MEU جامعة الشرق الأوسط
MIDDLE EAST UNIVERSITY

Intrusion Detection Model Inspired by Immune Using K-Means and Naive Bayes as Hybrid Learning Approach

By

Emad Fares Salim Islim

Supervised By

Dr. Hazim A. Farhan

Master Thesis

Submitted in Partial Fulfillment of the Requirements for the Master Degree In Computer Science

Department of Computer Science
Faculty of Information Technology
Middle East University
Amman – Jordan

December, 2011

I

Middle East University

Authorization Statement

I, Emad Fares Salim Islim, authorize Middle East University to supply hardcopies and electronic copies of my thesis to libraries, establishments, or bodies and institutions concerned with research and scientific studies upon request, according to the university regulations.

Name : Emad Fares Salim Islim

Date : 25/12/2011

Signature :

Middle East University

Examination Committee Decision

This is to certify that the thesis entitled **"Intrusion Detection Model Inspired by Immune Using K-Means and Naive Bayes as Hybrid Learning Approach"** was successfully defend and approved on January 23rd 2012.
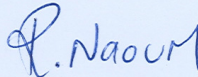
Examination Committee Member   **Signature**

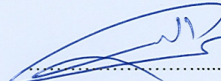1- Prof. Reyadh S. Naoum   Chairman

Professor
Dean of Faculty of Information Technology
Middle East University

2- Dr. Abdelfatah A. Tamimi   Member

Associate Professor
Dean of Faculty of Science and Information
Technology
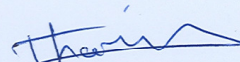Alzaytoonah University of Jordan

3- Dr. Hazim A. Farhan   Supervisor

Associate Professor   and
Department of Computer Science
Faculty of Information Technology   Member
Alzaytoonah University of Jordan

Acknowledgments

"In the name of Allah the Most Gracious the Most Merciful". My guidance can not come except from Allah, in Him I trust, to Him I repent, and to Him praise and thanks always go.

I offer my sincerest gratitude to my advisor Dr. Hazim Farhan for his valuable contributions, knowledge, encouragement and helpful advices. Also, I would like to express a very special thanks to Prof. Reyadh Naoum, for his vision which brought this work forward and for being there any time I knocked on his door. I wish both of them more and more success and giving.

I am highly indebted to my parents who taught me the right things, encouraged and gave me the hope and unconditional love. I wish both of them happiness and good health. Thanks for my brothers, sisters, relatives and friends for supporting me.

Also, I offer my sincerest gratitude to Mr. Ismail Najib for his valuable contributions, knowledge and for clarifying the ambiguity of the medical issues related to human immune system. I wish him more and more success and happiness.

Very special thanks belong to my wife and partner for life for being with me in happiness and sadness, giving me hope and strength and supporting me through this thesis. Also very special thanks go to my daughters and sons for their patience and smiles. To all above mentioned persons, this thesis couldn't have been done without your support.

Dedication

I dedicate this work to my father, my mother, my wife and partner for life, my children, my brothers and sisters; for their love, understanding and support, they were the light in my path. Without them nothing of this would have been possible. Thank you for everything. I love you!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CERT	Computer Emergency Response Team.
CSI	Computer Security Institute.
FN	False Negative Means that Attack Occur but No Alarm.
FP	False Positive Means that No Attack but There is Alarm.
HIDS	Host-based Intrusion Detection Systems.
HIS	Human Immune System.
IDS	Intrusion Detection System.
TN	True Negative Means that No Attack and No Alarm.
TP	True Positive Means that Attack Occur and There is Alarm.
VQ	Vector Quantization..
Weka	Waikato Environment for Knowledge Analysis.

Abstract

Intrusions in computer networks can be compared to human diseases with the difference that human body has an effective mechanism to deal with them. Human immune system can detect and defend against unseen intruders. Also, it is distributed and adaptive. Human immune system is the most powerful defense system which may be helpful to apply its mechanism and properties in computer security field. This thesis presents a model for intrusion detection system that consists of four components depending on innate/adaptive human immune system approaches and self/non-self theory of human immune system. The proposed model is divided into two subsystems; the first one is attack response system which is similar to innate human immune system and the second is learning system which is similar to adaptive immune system. Learning system is the core of the model; it presents a hybrid approach of machine learning through hybridization between k-Means clustering algorithm and Naive Bayes as a classifier. The model goal is keeping information systems environment safe against intrusions and attacks through applying human immune system mechanism and properties to intrusion detection system. Experimental results indecate that the proposed model provide a higher detection rate in both of DoS attacks and U2R attacks, which give the power to the proposed hybrid model and increase the security of information systems, especially in the critical environments.

Keywords : **Intrusion Detection System, Human Immune System, K-Means Clustering Algorithm, Naive Bayes, Computer Immunology.**

الملخص

يمكن مقارنة الاختراقات والهجمات على شبكات الحاسوب بالأمراض التي تصيب الإنسان مع الأخذ بعين الاعتبار بأن الجسم البشري يملك آلية فعالة للتعامل مع تلك الأمراض ، حيث يستطيع نظام المناعة في جسم الإنسان التعرف على مسببات الأمراض التي لم يسبق له التعرض لها ومقاومتها ، وذلك لكون هذا النظام يتمتع بخاصية التكيف. يعتبر نظام المناعة في جسم الانسان أقوى نظام دفاعي ، حيث يمكننا الاستفادة من دراسة هذا النظام و آليات عمله وخصائصه في مجال أمن الحاسوب و المعلومات. تهدف هذه الرسالة الى تقديم نموذج لكشف الاختراقات يتكون من أربعة مكونات رئيسية معتمداً على مفهومي نظام المناعة البشري الفطري و نظام المناعة البشري المتكيف وعلى نظرية التمييز بين المكونات الذاتية وغير الذاتية للجسم البشري. يحتوي هذا النموذج على نظامين فرعيين هما نظام الاستجابة للهجمات والذي يشبه نظام المناعة البشري الفطري، أما النظام الفرعي الثاني فهو نظام التعلم والذي يشبه نظام المناعة البشري المتكيف، ويعتبر نظام التعلم هو الجوهر الأساسي لهذه الرسالة حيث يقدم نموذجاً مهجناً لتعلم الآلة من خلال التهجين بين خوارزمية K-Means لتقسيم البيانات و خوارزمية Naive Bayes لتصنيف البيانات وذلك بهدف الحفاظ على انظمة المعلومات و بيئتها بشكل آمن و حمايتها من الاختراقات والهجمات ، وذلك بالاعتماد على آليات و خصائص نظام المناعة في جسم الانسان. لقد أشارت النتائج التجريبية بأن النموذج المقترح قد حقق نسب اكتشاف أعلى للاختراقات من نوعي DoS و U2R مما يعطي القوة للنموذج المهجن المقترح ويمكنه من الحفاظ على أمن المعلومات وبخاصة في بيئة المعلومات ذات الخصوصية والحساسية العالية.

Chapter One

Introduction

Chapter one

Introduction

1.1 Introduction.

Due to the revolution in telecommunications and information technology, the information systems security became an important problem and a major issue in computer security field in the past few years. It has transferred the world into a small digital village. The growth of information systems networks and their infrastructure led to growing attack rates against information systems environment. While a completely secured system is still a myth, a real need appeared for understanding the main challenge in computer security which is determining the difference between normal and abnormal activities in order to prevent intrusions from damaging or attacking the information system environment.

The increase in internet usage plays a main role in security problem. In 2010 there were 1.966 billion internet users in the world versus 360.985 million internet users in 2000 which means that internet users increased 81.86% from the end of 2000 until the mid of 2010, and 28.72% of world population are using internet. They are using internet services in different activities such as: browsing, e-mail, search, social networks and e-commerce (Miniwatts Marketing Group, 2010). Intruders and cyber crimes are expanding in parallel with the growth in information technology environment. These crimes have many forms such as: Viruses, Worms, Trojan horses, Malware, Identity theft and more.

Intrusion Detection System is the third form of defense after antivirus and firewall applications. (The Internet Society, 2000) defined this concept as a security service that monitors and analyzes system events for the purpose of finding, and providing real-time or near real-time warning of, attempts to access system resources in an unauthorized manner. From this definition, anyone conclude that this field of research is still an open problem.

Anderson (1980) put the first stone that was used for intrusion detection systems development later. Also, he defined the concept of threat and proposed an approach for auditing data in order to recognize the threats depending on audit trails which are taken relatively long term basis; this audit data is derived from records collected daily from all machines then consolidated into a data set called "dump" that reviewed by security officers and transferred to tape after reports are generated.

In 1983, Denning began working in a U.S. government project at Stanford Research Institute International (SRI international). The project goal was analyzing audit trials of government computers and creating profiles of users upon their activities. In 1984, she helped in developing Intrusion Detection Expert System (IDES) which is the first model of IDS (Innella, 2001).

In 1988, Haystack project occurred at Lawrence Livermore National Laboratory (LLNL) at the University of California Davis. The project produced a Distributed Intrusion Detection System that was developed for U.S. Air Forces to analyze data which

are based on events collected from all network nodes such as logins, system calls, file access and to compare these audit data with the predefined attack patterns in order to detect intrusions (Endorf, Schultz and Mellander, 2004). The developers from Haystack project established a commercial company in 1989 called Haystack Labs which became the first commercial vendor of host-based intrusion detection systems (Innella, 2001).

In 1990, Heberlein developed a Network Security Monitor (NSM) at the University of California Davis, which was the first network intrusion detection system. The idea of (NSM) is developing profiles of network resources usage then comparing the current usage pattern with the previously created profiles, in order to detect intrusions (Heberlein, et al., 1990).

In 1991, Automated Security Measurement System (ASIM) was developed by U.S. Air Force's Cryptologic Support Center. It was responsible for monitoring the traffic on the U.S. Air Force's network and became the first solution that incorporated hardware and software solution to network intrusion detection. The development group of (ASIM) established a commercial company in 1994 called Wheel Group that produced NetRanger product which is the first commercially network intrusion detection device (Innella, 2001).

IDS market became a good investment and generated revenues around 1997. The market leader, Internet Security Systems (ISS), developed a network intrusion detection system called RealSecure. In 1998 the development staffs from Haystack Labs and

CMDS team from SAIC were merged to form Centrax Corporation. In the same year Cisco purchased the Wheel Group because it recognized the importance of network intrusion detection. Intrusion detection market continues growing from that time until now (Innella, 2001).
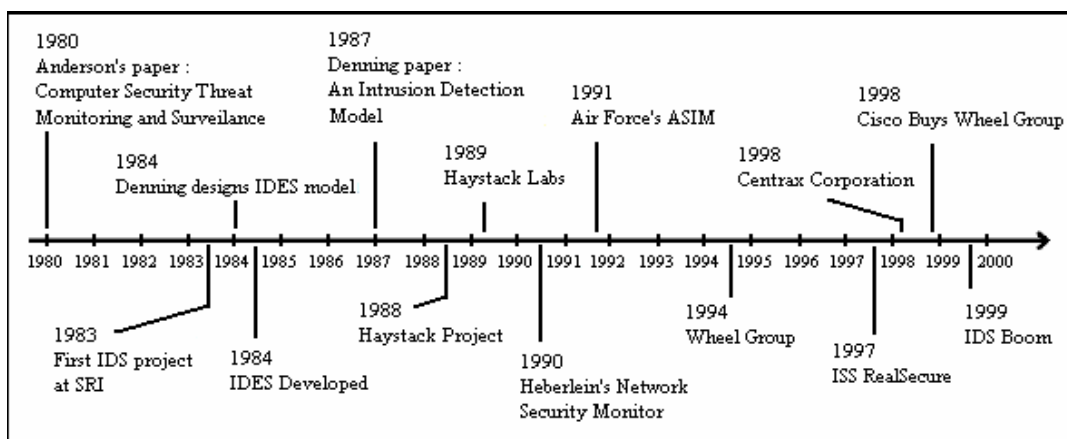
Figure 1.1: History of Intrusion Detection Systems (Innella, 2001).

Carnegie Mellon University (2011) presented a comparison between insider and outsider percentage attacks from 2004 until 2010 which showed that outsider attacks still had the lion's share versus insider attacks, as shown in Figure 1.2.
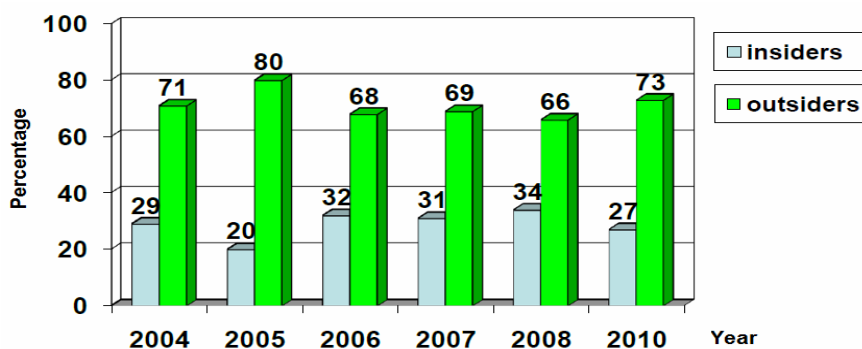
Figure 1.2: Percentage of insiders versus outsider attacks

(Carnegie Mellon University, 2011).

CERT (2009) indicated an increasing in the number of incidents reported from 1993 until 2003, as shown in Figure 1.3. Also, CERT mentioned that they stopped providing statistics at the end of 2003 because widespread uses of automated attack tools, attacks against internet-connected systems have become so commonplace that counts of the number of incidents reported provide little information with regard to assessing the scope and impact of attacks.
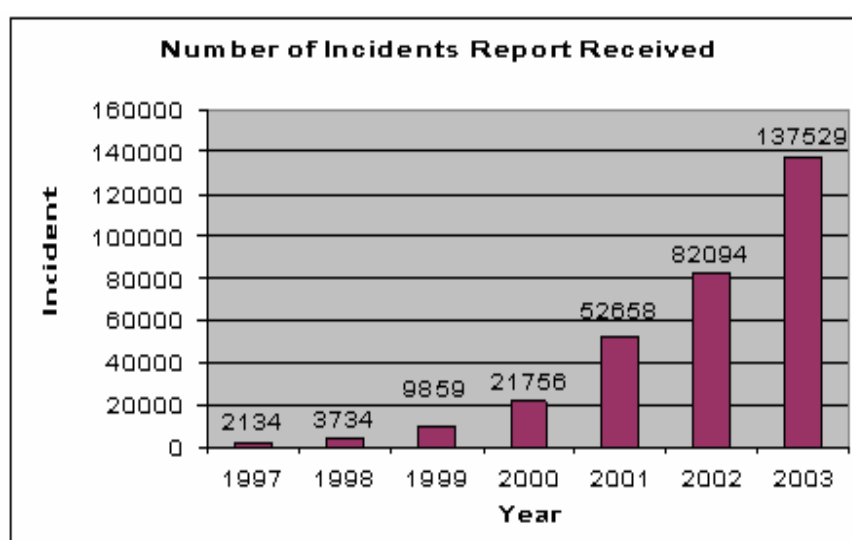
Figure 1.3: Number of Incidents Report Received (CERT, 2009).

The previous figure shows a dramatic growth in number of incidents, which represent a fact that is a very important for people, companies and organizations to keep their computers and networks safe against attacking. For an example, a person could lose his bank balance because of a successful hacking on his credit card. Also, a company could lose critical information which is very important for competitors, because of a successful internal or external attack.

CSI (2009) in its 14th annual edition shows some types of attacks experienced by percent of respondents between 2005 until 2009. By taking 2008 and 2009 as sample, it is noted that some types of attacks were increased while other types were decreased, as shown in Table 1.1.

Table 1.1. Types of attacks experienced by percent of respondents (CSI, 2009).

Type of Attack	2008	2009
Malware infection.	50%	64%
Password sniffing.	9%	17%
Denial of service.	21%	29%
Web site defacement.	6%	14%
Exploit of DNS servers.	8%	7%
Instant messaging abuse.	21%	8%
Theft of or unauthorized access to intellectual property due to all other causes.	5%	8%

Intrusion Detection System (IDS) is one of the computers and networks defense forms against attacking computer network or information systems environment. It can be defined as system monitoring a stream of information for occurrences of computer attacks (Rieck, 2009). The process of detection is reached by analyzing the collected information from different sources within a computer system or across the computer network to find signs of unusual system behavior or signs of intrusion.

These signs can be found by observing network traffic, system log files, systems or user activities. So, responding to attacks in real-time is very important to keep information systems environment safe. Also, Intrusion detection system is responsible for detecting and controlling of malicious network traffic or behavior to keep viruses, Trojan horses, worms and hackers far away from information system environment.

There are two types of intrusion detection systems; the first is Host-based intrusion detection systems (HIDS). It is small computer programs (agents) that are installed on a single host which can be any network computer or database server, as shown in Figure 1.2 These agents can monitor the operating system, system security log files, information system log file, memory, network traffic and registry in able to find any malicious activity (Moskovitch, et al., 2007).
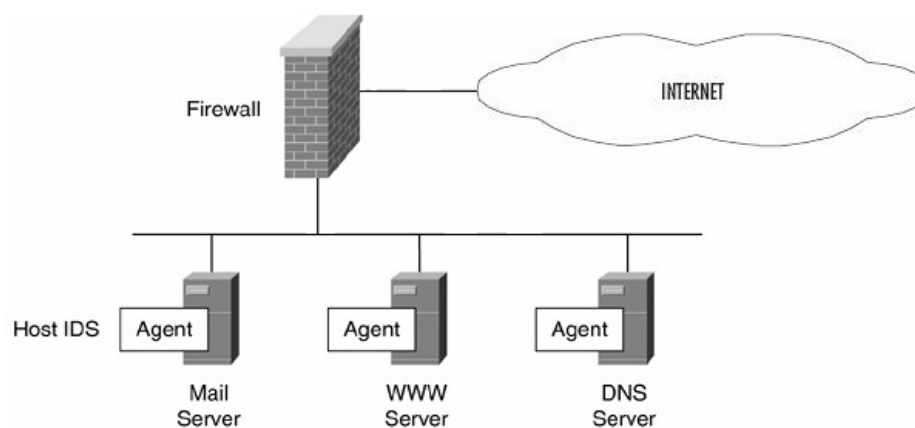
Figure 1.4: Host-based intrusion detection system (Moskovitch, et al., 2007).

The second type is network-based intrusion detection systems (NIDS), which is installed on dedicated machine within the network, as shown in Figure 1.2. Generally, it consists of computer programs (sensors or packet sniffers) that able to monitor all

network traffic by testing all packets one by one (Raghunath & Mahadeo 2008). It protects the whole network from threats by selecting points on network then start catching and analyzing network traffic at these selected points (Heberlein, et al., 1990).
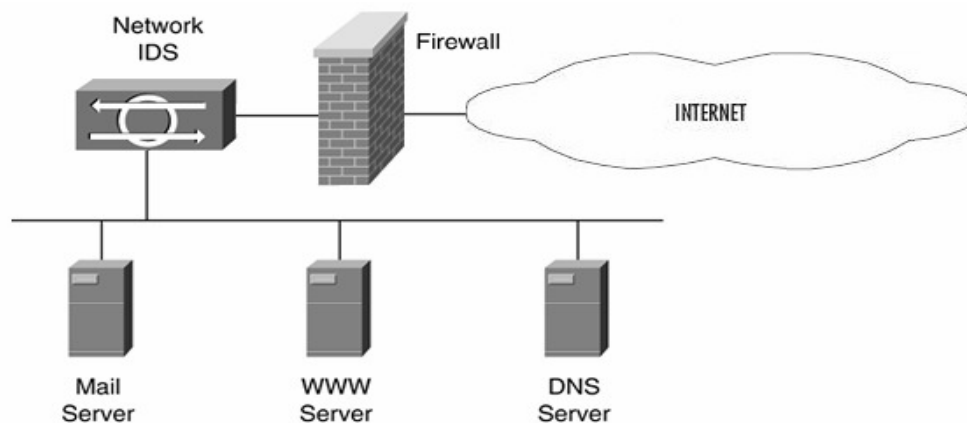
Figure 1.5: Network-based intrusion detection system (Moskovitch, et al., 2007).

There are two approaches used by all intrusion detection systems to detect attacks. The first approach is Misuse Detection (or Signature Detection) which is the most popular approach that was used in intrusion detection systems (Patil, et al., 2008). The Second approach is Anomaly Detection. It was proposed in 1985 by Dr. Dorothy Denning (Dinning, 1987).

1.2 Problem Statement.

Defense against intrusions, attacks and other types of threats is very important and highly recommended as a security issue, because intrusions and attacks are growing rapidly every day. There are two main approaches for intrusion detection systems. The researcher believes that human immune system is the most powerful defense system. It can be used to design a hybrid model depending on the mechanism and properties of the

human immune system to increase the security of information systems environment. Thus, this study was carried out to answer the following questions:

■ How to keep information systems environment safe against intrusions by combining two type methods of machine learning into one hybrid approach depending on innate and adaptive human immune system?

■ How to detect new attacks depending on immune network clustering and classification methods?

■ How to evaluate the performance of learning in the proposed model?

1.3 Objectives of the Study.

In fact it is impossible to build a complete secured system although when it use cryptographic because the encryption can be broken and passwords can easily be cracked. So, there is a real need to build a system that has the ability to learn by itself. To accomplish the aim of this research, some objectives have been identified, these objectives are:

■ Improve the accuracy of detecting attacks by selecting all features of data sample in the clustering and classification processes.

■ Investigate the ability of building a hybrid intrusion detection model based on innate and adaptive human immune system.

■ Evaluate the learning system, which is the core of the proposed hybrid model, in order to compute the accuracy of the proposed model.

1.4 Significance of the Study.

This study aims to distinguish between self and non/self system behavior depending on human immune system mechanisms and properties. It evaluates the performance of hybridization between the K-Means clustering algorithm and Naive Bayes classifier algorithm as a hybrid learning approach, and to check whether this approach can produce better performance as an intrusion detection model.

1.5 Limitations of the Study.

Similar to many studies there are some challenges that faced the intrusion detection systems. This study faced some limitations, which can be summarized as follows:

■ On-Line Intrusion detection systems suggest a periodic update to the training set and profiles, using a static training data might become outdated and deficient for prediction.

■ Intrusion detection system needs to integrate and interact with other components to form an intrusion and prevention system.

■ A completely secured system stills a myth; because the growth of information systems networks and their infrastructure led to growing in attack methods and rates against information systems environment.

■ The accuracy of classification is not 100%.

1.6 Thesis Organization.

This thesis consists of five chapters. The first chapter presents an introduction to thesis, problem statement, also it gives the objectives of research, discusses the significance and limitation of the study and finally it presents thesis organization. Chapter two reviews intrusion detection approaches, presents an overview of human immune system and gives an overview of attacks and machine learning methods, and finally it discusses the related work. Chapter three outlines research methodology used in this thesis. Also it presents the proposed hybrid model architecture and the software that have been used for evaluation the model. Chapter four describes the dataset used for experiments in this study, experiments environment and procedures and presents the evaluation measures and experimental results, finally a comparison with other studies results is made. Chapter five concludes the research and gives some future directions for future research.

Chapter Two

Intrusion Detection and Human Immune System (Literature Review)

Chapter Two

Intrusion Detection and Human Immune
System (Literature Review)

2.1 Introduction.

This chapter consists of five sections. Section 2.2 discusses the intrusion detection approaches; section 2.3 gives an overview about human immune system; section 2.4 discusses the categories of attacks; section 2.5 discusses machine learning methods and section 2.6 gives an overview of the works related to this thesis.

2.2 Intrusion Detection Approaches.

Allen et al. (2000) defined the intrusion detection system as a component of the information security framework. The main goal of intrusion detection system is to differentiate between the normal and abnormal activities of the system. It aims to automatically scan network activities and detect any attacks. According to (Patil, et al., 2008) there are two main types of IDS approaches: Misuse Detection (or Signature Detection) and anomaly detection.

2.2.1 Misuse Intrusion Detection System Approach.

The main idea for this approach is to detect intrusions by matching the observed behavior or event with a set of attack patterns (Patil, et al., 2008). This approach contains a set of signatures such as failed login, file access and network traffic, a list of unacceptable actions of users or packet contents. A block diagram of a typical misuse detection system is shown in Figure 2.1 below (Sundaram, 1996).
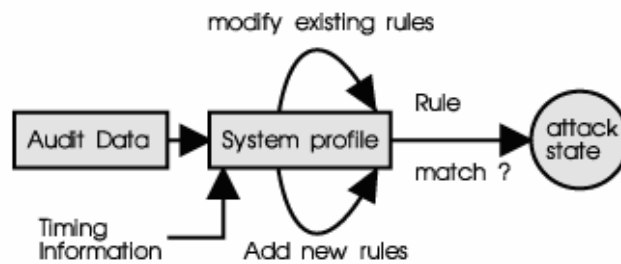
Figure 2.1: Block diagram of a typical misuse

detection system (Sundaram, 1996).

2.2.2 Anomaly Intrusion Detection System Approach.

The main idea for this approach is to detect anomalous activities that deviate from accepted thresholds that were created upon statistical evaluation of a collected data like user login times within an hour, number of login failures within a minute, network traffic and total consumed time by a program. These data are collected from a host or network within a period of time (Dinning, 1987). A block diagram of a typical anomaly detection system is shown in Figure 2.2 below (Sundaram, 1996).
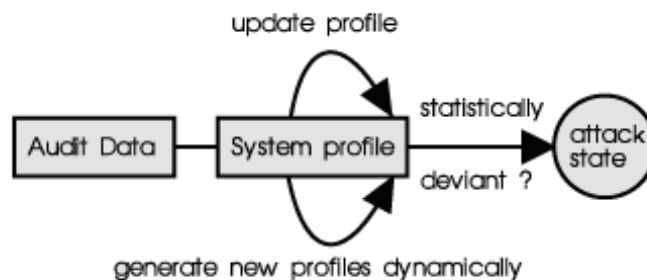
Figure 2.2: Block diagram of a typical anomaly

detection system (Sundaram, 1996).

Carter (2002) discussed some benefits and drawbacks of misuse and anomaly approaches in his published article as shown in Table 2.1.

Table 2.1. Benefits and Drawback of misuse and anomaly IDS approaches (Carter, 2002).

	Misuse Approach	Anomaly Approach
Benefits	- The user can examine the signature database for known intrusive activity.	- Easy detection of insider attacks or account theft, alarm generated when using a stolen account to perform actions that are outside the normal user profile.
	- Start protecting network upon installation.	- Very difficult for attackers to know what activity without triggering an alarm.
	- It is easy to understand, when an alarm triggered.	- It can detects an attack that first time it is used, because the alarm is generated when an activity deviates from normal activity.
Drawbacks	- Maintaining information for Signatures when intrusion encompasses multiple discrete events.	- Intrusion detection system must be trained to create the appropriate user profiles, that means defining normal activity is a challenge it self.
	- Misuse detection system must have signature defined for all possible attacks, this leads for a need for keep updating the signature database.	- The difficulty of associating an alarm with the event that triggered the alarm, means that if the intrusive activity is too close to normal user activity, then there will be no guarantee that alarm will be triggered.
		- Maintenance of the profiles can also become time consuming.

2.3 Human Immune System Overview.

Human immune system (HIS) is a very complex, stunning, tightly created and amazing system that protect body against harmful threats. The main job for HIS is to fight those threats to keep human body healthy by performing two tasks, the first one is detecting threats and second one is eliminating them.

U.S. National Institute of Allergy and Infectious Diseases (2007) described the human immune system as a network of cells, tissues and organs that work together to defend the body against attacks by foreign invaders. These foreign invaders (Pathogens) such as: bacteria, viruses, fungi and parasites, are shown in Figure 2.3.
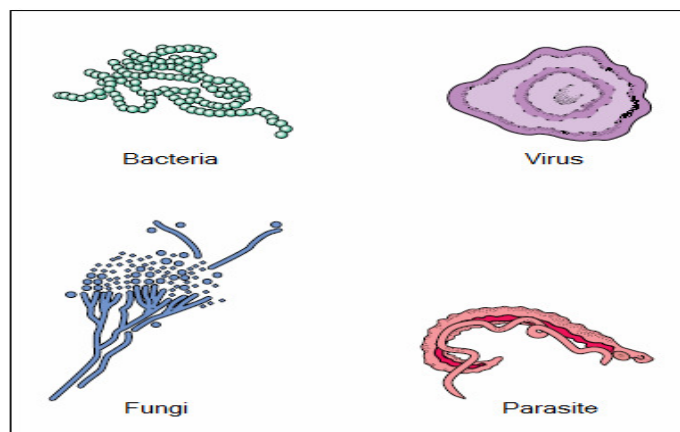
Figure 2.3: Foreign invaders that attacks human body

(U.S. National Institute of Allergy and Infectious Diseases, 2007).

Pathogens can cause irreversible infection and damage to human body if not stopped in time. Since pathogens are continually developing ways to avoid detection, the human immune system is still able to discover these attacks which happen because it can recognize and remember millions of enemy cells.

The human immune system consists of several defense mechanisms, as shown in Figure 2.4. The skin which is the first defense barrier prevents substances that don't have rights to pass through. So it is very effective in preventing many microorganisms from entering the human body. Some of these pathogens can pass through skin if it is compromised by wounds, then they will meet the physiological barrier of defense which is temperature and pH. Human immune system increases the human body temperature and decreases the pH to prevent pathogens from evolving. The farther mechanisms are the Innate Immune System and the Adaptive Immune System, both of them are based on big number of cells that circulating in the blood stream (Goldsby, et al., 2003).

2.3.1 Innate Human Immune System.

It is part of the immune system in which we acquire at birth. It is considered as the third defense mechanism and the first defense line for the already known pathogens but it can not detect new types of pathogens. Innate human immune system consists of roaming scavenger cells Macrophages such as Phagocyte, as shown in Figure 2.4, Natural Killer cells and Neutrophils. Phagocyte cells are able to engulf pathogens. Innate human immune system is based on a chemical response system called Complement System, and the complements also called Antibodies (Piel, 1993).
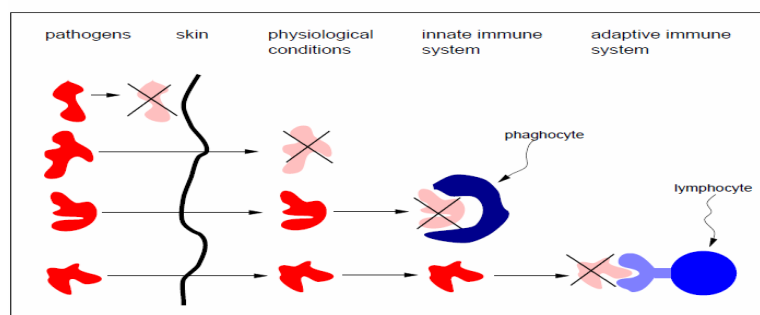
Figure 2.4: Human immune system defense mechanisms (Piel, 1993).

2.3.2 Adaptive Human Immune System.

It consists of lymphocytes, which are certain white blood cells that circulate in the blood stream. These cells have the ability to learn or adapt to new kinds of pathogens and retains them in a memory cells. The primary response for pathogen is a slow process because it can take up to three weeks to clear infection. After clearing the infection, HIS retains that pathogen in memory cells, in order to speed up recognition of them in future, this called secondary response (Piel, 1993).

Paul (1993) claims that Lymphocytes (B-Cells and T-Cells), are able to react for the strange foreign shapes such as virus and bacteria. These are called pathogen that carry antigen over its surface. Lymphocytes have the specific binding areas called receptors which have the complementary shapes which are used to determine antigens which are recognized by binding its marker molecule called epitopes to lymphocyte antibody receptor, as shown in Figure 2.5.
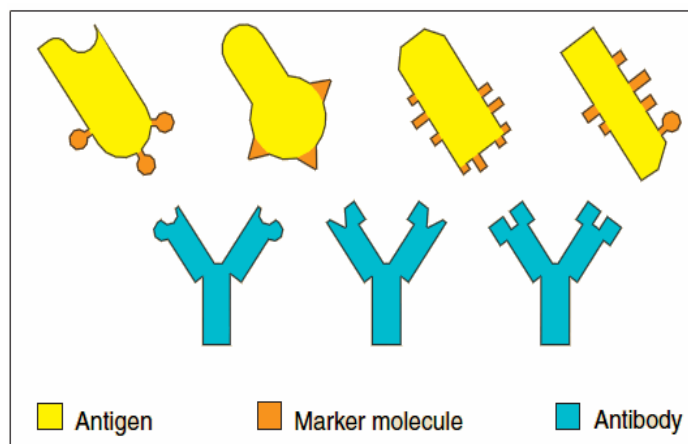
Figure 2.5: Antigen detection mechanism based on complementary shapes (U.S. National Institute of Allergy and Infectious Diseases, 2007).

Tizard (1995) claims that B-cells and T-cells receptors are created by randomly selection of gene segments from gene library (DNA). There is a possibility that these receptors can be bound with self cell epitopes. To prevent this from happening, B-cells and T-cells have to pass the last test called (Negative Selection) before leaving Bone Marrow and Thymus to the blood.

The B-cells and T-cells that fail in the test are killed, while the passed cells are released from bone marrow and thymus to the blood. With or without assistance of T-cells, B-cells become activated when it detects antigens. This activation is followed immediately by (Clonal Selection), as shown in Figure 2.6, to generate memory cells that make detection fast for the same antigens in future.
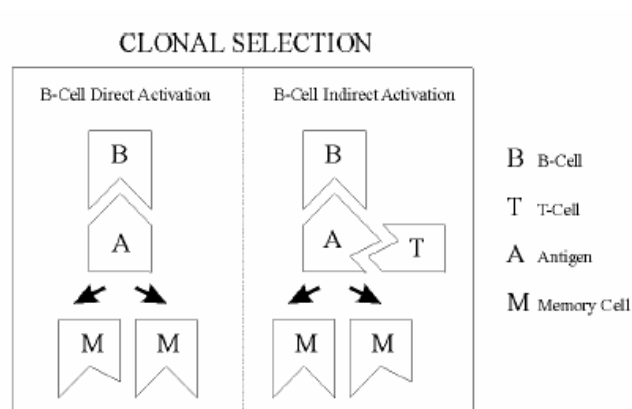
Figure 2.6: Clonal Selection.

2.3.3 B Lymphocytes (B-Cells).

It is created, matured and programmed in the Bone Marrow. Each B-cell produces a specific antibody which is carried over its surface. When B-cell encounters the antigen that matches its antibody receptor, it will still wait to become activated by T-Cell. When

that happens, B-cell multiply rapidly produces more similar cells. Some of these produced cells are transformed into plasma cells that produce antibodies which are similar to the one that is carried over B-cell surface. The rest of plasma cells will be transformed into memory cells that have the ability to become active, if the same antigen tries to attack the body in future (Nobelprize.org, 2011).

2.3.4 T Lymphocytes (T-Cells).

It is created in bone marrow and then migrates to Thymus to get matured. Each T-cell can distinguish between self cells and non-self; where self cells are the human body cells, while non-self cells are pathogens. T-cell has a one specific receptor that is carried over its surface. According to functions, T-cells can be divided into three types: Helper T-cells which have the responsibility to immunology process by activating B-cells. The second type is Cytotoxic T-cells which is responsible for attacking and killing the abnormal self body cells, like cells infected by cancer. The third type is Suppressor T-cells, which is responsible for stopping activation process in other cells after threats become released (Nobelprize.org, 2011).

2.3.5 Human Immune System Theories.

There are several theories that have discussed how HIS can be modeled and how it can recognize the different types of pathogens in order to prevent them from attacking human body.

2.3.5.1 Self/Non-Self Theory.

U.S. National Institute of Allergy and Infectious Diseases (2007) declared that the key to a healthy immune system is its remarkable ability to distinguish between self cells which are the human body's own cells, and non-self cells which are a foreign cells such as pathogens. This theory says that each cell or organism is carrying a marker molecule over its surface. This marker says "self" or "foreign", it will give the human immune system the ability to learn and defend body by detecting the previously unseen patterns.

2.3.5.2 Danger Theory.

This theory was proposed in 1994 by Polly Matzinger. The danger theory suggests that the immune system reacts to threats based on existence of danger signals which means that immune system does not attack foreign threats when it detects them, but it attacks foreign cells when it starts to cause troubles (Matzinger, 1994).

2.3.5.3 Immune Network Theory.

Jerne (1974) suggests that the immune system is not a set of discrete agents (antibodies) that react only when it is triggered by antigen, but it is a regulated network of cells and molecules that can recognize each other.

2.4 Attacks Overview.

MIT (1999) have collected attacks types and categorized them in four categories. These categories are: Denial of Service (DOS), Remote to User (R2L), User to Root (R2U) and Probes, as shown below.

2.4.1 Denial of Service (DOS) Attack.

In this type of attacks, the attacker makes some computer resources or memory resources too full, so it becomes not able to handle request. These attacks are Ping of death, Teardrop, Mailbomb, Smurf, Land, Apache2, SYN Flood, (MIT, 1999).

CNET News (2000) indicates that in the typical connection; the user sends an authentication request to the server which responses and sends authentication approval back to the user. In the next step, the user acknowledges this approval and then is allowed onto the server. While in denial of service, the user sends several authentication requests to the server which have false return addresses; the server can't find the user and become unable to send the authentication approval response and the server becomes waiting for a sometime before it closes the connection. The user (attacker) sends new batch of forged requests to the server and so on.

2.4.2 Remote to Local (R2L) Attack.

In this type of attack, the attacker dos not have an account on a remote machine, so he sends packets over a network to that machine to gain local access as a user of that machine. These attacks are: Xlock, Dictionary, Phf, Guest, Named, Imap and Ftp_write (MIT, 1999).

2.4.3 User to Root (U2R) Attack.

The attacker starts out with access to a normal user account on the system to gain root access to the system. This can be happen by sniffing passwords. These attacks are: Ps, Eject, Xterm, Perl, Loadmodule and Fdformat (MIT, 1999).

2.4.4   Probes Attack.

In this type, the attacker scans a network of computers to gather information or find vulnerabilities using a software program. He will have a map of machines and services that are available on the network, and he can use this information in order to look for exploits such as Nmap, Satan, Mscan, Saint and Ipsweep (MIT, 1999).

2.5   Machine Learning Methods.

The pioneer of machine learning, Tom M. Mitchell (1997) defined the machine learning as a process of training computer algorithm to properly classify future inputs after having trained the algorithm with sample data.

2.5.1   Unsupervised Learning.

It is also called (data clustering) because it is the separation of a set of objects into groups. Each group consists of similar objects that are dissimilar of objects in other groups (Dunham, 2003). K-means is one of the best simplest clustering techniques to partition *(n)* instances into *(k)* clusters in which each instance belongs to the cluster with the nearest mean (MacQueen, 1967). Boundaries between clusters are still linear in the implicit high-dimensional space, and they can become non-linear when projected back to the original space, thus allowing kernel k-means to deal with more complex clusters (Dhillon, et al., 2004).

Clustering and Vector Quantization are concerned with the grouping of unlabeled "feature" vectors into clusters. Usually, it is assumed that the number of clusters is known

in advance, but otherwise no prior information is given about the data, vector quantization is an application for k-means, the centroid index or cluster index is referred to as a "code" and the table mapping codes to centroids and vice versa is often referred as a "code book". The result of k-means is a set of centroids that can be used to quantize vectors. There are two methods of quantization that can be used to reduce a number of feature vectors. The first is (VQ1), it was developed by (Juang, *et al,.* 1982) and based on splitting every cluster into two clusters, while the second method is (VQ2). It was developed by (Lipeika, *et al.,* 1995) and based on splitting a cluster with largest average distortions into two clusters. Both of (VQ1) and (VQ2) aim to find an encoding of vectors that reduces the expected distortion.

2.5.2 Supervised Learning.

It is also called classifier because it aims to build a predictive model (classifier) to classify the incoming patterns. This classifier should be trained with labeled patterns so, it can be able to classify the new unlabeled pattern later. (Zhang, 2004) explains that Naive Bayes is one of the most efficient and effective inductive algorithms for machine learning and data mining. It is surprising because its competitive performance in classification is based on conditional independence assumption which is rarely true in real world applications.

2.6 Related Works.

Greensmith et al. (2004) published a paper that discussed the danger theory and how this concept could be useful in inspiring artificial immune systems, especially in the field of computer security. They discussed the intrusion detection systems and suggested

that improvements to IDS could be done by studying how danger signals can be identified in HIS and how it could be translated for detecting danger within computer systems environment.

Fabricio et al., (2004) presented an intrusion detection framework and a prototype called "ADENOIDS" was presented depending on that framework which takes its architecture from the human immune system to deal with application attacks and extract attack signature for remote buffer overflow attacks.

This prototype brought more features to intrusion detection system, based on human immune system, such as intrusion evidence detection, automated attack signature extraction, intrusion tolerance and system recovery mechanisms.

This framework presented an assumption, which is one of the most important aspects that successful attacks are inevitable and its strongest feature is its ability to deal with such situation, which is the same case with human immune system, Since some diseases antigens (viruses and bacteria) are successful in invading organism and causing a harm to it before human immune system can eliminate them by creating the suitable antibodies to cope with antigens. After that happened, the human immune system learns to cope with this type of antigens and take a repair strategy to recover the damaged parts.

The deep studying of "ADENOIDS" showed that its signature extraction mechanism was only covered buffer overflow attacks, so there is a real need to be extensible to cover other classes of attacks.

Jungwon et al. (2007) collected the algorithms used in intrusion detection systems which developed upon human immune system and discussed these algorithms and the outcome of their implementation. They provided an overview of intrusion detection systems based on human immune system to help researchers to identify suitable research problems in immune-based intrusion detection field.

Also, they summarized six immune features that are desirable in effective intrusion detection systems; these features are multi-layered, distributed, self-organized, lightweight, disposable and diverse. Through their careful examination of previous published researches in immune-based intrusion detection field they concluded that this field of research still has much room to grow and many areas to explore. On the other hand they presented a phylogenetic tree to show the research history in this area. The researcher believes that it is very important to read that research carefully for every immune-based intrusion detection system researcher before start his own research because it gives him a deep understanding of this area of research.

Dal et al. (2008) illustrated a technique that applied artificial immune system and genetic algorithm to develop a secondary immune response in an intrusion detection system depending on memory cells concept to make the system less predictive and enhance the detection process to trap similar anomalies.

Kotov & Vasilyev (2009) presented an intrusion detection system based on artificial immune system for MS Windows operating system. The main idea of that intrusion detection system is to trace the sequence of applications system calls to detect anomaly

changes in the normal system behavior using negative selection algorithm, as shown in Figure 2.7 and Figure 2.8.
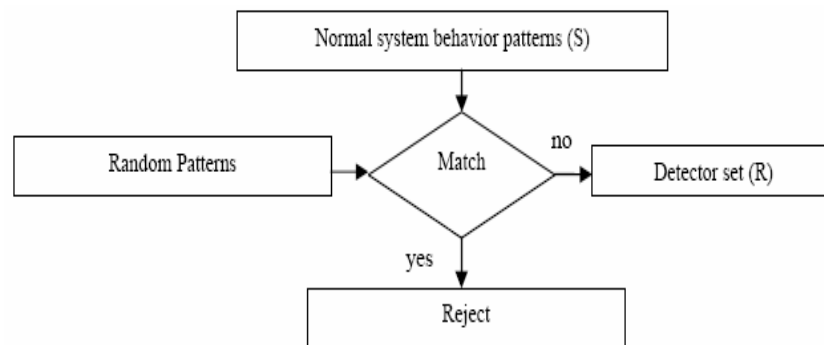
Figure 2.7: Negative selection algorithm (Kotov & Vasilyev, 2009).
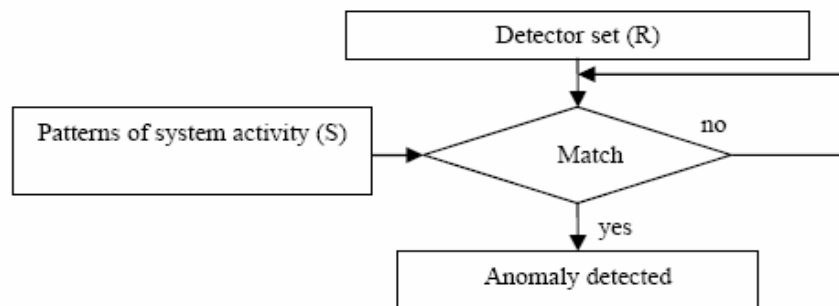
Figure 2.8. Anomaly detection (Kotov & Vasilyev, 2009).

Twycross, Aickelin and Whitbrook (2010) presented (tlr algorithm) to solve anomaly detection problem on an FTP sever, based on artificial immune system. It depended on runtime statistics (such as process memory and file usage) to detect intrusions that use runtime information as well as system call information. This algorithm inspired by the understanding of interaction between T-Cells (TCs) and Dendritic Cells (DCs), which are two classes of immune cells that are found in human immune system.

Interaction idea between two types of immune cells have been abstracted to form (trl algorithm) which is innate and adaptive algorithm that can be used in host-based intrusion detection systems.

Sunjun (2010) proposed an immune-based model for detecting and preventing network intrusions by using a scheme of bacteria for quickly detecting the similar intrusions in neighbor networks. This model consisted of intrusion detection agent that was responsible of monitoring all network packets and a center for treated bacteria which is a repository for prevention techniques.

The intrusion detection agent has memory cells which are patterns for intrusions. These patterns came from treated bacteria center which has the responsibility of sending memory cells (bacteria intrusion patterns) and the suitable prevention technique to neighbor networks as vaccine, and receiving memory cells and prevention technique from the neighbor networks. The model represented self-training approach and self-adapting approach. These approaches are from the characteristics of human immune system.

Mohamed et al. (2010) made integration between artificial immune system and pattern recognition algorithms that are used for intrusion detection. They introduced a model which is inspired by the mechanisms of dendritic cells, B-cells and T-cells.

Chapter Three

The Proposed Hybrid Model Architecture

Chapter Three

The Proposed Hybrid Model Architecture

3.1 Introduction.

The proposed model in this thesis is designed to depend on innate/adaptive human immune system approaches and self/non-self theory of human immune system. This model is divided into two subsystems; the first subsystem is attack response system and the second is learning system, which is the core of this thesis.

3.2 The Proposed Hybrid Model Architecture.

The proposed hybrid intrusion detection model architecture consists of four components, as shown in Figure 3.1. And are followed by a detailed descriptions.
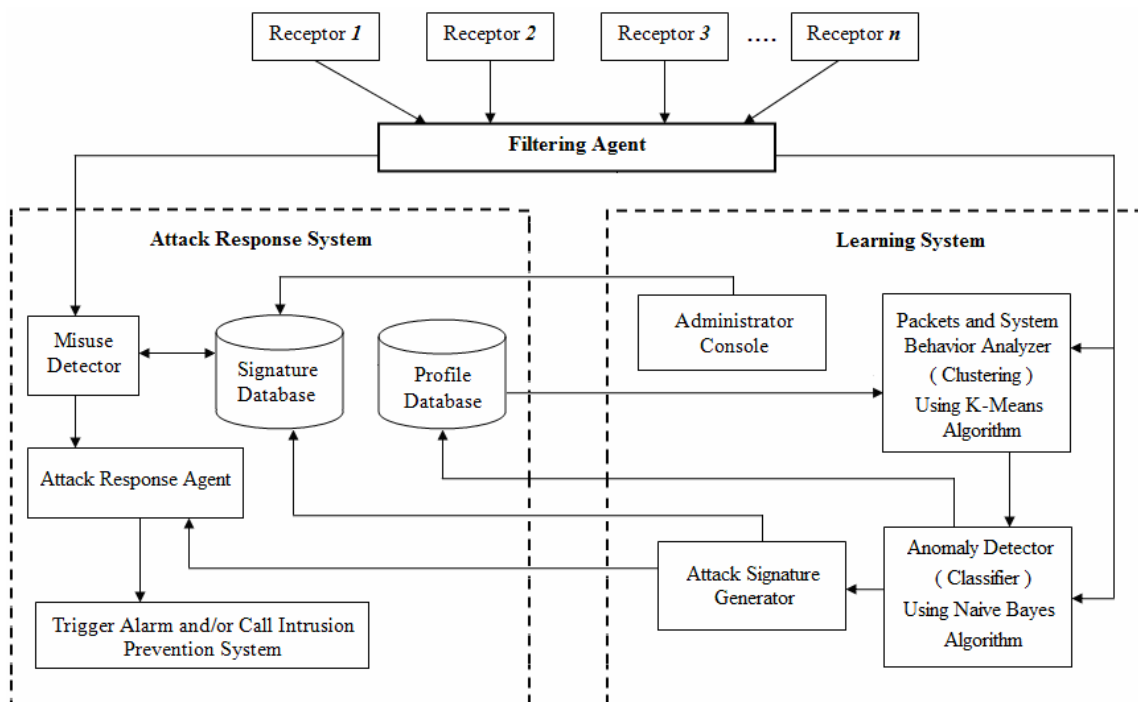
Figure 3.1: The proposed hybrid model based-on human immune system.

3.2.1 Receptors.

Receptors are counted from *1* to *n* receptors; they are located in different nodes (computers) among network. The main job of each receptor is to collect its node system behavior and the transferred packets, then passes these data to the filtering agent.

3.2.2 Filtering Agent.

This agent is responsible for auditing the data received from receptors and to make sure that there is no redundant data. When that happens; it can pass this information to misuse/anomaly detection system, packets and system behavior analyzer (clustering) and to anomaly detector (classifier).

3.2.3 Attack Response System.

Attack response system is responsible for testing the received information and discovering if it contains a misuse patterns. This system consists of five sub components, which are:

3.2.3.1 Signature Database.

It is used for storing the attack signatures (patterns) that can be used later by the misuse detector to match the received information with the stored attack signatures.

3.2.3.2 Profile Database.

It is used for storing profiles that describe the normal system behavior for each node (computer) among the network. This database records are very useful for packets and system behavior analyzer (clustering).

3.2.3.3 Misuse Detector.

This component receives information from filtering system and matching it with saved misuse attack signatures in signature database. When a match found; it activates the attack response agent.

3.2.3.4 Attack Response Agent.

When this component becomes activated, it takes several operations like disabling login, close connection and system files protection, and then it activates the trigger alarm and/or call the intrusion protection system component in order to prevent the attack damages.

3.2.3.5 Trigger Alarm and/or Call Intrusion Prevention System.

This component is activated by attack response agent in order to trigger an alarm for network administrator and/or calling the intrusion prevention system.

3.2.4 Learning System.

It is the core of the hybrid model which is responsible of making the hybrid model adaptive by storing new attack signatures automatically and manually in signature database. Also it collects and analyzes the packets and system behavior information in order to update normal behavior records in profile database, so to make intrusion detection faster in future. This system consists of four components, these are:

3.2.4.1 Packets and System Behavior Analyzer (Clustering).

This component is responsible of clustering the collected packets and system behavior information into clusters depending on K-Means algorithm in order to facilitate the anomaly detector classification job.

K-means is an iterative unsupervised learning clustering algorithm. The number of clusters *(k)* must be specified at the beginning, and then each cluster is associated with an initial centroid (mean) that is chosen randomly from the data points. The next step is assigning each data point to the closest centroid by computing Euclidean Distance between the point and each centroid and choosing the lowest distances, the Euclidean Distance formula, *(d),* is described in equation 3.1.

$$d\ (x\ ,c\ ) = \sqrt{\sum_{k\ =\ 1}^{n} \left(x_{\ k} -\ c_{\ k}\right)^2} \qquad \dotsc\dotsc\dotsc\dotsc\dotsc \quad 3.1$$

Where: x_k is a data point and C_k is the centroid of the cluster.

The next step is recomputing the centroid for each cluster and assigning the new point to the closest centroid depending on the Euclidean Distance between the new point and each centroid. Then by recomputing the centroid for each cluster again, one can note that these iterative steps will make centroids change their locations step by step until no more changes are done, as shown in Figure 3.2. In other words; these iterative steps will continue until centroids don't change (Tan et al., 2005).

```
1: Select K points as the initial centroids.

2: Repeat

3: Form K clusters by assigning all points to the closest centroid.

4: Recompute the centroid of each cluster.

5: Until The centroids don't change.
```

Figure 3.2: K-means clustering algorithm (Tan et al., 2005).

3.2.4.2 Anomaly Detector (Classifier).

Anomaly detector is a classifier that receives information from filtering agent and verifies if it is normal or not. When an abnormal activity is found; this component activates the attack response agent. Otherwise it updates profile data base with the normal activities.

Tan et al. (2005) explain that Naive Bayes is one of the most efficient and effective inductive algorithms for classification. It is based on conditional probabilities which uses Bayes Theorem. Let A denotes the class variable and B denotes the attribute set, if the class variable has a non-deterministic relationship with the attributes. Then A and B can be treated as random variables and capture their relationship probabilistically using $P(A|B)$. This conditional probability is also known as the posterior probability for A; In other words, it finds the probability of an event occurring given the probability of another event that already occurred. Bayes Theorem formula can be stated as described in equation 3.2.

$$\text{Posterior} = \frac{\text{Likelihood X Prior}}{\text{Evidence}} \quad , \quad P(A|B) = \frac{P(B|A)\ P(A)}{P(B)} \quad \cdots\cdots\ 3.2$$

The classification stage in this proposed model depends on Naive Bayes classifier algorithm because it is simple to implement and easy to train, (Mitchell, 1997) explains that Naive Bayes works as follows:

- Let **D** a data set (a set of Tuples), each Tuple **X** is an *n*-attribute or *n*-dimensional vector $X = \{ X_1, X_2, \ldots, X_n \}$.

- Let there are **k** classes: $C_1,\ C_2\ldots C_k$.

- Naive Bayes predicts X belongs to class C_i if and only if :

    $$P(C_i \mid X) > P(C_j \mid X) \qquad \text{For } 1 <= j <= m, j \neq i$$

- Maximum Posteriori Hypothesis

    $$P(C_i \mid X) = \frac{P(X \mid C_i)\ P(C_i)}{P(X)} \qquad i = 1, 2, 3, \ldots, n \ \cdots\cdots\ 3.3$$

- With many attributes, it is computationally expensive to evaluate $P(X \mid C_i)$. Naive assumption of "class conditional independence"

    $$P(X \mid C_i) = \prod_{k=1}^{n} P(x_k \mid C_i) \qquad \ldots\ldots\ldots\ldots\ldots\ 3.4$$

3.2.4.3 Signature Generator.

This component is responsible for receiving the new detected anomaly attack information from the anomaly detector (classifier) and generates a signature for this attack. Then it stores new signature in signature database automatically, in order to keep signature database up-to-date. This will facilitate detection of such an attack in future.

3.2.4.4 Administrator Console.

It is a manual that allows network administrator to feed the signature database with new attack signatures in order to keep signature database up-to-date.

3.3 Training and Testing the Proposed Model.

The hybrid intrusion detection model has been trained and tested by using randomly instances sample that was extracted from the 10% of a dataset called KDD'99 as shown in chapter four of this thesis. This dataset contains 22 attacks that are divided into four attack classes: DoS, R2L, U2R and Probe in addition to 92278 instances which represent the Normal connections class, as shown below in Table 3.1, Also it contains network connections with 41 features per connection, as shown in Table 3.2, Table 3.3 and Table 3.4, which forms the KDD'99 intrusion detection benchmark in the International Knowledge Discovery and Data Mining Tools Competition. (KDD, 1999).

Table 3.1. KDD'99 connection classes (KDD, 1999).

Class	Number of connections	Instance Percentage
Normal	97278	19.6911 %
Dos	391458	79.2391 %
R2L	1126	0.2279 %
U2R	52	0.0105 %
Probe	4107	0.8313 %
Total	**494021**	**100 %**

There are total 41 attributes in KDD'99 dataset for each network connection that has either discrete or continuous values and is divided into three groups. The first group of attributes is the basic features of network connection, as shown in Table 3.2; the second group of attributes in KDD99 is composed of the content features of network connections, as shown in Table 3.3 and the third group is composed of the statistical features that are computed either by a time window or a window of certain kind of connections, as shown in Table 3.4.

Stolfo et al. (2000) defined higher-level features that helped in distinguishing normal connections from attacks. There are several categories of derived features. For example, the "same host" features examined only the connections in the past 2 s that have the same destination host as the current connection, and calculated statistics related to service, protocol behavior, etc. These features are either continuous or discrete. For example "duration" is a continuous feature for the KDD'99 database.

Table 3.2. Basic features of individual TCP connections (KDD, 1999).

feature name	*Description*	*type*
duration	length (number of seconds) of the connection	continuous
protocol_type	type of the protocol, e.g. tcp, udp, etc.	discrete
service	network service on the destination, e.g., http, telnet, etc.	discrete
src_bytes	number of data bytes from source to destination	continuous
dst_bytes	number of data bytes from destination to source	continuous
flag	normal or error status of the connection	discrete
Land	1 if connection is from/to the same host/port; 0 otherwise	discrete
wrong_fragment	number of ``wrong'' fragments	continuous
urgent	number of urgent packets	continuous

Table 3.3. Content feature within a connection suggested by domain knowledge

(KDD, 1999).

feature name	*Description*	*type*
hot	number of ``hot'' indicators	continuous
Num_failed_logins	number of failed login attempts	continuous
logged_in	1 if successfully logged in; 0 otherwise	discrete
Num_compromised	number of ``compromised'' conditions	continuous
Root_shell	1 if root shell is obtained; 0 otherwise	discrete
su_attempted	1 if ``su root'' command attempted; 0 otherwise	discrete
Num_root	number of ``root'' accesses	continuous
Num_file_creations	number of file creation operations	continuous
Num_shells	number of shell prompts	continuous
Num_access_files	number of operations on access control files	continuous
Num_outbound_cmds	number of outbound commands in an ftp session	continuous
is_hot_login	1 if the login belongs to the ``hot'' list; 0 otherwise	discrete
is_guest_login	1 if the login is a ``guest''login; 0 otherwise	discrete

Table 3.4. Traffic features computed using a two second time window (KDD, 1999).

feature name	*Description*	*type*
count	number of connections to the same host as the current connection in the past two seconds	continuous
	Note: The following features refer to these same-host connections.	
serror_rate	% of connections that have ``SYN'' errors	continuous
rerror_rate	% of connections that have ``REJ'' errors	continuous
same_srv_rate	% of connections to the same service	continuous
diff_srv_rate	% of connections to different services	continuous
srv_count	number of connections to the same service as the current connection in the past two seconds	continuous
	Note: The following features refer to these same-service connections.	
srv_serror_rate	% of connections that have ``SYN'' errors	continuous
srv_rerror_rate	% of connections that have ``REJ'' errors	continuous
srv_diff_host_rate	% of connections to different hosts	continuous

3.4 Model Evaluation using Weka Software.

Witten & Frank, (2000) explained that "Weka" stands for the Waikato Environment for Knowledge Analysis; it is computer software that developed at University of Wikato in New Zealand as a collection of machine learning algorithms for solving real-world data mining problems. Weka is written in Java and runs on almost any platform; these algorithms can apply directly to a dataset or called from any Java code.

Weka contains tools for data pre-processing, clustering, classification, association rules and visualization. It can be used to apply a learning method to a dataset and analyze its output to extract information about this dataset.

Chapter Four

Evaluation and Experimental Results

Chapter Four

Evaluation and Experimental Results

4.1 Introduction.

This chapter deals with the experiments, its settings and results. The experiment was conducted on a dataset that was previously used in other related works. The dataset is presented in details in section 4.2. The experiment dataset samples and the experimental procedures are presented in Section 4.3. The performance measures that usually are used for evaluation the quality in this domain are described in Section 4.4. The experimental results obtained are presented in Section 4.5 and finally a comparison with other studies results is made in section 4.6.

4.2 KDD'99 Dataset.

It is a dataset that contains 22 attacks and divided into four attack classes in additional to 92278 instances which represent the normal connections class, as shown in Figure 4.1. It contains network connections with 41 features per connection, which formed the KDD'99 intrusion detection benchmark in the International Knowledge Discovery and Data Mining Tools Competition. (KDD, 1999).
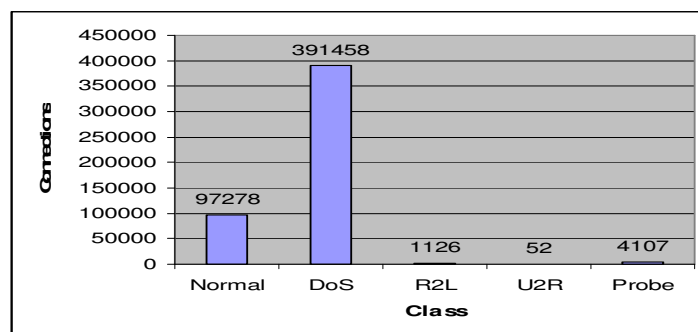

Figure 4.1: Distribution of classes in the KDD'99 Dataset.

Each class contains many different types of attacks that have different inside the class, the following Figures 4.2, 4.3, 4.4 and 4.5 show the different attacks and their distribution for each class.
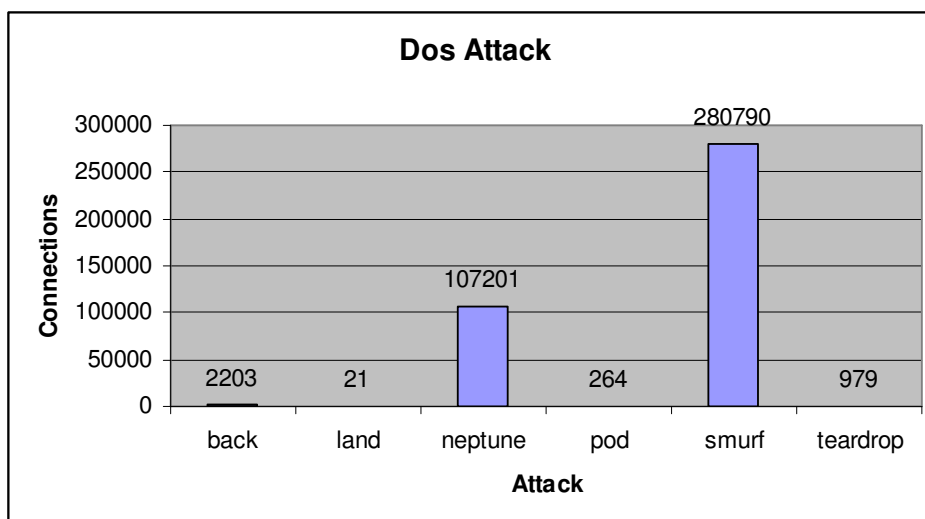
Figure 4.2: Distribution of attacks in DoS Class.
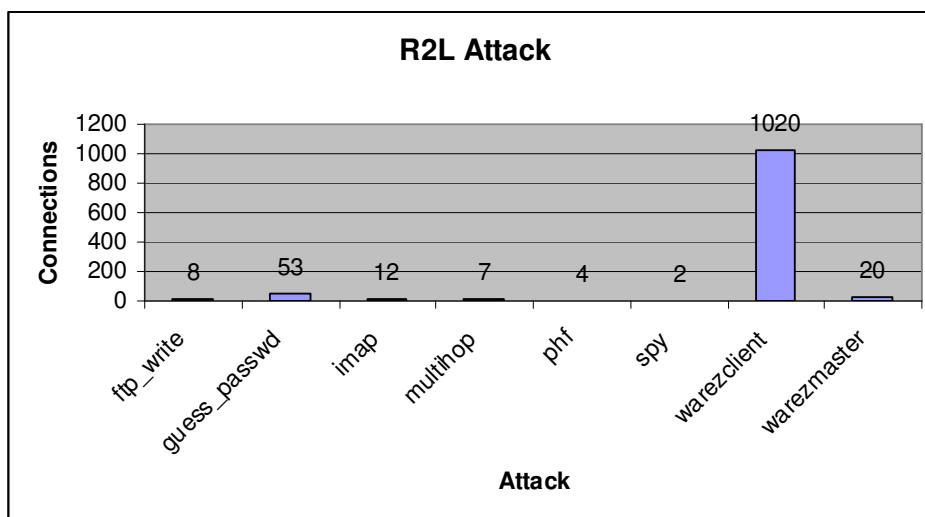
Figure 4.3: Distribution of attacks in R2L Class.
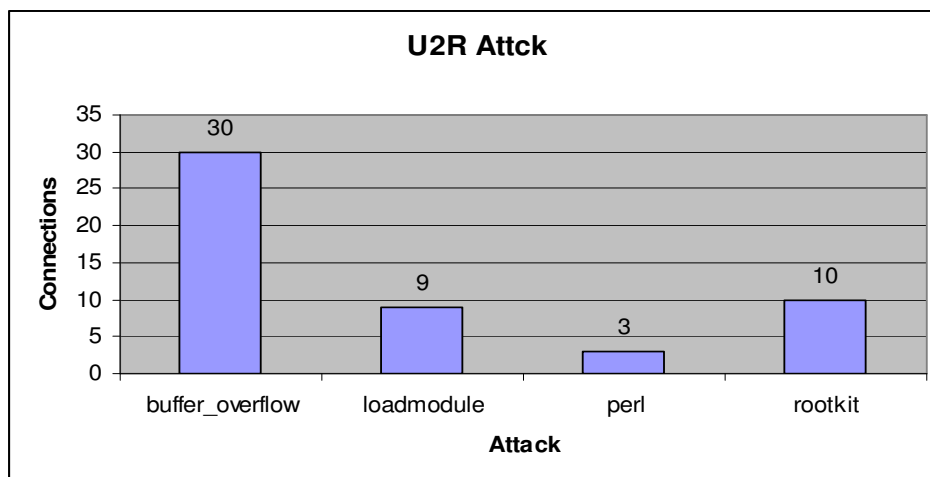
Figure 4.4: Distribution of attacks in U2R Class.
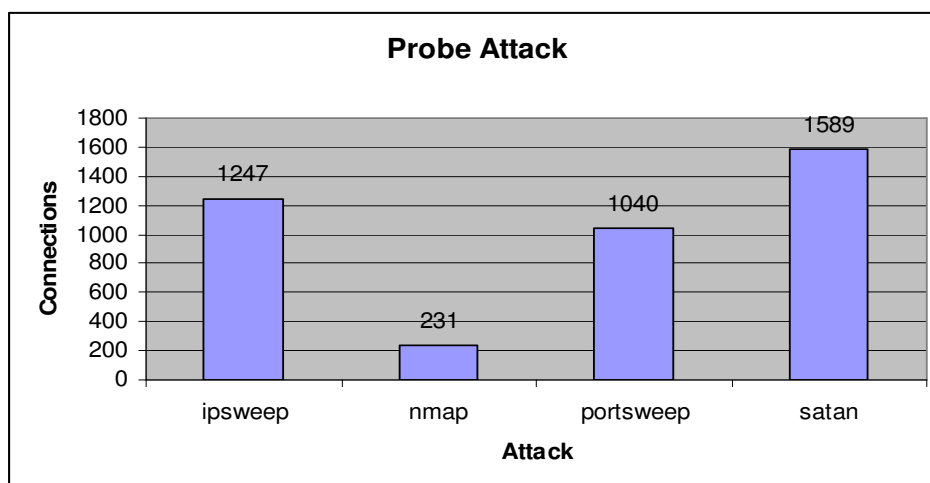
Figure 4.5: Distribution of attacks in Probe Class.

4.3 Experiments Environment and Procedures.

The data sample has been chosen from the 10% percent of the KDD'99 data set. The data sample contains 10000 instances which are divided into 5 classes, as shown in Table 4.1. These instances were picked up randomly from the 10% percent of the KDD'99 dataset.

Table 4.1. Data sample classes.

| Class	Number of instances	Number of attacks within class	Attacks within class
Normal	1967	0	
DoS	7780	6	back, land, neptune, pod, smurf, teardrop
R2L	111	8	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
U2R	52	4	buffer_overflow, loadmoudle, perl, rootkit
Probe	90	4	Ipsweep, nmap, portsweep, satan
Total	**10000**	**22**	

In addition to Normal class, the data sample contains four attack classes; the first attack class in the data sample is denial of service (DoS) which contains 7780 instances distributed over 6 attack types according to the 10% percent of the KDD'99 dataset. As shown in Table 4.2.

Table 4.2. Distribution of attacks in DoS class.

Class	Attack	Instances
DoS	back	40
	land	10
	neptune	2100
	pod	10
	smurf	5600
	teardrop	20
Total		**7780**

The second attack class in the data sample is User to Root (U2R), it contains 52 instances distributed over 4 attack types according to the 10% percent of the KDD'99 dataset. As shown in Table 4.3.

Table 4.3. Distribution of attacks in U2R class.

Class	Attack	Instances
U2R	buffer_overflow	30
	loadmoudle	9
	perl	3
	rootkit	10
Total		**52**

The third attack class in the data sample is Remote to Local (R2L). It contains 111 instances distributed over 8 attack types according to the 10% percent of the KDD'99 dataset. As shown in Table 4.4.

Table 4.4. Distribution of attacks in R2L class.

Class	Attack	Instances
R2L	ftp_write	8
	guess_passwd	40
	imap	10
	multihop	7
	phf	4
	spy	2
	warezclient	20
	warezmaster	20
Total		**111**

The forth attack class in the data sample is Probes Attack (Probe). It contains 90 instances distributed over 4 attack types according to the 10% percent of the KDD'99 dataset. As shown in Table 4.5.

Table 4.5. Distribution of attacks in Probe class.

Class	Attack	Instances
Probe	ipsweep	30
	nmap	10
	portsweep	20
	satan	30
Total		**90**

After data sample preparation, "Weka" software was used to split data sample into two parts. The first part is training set which represents 30% of the data sample; it contains 3000 instances of the data sample. The second part is testing set which represents 70% of the data sample; it contains 7000 instances of the data sample.

The next step is applying K-means algorithm with parameters "*numClusters = 5*" and "distancefunction = EuclideanDistance", on the training set using "Weka", As shown in Figure 4.6 to divide the data sample into five classes; these five classes are: Normal, DoS, R2L, R2U and Probe.
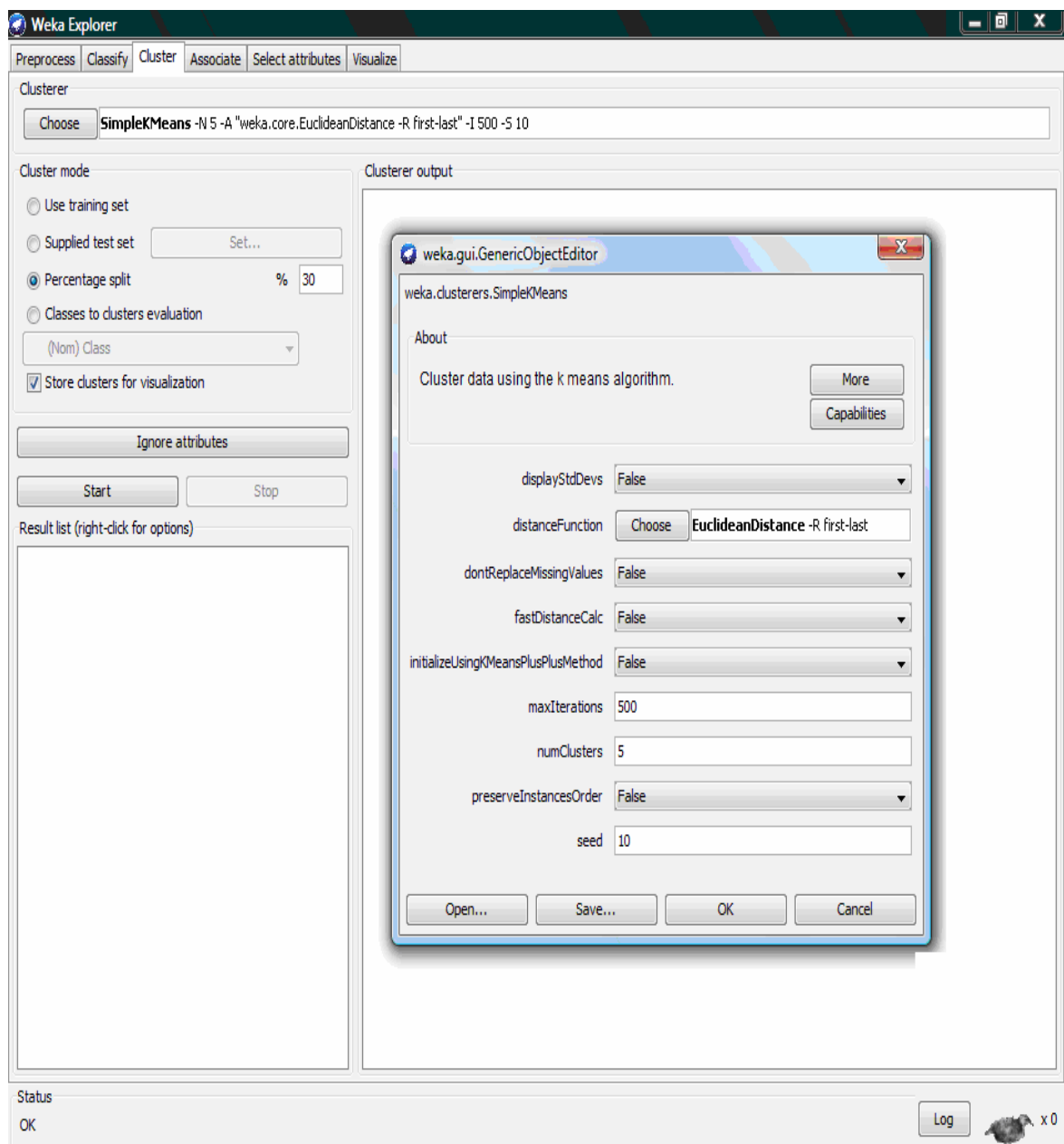
Figure 4.6: K-means clustering algorithm parameters.

Next phase is applying Naive Bayes classifier algorithm on the testing set using "Weka", as shown in Figure 4.7, where 7000 instances are used as a testing set. The correctly classified instances were 6852 which represent 97.8857% of the total testing set. On the other hand there were 148 incorrectly classifies instances that represent 2.1143% of the total number instances in the testing set.

Figure 4.7: Naive Bayes Classifier algorithm result.

4.4 Performance Evaluation Measures.

 This section aims to explain the measures that were used to measure the accuracy of the proposed hybrid intrusion detection model. These measures were calculated in order to get the efficiency of classification for each class of the testing data sample. Some of those measures will used later to make a comparison with other studies results.

Tan et al. (2005) explained that a confusion matrix can be used to summarize the number of instances predicted correctly or incorrectly by a classification model, the relation between actual class and predicated class shown in Table 4.6.

Table 4.6. The Relation between actual class and predicated class.

		Predicated Class	
		+	-
Actual Class	+	TP	FN
	-	FP	TN

Also, he explained that the count in confusion matrix can also be expressed in term of percentages as follows:

■	**TP Rate (TPR).**

It is the true positive rate and also called "Sensitivity". It is defined as a fraction of positive examples predicted correctly by the model. TP Rate formula can be stated as described in equation 4.1.

$$TPR = TP / (TP + FN)$$ …. 4.1

- **TN Rate (TNR).**

  It is the true negative rate and also called "Specificity". It is defined as a fraction of negative examples predicted correctly by the model. TN Rate formula can be stated as described in equation 4.2.

  TNR = TN / (TN + FP)   …. 4.2

- **FP Rate (FPR).**

  It is the false positive rate, which is defined as a fraction of negative examples predicted as positive class by the model. FP Rate formula can be stated as described in equation 4.3.

  FPR = FP / (TN + FP)   …. 4.3

- **FN Rate (FNR).**

  It is the false negative rate, which is defined as a fraction of positive examples predicted as negative class by the model. FN Rate formula can be stated as described in equation 4.4.

  FNR = FN / (TP + FN)   …. 4.4

- **Precision (P).**

  It determines the fraction of records that actually turns out to be positive in the group that the classifier has declared as a positive class, Precision formula can be stated as described in equation 4.5.

  Precision = TP / (TP + FP)   …. 4.5

■   **Recall (R).**

It measures the fraction of positive examples correctly predicated by the classifier. Precision formula can be stated as described in equation 4.6.

  Recall = TP / (TP + FN).   …. 4.6

■   **F-Measure.**

It is the harmonic mean of Precision and Recall. F-Measure formula can be stated as described in equation 4.7.

  F-Measure = 2 * ((Precision * Recall) / (Precision + Recall))   …. 4.7

4.5 Experimental Results.

This section aims to detect the performance of the proposed hybrid intrusion detection model by analyzing and discuss the experiment and represent the results visually by using tables and charts. The steps for the experiment started as follows:

The experiment started by preparing the 10000 instances data sample. Those instances were picked up randomly from the 10% percent of the KDD'99 dataset and were divided into five classes (Normal, DoS, R2L, U2R and Probe) as shown in details in Table 4.2, Table 4.3, Table 4.4 and Table 4.5. Then the 10000 instances data sample was divided into two parts; the first part is training set which represents 30% of the data sample, it contains 3000 instances of the data sample. The second part is testing set which represents 70% of the data sample; it contains 7000 instances of the data sample.

The next step is using "Weka" to apply K-means algorithm with parameters "*numClusters = 5*" and "distancefunction = EuclideanDistance", on the training set as shown in Figure 4.6. Then Naive Bayes classifier algorithm was applied on the testing set as shown in Figure 4.7.

Depending on the experimental results it is found that the correctly classified instances were 6852 which represent 97.8857% of the total testing set, on the other hand there were 148 incorrectly classifies instances that represent 2.1143% of the total number instances in the testing set. Table 4.7 illustrates the performance of the hybridization between K-means clustering algorithm and Naive Bayes classifier. It can be noticed that TP Rate of the hybrid intrusion detection model is 96.7% as a detection rate for normal instances, the detection rate for DoS attacks is 98.8% and 96.8% for Probe attacks, while the detection rate for R2L attacks is 61.3% on other hand the detection rate for U2R attacks is 82.1%. The total weighted average for TP Rate is 97.9% for the proposed model.

Table 4.7. Experimental results for the classification of data sample.

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
Normal	0.967	0.004	0.985	0.967	0.976
DoS	0.988	0.003	0.999	0.988	0.994
R2L	0.613	0.006	0.517	0.613	0.561
U2R	0.821	0.01	0.308	0.821	0.448
Probe	0.968	0.001	0.882	0.968	0.923
Weighted Average	**0.979**	**0.003**	**0.986**	**0.979**	**0.982**

Figure 4.8 shows the true positive rates (detection rate) for the proposed hybrid intrusion detection model, while Figure 4.9 shows the false positive rates (FP Rate) for the same model.

True Positive Rate (Detection Rate)

	Normal	DoS	R2L	U2R	Probe
TP Rate	0.967	0.988	0.613	0.821	0.968

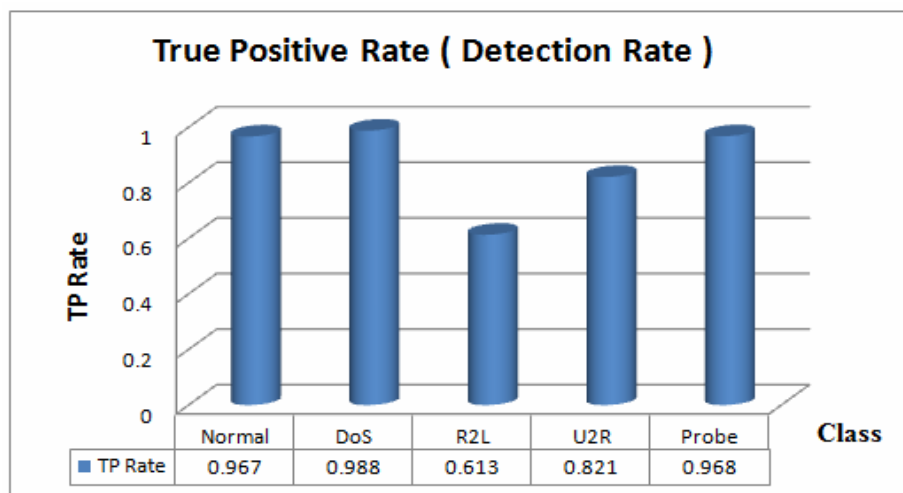**Figure 4.8: True Positive Rates (Detection Rate) for the proposed hybrid intrusion detection model.**

False Positive Rate

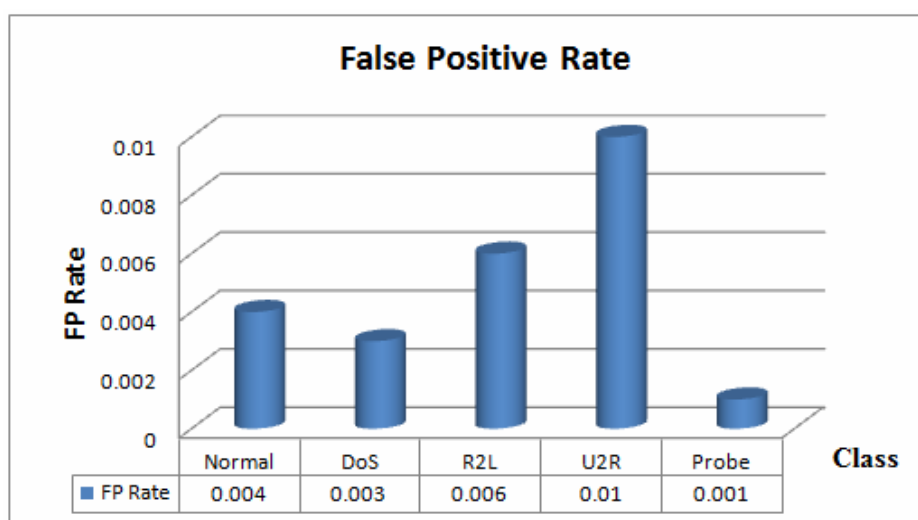|	Normal	DoS	R2L	U2R	Probe
FP Rate	0.004	0.003	0.006	0.01	0.001

Figure 4.9: False Positive Rates for the proposed hybrid intrusion detection model.

4.6 Comparison with Other Studies Results.

This section shows a comparison between the researcher's model results and other studies results depending on TP Rate which reflects the detection rate. The researcher explores results of other researches to define the rank of the proposed method. Table 4.8 Illustrates the True Positive Rate (TP Rate) for each class in the mentioned researches.

Table 4.8. Comparison between the proposed hybrid model experimental

results and other researches depending on TP Rate.

Model	Method	Normal	DoS	R2L	U2R	Probe
The Proposed Hybrid Model	K-Means + Naive Bayes	0.96700	**0.98800**	0.61300	**0.82100**	0.96800
Faraoun & Boukelif (2006)	Multi-Classifier NNet	0.96640	0.95000	0.09850	0.08520	0.87650
Al-Sharafat & Naoum (2009)	Genetic-based Machine Learning (GBML-NID)	0.96320	0.97600	0.83930	0.28850	0.37720
Yongzhong & Wang (2009)	Immune Agent based on Dynamic Clonal Selection	0.98127	0.97565	0.03710	0.71491	0.90494
Al-Rashdan (2011)	Hopfield + Kohonen SOM + Conscience Function (HNNMLM-IDS)	1.00000	0.84600	0.84300	0.73300	0.98000
Farzan et al. (2011)	K-Means + Bayes Rule + SVM	0.86400	0.96500	0.93500	0.75000	0.92200

From the previous Table 4.8, one can see that the method gains better results than others in detecting DoS and U2R attacks. By exploring TP Rates for each class, one can define the rank of the model with amongst other models, as shown bellow in Figure 4.10 and Figure 4.11. It is very clear that the hybrid model achieved the first rank in detecting DoS attack with 98.8% detection rate and 82.1% as a detection rate for U2R attack.
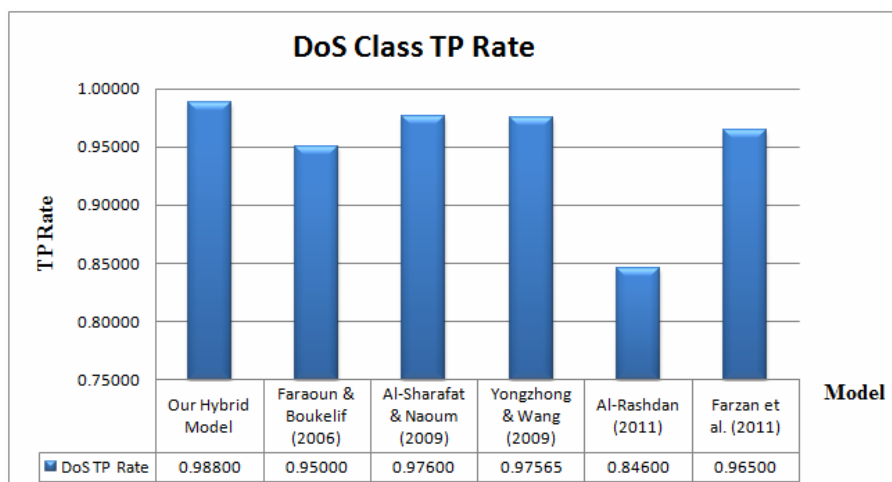
Figure 4.10: Comparison between models depending on

DoS Attack TP Rate.

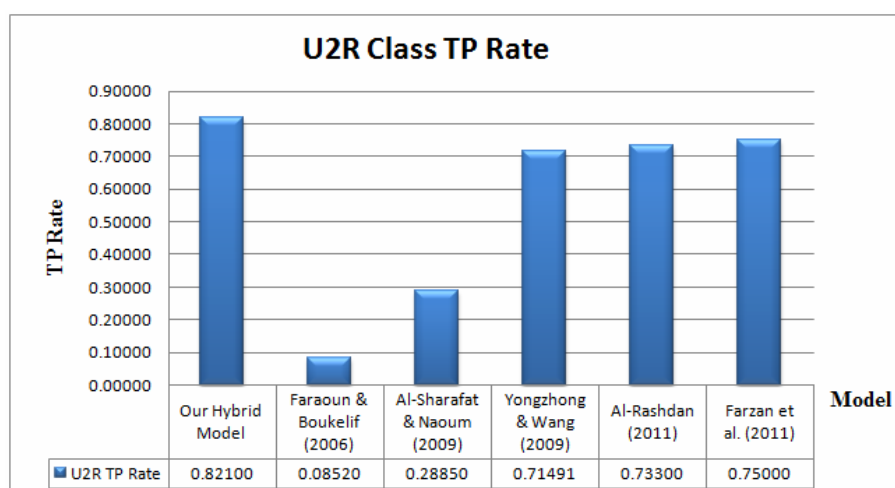**Figure 4.11: Comparison between models depending on**

U2R Attack TP Rate.

Also, from the same Table 4.8, one can see that the method achieved the second rank in detecting Probe attack, as shown in Figure 4.12, with 96.8% as a detection rate while the first rank in the same class was achieved by (Al-Rashdan, 2011) HNNMLM Model that based on combining Hopfield, Kohonen SOM and Conscience Function.

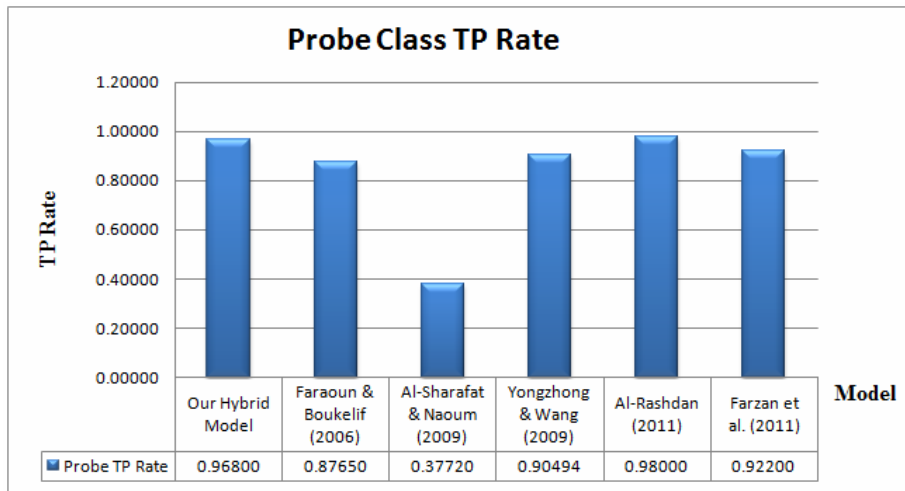**Figure 4.12: Comparison between models depending on**

Probe Attack TP Rate.

As shown in Figure 4.13, one can see that the proposed method achieved the third

rank in detecting Normal class with 96.7% and achieved the fourth rank in detecting R2L

attack with a detection rate of 61.3%, as shown in Figure 4.14.
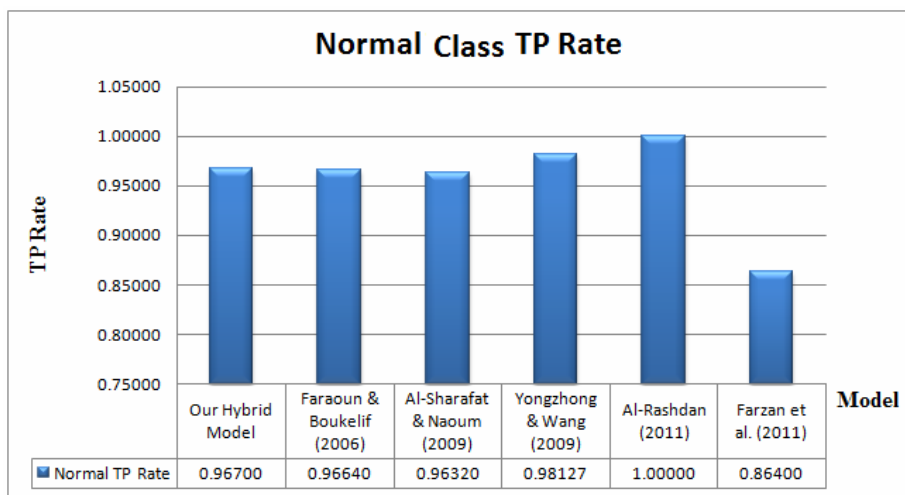
Figure 4.13: Comparison between models depending on

Normal Class TP Rate.
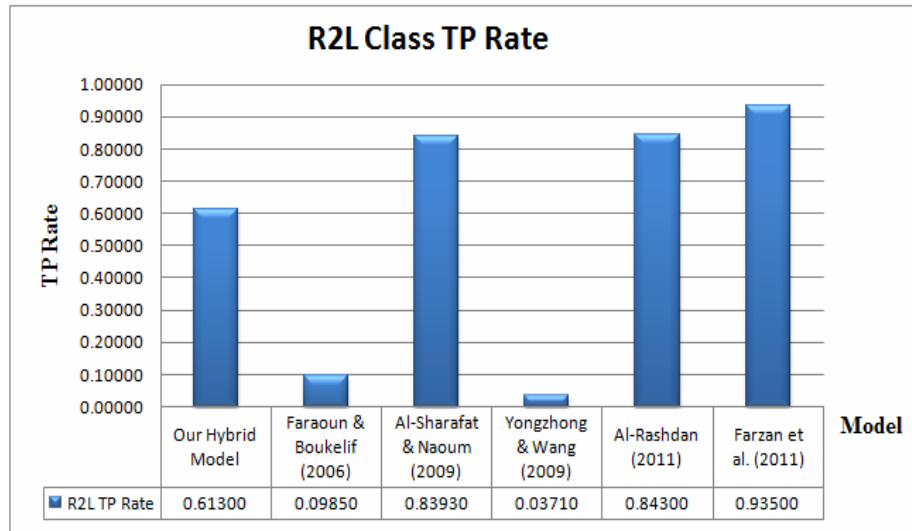
Figure 4.14: Comparison between models depending on

R2L Attack TP Rats.

According to experimental results; the proposed intrusion detection model achieved the first rank in detection both of DoS and U2R attacks comparing with other five studies that mentioned above in Table 4.8. The hybrid model rank can be summarized as seen bellow in Table 4.9.

Table 4.9. The Proposed Hybrid Model Rank.

Class	TP Rate	Rank
DoS	0.988	1
U2R	0.821	1
Probe	0.968	2
Normal	0.967	3
R2L	0.613	4

Chapter Five

Conclusion and Recommendations

Chapter Five

Conclusion and Recommendations

5.1 Introduction.

This chapter concludes thesis and gives some future directions for future work in order to make more improvement in intrusion detection domain.

5.2 Conclusion.

According to the experimental results; the hybridization between K-Means clustering algorithm and Naive Bayes Classifier presented a better TP Rate results for detection. It Achieved first rank in detecting both of DoS and U2R attacks, the detection rate for DoS attacks was 98.8%, the model detect U2R class with 82.1% as a detection rate. Also it detected Normal class with 96.7%, R2L class with 61.3% and it showed 96.8% detection rate for Probe class. One can note that the proposed model has the highest detection rate for DoS attack with 98.8% as a detection rate and U2R class with 82.1% depending on the comparison as shown in chapter four.

5.3 Recommendations for Future Research.

In order to improve the performance of intrusion detection systems, the researcher recommends the follows:

■   More investigations are needed in order to find the optimal way to determine the number of classes and the data sample for each class.

■   Other methods like K-Medoid can be used with supervised and unsupervised learning in intrusion detection field.

■   As mentioned before, KDD'99 dataset consists of instances with 41 features for each instance. More in depth studies are needed in order to reduce the number of these features.

■   Combining reinforcement learning and human immune system mechanisms and properties could produce higher TP Rate in the filed of intrusion detection.

References.

Allen, J., Christie, A., Fithen, W., McHugh, J., Pickel, J., & Stoner, E. (2000). State of the practice of intrusion detection technologies. Technical Report.: Carnegie Mellon University.

Al-Rashdan, W. K. (2011). *A hybrid artificial neural network model (Hopfield-SOM with Conscience) for effective network intrusion detection system,* (Unpublished doctorate dissertation), University of Banking and Financial Sciences, Amman, Jordan.

Al-Sharafat, W. S. & Naoum, R. S. (2009). Development of genetic-based machine learning for network intrusion detection (GBML-NID). *International Journal of Computational Intelligence,* 3(2), 161-168.

Anderson, J. P. (1980). Computer security threat monitoring and surveillance. Fort Washington, PA : James P. Anderson Co.

Carnegie Mellon University. (2011). **Cybersecurity watch survey**, USA : Software Engineering Institute.

Carter, E. (2002). *Intrusion Detection System*. *CISCO PRESS*, (On-Line), Available: http://www.ciscopress.com/articles/article.asp?p=25334

CNET news. (2000). *How a denial of service attack works*. (On-Line), Available: http://news.cnet.com/2100-1017-236728.html

Computer Emergency Response Team. (2009). *CERT statistics (historical).* (On-Lin), Available: http://www.cert.org/stats/

Computer Security Institute. (2009). *14th Annual SCI computer crime and security survey.* (On-Line), Available: http://www.pathmaker-group.com/whitepapers/CSISurvey2009.pdf

Dal, D., Abraham, S., Abraham, A., Sanyal, S., & Sanglikar, M. (2008). Evolution induced secondary immunity: An artificial immune system based intrusion detection system. *7th Computer Information Systems and Industrial Management Applications*, 65-70.

Denning, D. (1987). An intrusion-detection model. *Software Engineering IEEE Transaction,* SE-13(2), 222-232.

Dhillon, I., Guan, Y. & Kulis, B. (2004). Kernel k-means: Spectral clustering and normalized cuts. *Proceedings of the 10th ACM KDD Conference,* 551-556.

Dunham, M. (2003). *Data mining: Introductory and advanced topics.* New Jersey : Prentice Hall.

Endorf, C., Schultz, E. & Mellander, J. (2004). *Intrusion detection & prevention.* New York : McGraw-Hill.

Fabricio, P., Leandro, C. & Paulo, G. (2004). An intrusion detection system using ideas from the immune system. *Proceedings of the Congress on Evolutionary Computation (CEC),* Portland, Oregon, USA.

Farzan, A., Razavi, N., Balafar, M. & Arvin, F. (2011). Intrusion patterns recognition in computer network. *Proceedings of the World Congress on Engineering and Computer Science,* 1, 433-436.

Faraoun, K. & Boukelif, A. (2006). Neural networks learning improvement using the M-Means clustering algorithm to detect network intrusions. *International Journal of Computational Intelligence,* 3(2), 161-168.

Goldsby, R. A., Kindt, T. J. & Osborne, B. A. (2003). *Immunology.* (5th Ed,). New York : W. H. Freeman and Company.

Greensmith, J., Aickelin, U. & Twycross, J. (2004). Detecting danger: Applying a novel immunological concept to intrusion detection systems. *Proceeding of the 6th nternational Conference in Adaptive Computing in Design and Manufacture (ACDM 2004).*

Heberlein, L. T., Gihan, V. D., Karl, N. L., Biswanath, M., Jeff, W. & David W. (1990). A network security monitor. **IEEE Computer Society Symposium,** 296-304.

Innella, P. (2001). *The evolution of intrusion detection systems,* Tetrad Digital Integrity, LLC., (On-Line), Available: http://www.symantec.com/connect/articles/evolution-intrusion-detection-systems

Jerne, N. K. (1974). Towards a network theory of the immune system. *Annales d'immunologie*, 125C(1-2), 373-389.

Juang, B., Wong, D. & Gray, A. (1982). Distortion performance of vector quantization for LPC voice coding. ***IEEE Transaction on Acoustic Speech and Signal Processing,*** 30(2), 294-304.

Jungwon, K., Bentley, P., Aickelin, U., Greensmith, J., Tedesco, G. & Twycross, J. (2007). Immune system approaches to intrusion detection – a review. ***Natural Computing Springer.*** 6(4), 413-466.

KDD, (1999).***The International Knowledge Discovery and Data Mining Tools Competition,*** (On-Line), Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

Kotov, V. D. & Vasilyev, V. I. (2009). Artificial immune systems based intrusion detection system. ***Proceedings of the 2nd International Conference on Security of Information and Networks.*** 207-212.

Lipeika, A. & Lipeikien˙e, J. (1995). Speaker identification using vector quantization. ***Informatica,*** 6(2), 167–180.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. ***Proceeding Of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,*** 1, 281-296.

Matzinger, P. (1994). Tolerance, danger, and the extended family. ***Annual review of immunology,*** 12, 991-1045.

Miniwatts Marketing Group. (2010). ***World internet usage and population stats***, (On-Line), Available: http://www.internetworldstats.com/stats.htm

MIT, (1999). ***Lincoln Laboratory,*** (On-Line), Available: http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/docs/attackDB.html

Mitchell, T. (1997). ***Machine learning.*** (1st Ed.), New York : McGraw-Hill.

Mohamed, M. , Smair, B. & Abdullah A. (2010). Immune multiagent system for network intrusion detection using non-linear classification algorithm. *International Journal of Computer Application,* 12(7). 7-12.

Moskovitch, R., Pluderman, S., Gus, I., Stopel, D., Feher, C., Parmet, Y., Shahar, Y. & Elovici, Y. (2007). Host based intrusion detection using machine learning. *IEEE Intelligence and Security Informatics,* 107-114.

Nobelprize.org, (2011). *The immune system – in more detail.* (On-Line), Available: http://nobelprize.org/educational/medicine/immunity/immune-detail.html

Patil, N., Das, C., Patankar, S. & Pol, K. (2008). Analysis of distributed intrusion detection systems using mobile agents. *Emerging Trend in Engineering and Technology ICETET'08.* 1255-1260.

Paul, W., (1993). *The Immune System : An Introduction.* (3rd Ed.). New York: Raven Press Ltd.

Piel, J. (1993). Life, death and the immune system, special issue, *Scientific American,* 20-269.

Raghunath, B. & Mahadeo, S. (2008). Network intrusion detection system (NIDS), *Emerging Trends in Engineering and Technology ICETET'08.* 1272-1277.

Rieck, K. (2009). *Machine learning for application layer intrusion detection,* (Unpublished doctorate dissertation), Berlin Institute of Technology, Berlin, Germany.

Stolfo, S., Lee, W., Prodromidis, A. & Chan, P. (2000). *Cost-based modeling for fraud and intrusion detection: Results from the jam project. Proceedings 2000 DARPA Information Survivability Conference and Exposition,* 30-144.

Sunjun, L., (2010). An intrusion protection model based on artificial immune. *Proceeding of the 2010 International Forum on Information Technology and Applications*, China.

Tan, P., Steinbach, M. & Kumar, V. (2005). *Introduction to Data Mining.* USA : Addison-Wesley.

The Internet Society, (2000). *RFC 2828 – internet security glossary,* (On-Line), Available: http://www.rfc-archive.org/getrfc.php?rfc=2828#top

Tizard, I., (1995). *Immunology : Introduction,* (4th Ed.) : Saunders College Publishing.

Twycross, J., Aickelin, U. & Whitbrook, A. (2010). Detecting anomalous process behavior using second generation artificial immune systems. *Int. Journal of Unconventional Computing*, 6, 301-326.

U.S. National Institute of Allergy and Infectious Diseases. (2007). *the immune system How it works.* Publication No. 07-5423, USA.

Witten, I. & Frank, E. (2000). *Data mining: Practical machine learning tools and techniques with java implementation*, 265-267 : Morgan Kaufmann Publishers.

Yongzhong, L. & Wang, J. (2009). A novel distributed intrusion detection model based on immune mobile agent. *Proceeding of the 2009 International Symposium on Web Information Systems and Applications International,* 22-24.

Zhang, H. (2004). The Optimality of Naive Bayes. *Proceeding Of the 17th International Florida Artificial Intelligence Research Society Conference,* USA: American Association for Artificial Intelligence Press.