AN ENHANCEMENT OF ALERT CORRELATION RULE SET FOR MALWARE ALARM REDUCTION

ABSTRACT

Security issue is becoming one of the focal points of research because of the rising popularity of internet and most organizations solely depend on their internet services in day-to-day business management. However, malware activity has increasingly threatened network security. Current detection techniques by Intrusion Detection System (IDS) are still not comprehensive enough as IDS can generate huge amounts of data. It can exhaust administrator in terms of time and implicate costs due to its inability to eliminate the false alarms which lead to alert report failures. In intrusion detection field, alert correlation can increase the alert detection rate. However, it is still incapable of reducing the amount of false alarm. Hence, this research is to enhance the process of reducing the false alarm by primarily analyzing and classifying the existing malware detection technique and generate an improved malware detection technique taxonomy integrated with alert correlation technique. This improved taxonomy acts as a primary guideline in developing an improved Alert Correlation Rule Set (**ACRS**) using specification-based technique that specifically deploys pre-requisites and consequences of individual attack's technique. This research introduces generic attack pattern which is used to construct the multi-step attack model. Both generic attack pattern and multi-step attack model are then applied in formulating the **ACRS**. The **ACRS** is implemented on an improved alert correlation framework and it is evaluated and validated to verify its effectiveness in identifying and reducing the amount of false alarm generated by IDS and other administrative logs. In evaluation and validation, the results show that **ACRS** has effectively obtained high alarm correlation rate, high alarm reduction rate, high identification rate, low misclassification rate and high total alarm reduction rate by 99.97%. Using statistical method for validation, high significance results are obtained in correlation alarm analysis, alarm reduction analysis and identification rate analysis. This **ACRS** helps the network administrator to identify the perspectives of the attack: *attacker, victim or multi-step attacker (victim/attacker).* This research can also be extended into research areas in alert correlation and computer forensic investigation.

CHAPTER 1

INTRODUCTION

1 Introduction

Internet is considered as one of the important communication services. Thus, companies have increasingly put critical resources online for effective business management. This has given rise to activities of cyber criminals which are related to malicious software (malware) as mentioned by Lee *et.al*. (2004), Andreas *et. al.* (2007) and Chenfeng Vincent *et. al.* (2009b). A very large volumes of malware can also be found with extreme variety and sophisticated features as reported by (GTISC, 2011) .

There are three main objectives in computer network security which are confidentiality, integrity and availability (Sundaram, 1996). Virtually, all organizations face increase threats to their networks and the services that they provide and this will lead to network security issues as mentioned by Chiu *et. al.* (2010) and Xu *et. al.* (2010). This statement has been proven by the increasing number of computer security incidents related to vulnerabilities from 171 in 1995 to 7,236 in 2007 and 6,058 in Q3, 2008 as reported by Computer Emergency Response Team (CERT/CC, 2009). CSI (2009) has claimed that the malware infections are still at the highest rank which is at 64.3% compared to others security incidents. Meanwhile, CyberSecurity Malaysia (MyCert, 2010) has also reported that the malicious code incident has the third highest percentage of incidents which is at 11%. The vulnerability report by CERT/CC has stated that malware attack has generated significant worldwide epidemic to network security environment and also resulted in huge financial loss. This is due to the motivation of malware authors continuously shifting from a general desire to inflict damage, to the intention of gaining financial benefits through theft of personal information such as credit card data or access of financial accounts.

According to CSI (2009), financial fraud was a major concern and it costs enterprises approximately US$ 450,000 per incident. Figure 1.1 shows the financial impacts due to virus attack from 1995 to 2006 as reported by Computer Economics (2007) and the financial impacts are still above the figures of 2005 and 2006 as reported by CSI (2009).

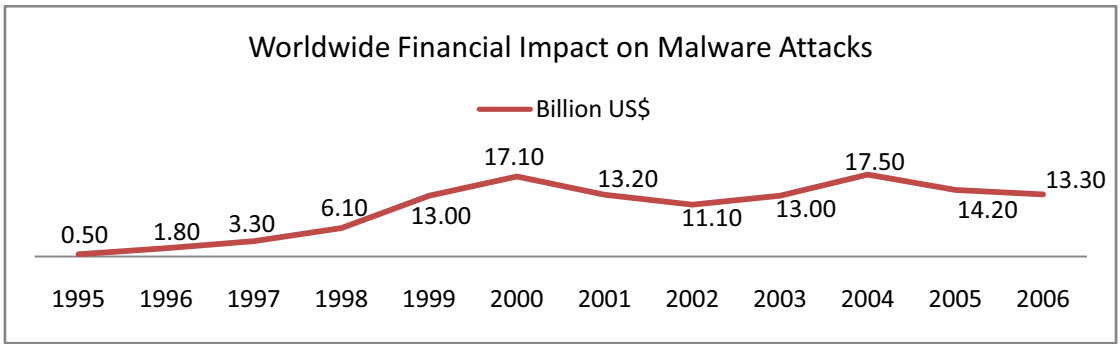

Figure 1.1 Financial Impact on Malware Attacks by (Computer Economics, 2007)

Referring to Figure 1.1, malware attacks are related to direct cost, for example cost in cleaning up malware infections. The figure shows that damages have been declining worldwide over the past two years since 2005. In 2006, direct cost damages were US$ 13.3 billion, down from US$ 14.2 billion in 2005 and from US$ 17.5 billion in 2004. However, this is not positive news as the indirect impact or secondary damages could be enormous due to, for example a hacker using a piece of spyware sniffing passwords and gained access to the corporate network. This shall cause multiple costs compared to cleaning up the malware infections. This fact has motivated research communities to do further research on malware issue.

It is important to build a security mechanism which is designed to prevent intrusion from hackers into system resources and data in enabling us to detect and research on intruders so that we can take action and improve the computer system's security (Sundaram, 1996). This research field is known as intrusion detection where malware detector tool is identified as one of the intrusion detection tool. Intrusion Detection System (IDS) is a system that continually monitors the dynamic behavior of the computer system to warn against actions that compromise the integrity, security and availability of any resource in the system (Razak *et. al.*, 2002) and malware detector tool is a complement of IDS (Benjamin *et. al.*, 2009).

1.1 Research Problem

Malware is considered as worldwide epidemic due to the malware author's activity in achieving financial gain through theft of personal information such as gaining access to financial accounts. Hence, this kind of activity can be captured by the wide deployment of IDSs and it can also process large amount of traffic which can generate a huge amount of data as stated by Ning and Xu (2003), Barford *et. al.* (2004), Abdulrahman and Hideki

(2005), Peyman and Ali (2007), Tjhai *et. al.* (2008), Benjamin *et. al.* (2009), Maggi *et. al.* (2009), Al-Mamory and Zhang (2010), Guenther *et. al.* (2010), Katipally *et. al.* (2010) and Massimo and Luigi (2010). However, this huge amount of data can also exhaust the network administrator's time and implicate cost as mentioned by Thonnarda and Dacier (2008), Nehinbe (2009a), Morin *et. al.* (2009), Sadoddin and Ghorbani (2009) and Ahmadinejad *et. al.* (2011). The data can also be used to find the intruder if new outbreak attack happens, especially involving malware attack.

In the real environment, IDS has problems in detecting intruder's accurate manner due to its inability to distinguish between the false positive alarm and false negative alarm (Song *et. al.*, 2011). An important problem in the field of intrusion detection is the management of alerts (Moon & Kyeong, 2006) as IDS tends to produce high number of false positive alerts as stated by Emmanuel (2006), Tjhai *et. al.* (2010) and Subbulakshmi *et. al.* (2010a).

Reducing false alarms is a serious problem in ensuring IDS efficiency and usability as mentioned by Tjhai *et. al.* (2008) and Lin *et. al.* (2009a). In order to increase the detection rate, the use of multiple IDSs can be used to correlate the alert, but in return, it increases the number of alerts to process (Autrel & Cuppens, 2005). The increment of the traffic volume can cause the IDS to produce large number of alarms as discussed by Ren *et. al.* (2010). These alarms shall consist of the false positive alarm as stated by Cheung *et. al.* (2003), Al-Mamory and Zhang (2009), Meharouech *et. al.* (2009) and Nashat *et. al.* (2010); and false negative alarms can overwhelm the network security officer as mentioned by Nehinbe (2009b) and Subbulakshmi *et. al.* (2010b). In order to reduce this alarm, threshold for intrusion detection is raised; however this can reduce the overall detection rate. Due to this reason, some researchers have attempted to develop a new generation of IDS that correlate information from several IDS and other heterogeneous devices that can generate *alert log* to identify intrusions as suggested by Lazarevic *et. al.* (2005). Moreover, Tian *et. al.* (2009) and M. M. Siraj *et. al.* (2009) had stated that alert correlation is the promising technique in intrusion detection.

Hence, this research has developed improved alert correlation rule set based on the generic attack pattern and attack model. This improved alert correlation rule sets formulation is integrated with the proposed improved generic taxonomy of malware detection technique. This rule set is to improve the process of identifying the intruder by reducing the number of false alarm in IDS

Therefore the statements of this research problem are:

Malware is considered as an epidemic. Current detection techniques by any IDS are still not comprehensive enough and IDS can generate huge amount of data that can exhaust administrator in terms of time and implicate cost due to its inability to eliminate the false alarm which lead to alert report failure. In intrusion detection field, alert correlation can increase the alert detection rate; however it is still incapable to reduce the amount of false alarm.

The Research Problem (RP) is divided into three sub-problems and the summary of the above statements are illustrated in Table 1.1.

Table 1.1 Summary of research problems

No	Research Problem
RP1	Malware is an epidemic and current detection techniques are still not comprehensive enough.
RP2	IDS generates huge amount of data that can exhaust admin in terms of time and implicate cost due to its inability to eliminate the false alarm which lead to alert report failure.
RP3	In intrusion detection field, alert correlation can increase the alert detection rate; however, it is still incapable in reducing the amount of false positive alarm.

Thus, the existing malware detection technique, alert correlation technique and alert correlation rule set is further investigated to reduce the number of false alarm generated by IDS.

1.2 Research Questions

Four Research Questions (RQ) are constructed to identify the research problem as discussed in previous section.

RQ1: How can we improve the malware detection technique?

This research question is formulated by considering the malware detection technique's issue which is still not comprehensive enough as highlighted in RP1 in Table 1.1.

RQ2: How can we reduce the false alarm?

The huge amount of alarm generated by IDS stated in RP2 in Table 1.1 has initiated the investigation of the method used to reduce the false alarm in IDS.

RQ3: How can we improve the alert detection's false alarm?

The alert correlation problem related with false alarms issue highlighted in RP3 in Table 1.1 has motivated a further research on improving the alert detection's false alarm in IDS.

RQ4: How can we measure the improved alert correlation rule set to prove the improvement of its false alarm detection?

This research question is intended to evaluate the effectiveness of the proposed solutions in RP1, RP2 and RP3.

The summary of the research questions for the research problems is depicted in Table 1.2.

Table 1.2 Summary of research questions

RP	RQ	Research Questions
RP1	**RQ1**	How can we improve the malware detection technique?
RP2	**RQ2**	How can we reduce the false alarm?
RP3	**RQ3**	How can we improve the alert detection's false alarm?
RP1, RP2, RP3	**RQ4**	How can we measure the improved alert correlation rule set to prove the improvement of its false alarm detection?

These four research questions are the primary guides to formulate the research objectives of this research.

1.3 Research Aim and Objectives

Based on the research questions formulated in previous section, a research aim is developed as follows:

The aim of this research is to develop an improved alert correlation rule set which runs on an improved alert correlation framework by using selected malware detection technique integrated with selected alert correlation technique to assist the network administrator on administrating network system. By doing this, it should help the network administrator to identify the intruder and reduce the amount of false alarm generated by the IDS.

Appropriate Research Objectives (RO) related to each research questions is created to achieve the research aim. There are five research objectives identified for this research which are listed as follows:

RO1: To analyze and classify the existing malware detection technique

This research objective intends to investigate the possibility of improving malware detection techniques issue by analyzing and classifying the existing malware detection technique. The current approach used in detecting malware is also investigated to improve the existing malware detection technique.

RO2: To generate an improved generic taxonomy for malware detection technique

Based on the analyzing and classifying task, an improved generic taxonomy for malware detection technique is generated to overcome the false alarm issue. The improved generic taxonomy for malware detection technique is then integrated with the current approach used in detecting malware which is alert correlation technique for malware's optimal detection.

RO3: To construct malware attack pattern and malware attack model based on the improved generic taxonomy for formulating improved alert correlation rule set

The improved generic taxonomy of malware detection technique is the main guideline in generating an improved alert correlation rule set. However, a preliminary experiment is needed to be implemented in order to identify a generic malware's attack pattern and a generic malware's attack model, which is then used as a primary guideline for formulating the improved alert correlation rule set.

RO4: To formulate the improved alert correlation rule set implemented on improved alert correlation framework

The improved alert correlation rule set will use the generic malware's attack pattern and generic malware's attack model as its main guideline for detecting the malware and eliminating false alarm and then implement it on the improved alert correlation framework to improve the false alarm generated by IDS.

RO5: To evaluate and validate the improved alert correlation rule set that can identify intruder and reduce the amount of false alarm generated by IDS

This improved alert correlation rule set is evaluated and validated for its effectiveness in identifying the intruder whilst reducing the amount of false alarm generated by IDS.

Table 1.3 Summary of research objectives

RP	RQ	RO	Research Objectives
RP1	**RQ1**	**RO1**	To analyze and classify the existing malware detection technique.
RP2	**RQ2**	**RO2**	To generate an improved generic taxonomy for malware detection technique.
RP3	**RQ3**	**RO3**	To construct malware attack pattern and malware attack model based on the improved generic taxonomy for formulating improved alert correlation rule set.
		RO4	To formulate the improved alert correlation rule set implemented on an improved alert correlation framework.
RP1, RP2, RP3	**RQ4**	**RO5**	To evaluate and validate the improved alert correlation rule set that can identify intruder and reduce the amount of false alarm generated by IDS.

The mapping summary of the five research objectives against the research questions and research problems are as shown in Table 1.3. These five research objectives are the main issues that shall be discussed in the next few chapters. The main objective of this research is to formulate an improved alert correlation rule set and in order to achieve this objective; a few research scopes are set.

1.4 Research Scope

The scope of this research will focus on some issues as stated below:

1. This research is implemented only on specific type of malware attack, which is the traditional worm attack as this type worm are still persistent in internet as claimed by Bailey *et. al* (2005) and IBM (2011).
2. Open source IDS which is Snort is used in this research since it has become the de facto standard of IDS as claimed by Sourcefire (2010) and it is an open source network intrusion prevention and detection system developed by Sourcefire. .
3. The administrative logs are collected from selected IDS and other selected devices which can produce administrative logs.
4. This research is using specification-based technique which specifically deploys pre-requisites and consequences of individual attack's technique to reduce the false alarm and identifying the perspective of the intrusion.

1.5 Research Contribution

Based on the research objectives, this research will contribute to some of the issues listed below and the summary of the research contribution is depicted in Table 1.5:

Table 1.5 Summary of research contributions

RP	RQ	RO	Research Contributions
RP1	**RQ1**	**RO1**	• Classification of existing malware detection technique. • Classification of alert correlation technique.
RP2	**RQ2**	**RO2**	• Proposed an improved generic taxonomy for malware detection technique. • Proposed an integration of malware detection technique with alert correlation technique for malware's effective detection.
RP3	**RQ3**	**RO3**	• Proposed a generic malware's attack pattern. • Proposed a generic malware's attack model (Basic malware's attack model and multi-step malware's attack model).
		RO4	• Proposed an improved alert correlation framework. • Proposed an improved alert correlation rule set for effective malware detection.
RP1, RP2, RP3	**RQ4**	**RO5**	• Formulate an improved alert correlation rule set that will reduce the false alarm of malware detection in IDS whilst identify the perspective of the attack.

1. Classification of the existing malware detection and alert correlation techniques.
2. Generation of an improved generic taxonomy for malware detection techniques integrated with alert correlation technique for malware's optimal detection.
3. Construction of the generic malware's attack pattern and generic malware's attack model. The generic malware's attack model consists of two attack models that are basic malware's attack model and multi-step malware's attack model. These findings are used as primary guidelines for formulating improved alert correlation rule set.

4. Formulation of improved alert correlation rule set for effective malware detection on an improved alert correlation framework.
5. Reduction of false alarm generated by IDS which can help the network administrator to detect and identify the intruder that generates the malware attack.