information Systems5 December 2009).

[17] D. J. Brown, B. Suckow, and T. Wang, 2002. Surveyof Intrusion Detection Systems.

[18] http://csrc.nist.gov/publications/nistpubs/800-31/sp800-31.pdf