ENHANCED FRAMEWORK FOR ALERT PROCESSING USING CLUSTERING APPROACH BASED ON ARTIFICIAL IMMUNE SYSTEM

ASHARA BANU BINTI MOHAMED

A thesis submitted in fulfilment of the
requirements for the award of degree of
Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

AUGUST 2015

DEDICATION

Dedicated to :

My mother,
Hajjah Khatijah Bt Osman, without whom I would not have the courage to embark
on this journey and the perseverance to complete it.

My father,
Haji Mohamed Bin Maidin, for giving me balance and perspective.

My Brothers,
The charming men in my life, without fail supporting and encouraging me,
throughout this journey

My Sisters
The beautiful ladies in my life, especially Anita, Aida, Azlinah without whom I
would have been lost in the process of completing this phase of my life.

My Nephews and Nieces
Adorable darlings that made my stressful and hectic life bearable

Thank you for your support and Doa.

ACKNOWLEDGEMENT

Bismillah. With the Grace of Allah SWT, I would first and foremost thank the Lord Almighty for giving me the health, strength and ability to complete my thesis as a requirement for my Doctorate studies. I am very grateful to the Malaysian Government, through the Jabatan Perkhidmatan Awam (JPA) who has been the key instrument in enabling me to enhance my capabilities by pursing my studies to the highest level. I am truly blessed to be given this opportunity.

This thesis would not have been possible without the guidance and support I received from both supervisors, Dato' Professor Dr. Norbik Bashah Bin Idris and Dr. Bharanidharan Shanmugum. Your words of encouragement, patience, and guidance had inspired me and gave me the strength throughout this journey. I would also like to extend my gratitude to PRISMA, my previous work place for providing me with the confidential dataset to be applied in my experiments.

Life would be harder without the support of many good friends. Thanks to Rozelinda Bt Khalid, Jalajah Gopalan and Dr. Azalena Haron, for providing me solace when I needed a break and encouragement when I needed a boost of morale. Not forgetting my friends in CASE Lab UTM Kuala Lumpur and new colleagues in i-IMATEC INTAN, thank you for being great companions.

Finally, again to my beloved family who had always stood by me with undying love, encouragement and patience in seeing me through this phase of my life.

May ALLAH swt. bless and bestow his Rahmah to us all.

Amin

ABSTRACT

The Intrusion Detection System (IDS) is an industrial-driven technology that monitors the network infrastructure of an organization from malicious intent. Although the IDS technology has advanced tremendously, one of the main issues that still remains since its beginning is the huge amount of attack alerts that have to be processed immediately on a daily basis. To manage these alerts effectively, both techniques of data reduction and correlation have to be applied concurrently. Therefore, this research proposes a framework named Intelligent Alert Processing Framework (IAPF) that incorporates both techniques named Alert Reduction Module (ARM) and Alert Correlation Module (ACM) to produce an integrated result. The ARM consists of a new clustering algorithm inspired by the Artificial Immune System (AIS) approach which is the Clonal Selection principle, while the ACM is based on pattern recognition approach. The new clustering algorithm introduces a one-to-one clustering method that first and foremost creates cluster based on a perfect matching criterion and next calculates its vulnerability level. Clusters with 0 vulnerability level will be filtered while other clusters will than proceed to ACM for attack scenario formulation and its successful attack scenario probability. The IAPF was successfully experimented using a standard simulated dataset and a real-time dataset from PRISMA (Pemantauan Rangkaian ICT Sektor Awam). The result of the experiment indicated that ARM achieved accurate clustering output, with zero cluster error within an average of 6.36 seconds processing time and the reduction rate of alerts attained is 95.34%. Meanwhile ACM managed to detect all possible attack scenarios based on the predefined patterns. The proposed framework has reduced the number of alerts, creates attack scenarios and simultaneously produced vulnerability level for each clusters and the correlated successful attack scenario probability.

ABSTRAK

Sistem Pengesanan Pencerobohan (IDS) adalah teknologi yang dipacu oleh industri yang digunakan untuk memantau rangkaian infrastruktur organisasi daripada ancaman musuh. Walaupun teknologi IDS telah maju dengan pesat, salah satu isu utama yang masih kekal sejak awal adalah penjanaan jumlah amaran serangan yang besar yang mana perlu diproses dengan serta-merta setiap hari. Bagi mengurus amaran ini dengan berkesan, kedua-dua teknik, pengurangan data dan korelasi perlu digunakan serentak. Oleh yang demikian kajian ini mencadangkan satu rangka kerja bernama Rangka Kerja Pemprosesan Amaran Pintar (IAPF) yang menggabungkan kedua-dua teknik tersebut iaitu, Modul Pengurangan Isyarat (ARM) dan Modul Korelasi Isyarat (ACM) bagi menjana keputusan yang bersepadu. ARM mengandungi algoritma pengklusteran baru yang diilhamkan daripada Sistem Imun Tiruan (AIS) iaitu prinsip Seleksi Klonal, manakala, kaedah korelasi yang digunakan pula adalah berdasarkan kepada pendekatan Pencaman Corak. Algoritma pengklusteran baru memperkenalkan kaedah kluster satu persatu untuk mewujudkan kluster berdasarkan padanan criteria yang sempurna seterusnya mengira tahap kelemahan kluster tersebut. Kluster dengan tahap kelemahan '0' akan ditapis manakala yang selainnya akan disalurkan ke modul ACM bagi pembentukan senario pencerobohan dan kebarangkalian kejayaan senario pencerobohan. IAPF telah berjaya diuji dengan menggunakan pangkalan data simulasi yang standard dan pangkalan data terkini (masa nyata) dari PRISMA (Pemantauan Rangkaian ICT Sektor Awam). Keputusan eksperimen menunjukkan bahawa ARM berjaya menjana kluster tanpa ralat dalam masa pemprosesan 6.36 saat dengan kadar pengurangan isyarat sebanyak 95.34%. Manakala ACM berjaya mengesan semua senario pencerobohan berdasarkan corak yang telah ditetapkan. Rangka kerja yang dicadang telah berjaya mengurangkan bilangan isyarat, membentuk senario pencerobohan, pada masa yang sama menjana tahap kelemahan setiap kluster serta kebarangkalian senario pencerobohan yang berjaya.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AINE	-	Artificial Immune Network
AIS	-	Artificial Immune System
AOI	-	Attribute-Oriented Induction
APM	-	Alert Processing Method
BIRCH	-	Balanced Iterative Reducing and Clustering using Hierarchies
CAML	-	Correlated Attack Modeling Language
CBHFS	-	Correlation-based Hybrid Feature Feature Selection
CFS	-	Correlation-based Feature Selection
Chameleon	-	Hierarchical clustering using dynamic modeling
CI	-	Computational Intelligence
CLARAty	-	Clustering Alerts for Root Cause Analysis
CLUBS	-	CLustering Using Binary Splitting
CRG	-	Causal Relation Graph
CRQT	-	Causal Relation Queue Tree
CS	-	Computer Security
CURE	-	Clustering Using REpresentatives
DAG	-	Directed Acyclic Graphs
DDOS	-	Distributed Denial of Service
DOS	-	Denial of Service
DS	-	Design Science
DSRM	-	Design Science Research Methodology
FCM	-	Fuzzy C mean
GFMM	-	General Fuzzy Min-Max
HIS	-	Human Immune System
IMDEF	-	Intrusion Detection Message Exchange Format

IS	-	Information System
IT	-	Information Technology
IUR	-	Improved Unit Range
PAM	-	Partitioning Around Mediods
PCA	-	Principal Component Analysis
PRISMA	-	Pemantauan Rangkaian ICT Sektor Awam (Public Sector Network Security Operation Centre)
RAINE	-	Resource Limited Artificial Immune System
ROCK	-	Robust Clustering
SE	-	Software Engeneering
SOM	-	Self-Organizing Map
SSAIS	-	Self-Stabilizing Artificial Immune System
SVC	-	Support Vector Clustering
TMDP	-	Total Mean Distribution

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Issues of hacking and phreaking are as old as the history of computers, dating back to the first generation of computers in the late 1930's. These activities became more blatant and extensive with the advent of the Internet in 1969 and personal home and office computers in 1981. In view of this, protection against intruders or attackers has become essential although initially this of greater concern to the government, because the targets were mostly universities, laboratories and military organizations (Lawson, 2011). The next stage was the rise of corporate networks with companies being intricately linked via Internet and the nascent developments in e-commerce in the 90's with their online shopping and other services that deal with private and confidential material such as credit cards, bank accounts and tax revenue which require much personnel identification information.

Advancements in technology have made users extremely vulnerable and concerns for security and the need to protect private and confidential information has escalated into the public and private sectors domains. This is reflected clearly in the security report published in 2013 where cyber-attacks were noted to have branched to various fields as depicted in Figure 1.1. Based on this report the three largest targets of the cyber-attacks are the government (23%), industry (22%) and finance (14%) sectors along with various other anonymous entities such as organizations, the education sector and news groups (Passeri, 2013). Although this report is not exhaustive and only represents discovered attacks, it does provide a global view of the existing and continuously emerging threats to the security of computer data.

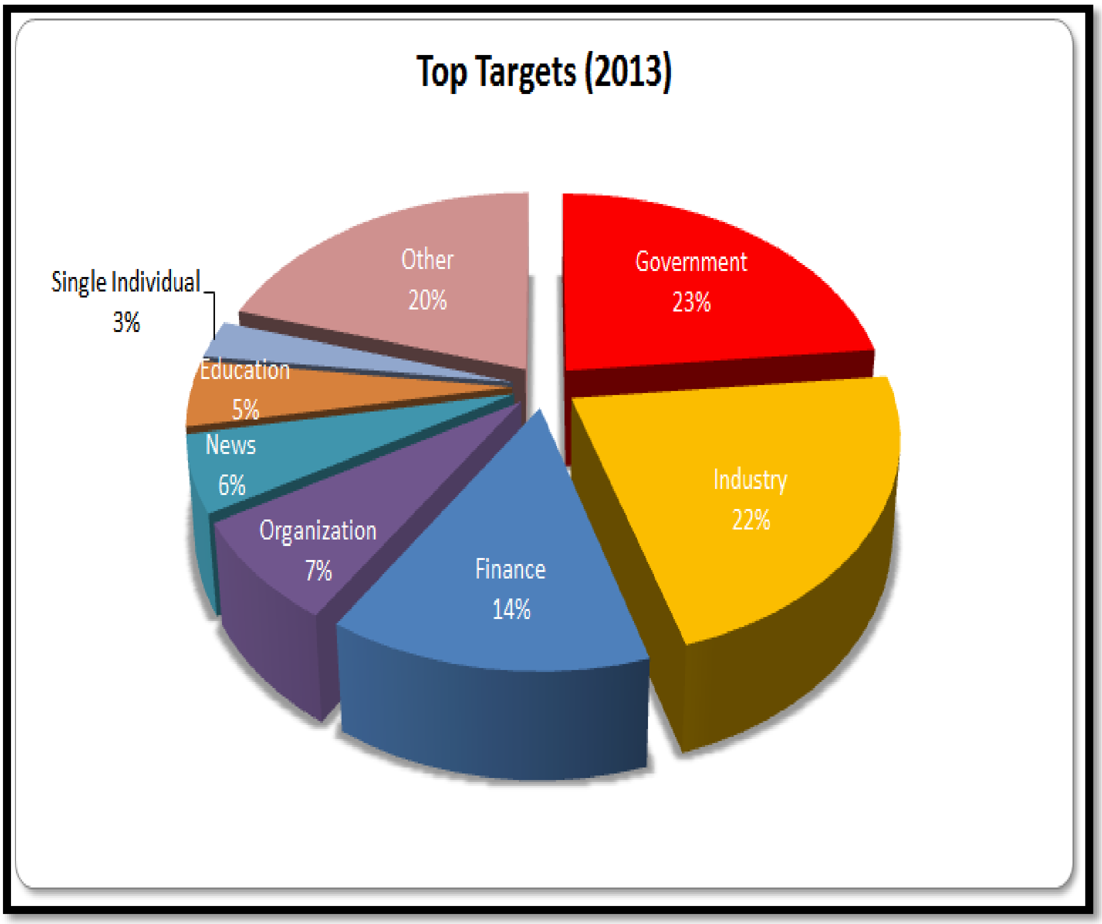


Figure 1.1: Distribution of Cyber-Attack Targets (Passeri, 2013)

The concept of security through monitoring user activity was first introduced in 1980 by James Anderson (1980). It was to protect information from being accessed by unauthorized external or internal users and also to protect information from ‘misfeasors’, users that misused their privilege (Lunt, 1988). This was the initial beginning of Host-Based Intrusion Detection System (HIDS) (Dewan and Mohammad, 2010) but it was limited to monitoring user activities for any malicious or unusual behaviors which is normally done manually using the printed audit logs (Kemmerer and Vigna, 2002). The Intrusion Detection System (IDS) is a form of security software used to identify unauthorized use of computer assets or facilities.

It operates as a secondary defense layer prior to conventional security technique such as authentication and access control (Peddabachigari *et al.*, 2007).

Detection is the initial aspect in IDS, this is followed by the manual processing of the detected alerts. Although there have been many new and innovative IDS softwares since the formation of IDES, due to the persistent problems mentioned in Section 1.2.2, this research focuses on the operational aspects of IDS which is the Alert Processing Method (APM). A framework that is able to refine process and manage the huge amount of alerts generated by IDS so as to finally produce an operational framework which analyze positive attack scenarios is proposed. The proposed framework introduces two most important features in alert processing; data reduction and alert correlation technique. Under the reduction component, feature selection is applied to choose the needed attributes to precede with the clustering and this is followed by the filtering of clusters that are not relevant or in this case clusters that are not a threat to the targeted computer. Next, results from the reduction component are then used as input for our alert correlation module to create attack scenarios and calculate attack scenarios probability. While the reduction solution is inspaired by Human Immune System (HIS) principles, correlation component applied pattern recognition method. By implementing the algorithm, we are able to achieve our objective and solve the problem stated at the initial stage of our research.

The rest of this chapter will explain further the importance of CS by stating the purpose and focus of this research. Firstly, Section 1.2 narrates the background of CS and the central problems addressed in this research. Section 1.3 describes the problem statement, followed by the motivation for performing this research in Section 1.4. The research objectives are presented in Section 1.5, while its scope and significance is in Section 1.6 and 1.7, respectively. For easy understanding and reference Section 1.8 provides an organizational layout of the thesis. Finally, this chapter concludes with the definition of term used in the entire research content in Section 1.9.

1.2 Background of Computer Security (CS)

Security has become an integral component of the digital world as humans are no longer safe just by locking their physical doors and windows. Nowadays security has evolved and mutated in synchronization with the ease and speed of hackers and attackers in designing new and sophisticated attacks. New security mechanisms have to be developed to defend and protect users from these attacks; but even as sophisticated security mechanisms are developed, there will be attempts to intrude and violate the protection established. While computer security is being developed and enhanced, attack activities are growing and spawned at an alarming rate with new holes, new attack strategies and new attackers surfacing almost daily. Unfortunately, security defenders are mostly reacting to attacks, instead of being able to prevent attacks before they occur. Nevertheless efforts must continue in the attempts to produce an improved security mechanism to assists security defenders in safe guarding organization assets.

Previously, research and development of security were mostly focused and based on military importance and needs (Ware, 1967; Landwehr, 1983). However, with the growth in computer crimes from simple hacking to financial and information theft, the need for security has spread broader and wider among users in the computer world. CS is a service provided to protect and maintain the confidentiality, integrity and availability (CIA) of an automated information system resources (NIST., 1995) which also includes the network infrastructures (Bishop, 2003) governing it. Consequently, the CIA triad has become the fundamental characteristics (Kesh and Ratnasingam, 2007) in Information Security (IS) (ISO/IEC, 2005.). In addition to the triad, non-repudiation, accountability, authenticity, and reliability are the characteristics of Information and Communication Technology (ICT) security (ISO/IEC, 2004). Meanwhile the functionality of Cyber Security (CBS) is to protect the CIA of information in the cyberspace (ISO/IEC, 2012); which indicates the permutation or intersection of IS and ICT security. The definition of security by ISO/IEC standards is shown in Figure 1.2.

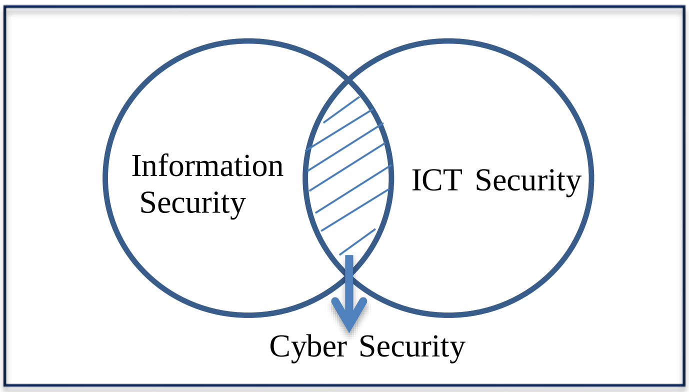


Figure 1.2: Computer Security

Standalone systems, applications and data are protected via information security ; wherelse network connection and transmission is protected by ICT security. Therefore the transmission or accessibility of data or information through the network will constitute the application of cyber security. There are various approaches in the development of CS mechanisms built to address or resolve a particular issue at a given time. Since there is no single solution to multiple problems, the different security mechanisms are used simultaneously as they complement each other in setting up a secure environment.

1.2.1 Challenges of Existing Security Mechanisms

The main goal in implementing CS is to protect classified material against malicious activity whether intentional or not and is a major challenge faced by security practitioners and developers. Security mechanisms have their limitations and vulnerabilities that are constantly being addressed by researchers and developers. Technically there are many different kinds of security mechanisms applied at various levels of the Open System Interconnection (OSI) layer, information and the network infrastructure remain vulnerable. Since existing security mechanisms are solutions to current problems, attackers are continuously designing new and sophisticated attacks to outsmart them. For example hackers have developed methods such as ‘SQL injection’ and ‘cross site scripting’ to penetrate the source code of an application or web portal to steal information. Meanwhile operating systems and anti-virus packages are constantly being updated with new patches to resolve vulnerabilities that are created. Although firewalls are a common element in a network infrastructure, they are still susceptible to circumvention by attackers besides being unable to detect new threats and viruses as they are not programmed to block malicious acts committed by insiders. Similarly, IDS faced similar limitations except that it is able to detect and monitor malicious activity from within and outside the network domain. However IDS has additional concern, which is the huge number of alert generated on a daily basis.

Providing security services is a complicated and difficult task where users, custodians and security experts need to be alert and prepared at all times. It is important to keep abreast of the current security developments and attackers' new exploits. Based on the challenges mentioned, it is obvious that existing security mechanisms require further improvement. This research, however, focuses on the IDS mechanism in managing the huge amounts of alerts generated.

1.2.2 Problem Background on IDS

Even as information security is being developed and enhanced, attack activities are growing and strengthening at a much faster and alarming rate. Although there currently exist new and sophisticated security mechanisms, the vulnerabilities and attack strategies are rampantly multiplying with constant attack and attempts to intrude and violate the established protection systems. To curb this situation it is important to have a mechanism that not only blocks and prevents intruders but also monitors human activities for abnormal or malicious behavior. The Intrusion Detection System (IDS) is a form of security software used to monitor and identify unauthorized use of a computer asset or facility. It operates as a secondary defense layer prior to the conventional security technique such as authentication and access control (Peddabachigari *et al.*, 2007). Unlike other security devices, IDS does not block traffic from entering or leaving a network, it merely detects and logs traffic that is suspected or considered malicious. Although IDS technology has evolved much since the formation of the Intrusion Detection Expert System (IDES) in 1980 (Denning, 1987), towards designing and producing better detection systems, there is still room for improvement as will be discussed briefly in Chapter 2.

IDS has become a major security application which is widely implemented among practitioners and its importance as a security feature to computer infrastructure is indisputable. There are two main segments in IDS applications: a fully automated detection mechanism and an alert processing mechanism which is human dependent (Chaboya *et al.*, 2006). The alerts produced to be processed are generated via the detection segment of IDS sensors resulting in huge amounts of daily alerts due to the efficiency and accuracy of the detection mechanism. The IDS-generated alerts are classified into four categories which are listed below:

i) True Positive (TP) indicating alert generated for a real intrusion
ii) False Positive (FP) indicating alert generated for normal activity
iii) False Negative (FN) indicating missing alert which is not generated
iv) True Negative (TN) indicating no alerts generated for normal activity

Although actions are continuously taken to create a better and more accurate detection mechanism, intruders and attackers are always introducing new and innovative ways to penetrate the user’s network. The single most notable limitation of IDS technology since the beginning of its formation is conceptual in nature which is the detection mechanism. This weakness, causes another major implementation problem for security operators, which is the enormous number of alerts to be manually processed daily. Although there are other shortcomings of the IDS technology, this research addresses issues concerning the alert processing segment which is mostly handled by human operators. These surrounding problems are listed below:

(a) **Huge number of alerts:** It is reported that, in 2012 there was a 42% increase in the number of attacks of which 40% were caused by hackers (Symantec-Corporation, 2013). This increase could be due to ‘overstimulation attacks’ which is a ploy developed to flood the network with false positive alerts and confuse or distract security experts from real attacks (Yurcik, 2002; Chaboya *et al.*, 2006; Corona *et al.*, 2013). Elusion Strategies , such as SYN Flood, is a fragmentation technique that sends useless data to IDS to swamp the IDS sensor with ambiguous data to masquerade the actual attack. Secondly, there are also cases where attacks are disguised through specific coding to evade recognition (Chaboya *et al.*, 2006). Also tools can be used to perform attacks (Giannetsos and Dimitriou, 2013; Mandelcorn, 2013; Prasad *et al.*, 2013), such as DDoS attack tolls that could be initiated randomly by armature hackers(Dayanandam *et al.*, 2013; Ozcelik *et al.*, 2013). Such activities have aggravated the issue further and could paralyze the whole network.

(b) **Processing Accuracy:** Practitioners Broderick (1998), Manganaris *et al.* (2000) and researchers Dain and Cunningham (2001) and Julisch (2001) have frequently highlighted the thousands of ‘innocence alerts’ or ‘mistakes’ (Abouabdalla *et al.*, 2009) generated daily by IDS sensors and the difficulty in sifting through them to find genuine threats (Vignesh *et al.*, 2010; Liao *et al.*, 2012). The act of manually analyzing and processing these alerts is said to be ‘labor-intensive/ human oriented base (Chaboya *et al.*, 2006) and error-prone (Broderick, 1998; Manganaris *et al.*, 2000; Dain and Cunningham, 2001)’. Although there are tools to automate alert analysis and processing (Dain and Cunningham, 2001; Debar and Wespi, 2001; Valdes and Skinner, 2001), there is still need for human involvement in the analysis and processing of alerts, since there is ‘no silver-bullet’ solution to this problem.

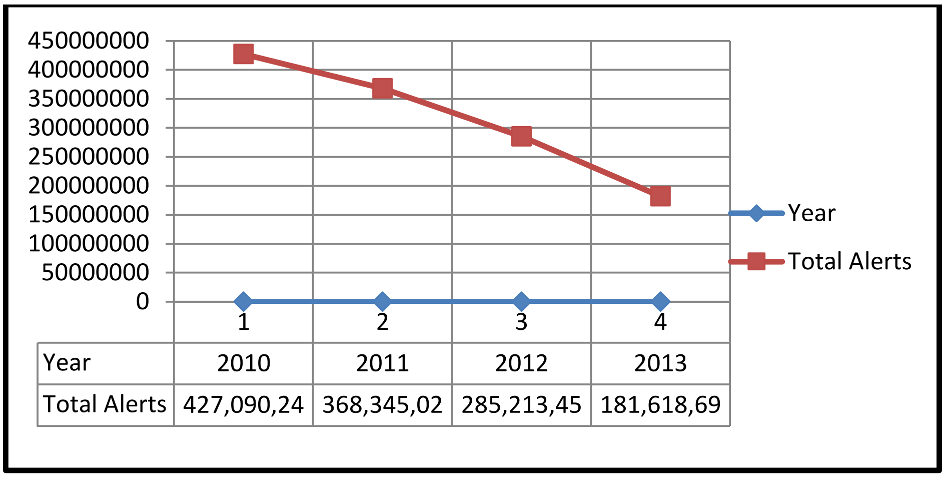
(c) **False Positive (FP):** It is reported that daily IDS’s detected thousands of attacks with more than 90% of them false positive (Julisch, 2003a).

Though the percentage of false positives has reduced overtime due to research and development work for creating better and more sophisticated detection mechanisms, detection inaccuracy is still an issue (Chun-Jen *et al.*, 2013; Tiwari and Alaspurkar, 2013). This false positive alerts dilemma is still a problem which has been extensively researched, as discussed in Chapter 2.

(d) **Processing Performance:** Owing to the above mentioned issues, the performance rate in terms of time and memory capacity of an IDS implementation decreases. As such, reducing the number of alerts and false positives will contribute to better performance rates.

(e) **Intelligent Decision Making Support:** IDS is a monitoring system that complements other security mechanisms. It is not a total integrated solution that could completely prevent attack from occurring and despite its strength, like other security system IDS lacks the ability to support decision making. Since the final assessment on the magnitude of damage caused by an attack is made by security experts therefore it is important that any IDS mechanism provide information to facilitate such decision making.

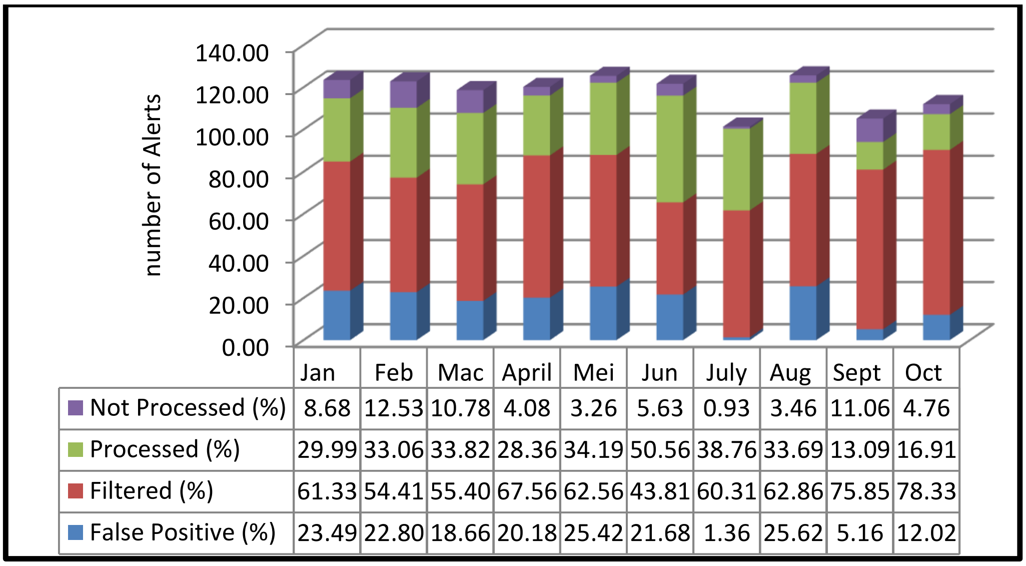
These are problems faced daily by PRISMA (Pemantauan Rangkaian ICT Sektor Awam) a unit created to monitor possible cyber-attacks on Malaysian government agencies. Due to the massive amount of alerts received daily, the 40-staff unit operates around the clock (24x7x365). Figure 1.3 shows the total amount of alerts captured since 2010 and the data up to October 2013 shows a continuous decline in alerts and is due mainly to PRISMA’s success in instilling awareness in the public sectors of the importance of CS and tightening its infrastructures security.



Year	2010	2011	2012	2013
Total Alerts	427,090,24	368,345,02	285,213,45	181,618,69

Figure 1.3: PRISMA Total Alerts Within Four Years

Although the number of alerts has reduced significantly with the implementation of extra security appliances such as firewalls and active maintenance of attack signatures used by IDS sensors that were inappropriate and obsolete, the 181.62 million alerts in 2013 indicates some degree of ineffectiveness. To illustrate further, a breakdown on the total alerts in 2013 into four categories is presented in Figure 1.4. The false positives, filtered, processed and not processed categories could explain the almost 182 million alerts. As shown more than 50% of the alerts were filtered and not considered a threat to the government agencies. This is a very large number and without proper automation process could cause missing alerts. Although the number of monthly false positive alerts is considerably small in percentage terms, in reality filtered alerts are also false positives alerts. The difference being that the filtering is done automatically, while the later were confirmed false positives after a series of manual processing by security experts. It is thus safe to assume that the overall average of false positive alerts for 2013 (up to October) was 79.88%. Another important issue that has to be highlighted is the low percentage of processed alerts that were tasked to the 40 security experts. Despite the large number of human resources and their non-stop efforts they were only able to process an average of 31.24% of the overall total alerts. Although the percentage of the unprocessed alerts looks small however, the potential damage that could occur from these unprocessed alerts which is considered as malicious threats is undeniable.



	Jan	Feb	Mac	April	Mei	Jun	July	Aug	Sept	Oct
🟪 Not Processed (%)	8.68	12.53	10.78	4.08	3.26	5.63	0.93	3.46	11.06	4.76
🟩 Processed (%)	29.99	33.06	33.82	28.36	34.19	50.56	38.76	33.69	13.09	16.91
🟥 Filtered (%)	61.33	54.41	55.40	67.56	62.56	43.81	60.31	62.86	75.85	78.33
🟦 False Positive (%)	23.49	22.80	18.66	20.18	25.42	21.68	1.36	25.62	5.16	12.02

Figure 1.4: Categorization of PRISMA Raw Alerts for year 2013

The issues relating to IDS implementation are not only in theory but very much a practical concern as demonstrated by the data gathered from PRISMA (refer Figure 1.4). The authenticity of mostly false positives alerts generated and filtered through the filtering process is questionable due to the lack of important information as a filtering criteria. Furthermore, even though a huge number of alerts were filtered, the processing performance is still inadequate owing the number of unprocessed alerts arising monthly. The tedious and time consuming manual process has made it impossible for security experts to process all the remaining alerts. Accordingly, the operative function which is supposedly online and in real time is impossible to implement because of the time spent processing useless and meaningless alerts. In addition, the equipment used is constantly in need of upgrading to cater to the huge amounts of data collected and processed.

Taking into consideration the specified problems mentioned above, the purpose of this research is to construct a holistic solution that focuses on the alert processing segment of the IDS application with the objective of minimizing human dependency and to assist security experts in focusing on diagnosing the impact of any attack. It is about managing these alerts and organizing them in a meaningful way to enable the experts to make appropriate decisions on each individual asset under attack. This is achieved by first, reducing the numbers of alerts collected and grouping them to produce attack scenarios. This is done by splitting tasks into smaller modules to create a holistic framework as explained in Chapter 4. Although the main design of this framework is to manage and organize the alerts, the most important element is the accuracy and performance factors in achieving these goals. As such, choosing the best technique to perform the tasks assigned in this framework is a crucial element of the decision-making process.

1.3 Problem Statement

This research focuses on providing a complete and enhanced solution in making alert processing efficient. Through automating alert processing activity human intervention is reduce and intellectual decision making element is incorporated. Based on the problem relating alert processing mentioned in Section 1.2.2, the main research question to be addressed is:

How to effectively refine, process and manage the huge number of alerts generated by IDS to eventually produce an operational framework which analyzes all vulnerable and possible attack scenarios?

To answer the above question, the following need to considered and addressed:

(a) RQ1: What are the methods and techniques used in the process of managing the alerts generated by IDS?

(b) RQ2: How to classify and enhance existing alert processing technique?

(c) RQ3: What are the types of data generated by the SNORT IDS sensor and how to assemble the data required?

(d) RQ4: What are the different known attack scenarios that are available and actively occurring in the datasets applied?

(e) RQ5: How to enhance the existing alert processing framework through the technique and approaches identified?

(f) RQ6: What are the required components involved in the process of constructing the proposed framework?

(g) RQ7: How to evaluate the framework and clustering technique introduced in the proposed solution ?

1.4 Research Motivation

In IDS implementation, alert processing is a component that is activated in response to the output produced by the detection module. The number of alerts produced and their accuracy is beyond the control of the alert processing component but that has to be managed and processed within its jurisdiction. Section 1.2.5 provides a full explanation on the problems originating from this component which is also the research problem of this study. Chapter 3 discusses the limitations of the existing alert processing techniques by earlier researchers, which either concentrate in reducing the number of alerts separately from correlating those alerts or performing both techniques. Based on the review conducted, calculations of vulnerability level assessment and successful attack probability estimation are not available in any of the previous frameworks. Therefore, the emphasis in this research is to address these limitations and enhance the existing framework. Improvements in the current solutions are essential because of the following:

(a) **To maximize accuracy:** The process of reducing or eliminating data in a particular data set is a very risky task as in maintaining its integrity, the method and technique applied has to be accurate. To maximize accuracy, the error rate has to be minimized or controlled. Although solutions that applied reduction techniques are commendable, the introduction of error in the process off data reduction will jeopardize the credibility of the data and ultimately present an unreliable end result.

(b) **To enhance automation**: Large amount of raw alerts are generated by IDS sensors on a daily basis. These alerts have to be processed immediately and efficiently and as such automation process is extremely crucial in this phase of IDS implementation. Even though human involvement in this process is unavoidable, the effort has to be made to try and enhance the automation process to its ultimate capacity using their implicit and explicit knowledge.

(c) **To support decision making:** For the raw alerts to have meaning and value they have to undergo a series of procedures and actions. These processed alerts will ultimately be a source of information for intelligent and decisive decision making. Currently, existing alert processing frameworks are focus on reducing the number of alerts and forming causal relationship among them, but, this is inadequate for sound and effective human decision making which requires more robust and reliable information inputs.

The above are the source of motivation of this research and will aid in answering the main research question:

How to effectively refine, process and manage the huge number of alerts generated by IDS to eventually produce an operational framework which analyzes all vulnerable and possible attack scenarios?

To accomplish this objective the following issues have to be addressed:

(a) Designing modules using best techniques in both spectrum of alert processing to produce zero error rate.

(b) Extracting implicit and explicit knowledge using the taxonomy derived based on previous studies and the scope of this research.

(c) Implementing the newly developed alert processing techniques using the documented knowledge collected.

1.5 Research Objectives

The research objectives based on the problem statement above, are as follows:

(a) To analyze existing clustering technique to present a novel method of reducing the amount of alerts generated by performing a new clustering technique based on Artificial Immune System.

(b) To analyze existing correlation technique to present a knowledge based correlation method based on pattern recognition technique to detect known attack pattern and successfully calculate its probability.

(c) To propose a framework on alert monitoring and processing operation that focuses on an individual machine under attack

1.6 Research Scope

The scope of this study is as follows:

(a) The overall focus of this research is the alert processing segment and the area highlighted is the reduction of the amount of alerts and correlating alerts into scenarios.

(b) This research is geared to the clustering methodology that is based on Artificial Immune System (AIS) approach and correlation technique based on pattern recognition technique.

(c) The data used for this research are the ones captured by SNORT engine. Though there is much information stored in a particular alert that is captured, for the purpose of this research five features which are destination IP, source IP, signature ID, time stamp and event ID are selected.

(d) Targeted assets under attack (destination IP) is the priority in this research; the alert cluster and the scenarios created are based on destination IP.

(e) This research focuses on producing a framework that is practicable and doable within a security monitoring environment.

1.7 Research Significance

The significance of the study lies in the following areas:

(a) Firstly, the most important contribution of this research is towards the ability to reduce the amount of alerts processed without jeopardizing the integrity of the alert.

(b) The clustering techniques are time efficient and computationally cost effective.

(c) The proposed approach is based on individual assets that are being monitored which is critical to any institution.

(d) Finally, the overall framework is flexible and suitable to an environment such as PRISMA.

1.8 Thesis Organization

This is a research that explores the computer security domain and focuses on a particular security application named Intrusion Detections Systems. It is a documentation of activities and processes conducted aimed at producing an artifact or object which is a solution to the issue or problem identified. This thesis consists of seven (7) chapters with each beginning with an introduction and ending with a summary to facilitate understanding of the thesis. The organization overview of this thesis is shown in Figure 1.5, which explains the flow of normal and alternative reading of this thesis. A summary of the succeeding chapters is as follows:

Chapter 2: Presents an overall review of the IDS technology and emphasizes on the limitation currently surrounding the detection mechanisms. This is followed by a survey of the comparative evaluation study performed on a previous Alert Processing Method. The chapter begins with an explanation of two important elements of the alert processing segment. Next is the discussion and comparative evaluation of the existing methods. This is followed with the implementation issues or limitations of each method. Finally, it highlights the elements that could contribute to overcoming the existing limitations.

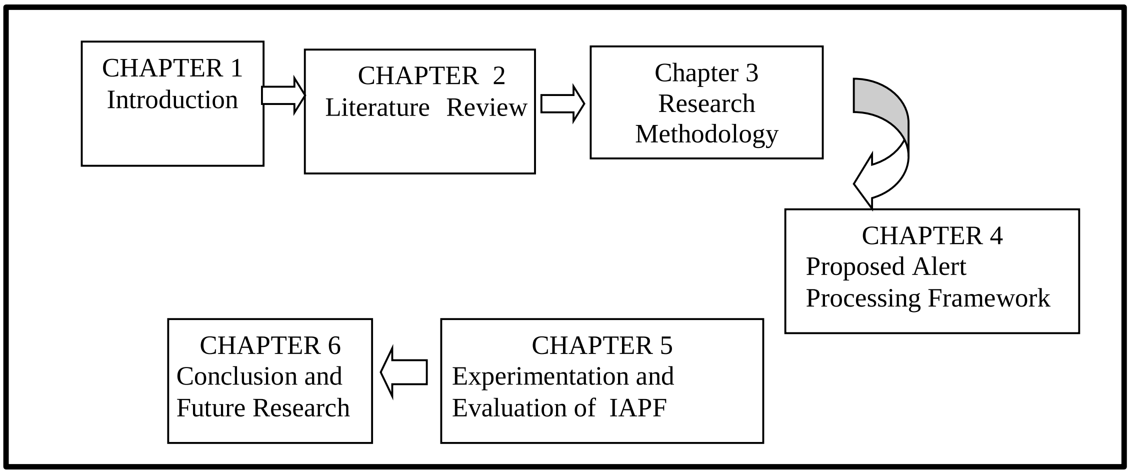


Figure 1.5: Overview of the Thesis Organization

Chapter 3: Constitutes the research methodology adopted in completing this research. The procedures and activities presented by the methodology, guided the research to its proposed objectives. Next, the operational framework presents deliverables of each objectives, followed by a brief presentation of the proposed solution. The data source and format applied for the experiment are discussed. Next the evaluation critera is discussed and lastly, the limititations and assuptions of the proposed solutions are listed.

Chapter 4: This chapter presents the detailed process of designing and developing the proposed Intelligent Alert Processing Framework (IAPF) solution. It begins by explaining the conceptual model of the IAPF and illustrates the designing of each module and the challenges involved in doing that. Lastly, the procedures and processes in building and developing the modules are explained .

Chapter 5: This chapter first describes the implementation procedures for the experiments to be conducted and the evaluation method. Next the experimental results from both data sets are evaluated using the model proposed. The performance of the IAPF prototype is measured based on the calculations mentioned in chapter 4. Next, the evaluation results are used to address the research purpose identified at the beginning of the research. Finally an overall analysis and discussion is conducted on the experimental results.

Chapter 7: Concludes the dissertation with a presentation of the research summary and findings. This is followed by the research contribution, limitations and future works.

1.9 Definition of Terms

The terms used in this research are as listed below:

Alert/Event/Attack – a notification of the occurrence of specific events that matches the signatures (in signature-based NIDS) or deviates from normal activities (for anomaly-based NIDS).

Alert correlation - multi steps process that receives raw alerts as input and acts as a platform to manage and understand the alerts.

Attack graph - is a relational/causal graph or Directed Acyclic Graph (DAG) that represents the causal relationship between attacks to reveal attack strategy. Edges represent action and nodes represent system's tate.

Attack step/stages - steps involved in an attack stage. Technically, it represents the clusters produced by clustering in ARM.

Attack strategy - a complete attack launched by attacker which consists of attack steps and attack stages.

DDoS - stand for Distributed Denial of Service. It referred to an attack which a multitude of compromised systems attack a single target, thereby causing denial of service for users of the targeted system. The flood of incoming messages to the target system essentially forces it to shut down, thereby denying service to the system to legitimate users.

True Positive (TP) - indicates alert that is generated for a real intrusion.

False Positive (FP) - indicates alert that is generated for normal activity

False Negative (FN) - indicates missing alerts which is not generated

False Negative (FN) - indicates no alerts generated for normal activity

False Positive Attack - indicates calculated vunerability level of a cluster.

Known alert - a labelled alert that has class information based on previous data or domain experts knowledge. It is usually used for training the machine learning algorithm.

Unseen/new alert - an unlabeled alert that has no class information. It is usually used for validation and testing the machine learning algorithm.

Missing alerts - attacks that are not captured by the IDS sensor.

Reduce alerts - alerts that are clustered or grouped together based on a sepecified criteria to reduce redundancy.

Reduce False Positive- alerts or clusters that are not malicious and not a threat to the targeted destination IP.

REFERENCES

Abouabdalla, O., El-Taj, H., Manasrah, A. and Ramadass, S. (2009). False positive reduction in intrusion detection system: A survey. *Broadband Network & Multimedia Technology, 2009. IC-BNMT'09.* 2nd IEEE International Conference, 463-466.

Abraham, A., Grosan, C. and Martin-Vide, C. (2007). Evolutionary Design of Intrusion Detection Programs. *IJ Network Security*. 4 (3), 328-339.

Ada, G. L. and Nosal, G. (1987). *The clonal selection theory.* Scientific American

Ahrabi, A. A. A., Feyzi, K., Orang, Z. A., Bahrbegi, H. and Safarzadeh, E. (2012). Using Learning Vector Quantization in Alert Management of Intrusion Detection System. *International Journal of Computer Science and Security*. 6 (2), 128.

Aickelin, U., Greensmith, J. and Twycross, J. (2004). *Immune System Approaches to Intrusion Detection – A Review*. In G. Nicosia, V. Cutello, P. Bentley and J. Timmis. *Artificial Immune Systems*. (pp. 316-329), Berlin Heidelberg:Springer.

Al-Mamory, S. O. and Zhang, H. L. (2007). A survey on IDS alerts processing techniques. *Proceedings of the 6th Wseas International Conference on Information Security and Privacy (Isp '07),*69-78.

Al-Mamory, S. O. and Zhang, H. L. (2009). Intrusion detection alarms reduction using root cause analysis and clustering. *Computer Communications.* 32 (2), 419-430.

Alserhani, F. (2013). A framework for multi-stage attack detection. *Electronics, Communications and Photonics Conference (SIECPC).* Saudi International. 1-6.

Alserhani, F., Akhlaq, M., Awan, I. U., Cullen, A. J. and Mirchandani, P. (2010). MARS: Multi-stage Attack Recognition System. *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference* 20-23 April 2010. 753-759.

Amrita, P. and Ahmed, P. (2014). A Hybrid-Based Feature Selection Approach for IDS. In *Networks and Communications (NetCom2013).* Springer, 195-211.

Aniello, L., Di Luna, G. A., Lodi, G. and Baldoni, R. (2011). A collaborative event processing system for protection of critical infrastructures from cyber attacks. In *Computer Safety, Reliability, and Security*. Springer, 310-323.

Anifowose, F. A. and Eludiora, S. I. (2011). Application of Artificial Intelligence in Network Intrusion Detection.

Asif-Iqbal, H., Udzir, N. I., Mahmod, R., & Ghani, A. A. A. (2011). Filtering events using clustering in heterogeneous security logs. *Information Technology Journal*, 10(4), 798-806.

Bakar, N. A., Belaton, B. and Samsudin, A. (2005). False positives reduction via intrusion alert quality framework. *IEEE 7th Malaysia International Conference on Communication.* . 16-18 November.Malaysia.6-11.

Balajinath, B. and Raghavan, S. V. (2001). Intrusion detection through learning behavior model. *Computer Communications.* 24 (12), 1202-1212.

Bansode, N. S., Pawar, A. B., & Parvat, T. J. (2014). Improve IDS Alert Result By Using Decision Support Techniques.

Bateni, M., Baraani, A. and Ghorbani, A. (2012). Alert correlation using artificial immune recognition system. *International Journal of Bio-Inspired Computation.* 4 (3), 181-195.

Bateni, M., Baraani, A., Ghorbani, A. and Rezaei, A. (2013). An AIS-inspired Architecture for Alert Correlation. *International Journal of innovative Computing, Information & Control.* 9 (1), 231-255.

Ben-Hur, A., Ong, C. S., Sonnenburg, S., Schokopf, B. and Ratsch, G. (2008). Support Vector Machines and Kernels for Computational Biology. *PLoS Computational Biology.* 4 (10), e1000173.

Berkhin, P. (2006). A survey of clustering data mining techniques.Springer-Verlag Heidelberg, 25-71

Biermanna, E., Cloeteb, E. and Venterc, L. M. (2001). A comparison of Intrusion Detection systems. *Computers & Security, Elsevier Science Ltd.* 20 (8), 676-683

Bishop, M. (2003). What is computer security? *Security & Privacy, IEEE.* 1 (1), 67-69.

Broderick, J. (1998). IBM Outsourced Solution. http://www.infoworld.com/cgibin/

displayTC.pl?/980504sb3-ibm.htm (Accessed: 6 January 2011).

Bundy, A. (1985). Incidence calculus: A mechanism for probabilistic reasoning. *Journal of Automated Reasoning.* 1 (3), 263-283.

Carpenter, G. A., Grossberg, S. and Reynolds, J. H. (1991). ARTMAP - Supervused real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks.* 4 (5), 565-588.

Castro, L. N. d. and Jonathan, T. (2002). *Artificial Immune Systems: A New Computational Intelligence Approach.* Great Britain: Springer-Verlag

Castro, L. N. d. and Timmis, J. I. (2003). Artificial immune systems as a novel soft computing paradigm. *Soft Computing.* 7(8), 526-544.

Chaboya, D. J., Raines, R. A., Baldwin, R. O. and Mullins, B. E. (2006). Network intrusion detection: automated and manual methods prone to attack and evasion. *Security & Privacy, IEEE.* 4 (6), 36-43.

Chebrolu, S., Abraham, A. and Thomas, J. P. (2005). Feature deduction and ensemble design of intrusion detection systems. *Computers & Security.* 24 (4), 295-307.

Chen, Y., Li, Y., Cheng, X.-Q. and Guo, L. (2006). Survey and taxonomy of feature selection algorithms in intrusion detection system. *Information Security and Cryptology*.153-167.

Cheng, B.-C., Liao, G.-T., Huang, C.-C. and Yu, M.-T. (2011). A Novel Probabilistic Matching Algorithm for Multi-Stage Attack Forecasts. *IEEE Journal on Selected Areas in Communications.*29 (7), 1438-1448.

Cheung, S., Lindqvist, U. and Fong, M. W. (2003). Modeling multistep cyber attacks for scenario recognition. *DARPA Information Survivability Conference and Exposition Proceedings*. 284-292.

Chou, T.-S. and Chou, T.-N. (2009). Hybrid Classifier Systems for Intrusion Detection. *Seventh Annual Communication Networks and Services Research Conference(CNSR'09).*286-291.

Chowdhary, M., Suri, S., & Bhutani, M. (2011). Comparative Study of Intrusion Detection System. *Journal of Computer Science International Journal of Computer Science International Journal of Computer Sciencesand Engineering and Engineering*, 2(4), 197-200.

Chun-Jen, C., Khatkar, P., Tianyi, X., Jeongkeun, L. and Dijiang, H. (2013). NICE: Network Intrusion Detection and Countermeasure Selection in Virtual Network Systems. *IEEE Transactions on Dependable and Secure Computing,* 10 (4), 198-211.

Corbató, F. J., Merwin-Daggett, M. and Daley, R. C. (1962). An experimental time-sharing system. *Spring Joint Computer Conference Proceedings.* May 1-3, 335-344.

Corbató, F. J., Saltzer, J. H. and Clingen, C. T. (1972). Multics: The first seven years. *Spring Joint Computer Conference Proceedings.* May 16-18, 1972, 571-583.

Corona, I., Giacinto, G. and Roli, F. (2013). Adversarial attacks against intrusion detection systems: Taxonomy, solutions and open issues. *Information Sciences.* 239 (0), 201-225.

Cuppens, F., & Miege, A. (2002). Alert correlation in a cooperative intrusion detection framework. *Proceedings of the IEEE Symposium on Security and Privacy*.202-215.

Cuppens, F. and Ortalo, R. (2000). LAMBDA: A language to model a database for detection of attacks. *Proceedings of Recent Advances in Intrusion Detectio.* 197-216.

Dadkhah, S., KhaliliShoja, M. and Taheri, H. (2013). Alert Correlation through a Multi Components Architecture. *International Journal of Electrical and Computer Engineering (IJECE).* 3(4), 461-466.

Dain, O. and Cunningham, R. K. (2001). Fusing a heterogeneous alert stream into scenarios. . In: *Proceedings of ACM Workshop on Data Mining for Security Applications.* Philadelphia,PA, 1-13.

Daley, R. C. and Neumann, P. G. (1965). A general-purpose file system for secondary storage. *Proceedings of the Joint Computer Conference.* November 30 - December 1, 213-229.

Dasgupta, D., Ji, Z. and Gonzalez, F. (2003). Artificial immune system (AIS) research in the last five years. *The Congress on Evolutionary Computation (CEC '03).* 8-12 December, 121, 123-130

Dayanandam, G., Rao, T., Reddy, S. P. K. and Sruthi, R. (2013). Password Based Scheme And Group Testing For Defending DDOS Attacks. *International Journal of Network Security & Its Applications,* 5(3).

Debar, H. and Wespi, A. (2001). Aggregation and Correlation of Intrusion-Detection Alerts. Heidelberg: Springer-Verlag, 85-103.

Denning, D. E. (1987). An Intrusion-Detection Model. . *IEEE Transactions On Software Engineering.*13(2).

Depren, O., Topallar, M., Anarim, E. and Ciliz, M. K. (2005). An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. *Expert Systems with Applications.* 29 (4), 713-722.

Després, S. and Zargayouna, H. (2009). Evaluation of knowledge based applications: benchmark and guidelines. *Fifth International Conference on Signal-Image Technology & Internet-Based Systems (SITIS).* 472-478.

Dewan, M. F. and Mohammad, Z. R. (2010). Anomaly Network Intrusion Detection Based on Improved Self Adaptive Bayesian Algorithm. *Journal Of Computers*. 5(1).

Du, H., Liu, D. F., Holsopple, J. and Yang, S. J. (2010). Toward ensemble characterization and projection of multistage cyber attacks. *Computer Communications and Networks Conference (ICCCN).* 1-8.

Dunn, J. (1974). A fuzzy relative of the isodata process and its use in detecting compact well-separated cluster. *Journal of Cybernetic.* 3 32.

Ebrahimi, A., Navin, A., Mirnia, M., Bahrbegi, H. and Ahrabi, A. (2011). Automatic attack scenario discovering based on a new alert correlation method. *Systems Conference (SysCon), IEEE International*. 52-58.

Elshoush, H. T. and Osman, I. M. (2011). Alert correlation in collaborative intelligent intrusion detection systems—A survey. *Applied Soft Computing.* 11 (7), 4349-4365.

Elshoush, H. T. and Osman, I. M. (2013). Intrusion Alert Correlation Framework: An Innovative Approach. In *IAENG Transactions on Engineering Technologies*. Springer, 405-420.

Embar-Seddon, A. (2002). Cyberterrorism Are We Under Siege? *American Behavioral Scientist.* 45 (6), 1033-1043.

Estivill-Castro, V. (2002). Why so many clustering algorithms: a position paper. *ACM SIGKDD Explorations Newsletter.* 4 (1), 65-75.

Faccin, S., Purnadi, R., Hulkkonen, T., Rajaniemi, J., Tuohino, M. and Sivanandan, M.(2012). System And Method Of Controlling Application Level Access Of Subscriber To A Network. *US Patent Application* 13/430,779.

Fava, D. S., Byers, S. R. and Yang, S. J. (2008). Projecting cyberattacks through variable-length markov models. *IEEE Transactions on Information Forensics and Security.* 3(3), 359-369.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine.* 17(3), 37.

Forrest, S., Perelson, A. S., Allen, L. and Cherukuri, R. (1994). Self-nonself discrimination in a computer. *IEEE Computer Society Symposium on Research in Security and Privacy*. 202-202.

Frank, J. (1994). Artificial Intelligence and Intrusion Detection: Current and Future Direction. *Proc. 17th National Computer Security Conference (Baltimore, MD).*

Gaber, M. M., Zaslavsky, A. and Krishnaswamy, S. (2005). Mining data streams: a review. *ACM Sigmod Record.* 34 (2), 18-26.

Gabrys, B. and Bargiela, A. (2000). General fuzzy min-max neural network for clustering and classification. *Ieee Transactions on Neural Networks.*11 (3), 769-783.

Gaidhane, R., Vaidya, C. and Raghuwanshi, M. (2014). Survey: Learning Techniques for Intrusion Detection System (IDS).

Giannetsos, T. and Dimitriou, T. (2013). Spy-Sense: spyware tool for executing stealthy exploits against sensor networks. *Proceedings of the 2nd ACM workshop on Hot topics on wireless network security and privacy*. 7-12.

Gifty Jeya, P., Ravichandran, M. and Ravichandran, C. S. (2012). Efficient Classifier for R2L and U2R Attacks. *International Journal of Computer Applications.4*5(21).

Gionis, A., Mannila, H., & Tsaparas, P. (2007). Clustering aggregation. *ACM Transactions on Knowledge Discovery from Data (TKDD),* 1(1), 4.

Gollmann, D. (2010). Computer security. *Wiley Interdisciplinary Reviews: Computational Statistics.* 2 (5), 544-554.

Gonçalves, M., Andrade Netto, M. and Costa, J. (2011). Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information. *Self-Organizing Maps Applications and Novel Algorithm Design.* 1, 299-322.

Graham, R. M. (1968). Protection in an information processing utility. *Communications of the ACM.* 11(5), 365-369.

Guha, S., Rastogi, R. and Shim, K. (2000). ROCK: A Robust Clustering Algorithm for Categorical Attributes. *Information System.* 25(5), 345-366.

Guha, S., Rastogi, R. and Shim, K. (2001). Cure: An efficient clustering algorithm for large databases. *Information Systems.* 26(1), 35-58.

Haines, J. W., 2012. "DARPA Intrusion Detection Evaluation." 2013, from http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/data/2000/LLS_DDOS_1.0.html.

Han, J., Kamber, M. and Pei, J. (2006). *Data mining: concepts and techniques*San Fancisco, CA itd: Morgan kaufmann.

Hansman, S. and Hunt, R. (2005). A taxonomy of network and computer attacks. *Computers & Security.* 24(1), 31-43.

Hashim, M. and Zaiton, S. (2008). Network intrusion alert correlation challenges and techniques. *Jurnal Teknologi Maklumat.* 20 (2), 12-36.

Hathaway, R. J. and Bezdek, J. C. (2001). Fuzzy c-means clustering of incomplete data. *Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics.* 31 (5), 735-744.

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly.* 28 (1), 75-105.

Holsopple, J. and Yang, S. J. (2008). FuSIA: future situation and impact awareness. *Information Fusion, 2008 11th International Conference on*. 1-8.

Huan, L. and Lei, Y. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on.Knowledge and Data Engineering.* 17(4), 491-502.

Iivari, J. (2007). A Paradigmatic Analysis of Information Systems as a Design Science. *Scandinavian Journal of Information Systems.* 19(2) 87-92.

Ingham, K. and Forrest, S. (2002), *A History and Survey of Network Firewalls*. University New Mexico: Technical Report :TRCS-2002-37.

ISO/IEC,(2004). ISO/IEC TR 13335-1:2004 information technology security techniques management of information and communications technology security part 1: concepts and models for information and communications technology security management. , ISO/IEC, JTC 1, SC27, WG 1 2004.

ISO/IEC (2011). Information technology-security techniques-information security incident management. ISO/IEC 27035:2009.

ISO/IEC(2012). ISO/IEC 27032:2012 Information technology — Security techniques — Guidelines for cybersecurity

ISO/IEC, (2005). ISO/IEC 27002: code of practice for information security management

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters.* 3(8), 651-666.

Jain, A. K., Murty, M. N. and Flynn, P. J. (1999). Data clustering: A review. *Acm Computing Surveys.*31(3),264-323.

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

Jiang, W. (2012). Survey of network and computer attack taxonomy. *IEEE Symposium on Robotics and Applications (ISRA)*. 3-5 June 2012. 294-297.

Julisch, K. (2001). Mining Alarm Clusters to Improve Alarm Handling Eficiency. . *In 17th Annual Computer Security Applications Conference (ACSAC).* 12-21.

Julisch, K. (2003). Clustering intrusion detection alarms to support root cause analysis. *ACM Trans. Inf. Syst. Secur.* 6(4), 443-471.

Kabiri, P. and Ghorbani, A. A. (2005). Research on intrusion Detection and Response: A Survey. *International Journal of Network Security.* 1(2), 84-102.

Kabiri, P. and Ghorbani, A. A. (2007). A Rule-Based Temporal Alert Correlation System. *International Journal of Network Security.*5 (1).

Karypis, G., Han, E. H. and Kumar, V. (1999). Chameleon: Hierarchical clustering using dynamic modeling. *Computer.*32(8), 68-75.

Kavousi, F. and Akbari, B. (2013). A Bayesian network-based approach for learning attack strategies from intrusion alerts. *Security and Communication Networks.* 7(5),833-853.

Kayacik, H. G., Zincir-Heywood, A. N. and Heywood, M. I. (2005). Selecting features for intrusion detection: A feature relevance analysis on KDD 99 intrusion detection datasets. *Proceedings of the third annual conference on privacy, security and trust*.

Kemmerer, D. and Vigna, G. (2002). Intrusion detection: A brief history and overview. *Computer.*27-30.

Kephart, J., Sorkin, G., Swimmer, M., & White, S. (1999). *Blueprint for a Computer Immune System.*Springer Berlin Heidelberg.242-261.

Kesh, S. and Ratnasingam, P. (2007). A knowledge architecture for IT security. *Communications of the ACM.* 50 (7), 103-108.

Kitchenham, B., Linkman, S. and Law, D. (1997). DESMET: a methodology for evaluating software engineering methods and tools. *Computing & Control Engineering Journal.* 8(3), 120-126.

Kitchenham, B. A. (1996). Evaluating software engineering methods and tool part 1: The evaluation context and evaluation methods. *ACM SIGSOFT Software Engineering Notes.* 21(1), 11-14.

Kohonen, T. (2001). Self-organizing maps (Vol. 30). Springer-Verlag New York. Inc., Secaucus, NJ, *43*.

Kruegel, C., Robertson, W. and Vigna, G. (2004). Using alert verification to identify successful intrusion attempts. *Praxis der Informationsverarbeitung und Kommunikation.* 27 (4), 219-227.

Kruegel, C. and Toth, T. (2003). Using decision trees to improve signature-based intrusion detection. *Recent Advances in Intrusion Detection, Proceedings.* 2820 173-191.

Kumar, G., Kumar, K. and Sachdeva, M. (2010). The use of artificial intelligence based techniques for intrusion detection: A review. *Artificial Intelligence Review.* 34 (4), 369-387.

Kumar, M., Hanumanthappa, M. and Kumar, T. S. (2011). Intrusion Detection System-False Positive Alert Reduction Technique. *ACEEE International Journal on Network Security.* 2 (3).

Kumar, M., Siddique, S. and Noor, H. (2009). Feature-based alert correlation in security systems using self organizing maps. *SPIE Defense, Security, and Sensing*. 734404-734407.

Landwehr, C. E. (1983). Best available technologies for computer security. *Computer.* 16 (7), 86-95.

Lawson, S. (2011). Beyond cyber-doom: Cyberattack Scenrios and the Evidence of History. *Mercatus Center George Mason University Working Paper.* 01-11.

Lee, S., Chung, B., Kim, H., Lee, Y., Park, C. and Yoon, H. (2006). Real-time analysis of intrusion detection alerts via correlation. *Computers & Security.* 25 (3), 169-183.

Li, M., Deng, S., Wang, L., Feng, S., & Fan, J. (2014). Hierarchical clustering algorithm for categorical data using a probabilistic rough set model. *Knowledge-Based Systems*, 65, 60-71.

Li, X., Lu, T., Wang, Z. and Gao, C. (2008). ICAIS: a novel incremental clustering algorithm based on artificial immune systems. *International Conference on Internet Computing in Science and Engineering(ICICSE'08).* 85-90.

Li, Y., Xia, J., Zhang, S., Yan, J., Ai, X., & Dai, K. (2012). An efficient intrusion detection system based on support vector machines and gradually feature removal method. *Expert Systems with Applications*, 39(1), 424-430.

Liao, H. J., Lin, C. H. R., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36(1), 16-24.

Liu, C. L. (1968). *Introduction to combinatorial mathematics.* New York: McGraw-Hill.

Liu, H., Dougherty, E. R., Dy, J. G., Torkkola, K., Tuv, E., Peng, H., Ding, C., Long, F., Berens, M., Parsons, L., Zhao, Z., Yu, L. and Forman, G. (2005). Evolving feature selection. *Intelligent Systems, IEEE.* 20 (6), 64-76.

Liu, H. and Motoda, H. (1998). *Feature extraction, construction and selection: A data mining perspective.* Springer Science & Business Media.

Liu, X., Xia, Y., Wang, Y. and Ren, J. (2013). Discovering anomaly on the basis of flow estimation of alert feature distribution. *Security and Communication Networks. 7*(10), 1570-1581.

Liu, Z., Jin, X., Bie, R. and Gao, X. (2007). FAISC: a Fuzzy Artificial Immune System Clustering Algorithm. *Third International Conference on Natural Computation(ICNC).* 657-661.

Lunt, T. F. (1988). Automated audit trail analysis and intrusion detection: A survey. *11th National Computer Security Conference.*

Lunt, T. F., Tamaru, A., Gilham, F. (1992). A Real-Time Intrusion Detection Expert System (IDES). SRI International, Computer Science Laboratory.

Mahboubian, M., Udzir, N. I., Subramaniam, S. and Hamid, N. A. W. A. (2012). An AIS Inspired Alert Reduction Model. *International Journal of Cyber-Security and Digital Forensics (IJCSDF).* 1(2) 130-139.

Maimon, O. Z. and Rokach, L. (2005). *Data mining and knowledge discovery handbook.*(Eds).New York: Springer.

Man, D., Yang, W., Wang, W. and Xuan, S. (2012). An alert aggregation algorithm based on iterative self-organization. *Procedia Engineering.* 29, 3033-3038.

Mandelcorn, S. M. (2013). An Explanatory Model of Motivation for Cyber-Attacks Drawn from Criminological Theories.

Manganaris, S., Christensen, M., Zerkle, D. and Hermiz, K. (2000). A Data Mining Analysis of RTID Alarms. *Computer Networks.* 34 (4) 571-577.

March, S. T. and Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems.* 15 (4), 251-266.

Marchetti, M., Colajanni, M. and Manganiello, F. (2011). Framework and Models for Multistep Attack Detection. *International Journal of Security and Its Applications.* 5 (4), 73-90.

Masciari, E., Mazzeo, G. M., & Zaniolo, C. (2013). A New, Fast and Accurate Algorithm for Hierarchical Clustering on Euclidean Distances. In *Advances in Knowledge Discovery and Data Mining* (pp.111-122). Berlin, Heidelberg: Springer.

McHugh, J. (2000a). The 1998 Lincoln Laboratory IDS evaluation a critique. *Recent Advances in Intrusion Detection..* 145-161.

McHugh, J. (2000b). Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory. *ACM transactions on Information and system Security.* 3 (4), 262-294.

Mirheidari, S. A., Arshad, S. and Jalili, R. (2013). *Alert Correlation Algorithms: A Survey and Taxonomy*. In *Cyberspace Safety and Security*. (pp.183-197) Springer International Publishing.

Morin, B., Me, L., Debar, H. and Ducasse, M. (2009). A logic-based model to support alert correlation in intrusion detection. *Information Fusion.* 10 (4), 285-299.

Nasraoui, O., Gonzalez, F., Cardona, C., Rojas, C. and Dasgupta, D. (2003). A scalable artificial immune system model for dynamic unsupervised learning. In *Genetic and Evolutionary Computation—GECCO*.(pp. 219-230). Berlin, Heidelberg: Springer.

NCBI (2007), *MICROARRAYS: CHIPPING AWAY AT THE MYSTERIES OF SCIENCE AND MEDICINE*. National Center for Biotechnology Information.

Neal, M. (2003). *Meta-stable memory in an artificial immune network*. In *Artificial Immune Systems*.(pp.168-180). Berlin, Heidelberg: Springer..

Nehinbe, J. O. (2010). Automated Method for Reducing False Positives. *Uksim-Amss First International Conference on Intelligent Systems, Modelling and Simulation*.54-59.

Ning, P., Cui, Y. and Reeves, D. S. (2002). Constructing Attack Scenarios through Correlation of Intrusion Alerts. *ACM Conference Computer and Communication Security*. 245-254.

Ning, P. and Xu, D. B. (2004). Adapting query optimization techniques for efficient alert correlation. *Data and Applications Security Xvii: Status and Prospects*. 142, 75-88.

NIST., 1995. An Introduction to Computer Security: The NIST Handbook .

Njogu, H. W. and Luo, J. (2010). Using Alert Cluster to reduce IDS alerts. *3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT*. 9-11 July 2010. 467-471.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization science*. 5 (1), 14-37.

Nonaka, I., & Konno, N. (2005). The concept of “5, 4”: building a foundation for knowledge creation. *Knowledge management: critical perspectives on business and management,* 2(3), 53.

Nunamaker, J. and Chen, M. P. (1991). Systems Development in Information Systems Research. *Journal of Management of Information Systems*. 7(3), 89-101.

Nunes de Casto, L. and Von Zuben, F. J. (2000). An evolutionary immune network for data clustering. *Sixth Brazilian Symposium on.Neural Networks*. 84-89.

Olson, D. and Delen, D. (2008). *Data Mining Process*. In Advanced Data Mining Techniques. (pp. 9-35) Berlin, Heidelberg: Springer.

Ozcelik, I., Fu, Y. and Brooks, R. R. (2013). DoS Detection is Easier Now. *Research and Educational Experiment Workshop (GREE).* 50-55.

Parashar , A., Saurabh, P. and Verma, B.,(2013). A Novel Approach for Intrusion Detection System Using Artificial Immune System. *Proceedings of All India Seminar on Biomedical Engineering (AISOBE 2012).* India: Springer, 221-229.

Pasa, L., Costa, J., Tosin, M. and Paiva, F. P. (2012). *Using SOM Maps for Clustering and Visualization of Diamond Films Deposited by HFCVD Process*. In J. Pavón, N. Duque-Méndez and R. Fuentes-Fernández. *Advances in Artificial Intelligence – IBERAMIA 2012*. (pp. 140-148) Berlin, Heidelberg: Springer.

Passeri, P., (2013). Cyber Attacks Statistics. *Hackmageddon.com*. 2014.

Paulauskas, N. and Garsva, E. (2008). Attacker skill level distribution estimation in the system mean time-to-compromise. *1st International Conference on.Information Technology(IT 2008).* 1-4.

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

Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007). A design science research methodology for Information Systems Research. *Journal of Management Information Systems.* 24 (3), 45-77.

Perdisci, R., Giacinto, G. and Roll, F., 2006. Alarm clustering for intrusion detection systems in computer networks. *Engineering Applications of Artificial Intelligence*. 19, 429-438.

Pfleeger, S. L. (1995). Experimental design and analysis in software engineering. *Annals of Software Engineering.* 1 (1), 219-253.

Pietraszek, T. and Tanner, A. (2005). Data mining and machine learning—towards reducing false positives in intrusion detection. *Information Security Technical Report.* 10 (3), 169-183.

Pikoulas, J., Buchan, W. J., Mannon, M. and Triantafyllopoulos, K. (2001). An Agent-based Bayesian Forecasting Model for Enhanced Network Security. *Proceedings of the International Symposium and Workshop on Engineering of Computer Based Systems.*247-254.

Porras, P. A., Fong, M. W. and Valdes, A. (2002). A mission-impact-based approach to INFOSEC alarm correlation. *Recent Advances in Intrusion Detection, Proceedings.*(95-114) Springer, Berlin Heidelberg.

Prasad, K., Reddy, A. and Rao, K. (2013). Discriminating DDoS Attack traffic from Flash Crowds on Internet Threat Monitors (ITM) Using Entropy variations. *African Journal of Computing & ICT.* 6 (2).

Qin, X. and Lee, W. (2003). Statistical causality analysis of infosec alert data. *Recent Advances in Intrusion Detection*. 73-93.

Qin, X. and Lee, W. (2007). *Discovering novel attack strategies from INFOSEC alerts*. In *Data Warehousing and Data Mining Techniques for Cyber Security*. (pp. 109-157). US: Springer.

Quinlan, J. R. (1986). *Induction of Decision Trees.* Boston. Kluwer Academic Publishers.

Rodgers, J. (1959). The meaning of correlation *American Journal of Science.* 257, 684-691.

Roesch, M. (1999). Snort: Lightweight Intrusion Detection for Networks. . *Proceedings of Usenix 13th Systems Administration Conference (LISA 99)*. November 7-12. Seattle, WA. 229-238.

Roschke, S., Cheng, F. and Meinel, C. (2012). An alert correlation platform for memory-supported techniques. *Concurrency and Computation: Practice and Experience.* 24 (10), 1123-1136.

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

Saad, S. and Traore, I. (2013). *Extracting attack scenarios using intrusion semantics*. In *Foundations and Practice of Security*. (278-292). Berlin, Heidelberg:Springer..

Sadoddin, R. and Ghorbani, A. A. (2009). An incremental frequent structure mining framework for real-time alert correlation. *Computers & Security.* 28 (3-4), 153-173.

Saha, I., Plewczynski, D., Maulik, U. and Bandyopadhyay, S. (2010). Consensus Multiobjective Differential Crisp Clustering for Categorical Data Analysis. In *Rough Sets and Current Trends in Computing.* (pp. 30-39). Berlin, Heidelberg: Springer.

Sandhu, R. S., Coyne, E. J., Feinstein, H. L. and Youman, C. E. (1996). Role-based access control models. *Computer.* 29 (2), 38-47.

Sapats, M. and Paulins, N. (2012). Evaluation methods of network intrusion detection systems.In *International Scientific Conference: Applied Information and Communication Technologies, 5, Jelgava (Latvia), 26-27 Apr*.

Sarita K. Tiwari and Alaspurkar, S. J. (2013). Review of Techniques Reducing False Positive In IDS. *International Journal of Research in Computer Engineering & Electronics.* 2 (2).

Schatzoff, M., Tsao, R. and Wing, R. (1967). An experimental comparison of time sharing and batch processing. *Communications of the ACM.* 10 (5), 261-265.

Shanmugam, B. and Idris, N. B. (2009). Improved Intrusion Detection System using Fuzzy Logic for Detecting Anamoly and Misuse type of Attacks. *International Conference of Soft Computing and Pattern Recognition.*212-217.

Shittu, R., Healing, A., Ghanea-Hercock, R., Bloomfield, R., & Rajarajan, M. (2015). Intrusion Alert Prioritisation and Attack Detection using Post-Correlation Analysis. *Computers & Security*. 50, 1-15.

Simon, H. A. (1996). *The sciences of the artificial*. MIT press.

Sindhu, S. S. S., Geetha, S., Sivanath, S. S. and Kannan, A. (2007). A neuro-genetic ensemble short term forecasting framework for anomaly intrusion prediction. *2006 International Conference on Advanced Computing and Communications,* 2,181-184.

Siqueira, A. F., Cabrera, F. C., Pagamisse, A. and Job, A. E. (2014). Segmentation of scanning electron microscopy images from natural rubber samples with gold nanoparticles using starlet wavelets. *Microscopy research and technique.* 77 (1), 71-78.

Siraj, M. M. (2013), *A Hybrid of Structural Causal and Statistical Model for Intrusion Alert Correlation*. PHD. Universiti Teknologi Malaysia. Skudai. Johor Bahru.

Smets, P. (1990). The combination of evidence in the transferable belief model. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 12 (5), 447-458.

Smith, R., Japkowicz, N., Dondo, M. and Mason, P. (2008). *Using Unsupervised Learning for Network Alert Correlation*. In S. Bergler. *Advances in Artificial Intelligence*. (pp. 308-319). Berlin, Heidelberg: Springer.

Snort, C., 2014. "Snort Web Page." from http://www.snort.org/snort-downloads.

Somayaji, A., Hofmeyr, S. and Forrest, S. (1998). Principles of a computer immune system. *Proceedings of the 1997 workshop on New security paradigms*. 75-82.

Spathoulas, G. and Katsikas, S. (2013). Methods for post-processing of alerts in intrusion detection: A survey. *International Journal of Information Security Science.* 2 (2), 64-80.

Spathoulas, G. P. and Katsikas, S. K. (2010). Reducing false positives in intrusion detection systems. *Computers & Security.* 29 (1), 35-44.

Stallman, R. 1989 GNU General Public License. .

Staniford-Chen, S., Cheung, S., Crawford, R., Dilger, M., Frank, J., Hoagland, J., Levitt, K., Wee, C., Yip, R. and Zerkle, D. (1996). GrIDS-a graph based intrusion detection system for large networks. *Proceedings of the 19th national information systems security conference*. 361-370.

Symantec-Corporation (2013), *Internet Security Threat Report 2013*. 18.

Tan, K. (1995). The application of neural networks to UNIX computer security. *IEEE International Conference on Neural Networks.* 476-481.

Tarakanov, A. and Dasgupta, D. (2002). An immunochip architecture and its emulation. *Conference on Evolvable Hardware, NASA/DoD*. 261-261.

Tavallaee, M., Bagheri, E., Lu, W. and Ghorbani, A.-A. (2009). A detailed analysis of the KDD CUP 99 data set. *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*.

Templeton, S. J. and Levitt, K. (2000). A Requires/Provides Model for Computer Attacks. *Proc. New Security Paradigms Workshop.* 31-38.

Thammasiri, D., Delen, D., Meesad, P. and Kasap, N. (2014). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. *Expert Systems with Applications.* 41 (2), 321-330.

Timmis, J. (2000), *Artificial immune systems: a novel data analysis technique inspired by the immune network theory*. Doctor of Philosophy (Ph.D.) thesis Department of Computer Science.

Timmis, J., Neal, M. and Hunt, J. (2000). An artificial immune system for data analysis. *Biosystems.* 55 (1–3), 143-150.

Tjhai, G. C., Furnell, S. M., Papadaki, M. and Clarke, N. L. (2010). A preliminary two-stage alarm correlation and filtering system using SOM neural network and K-means algorithm. *Computers & Security.* 29 (6), 712-723.

Tjhai, G. C., Papadaki, M., Furnell, S. M. and Clarke, N. L. (2008). Investigating the problem of IDS false alarms: An experimental study using Snort. *Proceedings of the Ifip Tc 11/ 23rd International Information Security Conference*: 253-267.

Torkaman, A., Javadzadeh, G. and Bahrololum, M. (2013). A hybrid intelligent HIDS model using two-layer genetic algorithm and neural network. *5th Conference on Information and Knowledge Technology (IKT).* 92-96.

Tsai, C. F., Hsu, Y. F., Lin, C. Y. and Lin, W. Y. (2009). Intrusion Detection by machine learning: A review. *Expert System with Application, Elsevier Ltd.* 36 11994-12000.

Tukey, J., 1977. Exploratory data analysis., Bd. 7616 von Behavioral Science: Quantitative Methods, Addison-Wesley.

Valdes, A. and Skinner, K. (2000). An approach to sensor correlation. . *In Recent Advances in Intrusion Detection (RAID 2000), Toulouse, France.*

Valdes, A. and Skinner, K. (2001). Probabilistic alert correlation. *Springer-Verlag Heidelberg.* vol. 3089 54–68.

Valeur, F., Vigna, G., Kruegel, C. and Kemmerer, R. A. (2004). A comprehensive approach to intrusion detection alert correlation. *Ieee Transactions on Dependable and Secure Computing.* 1 (3), 146-169.

van der Made, P. A., 2008. Computer immune system and method for detecting unwanted code in a P-code or partially compiled native-code program executing within a virtual machine. U.S. Patent No. 7,370,360. Washington, DC: U.S. Patent and Trademark Office.

Verwoerd, T. and Hunt, R. (2002a). Intrusion detection techniques and approaches. *Computer Communications.* 25 (15), 1356-1365.

Verwoerd, T. and Hunt, R. (2002b). Security architecture testing using IDS - a case study. *Computer Communications.* 25 (15), 1402-1412.

Vignesh, R., Ganesh, B., Aarthi, G. and Iyswarya, N. (2010). A Cache Oblivious based GA Solution for Clustering Problem in IDS. *International Journal of Computer Applications.* 1(11), 975 - 8887.

Viinikka, J. and Debar, H. (2004). *Monitoring IDS background noise using EWMA control charts and alert information*. In E. Jonsson, A. Valdes and M. Almgren. *Recent Advances in Intrusion Detection, Proceedings*. (pp. 166-187). Berlin: Springer-Verlag.

Viinikka, J., Debar, H., Me, L., Lehikoinen, A. and Tarvainen, M. (2009). Processing intrusion detection alert aggregates with time series modeling. *Information Fusion.* 10 (4), 312-324.

Wa'el, M. M., Hamdy, N. A. and Radwan, E. (2009). Intrusion Detection Using Rough Set Parallel Genetic Programming Based Hybrid Model. *World Congress on Engeneering and Computer Science (WCECS).*

Walls, J. G., Widmeyer, G. R. and El Sawy, O. A. (2004). Assessing information system design theory in perspective: How useful was our 1992 initial rendition. *Journal of Information Technology Theory and Application.* 6 (2), 43-58.

Wang, L., Liu, A. and Jajodia, S. (2006). Using attack graphs for correlating, hypothesizing, and predicting intrusion alerts. *Computer Communications.* 29 (15), 2917-2933.

Ware, W. H. (1967). Security and privacy in computer systems. *Proceedings of the Spring Joint Computer Conference.* April 18-20, 279-282.

Weijters, T., Van Den Herik, H. J., Van Den Bosch, A. and Postma, E. O. (1997). Avoiding overfitting with BP-SOM. *IJCAI.* 1140-1145.

Wierzchoń, S. and Kużelewska, U. (2002). Stable clusters formation in an artificial immune system. *First International Conference on Artificial Immune Systems (ICARIS).*

Williams, P. D., Anchor, K. P., Bebo, J. L., Gunsch, G. H. and Lamont, G. D. (2001). CDIS: Towards a computer immune system for detecting network intrusions. *Recent Advances in Intrusion Detection*. 117-133.

Wood, A. D. and Stankovic, J. A. (2004). A taxonomy for denial-of-service attacks in wireless sensor networks. *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems.*739-763.

Wu, S. X. and Banzhaf, W. (2010). The use of computational intelligence in intrusion detection systems: A review. *Applied Soft Computing.* 10 (1), 1-35.

Xiao, F., Jin, S. and Li, X. (2010). A Novel Data Mining-Based Method for Alert Reduction and Analysis. *Journal of Networks.* 5 (1).

Xiao, R.-B., Wang, L. and Liu, Y. (2002). A framework of AIS based pattern classification and matching for engineering creative design. *International Conference on Machine Learning and Cybernetics*. 1554-1558.

Xiao, X. (2013). An Artificial Immune Based Incremental Data Clustering Algorithm. *Applied Mechanics and Materials.* 256, 2935-2938.

Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *Ieee Transactions on Neural Networks.* 16 (3), 645-678.

Yeung, D.-Y. and Ding, Y. (2003). Host-based intrusion detection using dynamic and static behavioral models. *Pattern recognition.* 36 (1), 229-243.

Yue, X., Mo, H. and Chi, Z.-X. (2008). Immune-inspired incremental feature selection technology to data streams. *Applied Soft Computing.* 8 (2), 1041-1049.

Yurcik, W. (2002). Controlling intrusion detection systems by generating false positives: squealing proof-of-concept. *27th Annual IEEE Conference on Local Computer Networks (LCN 2002).* 134-135.

Zali, Z., Hashemi, M. R. and Saidi, H. (2013). Real-Time Intrusion Detection Alert Correlation and Attack Scenario Extraction Based on the Prerequisite-Consequence Approach. *The ISC International Journal of Information Security.* 4 (2).

Zhang, Q. L., Hu, G. Z. and Feng, W. Y. (2010). Design and Performance Evaluation of a Machine Learning-Based Method for Intrusion Detection. *Software Engineering, Artificial Intelligence, Networking and Parallel-Distributed Computing.* 295, 69-83.

Zhang, T., Ramakrishnan, R. and Livny, M. (1997). BIRCH: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery.* 1 (2), 141-182.

Zhihong, T., Baoshan, Q., Jianwei, Y. and Hongli, Z. (2008). Alertclu: A realtime alert aggregation and correlation system. *International Conference on Cyberworlds*. 778-781.

Zhou, C. V., Leckie, C. and Karunasekera, S. (2009). Decentralized multi-dimensional alert correlation for collaborative intrusion detection. *Journal of Network and Computer Applications.* 32 (5), 1106-1123.

Zurutuza, U. and Uribeetxeberria, R. (2004). Intrusion detection alarm correlation: a survey. *IADAT International Conference on Telecommunications and Computer Networks*.1-3.