UNIVERSITI PUTRA MALAYSIA

IMPROVING INTRUSION DETECTION FOR BETTER ANOMALY DETECTION BASED ON X-MEANS CLUSTERING AND MULTI-LAYER PERCEPTRON CLASSIFICATION

BORKAN AHMED ABBAS

FSKTM 2016 28

IMPROVING INTRUSION DETECTION FOR BETTER ANOMALY DETECTION BASED ON X-MEANS CLUSTERING AND MULTI-LAYER PERCEPTRON CLASSIFICATION

By

BORKAN AHMED ABBAS

Thesis Submitted to the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Computer Science

January 2016

To My Family

ABSTRACT

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Computer Science

IMPROVING INTRUSION DETECTION FOR BETTER ANOMALY DETECTION BASED ON X-MEANS CLUSTERING AND MULTI-LAYER PERCEPTRON CLASSIFICATION

By

BORKAN AHMED

January, 2016

Supervisor: Mrs Hjh Zaiton Muda

Faculty: Computer Science and Information Technology

Abstract: Due to excessive usage of network communication through the Internet with sensitive data in recent years, providing competent security medium to secure this data has become the most matters to be considered. One of the significant security mediums is an Intrusion Detection System (IDS) which offers anomaly detection with the proficiency to recognize unforeseen attacks. An IDSs should provide high accuracy, detection rates and low false alarm rate, but yet the majority of previous IDSs approaches suffered from the average rate of accuracy and detection as well as with high rate of false alarm .To enhance the capability of IDS, this thesis proposed a new hybrid machine learning approach based on X-Means and Multilayer perceptron called XM-MLP. X-Means used to cluster the data according to its behavior while multilayer perceptron (MLP) Neural Network classify those data into correct categories i.e. attack or normal. ISCX 2012 benchmark dataset has

applied to evaluate the proposed hybrid approach against single MLP classifier and previous hybrid approaches such as KM-MLP, XM-1R and XM-NB where the core detection method is based on clustering or classification technique. The performance of the proposed hybrid approach achieves better result from a single MLP classifier and other hybrid approaches in term of accuracy, detection and false alarm rate.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation and deepest gratitude to my supervisor Mrs Hjh Zaiton Muda and also Dr. Idawati Ahmad for their encouragement, valuable advices, and guidance throughout this research. In this project. I am grateful for their constant support and help.

My deeply appreciation to my father and mother who have been supportive and patiently waiting for me to complete my study. Finally, I owe my sincere thanks to my brothers and sisters, for their encouragement and affirmation, which made it possible for me to achieve this work.

Special thanks to my dearest friends who are always willing to help, share ideas and knowledge at times when they are busy with their own project. I will treasure their friendship.

APPROVAL SHEET

A thesis prepared by **Borkan Ahmed Abbas** with the title "**IMPROVING INTURSION DETECTION FOR BETTER ANOMALY DETECTION BASED ON X-MEANS CLUSTERING AND MULTI-LAYER PERCEPTRON CLASSIFICATIO**" submitted in partial to fulfilment of requirement of the master of Computer Science and Information Technology Universiti Putra Malaysia.

_____

Lecturer Mrs Hjh Zaiton Muda
Department of Communication Technology and Network
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Supervisor)

_____

Dr. Idawati Ahmad
Department of Communication Technology and Network
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Assessor)

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

BORKAN AHMED ABBAS

Date: 27\1\2016

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 Backgrounds

Recently, computers and the Internet are utilized almost in each part of our life. Since the personal computer and internet was investigated they have been unbelievable growing faster and faster and they become impossible to imagine universities, companies and even small shop without ensuring to save all the data for their customers securely (Debar, H et al., 2000).

With the high possibility of communication many computers and networks led to necessity of providing protection for network which they almost transmitting sensitive data from attackers (hackers) that would like to obtain some confidential data or information to use for them self-benefit or other purpose like destroy or modify valuable information. When the Internet evented, the security necessity started to increase as well and no way to get advantages of internet without providing protection for the systems and networks. Balancing between privacy and the resource of users is not an easy concept; the network also has to be flexible enough to cover the requirements needed to allow the pursuit of attackers (Fernandez, M. D. M et al., 2008).

There are many security measures used to protect the computer resources of a home user or a company, but in spite of all expert recommendations are applied, systems cannot be protected against possible successful attacks due to it is very difficult to have an invulnerable system which it might need to spend a lot of

money for designing and developing the system. In companies, an isolated system could massively decrease productivity and users of home whom are not having enough experience it may lead to a "hating technology" disease. Therefore, the security department or the user should be known what their values if they want to protect and how much it costs unless, doing Risk Analysis (Dieter Gollmann et al., 2002)

A well-educated users, a good security policy and a good risk analysis going to make the system better secure to intrusions. An intrusion in the system, or networks try to compromise one of the three major requirements in the network and computer security (Orchier, J et al., 2000).

- Confidentiality: Attacker gain access to confidential information.
- Integrity: Attacker tries to modify or alter information on the system or networks.
- Availability: Attacker blocks the system so it cannot be used normally by the system users.

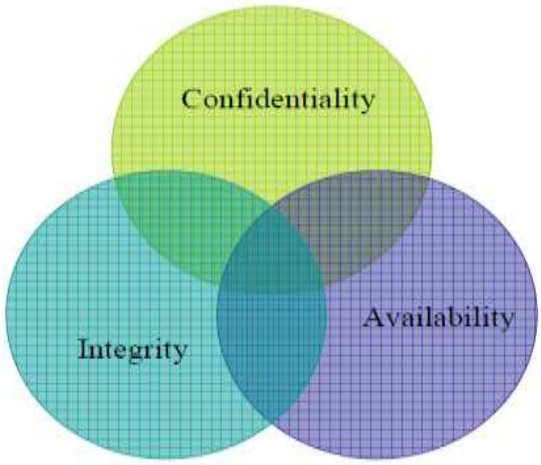


Figure 1.1: Major Requirements of Network and Computer System

An intrusion is any type of action that tries to compromise the confidentiality, integrity and availability of computer resources or some information. The attacker can utilize the weakness or defects in the system architecture as well as an internal knowledge of the operating system to cheat the authorization process or authentication (Fernandez, M. D. M et al., 2008).

Detecting way of intrusion can be done by the descriptions of some anomalous behavior for incoming data. Those type of detection tries to determine the normal behavior of a user during the process. For a correct differentiate it has to take into account three fundamental types of attacks as following:

- External penetration: The network here attacked by outside attackers.
- Internal penetration: an unauthorized user starts to attack from inside of network.

Resources abuse: authorized users utilize resources and/or data to which they gain access, in unwanted and unintended ways as shown in the figure 1.2.

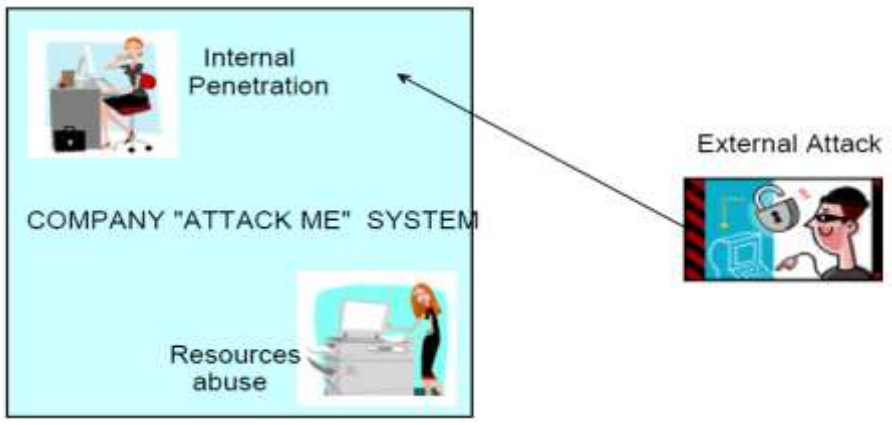


Figure 1.2: Three Main Type of Attack

The security in any system can be divided into two modes: active and passive.

The active security focuses on protecting and it is acting in case of attack. One of a very good example for active security is the firewall. To avoid access to some services or specific connections that might be abused by attackers if they gain access to some of them, filtering has been introduced to prevent the attackers to access systems. On the other hand, passive security is used in the system just to alarm if something abnormal is happening (Fernandez, M. D. M et al., 2008). It is not going to protect the system, it is just for alarming that abnormal is going on. An intrusion detection system is considered as a passive security, this concept also can be discussed since preventing of intrusion is possible but it calls Intrusion prevention systems (IPS).

1.2 An Intrusion Detection System (IDS)

It is one that permits to get information from various sources of the system where it has been implanted to alarm the administrator when the system or network is under attacks. The system will be alerted about how the attacker going to attack or who is trying to attack the system. However, sometimes it just informs if there is an attack and nothing else (Muda Z et al., 2011).

An IDS a passive measure and not such a measure to ensure system integrity. It assists the administrator or system security department to be aware of harm activities against the whole system.

Finally, in the case of a security checking, all the information gotten by the IDS going to help through taking decisions to protect the system and to create a suitable security policy.

The security audit is increasingly becoming more and more needed. These approaches not only identify and keep monitoring of intrusions; they even

enhance the reliability and stability of other security mechanisms. IDS is the process of monitoring the events that occur in a network or a system to analyze them, looking for security weaknesses.

IDS technology until now it does not considered as a mature, but almost. Since the 80s, a lot of effort has been introduced to find out a solution to this complex problem. From review in the 2013 **http://www.darkreading.com** Annual Security Survey and Computer Crime where covers around 400 computer security users in U.S corporations, universities, government agencies… it is shown that because of the high investigations in computer security, the losses caused by attacks are reducing (Fernandez, M. D. M et al., 2008). The use of IDS is in top 7 of equipment which it utilized in information security as shown in the figure 1.3. Firewalls and data encryption are already united in the first place.

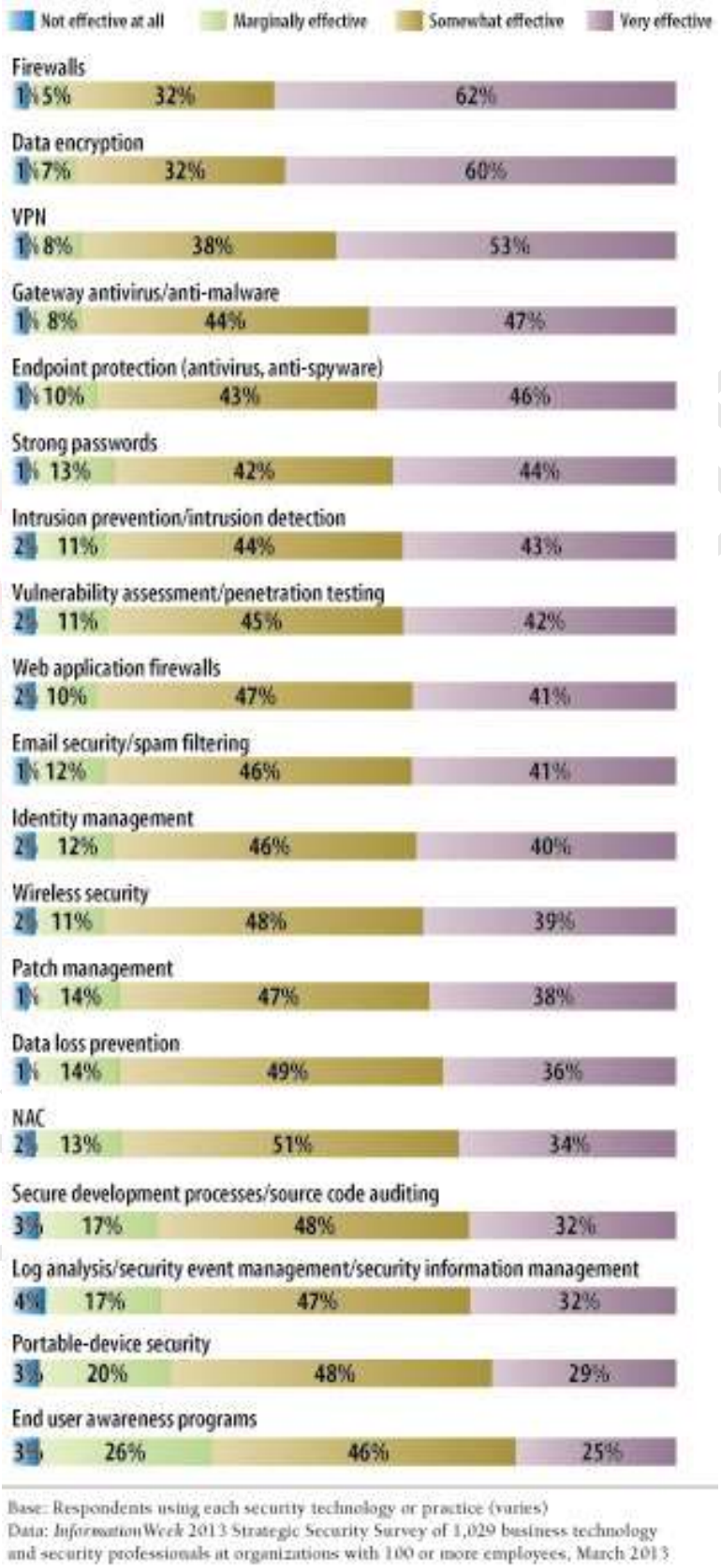


Figure 1.3: Security Technologies Effectiveness

1.3 IDS Overview

1.3.1 History

At 26th February 1980 the first paper published about IDSs and was prepared by James P. Anderson. It was under title of "Computer Security Threat: Monitoring and Surveillance"

James Anderson divides penetrations in internal and external ones, according on whether the user got an access license to the computer or not (Suduc, A. M et al 2010). The main objectives of the security audit are:

- Data could be gotten from various resources of the system.
- To obviate internal attacks, abnormal behavior of users of resources should be detected.
- Enough data should be gotten in order to find the issue.
- The security audit trail should be capable to differentiate the attacker strategy.

The initial IDS used to monitor network traffic was the Network System Monitor (NSM) and it evolved in the California University to work on a UNIX station for Sun.

The method that's been used was very close to the IDS of nowadays.

- The whole traffic was captured even though when it was not directed to the system.
- Network packets were getting.
- The protocol was identified in order to obtain the needed data.
- The data was checked and compared against rules and statistics therefore the violation or misuse could be noticed.

Todd Heberlein implements all that work. In Haystack project the IDS area changed and allowed the beginning of commercial applications.

Haystack Labs is the first commercial product was evolved by Haystack Labs, Also the Air Force Cryptologic Center discover the Automated Security Measurement System.

At the last decade many sellers (Internet Security Systems, Cisco Enterasys… have been continuously renewing because of the unbelievable growth of the internet and every company still believes that its computer-network system has to updated and present regularly (Fernandez, M. D. M et al., 2008). IDS also exists for home user products as well.

A large limit of offerings can be obtained, from very good open source ones to costly products such Snort.

1.3.2 Why Use an IDS?

Intrusion detection provides protection to organization systems against threats that occurs with rising of connectivity of network and the interconnection of information systems (Di Pietro, R et al 2008).

IDSs have got acceptance as a major part of the security infrastructure. There are many reasons for gaining and using an IDS:

1. Avoid issues by dissuading hostile individuals

When raise the possibility of detecting and penalizing attackers, the behavior going to change for some of them, decreasing the number of launched attacks. This can be a disadvantages as well as the attendance of a developed security system can rias the attacker curiosity.

Discovered attacks and other violations of security not prevented by other protection measures. Assailants, using known techniques, able to access non-

authorized systems, particularly those connected to public networks (Yue, W. T et al 2007). This usually appears when vulnerabilities which they known for others are still not fixed.

Although sellers and administrators take action to fix these vulnerabilities, where it is impossible with some situations:

- In some patrimonial systems, the operating systems impossible to be updated.
- Even when the systems able to be updated, sometimes administrators do not resources for setup new updates and might not have time. This is a common issue, basically within scenarios with a large number of nods with varied Hardware and operating systems.
- Administrators and users can make errors during configuring systems.

An IDS can be a very good protection tool. It can discover when an attacker has tried to access the system. In such this way IDSs could alert the administrator to run a backup to avoid the loss of information which they consider as a valuable (Fernandez, M. D. M et al., 2008).

2. Detect attack preambles

When the system attacked by an individual attack, it is done in expected steps. In the first step, the attacker examines and tests the network or system looking for the optimum point to breakthrough. In networks or systems where IDSs not found, the attacker has a big chance to test the system with minimum risk of being discovered. This assist him to find the weak point for the systems and networks as well (Kayacik, H. G.et al 2007).

Networks with an IDS tracking his activities, hinders him. In spite of, the attacker can test the network, the IDS going to capture these exams, defining them as

suspicious, block gaining access to the assailants, and administrators of security will be alerted.

3. Record the organization risk

When a strategy plans for arrangement of security policy or the security management is done. It is preferred to know the organization risk of threats that might be gotten, the probability of getting an attack or if it is already been attacked.

An IDS can assist us to discover the existing threats outside and inside the organization, assisting us to take decisions about resources of security that should be used in the network infrastructure (Fernandez, M. D. M et al., 2008).

4. Provide useful information about the intrusions currently taking place

Even though the IDS cannot prevent attacks, they can gather significant information about them. That information can, under some situations, can be used as a proof in legal actions. Furthermore, it can be used to make correction to failures in security in the organization or configuration security policy.

1.3.3 IDS Classification

There are several ways to classify IDSs relaying on some measurements, such as analysis type, information source, detection time and type of response. These ones are the most famous criteria (Gabra, H. N et al 2014).

1.3.3.1 Information Sources

Information sources are one of the major problems to focus on during an intrusion

detection system designs. The classification of these sources can be done in many ways. The intrusions classified in term of positions because some IDSs analyzing packages of network, captured from the backbone of network or LAN segments while other IDSs analyze events created by the application software or operating systems for signs of intrusion.

NIDS (Network-based)

Most of the IDSs are Network-based. These IDSs by capturing and analyzing network packets can detect attacks. A NIDS have ability to monitor traffic impacting multiple nods connected to that network segment, thus protecting these nods.

The network-based IDSs are usually formed by a group of sensors posed at various points in the network (Vigna, G et al 1999). These sensors surveillance traffic by representing such local analysis and reporting attacks implemented to the management console.

Advantages and disadvantages of a NIDS as following:

Advantages

- If the location of IDS is suitable, it can monitor a wide network as long as it has good capacity to analyze the traffic.
- The NIDSs have a small effect on the network, usually it is keeping passive and not Intervention with normal operations of the latter.
- NIDSs can be organized to be seen to the network in order to enhance the security against attacks.

Disadvantages

- The IDSs not only analyze the packages header, they also analyze their whole includes, so they may get some processing difficulties of all packages in a wide network or with more traffic and might fail to distinguish attacks during high traffic. Some sellers are trying to come over this issue by carrying out IDSs as a whole in hardware, which increase their speed of processing.
- The network-based IDSs do not analyze the encoded information. As an example, environments where communication is encoded it is not useful to test the contents of package and therefore incapable to evaluate whether this is a package with virulent contents or not. This issue is increased when the organization utilize encoding for security purpose in the network layer (Internet Protocol Security: IPSec) among hosts, but can be solved with a more relaxed security policy (eg, IPSec in tunnel mode).
- The network-based IDSs do not have idea if the attack was successful or not, the only known thing is that it was attacked. This means that after a NIDS discovers an attack, administrators have to manually verify every host attacked to decide if the attempt was successful or not.
- Some NIDSs have issues coping with network-based attacks transporting in fragmented packages. These packages make the IDS not recognize the attack or be fickle and might even get to fail.
- Because of their general configuration, NIDSs might have false positive rate or a high false acceptance. Normal activities can be identified as attacks in the report of analyzing. The issue appears when the number of such alarms is very high.

HIDS (Host-based)

HIDS were the first kind of IDSs evolved and implemented. They works on the information appeared from inside a computer, such as the operating system's audit files. The IDS by this way able to analyze actual activities with great accuracy, identifying exactly which processes and users are implicated in a specific attack within the operating system (Kozushko, H. et al., 2003).

It similar to any intrusion detection system, HIDSs also create many false positives. When the system is adjusted, it is remarkable the reducing of false positives and then also this kind of IDSs cannot cover very few attacks against the system.

In compare to NIDSs, HIDSs can see the outcome of an attempted attack, at the same time it directly access and observe data files and also processes of the system which it already attacked (Brenton et al., 2002).

Moreover, NIDSs have great rule to enhance the security and recently are more accepted, HIDS have certain Benefits over them:

Advantages

- The host-based IDSs, have ability to observe local events of a host, they can discover attacks that cannot be seen by IDS.
- They can usually operate in such environment which the network traffic travels encoded, since the root of information is analyzed before the data is encoded on the source host and /or after the data is decoded on the destination host.

Disadvantages

Host-based IDSs are costly (in money and time) to manger as they have to be managed and configured at each observed host. While the NIDSs have an IDS sensor for many observed systems, HIDSs each of them have an IDS.

- If the analysis station constructed within the monitored host, the IDS can be handicapped when an attack gets success on the machine.
- They are not accurate for detecting attacks on whole network (for instance, port scans) since the IDS only analyses the packets of network that been delivered to it.
- They can be handicapped by specific Denial of Service attacks.
- HIDSs use resources of the host that they are observing, impacting its performance.

1.3.3.2 Analysis Type

There are two methods of analyzing events for detecting attacks: Misuse detection (signature) and anomaly detection.

The misuse detection is the technique used widely by most commercial systems. Where anomalies detection, in which the analysis investigates strange or unusual patterns of activity.

Signature-based Detection here detectors analyze system activities to find events matching a pattern that already been identified and documented by the researchers as an attack. They collect network traffic and then proceed to analyze it.

The analysis is depending on a comparison of patterns. The system has a database of attack patterns and it goes to compare the incoming data with database of attacks when a match is reached then warning will go off (Fernandez, M. D. M et

al., 2008).

These systems play a big deal in detecting attacks but they still generate a large number of false positives (FP). Therefore, it is necessary to get regulated (tuning period) as short as possible. The suitable operation of such a system relies not just on a good installation and configuration, but also on the fact that attack's pattern in database should be updated.

Advantages

- Signature detectors are effective because they don't generate a large number of false alarms.
- They can diagnose quickly and accurately the use of a specific attack technique.

Disadvantages

- Signature detectors can detect only the attacks which they already known previously, so they must be continually updated with a new attacks signatures.
- Most of signature detectors are designed in such to be used with very tight patterns that prevent them from detecting common attacks.

Anomaly Detection

Unusual behavior identification is focusing point for anomaly detection withier in a host or a network. The basic consideration with anomaly detection is assuming that the activities of attacks are different from the normal. Anomaly detectors create profiles representing the normal behavior of hosts, users, or connections of Networks.

These profiles are created from historical data gathered within normal operation. The IDS collect data from the events and use a variety of measures to specify when the monitored activity differentiated from normal activity (Chandola, V et al., 2009). The measures and techniques used in anomalies detection covers:

- Detecting a threshold on specific features of user behavior. For instance, behavior attributes may have the number of files that can be accessed by a user in a given time period, the number of failed attempts to enter the system and the number of CPU used by a process.
- Statistic measures that can be parametric, where it is supposed that the distribution of the profiled attributes fits for a specific pattern, or non - parametric, where the distribution of the profiled features is learnt by historical values extracted over time.

Advantages

- The IDSs based on anomaly distinguishes detect abnormal behavior. Thus they have the capability to detect malicious (attack) for which they have no determined knowledge.
- Anomaly detectors can investigate information that is very useful to identify new patterns for Misuse detection.

Disadvantages

- The detection of anomalies might lead to produce a high number of false alarms because the unpredictable behavior of networks and users as well.
- Hard training is required to characterize patterns of normal behavior in anomaly detector.

1.3.3.3 Response

An IDS reacts, once the events have been analyzed and an attack has been revealed. Responses mostly cluster into two categories: active and passive. The passive IDSs introduce serves for others by sending reports to some others who then going to take action in that case, if it is appropriate. Where active IDSs launch replies to such attacks automatically (Maiwald et al., 2001).

Passive Response

In this type of IDS, the system users or the security manager are notified on what is happen. It is also good to alert the administrator of the site/system on which type of attack was launched, but the attacker has ability to monitor the email of the organization or that he has used unauthorized IP for attack purpose.

Active Response

It is automatic actions that can be taken when specific types of intrusions are detected. Two different categories can be set:

- Collection of additional information: The sensor's sensitivity level are going to be increased in order to obtain more evidences of the possible attack (e.g. all packets going to be caught by the source that starts the attack, during a certain of time).
- Changing the environment: In this case also active response can stop the attack; For example, with TCP connection, the session can be closed while injecting TCP RST segments to the victim and the attacker, or filter the IP address of the intrusion.

1.3.3.4 Detection Time

Two fundamental groups can be defined, those which detect intrusions within real time (in-line) and those which process audit data with off-line situation.

Some systems that utilize in-line detection IDS can also deal with offline detection (Lee, W et al 2001). Therefore, this kind of systems call hybrid which they can detect intrusions in two cases on-line and off-line. Figure 1.4 below shows the classifications of IDS

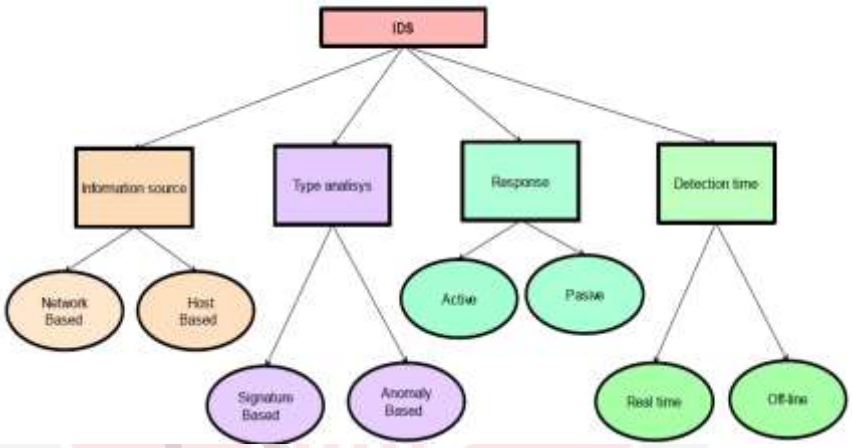


Figure 1.4: IDS Classification

1.3.4 IDS Architecture

In practical case whole intrusion detection systems have some well-defined parts that are explained in the next points:

- Application data collection sources: The space for collection of data for current or future analysis. It can be a system, or elements situated in the system itself, or a network (Garfinkel, T, et al., 2003)
- Rules: Are often those that categorize the violations that may be promised.
- Filter: This part is in charge of adjusting rules used for the getting data.

- Alarm or report generator: As soon as the data has processed by using filter rules, if there is any case that lead to give an impression that the system security has been known for others, this part of the intrusion detector sent reports to the administrator about this fact (by mail, sms, alerts...) as shown in figure 1.5.

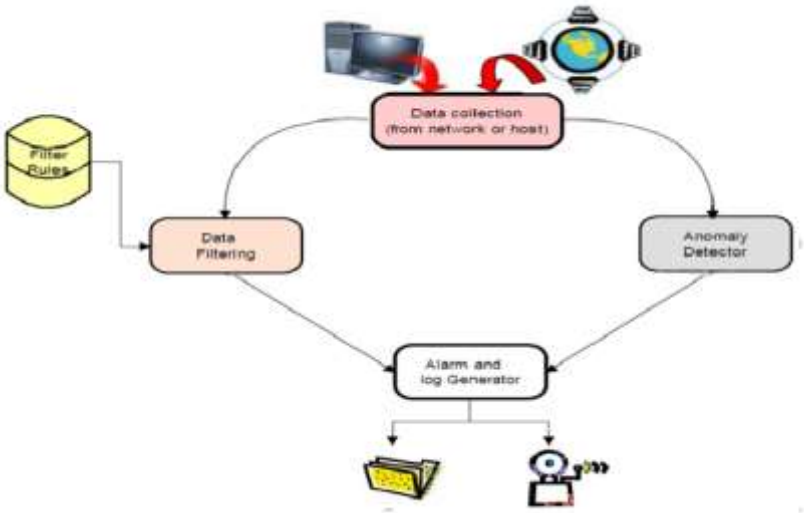


Figure 1.5: Typical IDS Architecture

1.3.5 Where to Place an IDS

There are a lot of ways to add the IDS tools to any kind of network; each has its advantages and drawbacks. The best selection would be a compromise between cost and wanted features, while remaining a high level of advantages and a controlled amount of disadvantages, all depends on the needs of the organization.

For this reason, IDSs position within a network provide different features. That is why we will see various possibilities in the same network.

For instance, demilitarized zone (DMZ) which is the area between the internal network and internet. It designed to supply public facilities without getting access to the private network of the organization as shown in figure 1.6. This subnet are

always located the major services like, DNS, HTTP and other facilities (Stewart, J.m et al., 2006).

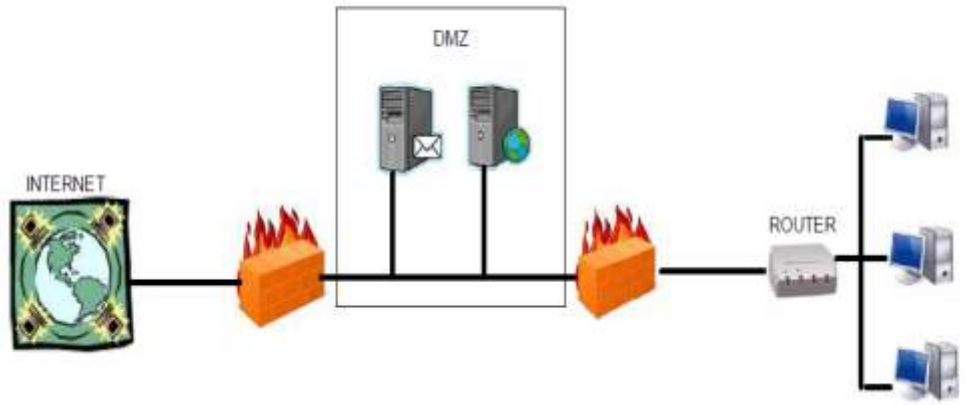


Figure 1.6: Typical Network Scenario

1.3.5.1 In Front of the External Firewall

In this design position, IDS will monitor all the incoming and outgoing traffic of the network, this means it also will monitor the number and type of attacks against the organization architecture, and the external firewall. IDSs of this location should be created with a low sensitivity since false alarms number is high.

The major disadvantages of this location are that the IDSs doesn't have ability to detect attacks which they use in their communications some technique to hide information such as encryption algorithms (Tayrani, R, et al., 2003). Figure 1.7 shows the general design when the IDS located in front of external firewall.

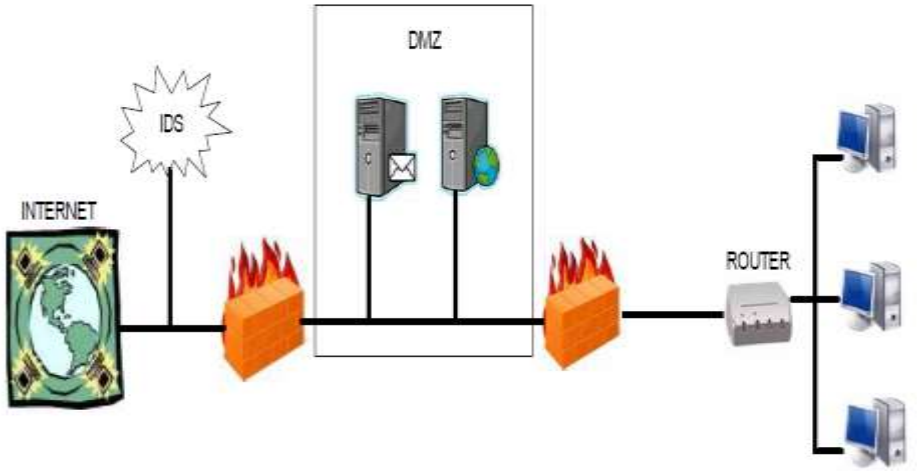


Figure 1.7: IDS in Front of External Firewall

1.3.5.2 Behind the External firewall

Another choice is to place the IDS in the demilitarized zone (DMZ), between two firewalls.

Intrusions that can pass through the major firewall are monitored. Attacks on servers that supply public services can be detected easily. The recognition of the most common attacks improves the major firewall configuration to be capable to prevent them next time.

Such in previous case, the drawbacks increase if attacker use encrypted technique for attacks. This area has less false alarms (Maiwald et al., 2001). To know the exact position of this design, figure 1.8 shows the location of IDS within DMZ.

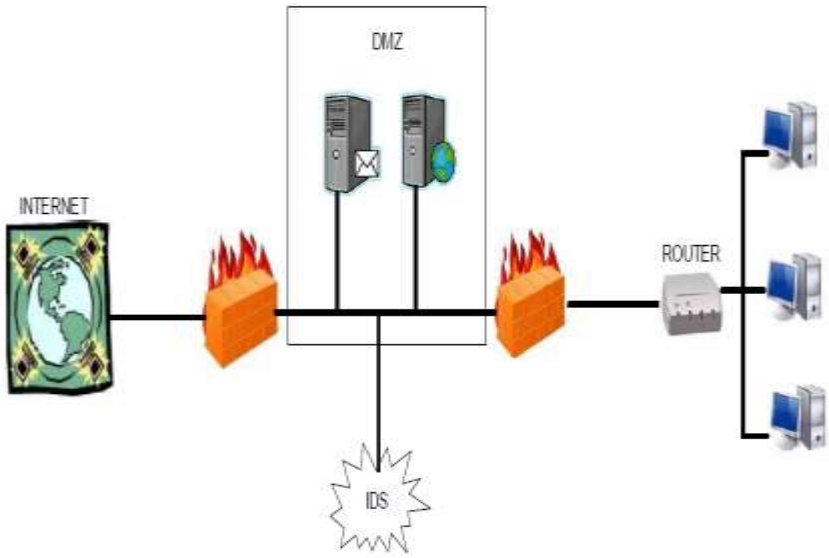


Figure 1.8: IDS in the DMZ

1.3.5.3 Behind the Second Firewall

IDS is located between the internal network and second firewall in this case. It will not listen to any internal traffic due to it is not inside the internal network.

This IDS should be less effective than those mentioned previously, because the volume of traffic is less at this point.

A fewer false alarms will occur in such this point of the network, so it is necessary to study immediately any alarm from the IDS (Fernandez, M. D. M et al., 2008).

This implementation might make these systems particularly suffer from attacks, not only form outside of network system but as well inside their own infrastructure. Figure 1.9 shows an example for the location of IDS which it located after second firewall.

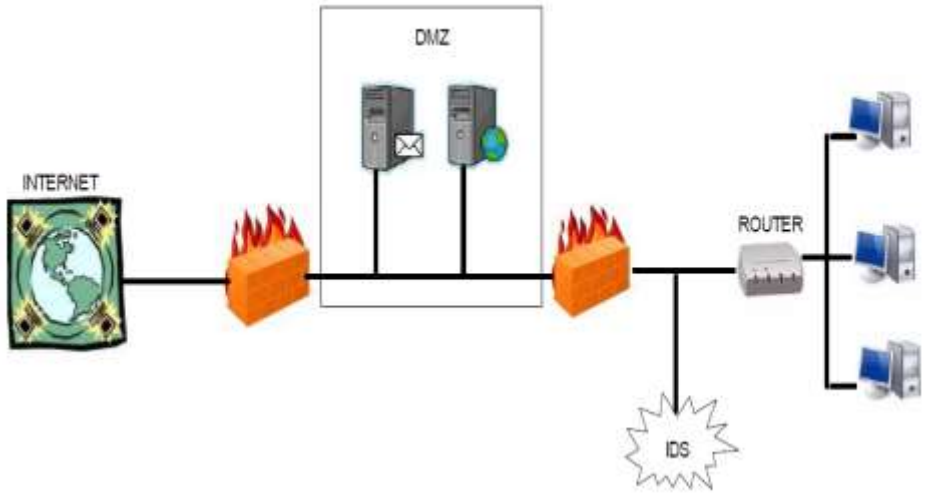


Figure 1.9: IDS After the second Firewall

1.3.6 Common Types of Network and Computer Attacks

Without doubt, security measures and controls in place are very significant for the systems and without them information or data might be compromised to attackers. Some attacks are passive which means, information is monitored; others are active, meaning the information is changed or modified with intent to destroy the data or the network itself or they corrupt both data and network (Needham et al., 2008).

Networks and the data within any system are vulnerable to any of the following types of attacks when they do not have a security plan in place.

Eavesdropping

Generally, the most of network communications occur in an unsecured environment that allows an attacker who has got access to paths of data in in such network to "listen in" or read the traffic. When the communications been eavesdropped by an attacker, it is indicating to as sniffing or snooping (Dai, H, N.,

2013). The capability of an eavesdropper to monitor the network is usually the biggest security issue that administrators might face in most of systems. Therefore, the strong encryption services that are based on cryptography are required to prevent the eavesdropper to read the data by others.

Data Modification

After data been read by an attacker, the following logical step is to change it. An attacker can alter the data in the packet and the sender or receiver doesn't notice the changing. Some of systems do not require confidentiality for all communications but no one want has messages to be modified during transit (Al-Ofeishat et al., 2012)

Identity Spoofing (IP Address Spoofing)

Most operating systems and network use the IP address of a computer or any connection device to identify a valid entity. However, it is possible for an IP address to be incorrectly assumed (identity spoofing). Surprisingly, an attacker may also use certain programs to create IP packets that appear made from valid addresses (Yaar et al., 2006).

Password-Based Attacks

A common similarity of most network and security plans operating system is password-as access control technique. So any one can get users user name and password he\she act as the actual user (Abdalla et al., 2005).

Traditional applications do not always protect information of identity as it is passed among the network for validation. This might allow an eavesdropper to gain access to the network by posing as a valid user.

After getting access to the network with a correct account, an attacker has ability to do any of the following:

- Get lists of valid user, computer names and network information.
- Alter network configurations and server, including routing tables and access controls.
- Delete users data, modify, reroute.

Denial-of-Service Attack

It differs from a password-based attack, the denial-of-service attack not allow normal use of users computer or network by valid users (Mirkovic et al., 2004).

The attacker can do any of the following after getting access to the network:

- Randomize the attention of users internal staff's Information Systems so they do not notice the intrusion directly, which lead to the attacker make more attacks within the diversion.
- Send invalid data to network or services applications, which result to abnormal termination or behavior of the services or applications.
- Flood the entire network or a computer with traffic to shut down by an overload.
- Lead to Loss of access to network resources by authorized users due to traffic block by DOS attack.

Man-in-the-Middle Attack

As the name refers, a man-in-the-middle attack occurs when someone is in the middle between two communicating persons is actively capturing, monitoring and communication transparently controlling. For example, re-routing is one of the DOS possible options while data exchanging. When computers are communicating at low levels of the network layer, the computers may not be able to determine whom is on the other side of communications.

Man-in-the-middle attacks are such someone assuming your identity in order to read your message (Desmedtet al., 2011). The person on the other end might trust it is you because the attacker might be replying actively like you to keep the exchange of messages going on and getting more information.

Compromised-Key Attack

A key is number or a secret code necessary for information security. Although having a key is a difficult and an attacker needs to resource-intensive process, it is possible. After an attacker getting a key, that key is indicated to as a compromised key.

An attacker utilize the compromised key to get access to a secured communication without noticing the sender or receiver of the attack. With the compromised key, the attacker can modify or decode data, and try to use the compromised key to compute others keys, which may lead the attacker to access to other secured communications (Newsome et al., 2004).

Sniffer Attack

A sniffer is a device or an application that can monitor, read and capture network information or data exchanges and read network packets. If the packets are not encoded, then a sniffer able to view of the data inside the packet. Even though, tunneled packets can be broken, open and read, but the attacker cannot have access to the keys because they are encrypted (Trabelsi et al., 2005).

Any of the following can be done by a sniffer attack:

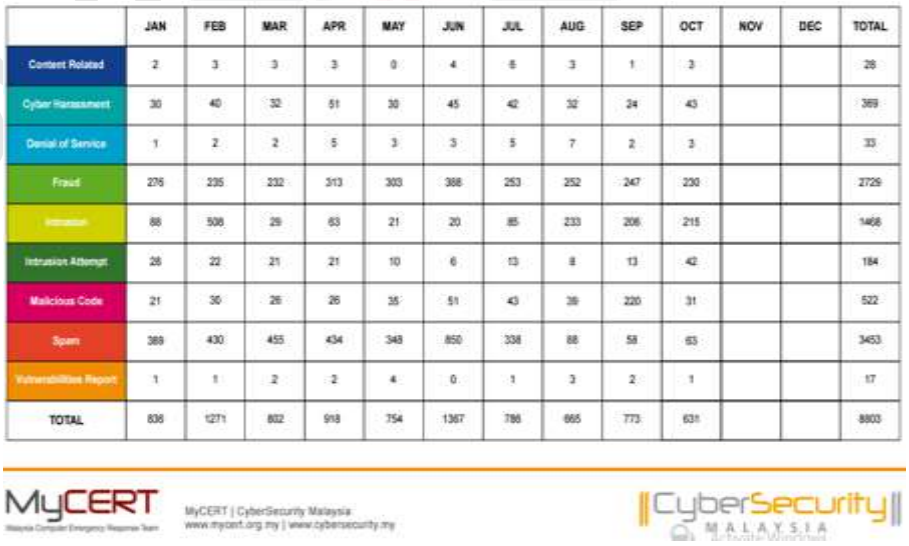
- Analyze the network and obtaining the information which it cause the network to be corrupted or to become crashed.
- Read the whole communications.

Application-Layer Attack

An application-layer attack goals application servers exactly causing an error in applications or a servers of operating system. This lead to the attacker winning the ability to bypass normal access controls (Xie et al., 2009). The attacker utilize of this situation, gaining control of the application, network, or system, and able to do any of the following:

- Read, delete, add, or modify the operating system or data.
- Send a virus program that able to control software applications and computers then it can copy whatever data they want.
- Send a sniffer program for analyzing the network or system and obtain information that can be used to break or to corrupt the systems and network.
- Abnormally terminate operating systems or the user's data applications.
- It tries to stop other security controls to be able for future attacks.

According to the researches; we find out that users of computer who are almost connected across network physical or wireless environment are insensible of the fact that they are vulnerable to the hazard of menace. Because of the growth in high-speed evolution of the internet day by day. Valuable and sensitive information separated everywhere in the network. In addition, real-time is almost provided by the internet to the users, there are problems in security information which some of them are visible to threats. Nowadays, Servers are the most possible part of systems that are under threats of attackers and might make them paralyzed which costs a large amount of financial loss and a viability of business. Yahoo suffered from DDos attack and was disabling to continue serve around one million users for three hours, at 7th February 2000. One day after the incident appeared on the other online providers such as Buy.com, Amazon, eBay and CNN which lead these providers to loss near to 1.1 million USD $. Figure 1.10 is a statistic from Malaysia Computer Emergency Response Team (MYCERT) sourced from http://www.mycert.org.my, showing raise the number of attack reports growth on monthly basis throughout the year 2015.



	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
Content Related	2	3	3	3	0	4	6	3	1	3			28
Cyber Harassment	30	40	32	51	30	45	42	32	24	43			369
Denial of Service	1	2	2	5	3	3	5	7	2	3			33
Fraud	276	235	232	313	303	368	253	252	247	230			2729
Intrusion	88	508	29	63	21	20	85	233	206	215			1468
Intrusion Attempt	28	22	21	21	10	6	13	8	13	42			184
Malicious Code	21	30	26	26	35	51	43	39	220	31			522
Spam	389	430	455	434	348	850	338	88	58	63			3453
Vulnerabilities Report	1	1	2	2	4	0	1	3	2	1			17
TOTAL	836	1271	802	918	754	1367	786	665	773	631			8803

Figure 1.10: Statistic of Reported Incident in 2015

Figure 1.10 shows the statistic of the reported incident based on general incident in 2015, source from http://www.mycert.org.my.

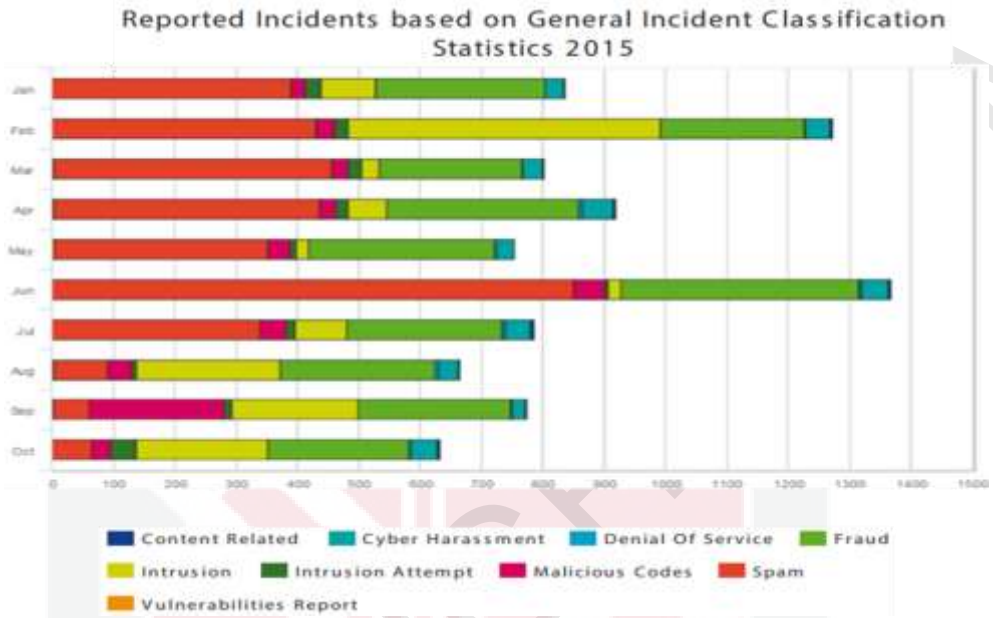


Figure 1.11: Graph of Reported Incident in 2015

1.4 Problem Statement

Without doubt, good anomaly detection in IDS must achieve high accuracy and detection rate, at the sometime false alarm rate have to be low as much as possible to be reliable IDS. At the recent year, there has been a huge effort to enhance the existing technique for anomaly detection because of high false alarm rate as well as average rate of accuracy and detection rate. In addition, the performance of single classifier has the weakness of classification the incoming data due to it classifies the unknown attacks during the detection process by wrong way which lead to reduce the performance of detection in term of accuracy and detection rate. To improve the efficiency of anomaly detection, we need to solve some

issues like predicating attacks as a normal data and predicating normal data as an attack.

In specific, many hybrid approaches have been introduced for anomaly detection in IDS and they have been successfully determined a number of intrusions correctly like , combining K-Means clustering with MLP classifier (Lisehroodi, M. M et al., 2014), gathering X-Means clustering with RF classifier (Juma, S, et al., 2014), combining K-means clustering with Naïve Bayes Classifier (Yassin, W et al., 2013) random forest sand weighted K-Means (Reda M. el basiony et at., 2013), K-Means clustering technique and 1R classifier (Muda, Z et at .,2011), combining BIRCH Clustering with SVM classifier (Horng et al., 2011), selection of features with SVM (Amiri et al., 2011), Fuzzy Multi-Class Support Vector (Haibin Zhuet et al., 2010); but there are still chance to enhance the accuracy and detection rate at the same time decrease false alarm rate .

The presence of false alarm rate with moderate accuracy and detection rate can be a potential issue for most of proposed approaches. To solve these drawbacks, we suggest a new hybrid approaches that combines X-Means clustering (XM) and Multilayer Perceptron classification technique (MLP) for anomaly detection in IDS.

1.5 Objective of Research

The main goal of this research is to raise the rate of accuracy and detection and lower the rate of false alarm by proposing an enhanced method (XM-MLP). The X-Means (XM) clustering technique has chosen to gather each data according to their behavior. Then, classifier technique (MLP) goes to classify clustered data

into attack and normal. The outcome consists of four categories: true positive, true negative, false positive and finally false negative.

1.6 Scope of Research

The domain of this research is in hybrid mining approach, which can be utilized for analyze the data to find the suitable patterns to separate normal instance and an attack instance in correct way. For this purpose, we combine two kind of techniques; one of them uses for clustering incoming data to the system which it calls X-Means clustering and the second one used for classifying clustered data which it calls Multi-layer perceptron (MLP) classification. Moreover, the proposed hybrid approach will be evaluated by using ISCX 2012 benchmarking which it created to assess the new approaches for IDS in term of accuracy detection rate and false alarm rate.

1.7 Thesis Structure

This thesis organized based on the standard organization and arrangement of thesis for Universiti of Putra Malaysia (UPM). It divided into six chapters as follow:

Chapter 1- Introduction. This chapter explains the important information on computer and network security. Then, it focuses on IDSs in general as a security tool at the same time it clarifies the current security issues forms the problem statement and research objective.

Chapter 2- literature Review. This chapter introduces review study on related work to Intrusion detection system in general. Then, it specifies the review to hybrid approaches which they are widely used by the researcher and they consider as research area for this thesis. After that, it goes to compare between some of reviewed hybrid approaches.

Chapter 3- Research Methodology. This chapter demonstrates the research steps, which consist of problem identification, dataset selection, and design of proposed approach, implantation of proposed hybrid approach and finally experiment and analysis.

Chapter 4- Proposed Hybrid Mining Approach. This chapter demonstrates the combination of two data mining approaches and clarifies the process and functions each of them as single approach. Then, it shows steps of hybrid algorithm that combines both clustering and classification.

Chapter 5-Result and Discussion. This chapter shows the result that been obtained by single classifier, proposed hybrid approach and also the result of previous hybrid approach by (Lisehroodi, M. M et al., 2014) and other previous hybrid

approaches. Then it shows the comparison of single and previous hybrid approach against proposed hybrid approach.

Chapter 6-Conclusion and Future Work. This chapter summarizes the whole research concept and goes to introduce the plane and concept of future work.

REFERENCES

Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J. (1999, June). OPTICS: ordering points to identify the clustering structure. In ACM Sigmod Record (Vol. 28, No. 2, pp. 49-60). ACM.

Amor, N. B., Benferhat, S., & Elouedi, Z. (2004, March). Naive bayes vs decision trees in intrusion detection systems. In Proceedings of the 2004 ACM symposium on Applied computing (pp. 420-424). ACM.

Abdalla, M., & Pointcheval, D. (2005). Simple password-based encrypted key exchange protocols. In Topics in cryptology–CT-RSA 2005 (pp. 191-208). Springer Berlin Heidelberg

Al-Ofeishat, H. A., & Mohammad, A. A. A. L. (2012). Near field communication (NFC). International Journal of Computer Science and Network Security, 12(2), 93-99

Aziz, A. S., Hassanien, A. E., Hanaf, S. E. O., & Tolba, M. F. (2013, December). Multi-layer hybrid machine learning techniques for anomalies detection and classification approach. In Hybrid Intelligent Systems (HIS), 2013 13th International Conference on (pp. 215-220). IEEE.

Abou Haidar, G., & Boustany, C. (2015, July). High Perception Intrusion Detection System Using Neural Networks. In Complex, Intelligent, and Software Intensive Systems (CISIS), 2015 Ninth International Conference on (pp. 497-501). IEEE

Bace, R. and Mell, O.(2001). Intrusion detection system. NIST Specieal Publications SP:800-31.

Biermann, E. (2001). A comparison of Intrusion Detection systems. Computers & Security, 20, 676–683.

Brenton, Chris, and Hunt Cameron.(2002) “Mastering Network Security.

Burton, J. D. (2003). Cisco security professional's guide to secure intrusion detection systems. Syngress Publ.

Brugger, S. T. (2004). Data mining methods for network intrusion detection.*University of California at Davis*.

Chan, P. K., Mahoney, M. V., & Arshad, M. H. (2005). Learning rules and clusters for anomaly detection in network traffic. In *Managing Cyber Threats*(pp. 81-99). Springer US.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, *41*(3), 15.

Chih-Frong Tsai, Hsu, Y. F., Lin, C. Y., & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. Expert Systems with Applications, 36(10), 11994-12000.

Chandrashekhar, A. M., & Raghuveer, K. (2013). Fortification of hybrid intrusion detection system using variants of neural networks and support vector machines. International Journal of Network Security & Its Applications, 5(1), 71-90.

Callegari, C., Casella, A., Giordano, S., Pagano, M., & Pepe, T. (2013, October). Sketch-based multidimensional IDS: A new approach for network anomaly detection. In Communications and Network Security (CNS), 2013 IEEE Conference on (pp. 350-358). IEEE.

Debar, H. (2000). An introduction to intrusion-detection systems. *Proceedings of Connect, 2000*.

Dieter Gollmann. (2002). Computer Security, Second Edition. Wiley, New Jersey**,**

Dokas, P., Ertoz, L., Kumar, V., Lazarevic, A., Srivastava, J., & Tan, P. (2002). Data Mining for Network Intrusion Detection. *In Proceedings of the NSF Workshop on Next Generation Data Mining, Baltimore, MD*, 21–30.

Di Pietro, R., & Mancini, L. V. (2008). *Intrusion detection systems* (Vol. 38). Springer Science & Business Media.

Desmedt, Y. (2011). Man-in-the-middle attack. In Encyclopedia of Cryptography and Security (pp. 759-759). Springer US

Day, D. J., Flores, D., & Lallie, H. S. (2012, June). CONDOR: A Hybrid IDS to Offer Improved Intrusion Detection. In Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on (pp. 931-936). IEEE.

Dai, H. N., Wang, Q., Li, D., & Wong, R. C. W. (2013). On eavesdropping attacks in wireless sensor networks with directional antennas. International Journal of Distributed Sensor Networks, 2013.

Eskin, E., Arnold, A., Prerau, M., Portnoy, L., & Stolfo, S. (2002). A geometric framework for unsupervised anomaly detection. In *Applications of data mining in computer security* (pp. 77-101). Springer US.

Ertöz, L., Steinbach, M., & Kumar, V. (2003). Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data. *Technical Report*. Retrieved from

Erman, J., Arlitt, M., & Mahanti, A. (2006, September). Traffic classification using clustering algorithms. In *Proceedings of the 2006 SIGCOMM workshop on Mining network data* (pp. 281-286). ACM.

Engen, V., Vincent, J., & Phalp, K. (2008). Enhancing network based intrusion detection for imbalanced data. International Journal of Knowledge-Based and Intelligent Engineering Systems, 12(5, 6), 357-367

Elbasiony, R. M., Sallam, E. A., Eltobely, T. E., & Fahmy, M. M. (2013). A hybrid network intrusion detection framework based on random forests and weighted k-means. Ain Shams Engineering Journal, 4(4), 753-762.

Fernandez, M. D. M., & Porres, I. (2008). An Evaluation of current IDS.

Guha, S., Rastogi, R., & Shim, K. (1998, June). CURE: an efficient clustering algorithm for large databases. In ACM SIGMOD Record (Vol. 27, No. 2, pp. 73-84). ACM.

Gollmann, D. (2002). An Immunogenetic Technique to Detect Anomalies in Network Traffic. *In Proceedings of the Genetic and Evolutionary Compuation Conference (GECCO)*, 1081–1088.

Garfinkel, T., & Rosenblum, M. (2003, February). A Virtual Machine Introspection Based Architecture for Intrusion Detection. In *NDSS* (Vol. 3, pp. 191-206).

Gaddam, S. R., Phoha, V. V., & Balagani, K. S. (2007). K-means+ id3: A novel method for supervised anomaly detection by cascading k-means clustering and id3 decision tree learning methods. Knowledge and Data Engineering, IEEE Transactions on, 19(3), 345-354.

Gabra, H. N., Bahaa-Eldin, A. M., & Korashy, H. (2014). Classification of ids alerts with data mining techniques. *arXiv preprint arXiv:1401.4872*.

Hourdakis, N., Argyriou, M., Petrakis, E. G., & Milios, E. E. (2010). Hierarchical Clustering in Medical Document Collections: the BIC-Means Method. JDIM, 8(2), 71-77.

Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley & Sons, Inc..

Holte, R. C. (1993). Very simple classification rules perform well on most com.

HOURDAKIS, N. (2006). DTESIGNT AND EVALUATION OF CLUSTERING APPROACHES FOR LARGE DOCUMENT COLLECTIONS, THE "BIC-MEANS" METHOD (Doctoral dissertation, Technical University of Crete).

Han, L. I. (2010, October). Research and implementation of an anomaly detection model based on clustering analysis. In Intelligence Information Processing and Trusted Computing (IPTC), 2010 International Symposium on (pp. 458-462). IEEE.

Horng, S. J., Su, M. Y., Chen, Y. H., Kao, T. W., Chen, R. J., Lai, J. L., & Perkasa, C. D. (2011). A novel intrusion detection system based on hierarchical clustering and support vector machines. *Expert systems with Applications*, *38*(1), 306-313.

Ishioka, T. (2005, July). An expansion of X-means for automatically determining the optimal number of clusters. In Proceedings of International Conference on Computational Intelligence (pp. 91-96).

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice-Hall, Inc.

John, G. H., & Langley, P. (1995, August). Estimating continuous distributions in Bayesian classifiers. In Proceedings of the Eleventh conference on Uncertainty in artificial intelligence (pp. 338-345). Morgan Kaufmann Publishers Inc.

Javidi, M. M., & Nattaj, M. H. (2013). A New and Quick Method to Detect DoS Attacks by Neural Networks. The Journal of mathematics and computer Science, 6, 85-96.

JUMA, S., MUDA, Z., & YASSIN, W. (2014). REDUCING FALSE ALARM USING HYBRID INTRUSION DETECTION BASED ON X-MEANS CLUSTERING AND RANDOM FOREST CLASSIFICATION. Journal of Theoretical & Applied Information Technology, 68(2).

JUMA, S., MUDA, Z., MOHAMED, M., & YASSIN, W. (2015). MACHINE LEARNING TECHNIQUES FOR INTRUSION DETECTION SYSTEM: A REVIEW. Journal of Theoretical & Applied Information Technology, 72(3).

Kaufman, L. and Rousseeuw, P. (1990). Finding groups in data: An introduction to cluster analysis. John Wiley and Sons, New york, NY.

Kozushko, H. (2003). Intrusion detection: host-based and network-based intrusion detection systems. *on September*, *11*.

Kruegel, C., Valeur, F., & Vigna, G. (2004). Intrusion detection and correlation: challenges and solutions (Vol. 14). Springer Science & Business Media.

Kayacik, H. G. (2007, May). On the contribution of preamble to information hiding in mimicry attacks. In *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on*(Vol. 1, pp. 632-638). IEEE.

Khor, K. C., Ting, C. Y., & Amnuaisuk, S. P. (2010, March). Comparing single and multiple Bayesian classifiers approaches for network intrusion detection. In Computer Engineering and Applications (ICCEA), 2010 Second International Conference on (Vol. 2, pp. 325-329). IEEE.

Kavitha, B., Karthikeyan, D. S., & Sheeba Maybell, P. (2012). An ensemble design of intrusion detection system for handling uncertainty using Neutrosophic Logic Classifier. *Knowledge-Based Systems*, *28*, 88–96. doi:10.1016/j.knosys.2011.12.004

Lee, W., Stolfo, S. J., Chan, P. K., Eskin, E., Fan, W., Miller, M., ... & Zhang, J. (2001). Real time data mining-based intrusion detection. In *DARPA Information Survivability Conference & Exposition II, 2001. DISCEX'01. Proceedings* (Vol. 1, pp. 89-100). IEEE.

Liu, F., & Luo, L. (2005). Immune Clonal Selection Wavelet Network Based. *Artificial Neural Networks: Biological Inspirations–ICANN 2005*, 331–336.

Lu, J., Plataniotis, K. N., Venetsanopoulos, A. N., & Li, S. Z. (2006). Ensemble-based discriminant learning with boosting for face recognition. Neural Networks, IEEE Transactions on, 17(1), 166-178.

Latifur, Mamoun Awad, and Bhavani Thuraisingham. "A new intrusion detection system using support vector machines and hierarchical clustering." The VLDB Journal—The International Journal on Very Large Data Bases 16, no. 4 (2007): 507-521.

Long, J., Schwartz, D., & Stoecklin, S. (2008). Case-oriented alert correlation. *Journal of Computer Security*, *7*(3). Retrieved from http://www.wseas.us/e-library/transactions/computers/2008/30-393N.pdf

Li, W., & Li, Q. (2010, November). Using Naive Bayes with AdaBoost to Enhance Network Anomaly Intrusion Detection. In Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on (pp. 486-489). IEEE

Lisehroodi, M. M., Muda, Z., Yassin, W., & Udzir, N. I. (2014). KM-NEU: an efficient hybrid approach for intrusion detection system. Research Journal of Information Technology, 6(1), 46-57.

Maiwald, E. (2001). Network security: a beginner's guide. McGraw-Hill Professional.

Moore, V. Paxson, S. Savage, C. Shannon, S. S. and N. W. (2003). Inside the Slammer Worm. IEEE Security and Privacy, 33–39

Mirkovic, J., Dietrich, S., Dittrich, D., & Reiher, P. (2004). Internet Denial of Service: Attack and Defense Mechanisms (Radia Perlman Computer Networking and Security). Prentice Hall PTR

Moradi, M., & Zulkernine, M. (2004, November). A neural network based system for intrusion detection and classification of attacks. In Proceedings of the 2004 IEEE international conference on advances in intelligent systems-theory and applications.

Meera Gandhi, G., & Srivatsa, S. K. (2010). Classification Algorithms in Comparing Classifier Categories to Predict the Accuracy of the Network Intrusion Detection-A Machine Learning Approach. Advances in Computational Sciences & Technology, 3(3).

Muda, Z., Yassin, W., Sulaiman, M. N., & Udzir, N. I. (2011, December). Intrusion detection based on k-means clustering and OneR classification. In Information Assurance and Security (IAS), 2011 7th International Conference on (pp. 192-197). IEEE.

Newsome, J., Shi, E., Song, D., & Perrig, A. (2004, April). The sybil attack in sensor networks: analysis & defenses. In Proceedings of the 3rd international symposium on Information processing in sensor networks (pp. 259-268). ACM.

NEEDHAM, A. B. R., & Lampson, B. (2008). Network Attack and Defense. whit paper.

Nguyen, H. A., & Choi, D. (2008). Application of data mining to network intrusion detection: classifier selection model. In Challenges for Next Generation Network Operations and Service Management (pp. 399-408). Springer Berlin Heidelberg.

Orchier, J., Soriano, R., Salvaterra, L., Ardito, D., & Byreddy, A. (2000).*U.S. Patent No. 6,070,244*. Washington, DC: U.S. Patent and Trademark Office.

Owens , R.R.Levary, S. F. (2006). An adaptive expert system approach for intrusion detection. *S.F. Owens and R.R. Levary*, 1, 206–217. doi:ISSN 1747-8405

Pelleg, D., & Moore, A. W. (2000, June). X-means: Extending K-means with Efficient Estimation of the Number of Clusters. In ICML (pp. 727-734).

Popescu, B. E. (2004). Ensemble learning for prediction.

Peddabachigari, S., Abraham, A., Grosan, C., & Thomas, J. (2007). Modeling intrusion detection system using hybrid intelligent systems. *Journal of network and computer applications*, *30*(1), 114-132.

Panda, M., & Patra, M. R. (2007). Network intrusion detection using naive bayes. *International journal of computer science and network security*, *7*(12), 258-263.

Panda, M., Abraham, A., Das, S., & Patra, M. R. (2011). Network intrusion detection system: A machine learning approach. Intelligent Decision Technologies, 5(4), 347-356.

Pilabutr, S., Somwang, P., & Srinoy, S. (2011, December). Integrated soft computing for Intrusion Detection on computer network security. In

Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on (pp. 559-563). IEEE.

Pevny, T., Komon, M., & Rehaky, M. (2013, May). Attacking the IDS learning processes. In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on (pp. 8687-8691). IEEE

Robinson, J. T. (1981, April). The KDB-tree: a search structure for large multidimensional dynamic indexes. In Proceedings of the 1981 ACM SIGMOD international conference on Management of data (pp. 10-18). ACM.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation (No. ICS-8506). CALIFORNIA UNIV SAN DIEGO LA JOLLA INST FOR COGNITIVE SCIENCE.

Ryan, J., Lin, M. J., & Miikkulainen, R. (1998). Intrusion detection with neural networks. *Advances in neural information processing systems*, 943-949.

Selim, S. S. Z., & Ismail, M. a. (1984). K-Means-Type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *PAMI-6*(1), 81–87. doi:10.1109/TPAMI.1984.4767478

Tayrani, R., Teshiba, M., Sakamoto, G. M., Chaudhry, Q., Alidio, R., Kang, Y., ... & Hauhe, M. (2003). Broad-band SiGe MMICs for phased-array radar applications. *Solid-State Circuits, IEEE Journal of*, *38*(9), 1462-1470.

Tittle, E., Stewart, J. M., & Chapple, M. (2006). *CISSP: Certified information systems security professional study guide*. John Wiley & Sons.

Shamsuddin, S. B., & Woodward, M. E. (2008). Applying knowledge discovery in database techniques in modeling packet header anomaly intrusion detection systems. *Journal of Software*, *3*(9), 68-76.

Suduc, A. M., Bîzoi, M., & Filip, F. G. (2010). Audit for information systems security. *Informatica Economica*, *14*(1), 43-48.

Shahbaba, M., & Beheshti, S. (2012, July). Improving x-means clustering with mndl. In Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on (pp. 1298-1302). IEEE.

Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. A. (2012). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Computers & Security, 31(3), 357-374.

Sato, M., Yamaki, H., & Takakura, H. (2012, July). Unknown attacks detection using feature extraction from anomaly-based ids alerts. In Applications and the Internet (SAINT), 2012 IEEE/IPSJ 12th International Symposium on (pp. 273-277). IEEE.

Trabelsi, Z., & Rahmani, H. (2005). An Anti-Sniffer Based on ARP Cache Poisoning Attack. Information Systems Security, 13(6), 23-36.

Tittle, E., Stewart, J. M., & Chapple, M. (2006). CISSP: Certified information systems security professional study guide. John Wiley & Sons.

Tsang, C.-H., Kwong, S., & Wang, H. (2007). Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection. *Pattern Recognition*, *40*(9), 2373–2391. doi:10.1016/j.patcog.2006.12.009

Tsai, C. F., & Lin, C. Y. (2010). A triangle area based nearest neighbors approach to intrusion detection. Pattern Recognition, 43(1), 222-229.

Vigna, G., & Kemmerer, R. A. (1999). NetSTAT: A network-based intrusion detection system. In *Journal of computer security*.

Vapnik, V .The Nature of Statistical Learning Theory. Springer-Verleg.

Vigna, R. A. K. and G. (2002). Intrusion Detection: A Brief History and Overview. *computer*, 35, 27–30.

Varshovi, A., Rostamipour, M., & Sadeghiyan, B. (2014, May). A fuzzy Intrusion Detection System based on categorization of attacks. In Information and Knowledge Technology (IKT), 2014 6th Conference on (pp. 50-55). IEEE.

Wu, T. F., Lin, C. J., & Weng, R. C. (2004). Probability estimates for multi-class classification by pairwise coupling. *The Journal of Machine Learning Research*, *5*, 975-1005.

Wang, G., Hao, J., Ma, J., & Huang, L. (2010). A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. *Expert Systems with Applications*, *37*(9), 6225–6232. doi:10.1016/j.eswa.2010.02.102

Xie, Y., & Yu, S. Z. (2009). Monitoring the application-layer DDoS attacks for popular websites. Networking, IEEE/AcM Transactions on, 17(1), 15-25.

Xue, M., & Zhu, C. (2009, April). Applied research on data mining algorithm in network intrusion detection. In Artificial Intelligence, 2009. JCAI'09. International Joint Conference on (pp. 275-277). IEEE.

Xiuqing, C., Ongping, Z. Y., & Jiutao, T. (2010, August). HMM-based integration of multiple models for intrusion detection. In Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on (Vol. 2, pp. V2-137). IEEE.

Yaar, A., Perrig, A., & Song, D. (2006). StackPi: New packet marking and filtering mechanisms for DDoS and IP spoofing defense. Selected Areas in Communications, IEEE Journal on, 24(10), 1853-1863.

Yu, Y., Wei, Y., Fu-Xiang, G., & Yu, Y. (2006, October). Anomaly intrusion detection approach using hybrid MLP/CNN neural network. In Intelligent Systems Design and Applications, 2006. ISDA'06. Sixth International Conference on (Vol. 2, pp. 1095-1102). IEEE.

Yue, W. T., & Cakanyildirim, M. (2007). Intrusion prevention in information systems: Reactive and proactive responses. *Journal of Management Information Systems*, *24*(1), 329-353.

Yasami, Y., & Mozaffari, S. P. (2010). A novel unsupervised classification approach for network anomaly detection by k-Means clustering and ID3 decision tree learning methods. *The Journal of Supercomputing*, *53*(1), 231-245.

Yassin, W., Udzir, N. I., Muda, Z., & Sulaiman, M. N. (2013, August). Anomaly-Based Intrusion Detection through K-Means Clustering and Naives Bayes Classification. In Proceedings of the 4th International Conference on Computing and Informatics (ICOCI), Sarawak, Malaysia (pp. 298-303).

Yassin, W., Udzir, N. I., Abdullah, A., Abdullah, M. T., Muda, Z., & Zulzalil, H. (2014, January). Packet Header Anomaly Detection Using Statistical Analysis. In International Joint Conference SOCO'14-CISIS'14-ICEUTE'14 (pp. 473-482). Springer International Publishing.

Yassin, W., Udzir, N. I., Abdullah, A., Abdullah, M. T., Zulzalil, H., & Muda, Z. (2014, August). Signature-Based Anomaly intrusion detection using Integrated data mining classifiers. In Biometrics and Security Technologies (ISBAST), 2014 International Symposium on (pp. 232-237). IEEE.

Zhang, C., Jiang, J., & Kamel, M. (2005). Intrusion detection using hierarchical neural networks. Pattern Recognition Letters, 26(6), 779-791.

Zhang, W., Teng, S., Zhu, H., Du, H., & Li, X. (2010, July). Fuzzy Multi-Class Support Vector Machines for Cooperative Network Intrusion detection. In Cognitive Informatics (ICCI), 2010 9th IEEE International Conference on (pp. 811-818). IEEE.

BIODATA OF STUDENT

Borkan Ahmed Abbas was born in Iraq on 8th April 1987. He obtained Degree in Computer Science from Tikret University College on 2011. He peruses his Master of Computer Science and Information Technology majoring in Distributed Network Department at Universiti Putra Malaysia by focusing in Intrusion detection System for Network Security. His research interests include investigating new approach for Intrusion Detection System (IDS) as he did his Master project in the same area. Recently, he did some publications in these areas.