

AN INTRUSION DETECTION SYSTEM (IDS) FOR INTERNET NETWORK

MOHAMAD FAUZAN BIN SAFIEE

A project report submitted in partial fulfilment of the
Requirements for the award of the degree of
Master of Engineering (Electrical - Electronic and Telecommunication)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

MAY 2007

To my beloved wife, mum and all my siblings
thanks a lot for your patience and prayer for my success.

Al-Fatihah to my late father.

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my thesis supervisor, Dr. Sharifah Hafizah bt. Syed Ariffin, for encouragement, guidance, critics and friendship. I am also very thankful to Assoc. Prof. Liza bt. Abdul Latif and my friend Br. Muhammad Sadry bin Abu Seman for their guidance, advices and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

I am also indebted to Ministry of Defence (MINDEF) for funding my Master study. Librarians at UTM and MINDEF also deserve special thanks for their assistance in supplying the relevant literatures.

My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members especially my beloved wife.

ABSTRACT

An Intrusion Detection System (IDS) is detects and blocks unwanted attacks to the civilian or the military systems. These attacks can be an internal attack or external attack. The traffic or the normal networks is heterogeneous where else the military network has more homogeneous traffic. Even though the internet security includes firewall and other system security it usually failed to filter out the unwanted attack to the system and allows system breakdown and system failures. The major problems in developing the IDS method is the evolving growth of the internet topology and the growth of the internets users which makes the modeling of the network with attack free data is difficult. Real world test has shown overwhelming numbers of false alarms of attack and little success in filtering them out. This project is to analysis the network with data free attacks in a simulator that involved self-similar traffic that ideally represents the internet traffic modeling as well as the Poisson traffic modeling for the non peak hours periods. With templates of data free attacks a system will reduce the complexity in detecting the attacks during peak hours and non peak hours. The network system was simulated in NS-2 simulator.

ABSTRAK

Sistem Pengesanan Pencerobohan (SPP) ini mengesan dan menyekat serangan yang tidak diingini kepada sistem orang awam atau tentera. Serangan ini boleh dilaksanakan samada serangan secara dalaman atau luaran. Bagi aliran trafik rangkaian biasa, ia adalah berbentuk pelbagai jenis yang mana rangkaian tentera aliran trafiknya lebih berbentuk sama jenis. Walaupun ia dilengkapi dengan sistem keselamatan internet termasuk 'firewall' dan sistem keselamatan yang lain, yang mana kebiasaannya ia gagal untuk menapis serangan yang tidak dikehendaki kepada sistem dan menyebabkan sistem terganggu dan gagal beroperasi. Masalah utama dalam membangunkan kaedah bagi sistem ini ialah perkembangan dalam topologi internet dan juga pengguna internet yang mana permodelan rangkaian ini lebih sukar untuk serangan data bebas. Ujian sebenar yang dilaksanakan ia menyatakan bahawa masalah ini berpunca kerana banyak serangan-serangan palsu dilakukan ke atas sistem ini dan ia tidak berupaya untuk menapis semua serangan tersebut. Projek ini bertujuan menganalisis rangkaian dalam simulator dengan serangan data bebas yang mana melibatkan trafik 'self-similar' yang menunjukkan permodelan trafik internet dan sebaik-baiknya permodelan trafik 'Poisson' digunakan bagi luar jangka waktu puncak. Dengan templet serangan data bebas, sistem akan mengurangkan kesulitan dalam mengesan serangan semasa waktu puncak dan luar waktu puncak. Sistem rangkaian ini disimulasikan dalam perisian 'NS-2 Simulator'.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENTS	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	xi
	LIST OF FIGURES	xii
	LIST OF ABBREVIATIONS	xiv
	LIST OF SYMBOLS	xvi
1	INTRODUCTION	1
	1.1 Overview	1
	1.2 Problem Statement	2
	1.3 Objective	2
	1.4 Scope of Work	3
	1.5 Thesis Organization	3
2	LITERATURE REVIEW	4
	2.1 Background	4
	2.2 Intrusion Detection System (IDS)	5
	2.2.1 Types of intrusion detection systems	6
	2.2.1.1 Host-based intrusion detection systems (HIDS)	7

2.2.1.2	Network-based intrusion detection systems (NIDS)	8
2.2.2	Methods and modes of intrusion detection	10
2.2.2.1	Anomaly detection	10
2.2.2.2	Misuse detection or pattern matching	12
2.2.3	Detection issues	12
2.2.4	Responses to Intrusion Detection	14
2.2.5	Common Attacks	15
2.3	Internet Network	16
2.3.1	Differentiated Services (DiffServ)	18
2.3.2	DiffServ Vulnerabilities	21
2.4	Internet Traffic	24
2.5	Poisson Traffic	27
2.5.1	Poisson Law	27
2.5.2	Poisson Process	29
2.5.3	Traffic Analysis	32
2.6	Self-similar Traffic	33
2.6.1	Self-similarity	33
2.6.2	Stochastic self-similarity and network traffic	36
2.6.3	Traffic Research	38
2.6.3.1	Measurement-based traffic modeling	38
2.6.3.2	Physical modeling	39
2.6.3.3	Queueing analysis	42
2.6.3.4	Traffic control and resource provisioning	44
2.6.4	Issues and Remarks	46
2.6.4.1	Traffic measurement and estimation	46
2.6.4.2	Traffic modeling	48
2.6.4.3	Performance analysis and traffic control	50
2.7	Literature Finding	53

3	METHODOLOGY	55
3.1	Simulation and Analysis	55
3.1.1	Simulation	55
3.1.2	Analysis	56
3.1.2.1	Analysis Process	56
3.1.2.2	1st Level Analysis	58
3.1.2.3	2nd Level Analysis	59
3.1.2.4	Response	60
3.2	Sequential Discrete Event Simulation	62
3.2.1	Simulation	62
3.2.2	Event	62
3.2.3	Simulation Time	63
3.2.4	Random Variables	64
3.2.5	Event Relations	64
3.3	Statistical Anomaly Detection	65
3.3.1	The Concept	65
3.3.2	Justification For Using the NIDES	67
3.4	The NS-2 Simulator	70
4	SIMULATION AND ANALYSIS	72
4.1	Simulation	72
4.1.1	DiffServ Network Topology Setup	72
4.1.2	QoS Application Traffic Setup	77
4.1.3	Background Traffic Setup	77
4.1.4	Sensors Setup	77
4.1.5	Attacks Setup	78
4.2	Validation of DiffServ and Traffic Simulation	78
4.3	Result of Simulation	80
4.4	Analysis of Simulation	88

5	DISCUSSION AND CONCLUSION	90
	5.1 Discussion	90
	5.2 Recommendation and Future work	90
	5.3 Conclusion	91
	REFERENCES	92

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Likelihood, Impact and Difficulty-of-detection for Attacks	23
4.1	NS-2 Traffic Service Differentiation	80
4.2	Average packets rate and peak packet received at Node 5 (Sink).	87

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	A centralized IDS.	8
2.2	Distributed IDS.	9
2.3	Detection issues in IDSs.	13
2.4	Relation between hosts on LANs and the subnet.	17
2.5	A Simplified DiffServ Architecture.	18
2.6	DiffServ Classification and Conditioning/Policing at Ingress.	19
2.7	Differentially Serving Scheduler	19
2.8a	The histogram of the Poisson distribution ($\lambda = 5$).	28
2.8b	Smoothed shape of the Poisson distribution for different parameter values.	28
2.9	2-dimensional Cantor set.	34
2.10	Left: 1-dimensional Cantor set interpreted as on/off traffic. Middle: 1-dimensional non uniform Cantor set with weights $\alpha_L = 2/3$, $\alpha_R = 1/3$. Right: Cumulative process corresponding to 1-dimensional on/off Cantor traffic.	35
2.11	Stochastic self-similarity - in the ‘burstiness preservation sense’ - across time scales 100s, 10s, 1s, 100ms (top-left, top-right, bottom-left, bottom-right).	37
2.12	Mean queue length as a function of buffer capacity for input traffic with varying long-range dependence ($\alpha = 1.05, 1.35, 1.65, 1.95$).	51

2.13	Performance gain of TCP Reno, Vegas, Rate, when endowed with multiple time scale capabilities as a function of RTT.	52
3.1	IDS Data Analyses	57
3.2	Algorithm of discrete-event simulator	63
4.1	Network Topology for the Simulations	73
4.2	Network topology for data free attack (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	74-75
4.3	Network topology for with attack data (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	75-76
4.4	CBR Traffic with 1000, 100 and 10 Second Time Scales	77
4.5	Result data free attack for self-similar traffic (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	81-82
4.6	Result data free attack for Poisson traffic (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	82-83
4.7	Result with attack data for self-similar traffic (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	84-85
4.8	Result with attack data for Poisson traffic (a) Light network (10 nodes) (b) Medium network (20 nodes) (c) Heavy network (40 nodes)	85-86

LIST OF ABBREVIATIONS

ACK	-	acknowledgement
AF	-	Assured Forwarding
ATM	-	Asynchronous Transfer Mode
BE	-	Best-effort
CBR	-	Constant Bit Rate
CBS	-	Committed Burst Size
CIR	-	Committed Information Rate
CPU	-	central processing unit
DiffServ	-	Differentiated Services
DES	-	Discrete Event Simulation
DNS	-	Domain Name Server
DoS	-	Denial of Service
DSCP	-	Differentiated Service Code Point
EF	-	Expedited Forwarding
HIDS	-	host-based intrusion detection system
IDS	-	Intrusion Detection System
IntServ	-	Integrated Service
IP	-	Internet protocol
ISP	-	Internet service provider
FARIMA	-	(fractional) autoregressive integrated moving average
FTP	-	File transfer protocol
GUI	-	graphical user interface
HTTP	-	Hypertext transfer protocol
LAN	-	Local Area Network
MPEG4	-	
NIDES/STAT	-	statistical anomaly detection engine/algorithm

NIDS	-	network-based intrusion detection system
NNTP	-	network news transfer protocol
NS-2	-	Network Simulator, version 2
PDF	-	probability distribution function
PHB	-	Per Hop Behavior
QoS	-	Quality of Service
RED	-	Random Early Dropping
RFC	-	Request For Comment
RNG	-	random number generators
RTT	-	round-trip times
RULE	-	Rule-based detection engine
SLA	-	Service Level Agreement
SMTP	-	simple mail transfer protocol
TBF	-	Token Bucket Filter
TCA	-	Traffic Conditioning Agreement
TCP	-	transmission control protocol
UDP	-	user datagram protocol
URL	-	uniform resource location
VLL	-	Virtual Leased Line
VoIP	-	Voice over Internet protocol
WAN	-	Wide Area Network
WRR	-	Weighted Round Robin
WWW	-	World Wide Web

LIST OF SYMBOLS

Gbps	-	Giga bits per second
KB	-	Kilobyte
Kbps	-	Kilobit per second
Mbps	-	Megabit per second

CHAPTER 1

INTRODUCTION

1.1 Overview

In the context of physical security, intrusion detection systems mean tools used to detect activity on the boundaries of a protected facility. When we commit to physically protecting the premises on which our staff work and which house our information processing equipment, we should carry out an exhaustive risk analysis and, where the threat requires, consider installing a perimeter intrusion detection system (IDS).

The simplest IDS are a guard patrol. Guards who walk on the corridors and perimeter of a facility are very effective at identifying attempts of break-in on the premises. If anything goes wrong, they will either raise the alarm or attempt to challenge the intruder. Of course, the most obvious shortcoming of a guard patrol is that the patrol cannot be at all points of the facility at the same time.

This leads to the next simplest IDS and that is video monitoring. Video camera can be place at locations in the facility where all points in the perimeter can be monitored simultaneously. If there is an intrusion attempt it will be detected and the alarm will be raise by the person in charged with monitoring the video an alarm.

IDS are designed to function like a burglar alarm on your house where these systems should record suspicious activity against the target system or network, and should alert the information security manager or support staff when an electronic

break-in is underway. The biggest downfall with IDS products is the necessary level of customization 'of the box'. Without significant amounts of customization, the IDS will produce a large number of false-positive alerts. A false positive is created when the IDS alerts the support staff to an event that will not have an impact on the target system. For example, a Code Red attack against an Apache Web server will not work, but the IDS may still sound the alarm.

1.2 Problem Statement

The major problems in developing the IDS method is the evolving growth of the internet topology and the growth of the internet users which makes the modelling of the network with attack free data is difficult. Real world test has shown overwhelming numbers of false alarms of attack and little success in filtering them out.

1.3 Objective

Objective of this thesis are:

- (a) To analysis the network with data free attacks in that involved the complex Internet Traffic as well as data traffic for the peak hours and non peak hours periods for the peak packet and average packet received.
- (b) To analysis the network with attacks in that involved the complex Internet Traffic as well as data traffic for the peak hours and non peak hours periods for the peak packet and average packet received.

1.4 Scopes of Work

The scopes of this project consist of:

- (a) Discussion about the concept and application of intrusion detection system (IDS).
- (b) Study on self-similar traffic which represent the Internet traffic source and Poisson traffic modelling which represent voice traffic in the analysis.
- (c) Simulate the attacks during peak hours and non peak hours in Network Simulator, version 2 (NS-2) simulator.
- (d) Analyses the result to determine the performance of Internet network and propose the proactive action to solve the weakness that going happened.

1.5 Thesis Organization

This thesis has the following structure. Chapter 2 give some literature review background information. Chapter 3 discusses the methodology simulation and analysis that to be used. Then, Chapter 4 explains the simulation and analysis. Finally, Chapter 5 concludes the thesis and gives possible directions for future research.