BORANG PENGESAHAN STATUS TESIS*

JUDUL: PACKET ANALYZER WITH MULTIPLE ARTIFICIAL NEURAL NETWORK BACK-PROPAGATION TRAINING

SESI PENGAJIAN: 2008

Saya LIM SIEW YUEN

(HURUF BESAR)

mengaku membenarkan tesis(PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan(/)

_______ SULIT (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub didalam AKTA RAHSIA RASMI 1972)

_______ TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

_______ TIDAK TERHAD

(TANDATANGAN PENULIS)

Alamat tetap: No 1, Lorong 3,
Jalan PE 6, Taman Paya Emas,
76450 Melaka.

Tarikh: 30/03/2008

(TANDTANGAN PENYELIA)

En. Nazrulazhar bin Hj Bahaman

Nama Penyelia

Tarikh: 2 - 5 - 08

CATATAN: *Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda(PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa

PACKET ANALYZER WITH MULTIPLE ARTIFICIAL NEURAL NETWORK BACK-PROPAGATION TRAINING

LIM SIEW YUEN

This report is submitted in partial fulfillment of the requirements for the
Bachelor of Computer Science (Computer Network)

FACULTY OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2008

DECLARATION

I hereby declare that this project report entitled
PACKET ANALYZER WITH MULTIPLE ARTIFICIAL NEURAL NETWORK
BACK-PROPAGATION TRAINING

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : ______________________ Date : 02/05/08

(LIM SIEW YUEN)

SUPERVISOR : ______________________ Date : 2-5-08

(EN. NAZRULAZHAR BIN HJ.BAHAMAN)

DEDICATION

Specially dedicated to my beloved parents, Mr. Lim See Juan and Mrs. Choo An Na; and my belove husband,Mr. Tan Wei Tak. Thank You for Your fully supported.

For my supervisor, En. NazrulAzhar bin Hj.Bahaman
(UteM)

And lastly to my entire friend who have encouraged, guided and inspired me throughout the journey for my project and also in my study.

ACKNOWLEDGEMENT

First and foremost, I would like to express my appreciation to Universiti Teknikal Malaysia Melaka (UTeM) for offering this courses, BITU 3973 Projek Sarjana Muda I (PSM I) and BITU 3983 Projek Sarjana Muda II (PSM II). The Projek Sarjana Muda (PSM) is compulsory for a UTeM student before being awarded the degree. The purposes of this PSM are to provide guidance and chances to students to develop a complete and individual project whereby students will encounter many new problems and challenges. Throughout this project, PSM will enhance the students' ability and skills in literature research, ability to analyse problems in various views and able to propose alternative solutions or models, ability to manage and utilize available resources in accomplishing the project and present the output effectively.

Besides, I would like to thank gracefully to Mr. NazrulAzhar Bin Hj.Bahaman, my supervisor who leads and guides me during this semester. The patience and generosity in guiding me through are much welcomed and appreciated. A note of thanks is especially dedicated to committees PSM on their kindness in organized briefing, seminar and talk to students who took this subject. Further more, I would like to thank all of the lecturers in UTeM that gives their cooperation during my PSM by providing me with the required needs. Without their cooperation, I would not be able to go through the phase smoothly.

Last but not least, I would like to thanks to my parents and my husband and also to all of my friends who are being kind and helpful throughout this special semester.

ABSTRACT

The packet analyzer with multiple artificial neural network back-propagation training is basically a system that integrated with the feed-forward back-propagation neural network to analyze the packet from the network. This system is proposed to overcome the intrusion that expands widely nowadays. This system using a system solution, which is the output of the training and testing process using MATLAB, to integrated with the simple system developed using Visual Basic 6.0. The system can automatically identified and categorizes the packet captured into two categories, that is either normal or misuse packet. Neural networks are widely considered as an effective approach to classify patterns. In this project, the packet analyzer system will receive data packet from the network and the neural network will classify it into normal packet or misuse packet. For any occurrence of the possibility misuse packet, the system will alert the system administrator. Since neural network have a high learning rate, it will recognize most of the characteristics of attacks packet and by this, it can recognize, and block the packet before the misuse packet can get through the network. The advantage of using neural networks over statistics resides in having a simple way to express nonlinear relationships between variables, and in learning about relationships automatically. With few exceptions, behavior of most other users is also predictable. Neural networks are still a computationally intensive technique, and are not widely used in the packet analyzer community.

ABSTRAK

"Packet Analyzer with Multiple Artificial Neural Network Back-Propagation Training" adalah satu sistem yang mengesan penggemar computer yang melakukan "hacking process" ke atas rangkaian komputer. Sistem ini diperkenalkan dengan tujuan untuk mengesan penggemar komputer ini dan mengurangkan aktiviti "hacking" ke atas rangkaian komputer. Sistem ini menggunakan sistm penyelesaian yang didapati daripada pelajaran dan ujian yang dilakukan dengan menggunakan MATLAB. Sistem ini akan automatic mengenal dan mengkategorikan "packet" sama ada dalam keadan normal atau tidak normal. Keadaan adalah tidak normal dimana penggemar komputer melakukan "hacking" kepada rangkaian komputer. Dalam sistem ini, "packet analyzer system" akan menerima "packet" dari rangkaian dan "neural network" yang digabungkan dengan sistem tersebut akan menkategorikannya kepada biasa atau tidak biasa. Untuk keadaan di mana packet yang tidak biasa memasuki rangkaian sesebuah organisasi, system tersebut akan memberitahu pentadbir organisasi tersebut. "Neural network" mempunyai kadar belajar yang tinggi, oleh itu, system tersebut akan mengecam ciri-ciri "packet" yang disyaki daripada penggemar komputer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Project Background

Due to a growing number of intrusion events and also because the Internet and local networks have become so ubiquitous, organizations are increasingly implementing various systems that monitor IT security breaches. Unfortunately, computer and network intrusions have become more common and more complicated, challenging the intrusion detection systems. As a consequence, the amount of data to be processed by an intrusion detection system has been growing, making it difficult to efficiently detect intrusions online. This condition leads most of the system administrators and information security researchers' goal to solve the problem by develops more and more effective intrusion detection system.

In an information system, intrusions are the activities that violate the security policy if the system, and intrusion detection is the process used to identify intrusions. Commonly, intrusion detection techniques composed of one of two methodologies which is anomaly detection or misuse detection. Anomaly detection is based on the normal behavior of a subject (E.G., a user or a system); any action that significantly deviates from the normal behavior is considered intrusive. While misuse detection catches intrusion in terms of the characteristics of known attacks or system vulnerabilities; any action that conforms to the pattern of a known attack or vulnerability is considered intrusive.

Intrusion Detection System (IDS) monitors network traffic and monitors for suspicious activity and alerts the system or network administrator. In some cases, the intrusion detection system may also respond to anomalous or malicious traffic by taking action such as blocking the user or source IP address from accessing the network. Intrusion detection system approaches the goal of detecting suspicious traffic in different ways. There are network based (NIDS) and host based (HIDS) intrusion detection systems. There is also intrusion detection systems that detect based on looking for specific signatures of known threats similar to the way antivirus software typically detects and protects against malware and there are intrusion detection system that based on comparing traffic patterns against a baseline and looking for anomalies.

Besides, intrusion detection system also differs in whether they are online or offline. Offline intrusion detection systems are run periodically and they detect intrusions after the fact based on system logs. While online systems are designed to detect intrusions while they are happening, thereby allowing for quicker intervention. Online intrusion detection systems are computationally very expensive because they require continuous monitoring. Decisions need to be made quickly with less data and therefore they are not as reliable.

Host based intrusion detection system (HIDS) can be used to determine if a system has been compromised and can warn administrator if that happens. There are four methods of host-based intrusion detection which includes file system monitoring (systems checking the integrity of files and directories), log file analysis (systems analyzing log files for patterns indicating suspicious activity), connection analysis (systems that monitor connection attempts to and from a host), and kernel-based intrusion detection (systems that detect malicious activity on a kernel level). Implementations of intrusion detection systems generally use one of the four methods to detect intrusions.

A network intrusion detection system (NIDS) is an intrusion detection system that tries to detect malicious activity such as denial of service attacks; port scans or even attempts to crack into computers by monitoring network traffic. The network intrusion detection system does this by reading all the incoming packets and trying to find suspicious patterns. Besides, a network intrusion detection system is not limited to inspecting incoming network traffic only. Network intrusion detection system work well with other system, such as update some firewall's blacklist with the IP address of computers used by suspected crackers.

The artificial neural networks provide a number of advantages in the detection of network intrusions. The application of the neural network techniques has been considered for both the misuse detection model and the anomaly detection model. An increasing amount of research has been conducted on the application of neural networks for detecting network intrusions. The artificial neural networks have the potential to resolve a number of problems encountered by the other current approaches in intrusion detection. The neural networks gain experience by training the system to correctly identify the pre-selected examples of the problem.

1.2 Problem Statements

A computer system should provide confidentiality, integrity and assurance against denial of service. However, due to increased connectivity (especially on the Internet), and the vast spectrum of financial possibilities that are opening up, more and more systems are subject to attack by intruders. These subversion attempts try to exploit flaws in the operating system as well as in application programs and have resulted in spectacular incidents like the Internet Worm incident of 1988.

However, intrusions can be divided into the main six types, which composed of attempted break-ins, which are detected by atypical behavior profiles or violations of security constraints; masquerade attacks, which are detected by atypical behavior profiles or violations of security constraints; penetration of the security control system, which are detected by monitoring for specific patterns of activity; leakage, which is detected by atypical use of system resources; denial of service, which is detected by atypical use of resources; and last but not least, malicious use, which is detected by atypical behavior profiles, violations of security constraints, or use of special privileges.

Host-based intrusion detection system can analyze activities on the host it monitors at a high level of detail, and it can often determine which processes and users are involved in malicious activities. Though they may focus on a single host, many host-based intrusion detection systems use an agent-console model where agents run on individual hosts but report to a single centralized console so that a single console can configure, manage, and consolidate data from numerous hosts. Host based intrusion detection system can detect attacks undetectable to the network based intrusion detection system and can gauge attack effects quite accurately. Host based intrusion detection system can use host based encryption services to examine encrypted traffic, data, storage, and activity. Last but not lease, host based intrusion detection system have no difficulties operating on switch based networks. On the other hand, there are some disadvantages of host based intrusion detection system which is data collection occurs on a per-host basis, writing to logs or reporting activity requires network traffic and this can decrease network performance. In addition, a host based intrusion detection system does consume processing time, storage, memory, and other resources on the hosts where such systems operate.

With the growth of computer networking, electronic commerce, and web services, security of networking systems has become very important. Many companies now rely on web services as a major source of revenue. Computer hacking poses significant problems to these companies, as distributed attacks can render their cyber-storefront inoperable for long periods of time. As a result, intrusion detection system is devoted to detecting this activity by detecting this attack by a careful analysis of network data.

1.3 Objective

Project objective define target status at the end of the project, reaching of which is considered necessary for the achievement of planned benefits. There are several objectives in order to develop this project and this including: -

- ✓ To provide an accurate, extendable and adjustable to make sure the security of the computer system are confidentiality, integrity and able to prevent intrusion or attacking from the intruders.
- ✓ To propose and investigate a neural network based packet analyzer so that it can analyze the packet whether it is misuse or normal packet.
- ✓ To uses a feed-forward back-propagation neural network to learn users' behavior.
- ✓ To train and test the data by using feed-forward back-propagation neural network.

1.4 Scope

Project scope is the sum total of all projects products and their requirements or features. This project's scope including: -

- ✓ Developed the mathematical function using the software called Math Works MATLAB to create an algorithm which will train the data to be use in packet analyzer with multiple artificial neural network back-propagation training.

- ✓ Classified data from DARPA into two categories which is normal packet and misuse packet and train both of the packet.

- ✓ Convert the artificial neural network algorithm in MATLAB to VB language so that it can integrate in the intrusion detection system that have been created using VB so that the system can identify the normal and misuse packet.

1.5 Project Significance

This project is significance in nowadays computer's world because of the security incidents that occur over the network. These attacks may consist of otherwise authorized users who are disgruntled employees. The remainder comes from the outside, in the form of denial of service attacks or attempts to penetrate a network infrastructure. Intrusion detection systems remain the only proactive means of detecting and responding to threats that stem from both inside and outside a corporate network.

The aim of packet analyzer using multiple artificial neural network back-propagation training is to integrate the neural network techniques in the intrusion detection system. As a result, the system can automatically identified and categorizes the packet captured into two categories, that is either normal or misuse packet. Neural networks are widely considered as an effective approach to classify patterns. In this project, the intrusion detection system will receive data packet from the network and the neural network will classify it into normal packet or misuse packet. For any occurrence of the possibility misuse packet, the system will alert the system administrator. Since neural network have a high learning rate, it will recognize most of the characteristics of attacks packet and by this, it can recognize, and block the packet before the misuse packet can get through the network. The advantage of using neural networks over statistics resides in having a simple way to express nonlinear relationships between variables, and in learning about relationships automatically. With few exceptions, behavior of most other users is also predictable. Neural networks are still a computationally intensive technique, and are not widely used in the intrusion detection community.

1.6 Expected Output

This project is design to suite the requirements in the intrusion detection system that needed by most of the organizations and corporate since the intrusion scenarios becomes more and more strictness. The packet analyzer using multiple artificial neural network back-propagation training will function in three steps which consist of packet capturing, packet preprocessing and intrusion or misuse detection.

First of all, third party software will be used to capture packets from network. In this project, Packet Sniffer and Wireshark will be used as the third party software. The TCP/IP streams of packet will be saving in a text file (.txt) and can be view using notepad or text editor.

The second step is preprocessing. Neural network training can be made more efficient by performing preprocessing steps on the network inputs and targets. Network-input processing functions transform inputs into a better form for the network use. Processing functions associated with a network output transform targets into a better form for network training, and reverse transformed outputs back to the characteristics of the original target data. In this project, the packet that has been captured will get through the preprocessing process and the packet will be analyzed to characterize the normal packet and the misuse packet. Normal packet is the packets with the normal data and information while misuse packet is the packet that containing suspicious intrusive information.

In order to prepare the data for training and testing process in neural network, the data will get through three levels of preprocessing. First, ten of the event record data elements will be selected in the data set. The elements are selected because its can supply a complete description of the information that transmitted by the packet. Next, a few elements in the network data packets will be converts into a standardized numeric representation. The data types and the sequential numbers that assigned to each element are categorized in the relational tables. Last, the results of the query will be converting into an ASCII comma delimited format that is needed in order to train the neural network.

The training process of artificial neural network will produced the output in binary form that is 0 and 1. The 0 represents the normal packets while 1 represents the misuse packets. The diverse connection weights were frozen and the network was interrogated after the training and testing process of the neural network. Sample data from DARPA with one normal packet and one misuse packet will be used to test the neural network. As a result of the testing process, a simulate graph will be shown regarding the normal and misuse network. For the result below 0.1, the packet is considered as normal and if the result is above 0.1, the packet will be consider as misuse.

After the test is prove correct, the solution will be translated into Visual Basic language. It will be merging with a system interface that developed using Visual Basic to deliver an alert system that can identified misuse in network.

1.7 Conclusion

Overall, this chapter is mainly explaining on the project background of developing an artificial neural network for intrusion detection system. The objectives, project scopes, project significance also covered in this chapter. The aim of this project is to develop a system using artificial neural network which can detect the attacking of intruders into the