

**ANALYZING PATTERN MATCHING ALGORITHMS APPLIED ON SNORT
INTRUSION DETECTION SYSTEM**

ABDIFATAH ABDIRAHMAN ABDULLAHI

A project submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Computer Science (Information Security)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

JANUARY 2013

This project is dedicated to my lovely mother who always prayer
that has granted me to witness the end of my successful journey in
the study, and to all family for their endless support and
encouragement

ACKNOWLEDGEMENT

Firstly , I would Like to use this opportunity to express my sincere gratitude to my supervisors in persons of **Professor Aizani Maaruf Dean of Computer Science and Information System** and **Assoc. Professor Dr. Suleiman Ibrahim Head of Information Security Unit UTM CIT Johor Bahru Skudai Campus** for their countless and constant support and guidance during my project work, they stimulated me greatly to work in this project, I also like to express my sincere gratitude to the course coordinator I am also thankful to Dr. Anazida for her guidance, advices and motivation.

Besides, I would like to express my sincere gratitude to my parents, my mother **Hawa Mohamed Ali** and my father **Abdirahman Abdullahi** for their constant courage, support and prayers which have no doubt grantee my success in my studies. Also I acknowledge the effort of my friends, colleagues and others who have provided assistance at various occasions during my studies. Their views and tips are useful indeed. Unfortunately, it is not impossible to list all of them in this limited space. Finally, I am grateful to all my family members.

ABSTRACT

Currently, intrusion detection system has become widely used as a network perimeter security. The used of IDS to prevent the extremely sophisticated attacks in most of our industries, governmental organization and educational institutions .However ,Intrusion detection system can be either host-based or network based intrusion detection system, in a host-base intrusion it monitors the host where its configured while the network-based IDS it monitors both inbound and outbound traffic network. Furthermore, signature based or anomaly based detection techniques are used to detect malicious packets or attack in both network and host-based intrusion detection systems. Therefore, the challenges faced by most of the signature based detection systems like Snort tool is incapability to detect malicious traffic at higher traffic network, which resulted in a packet drooping and subjected the network where this signature based system is configured as a network perimeter security. The challenges resulted as a result of inefficiency of the pattern matching algorithms to efficiently perform pattern matching. Moreover, this project research work aim to compare the current Boyer-Moore pattern matching algorithm applied by the snort IDS with the Quick Search pattern matching algorithm in order to evaluate their performance and recommend for the implementation of the new pattern matching algorithm that will enhance snort detection performance

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	ABSTRACT	I
	TABLE OF CONTENTS	II
	LIST OF TABLES	VI
	LIST OF FIGURES	VII
1	INTRODUCTION	
	1.1 Overview	1
	1.2 Problem Background	3
	1.3 Statement of the Problem	4
	1.4 Research Goals	5
	1.5 Objectives of the Project	5
	1.6 Scope of the project	6
	1.7 Organization of the project	6
2	LITERATURE REVIEW	
	2.1 Introduction	7
	2.2 Understanding of the Intruder	7
	2.3 Intrusion	8
	2.3.1 Categories of Intrusion	8
	2.4 Intrusion Detection System	8
	2.4.1 Type of intrusion detection system	10
	2.4.2 Host-based intrusion detection system	10

	10	
2.4.3	Network-based intrusion detection system	11
2.4.4	Hybrid Intrusion detection system (HIDS)	12
2.4.5	Anomaly- based intrusion detection Verses Signature-based detection	13
2.4.6	Limitations of Anomaly detection systems	14
2.4.7	The limitations of Misuse detection system	15
2.4.8	Limitations of Intrusion Detection Systems	15
2.5	History of Snort	16
2.5.1	Snort Components	18
2.5.1.1	The Packet Decoder	19
2.5.1.2	Preprocessor	20
2.5.1.3	The Detection Engine	21
2.5.2	The limitation of Snort Performance under higher traffic network	23
2.5.3	An improved snort intrusion detection system based on self- similar traffic model	24
2.6	Deep Packet Inspection Challenges	25
2.6. 1	Deep Packet Inspection Challenges	26
2.7	Matching Algorithms	27
2.7.1	Enhancement of snort IDS	28
2.7.2	Overview of the pattern matching algorithms	29
2.8	Pattern Matching Algorithm and Detection Engine	32
2.8.1	Multiple String Matching Algorithms	32
2.8.2	Single Pattern Matching Algorithms	33
2.8.2.1	Boyer-Moore algorithm	34
2.8.2.2	Limitations of Boyer-Moore Algorithm	34
2.8.2.3	Quick search algorithm	35
2.8.2.4	Quick Search Algorithm Features	35
2.8.2.5	Description of Quick Search Algorithm	36
2.8.3	Hybrid algorithm for exact string matching algorithm	36
2.9	Summary	38

3 RESEARCH METHODOLOGY

3.1	Introduction	39
3.2	Research improvement phase	39
3.3	Data preparation	41
3.4	Testing Quick Search	41
3.5	Testing Boyer-Moore	42
3.6	Comparative Analysis	43
3.7	Tools required	43
3.8	Summery	43

4 DATA PREPARATION AND IMPLEMENTATION ALGORITHMS

4.1	Introduction	44
4.2	Phase 1: Data Preparation	44
4.3	The Source of Data	44
4.3	Sample Obtained	46
4.5	Phase 2: Algorithms Implementations	48
4.5.1	Implementation of Boyer-Moore Algorithm	48
4.5.2	Algorithm working procedure	50
4.5.3	Implementation of Quick Search Algorithm	50
4.5.4	Algorithm working procedure	52
4.6	Integration of Modules	52
4.7	Phase 3: Procedures of Experiment	55
4.8	Performance Metrics	56
4.9	Theoretical Performance Analysis	56
4.10	Terms used	59
4.11	Summary	60

5 DISCUSSION AND ANALYSIS

5.1	Introduction	61
5.2	Structure analysis	61
5.3	Experiment using hexadecimal representation of TCP	62
5.4	Experiment using English Alphabet	67
5.5.	Summary	70

6 CONCLUSIONS AND FUTURE WORKS

6.1	Introduction	71
6.2	Result and Achievement	71
6.3	Objectives Revisited	72
6.4	Research Contribution	73
6.5	Research Limitation	73
6.6	Future work	74
6.7	Constraints and Challenges	74
6.8	Concluding Notes	74
	References	107
	APPENDICES	

LIST OF TABLES

LIST OF TABLES	TITLE	PAGE
4.1	Represent a table with a varying pattern lengths and character strings of varying lengths	53
4.2	Represents a table with a varying pattern lengths and character strings of varying lengths	54
4.3	Represent a table with a varying pattern lengths and character strings of varying lengths	54
4.4	Represent a table with a varying pattern lengths and character strings of varying lengths	55
5.1	The average number of character comparisons of generated hexadecimal representation of network packet	69
5.2	the average number of character comparisons of generated hexadecimal representation of network packet	70
5.3	the average number of character representation of network packet	72
5.4	The Average Number of Character Comparisons of generated hexadecimal representation of network packet	73
5.5	Illustrates the performance of the two algorithms using English Alphabets as experiments	75

LIST OF FIGURES

LIST OF FIGURES	TITLE	PAGE
2.1	The process of network intrusion detection systems	19
2.2	The combination of Misuse and anomaly detection engines to monitor all network connections	20
2.3	The components of Snort	25
2.4	Shows the process of packet Decoding	27
2.5	Illustrates Snort Basic Software Components	28
2.6	Explains the structure of Deep Packet Inspection	32
2.7	Multiple string matching (Aho-Corasick algorithm)	39
2.8	Quick Search shift in the hybrid algorithm	44
3.1	Project Framework	47
3.2	Show data preparation	48
3.3	Testing the Quick Search Algorithm	49
3.4	Testing the Boyer-Moore Algorithm	49
4.1	The source code of VB Implementation of Boyer-Moore Algorithm	56
4.2	Shows the source code of Visual basic Implementation of Quick Search Algorithm	58

		15
4.3	The Integration Modules	60
4.4	The flowchart of implementation Algorithms	61
5.1	the average number of character comparisons of generated hexadecimal representation of network packet	69
5.2	the average number of character comparisons of generated hexadecimal representation of network packet	71
5.3	the average number of character representation of network packet	72
5.4	the performance of Boyer-Moore and Quick Search based on difference string lengths	74
5.5	illustrates the performance of various pattern lengths of both two algorithms	76
5.6	the performance of different string length of both two algorithms	77

CHAPTER 1

INTRODUCTION

1.1. Overview

All the time face to face communication has been an essential role of social development and gives capability of spoken languages. After invented the telephone and built its network, the way of communication have been developed and served among long distances.

Even though the advanced technology reduces the distance of people significantly, among the distance and additional broadcasts also increases the security issues for the communicators.

Attacking technology initiating the developing of telephone networks, for the purpose of stealing calls, and attempts the bugging some telephone

systems, professionals sometimes attempt to broke the system of the administrator or provider to modify and control their accounts. Worm, viruses, backdoors, and spywares, were invented by hacker to steal, destroy data, and block the network, control remote, and advertise on computer, in the era of computer network.

Attacker can do whatever they want from the other side of the world, and almost everything they want they can do on the internet easily and they have techniques to control one or a set of zombie PCs (typically known as “Botnet”).

In contrast, most of intruder’s aims are begins from individual interests and gives to gain higher skills, because of that the attackers are attracting a new built-up chain. Professional attackers sells on the internet their products and services (Zhang, 2011).

In the history, viruses had caused huge universal damages of several computers that are the reasons to start and realize the essential of anti-virus software. How do they obligate the virus? Several computers need to browse webs, sent e-mail and malware. Nevertheless, attackers can access some others like servers or computers without internet get backdoors to steal the data. Data link-layer and IP addresses on the network layer can only block the firewall; unfortunately, on application layer firewalls cannot detect anything. It is not easy to discriminate or decide among threats and normal activities. (Geddes, 2009)

Intrusion detection system (IDS) is a tool that has capability to understand every aspect of the network packets. It can investigate the packets of the network and captures the identifications of the network attacks. Now day, it is essential most organizations to use intrusion detection systems (I DS) as their security infrastructure (Jajodia, 2001).

Fundamentally, IDS classifies into host-based and network-based IDS. Host-based IDS operates as traditional antivirus software collecting information from each individual computer system, and it can be blocked or detect attacks. Though, HIDS has two limitations. First HIDS requires too many resources of CPU to spend on the host computer, and the assist can only be received by itself. Second during the data transmission in the network, HIDS could not detect the network devices like switches from attacks, and the packet sniff.

A network-based IDS (NIDS) analyses and captures the packets of the network, and detects all attacks inside the network on the scope of the network; also all belief domains can obtain the advantage (Zhang, 2011).

Altogether, intrusion detection techniques are mainly classified into Misuse detection and anomaly detection. Misuse detection technique matches and identifies the evidence of malicious behavior attacks using against signature and predefined statements. Misuse detection has a high miss detection rate, and low false alarm rate. Anomaly detection expresses as normal behavior and tries to attempt the identification of abnormal modification; perhaps, anomaly detection has low miss detection rate, and can detect new attacks, though anomaly detection has very high false alarm rate.

Generally, misuse detection comprises a set of signatures that identifies different attacks, applying misuse detection needs high load processing during the traffic detection, which source forwarding rate of lower packet in the network than the normal layer-3 switch (Geddes, 2009).

Snort is an open source lightweight network intrusion detection system (IDS) established by source-fire and it essentially based on misuse detection approach. It is an ingenious accomplished IDS the solution of Snort inline

mode for all purpose CPU. Snort supports both IPV6 and IPV4. But, Snort is uncontaminated software design, comparing ASIC and FPGA, Snort has quite slower processing speed (Geddes, 2009).

1.2. Problem Background

The movement towards more secured computing system continues to rise as management become mindful of numerous threats that exist to their organizations. Today intrusion detection system (IDS) has become a standard component of network security. Network intrusion detection system (NIDS) has been widely implemented in order to build layered information security infrastructure. Snort tool is a real time packet analyzer and packet logger that perform packet payload inspection by using pattern matching algorithms.

1.3. Statement of the Problem

Most of the intrusion detection systems deployed by the IT-based professional enterprises or educational organizations are either used signature based techniques to detect anomalous network packet, or statistical anomaly based detections techniques. One of the well-recognized signature-based techniques is the snort tool. Though, using Snort tool as malware detections has several related aspects. Here are some of them.

Hence, this project will concentrate the existing pattern matching algorithm used by Snort signature-based detection system, and the performance of these applied algorithms with Snort poverty of dropping packets. This project will analyze a selected pattern matching algorithms and propose a single if applied being better performing in the Snort tool.

SNORT detection performance in Linux and Windows platforms.

1. Snort tool skills performance poverty in higher rate traffic network which lead to the packet dropping.
2. The powerlessness of the snort tool to detect and log unknown and known malware variants, and also obfuscated malware.
3. The lag time between snort repository rule update and release of new malware has contributed immensely in the downward performance of snort tool.
4. Also problem of false alarms appears to be most pressing one.

1.4. Research Goals

In this project presents the pattern matching algorithm of network deep packet inspection program. By way of Snort is the most open source typical deep packet inspection based on intrusion detection system this project will use the comparing some of pattern matching algorithm applying in Snort. Currently, Snort is still a sequence program. Some types of pattern matching algorithm have been implemented by researchers on some special network processors. However, almost none of them are based on a general processor. This project attempt to find and evaluate some possible methods of Snort pattern matching algorithm on general network system, and also, will come up analyzing the performance according to the accuracy and speed in Linux and windows OS platforms. At the end of this project, the analyzing performance of pattern matching algorithms using the current Snort is the expected outputs. However this project will not focus on the parallelization processing, and the efficiency of the memory usage.

1.5. Objectives of the Project

This project aims at studying different components of snort tools and compare between of pattern matching algorithms and snort Quick search. This is in order to analyze the performance of snort tool in UTM network. Therefore the following objectives are set to be achieved.

1. To study and compare between two snort pattern matching Algorithms
2. To identify limitations and strength of snort tool Network Intrusion Detection Systems (NIDS).
3. To analyze the performance of two pattern matching algorithms applied in Snort.

1.6. Scope of the Project

The project has the following limitation to be considered in the study.

1. The study is limited to snort components and snort pattern matching algorithms
2. The study uses open source snort tool application.
3. The study will aim the analyzing two of pattern matching algorithms applied in Snort
4. Hexadecimal representation of TCP and Plain of English text were used as the data of this project work to experiment the result of the algorithms

1.7. Organization of the Project

This project is charted in chapters as follows: Chapter 1 delivers the general impression of the research topic and statement of the problems, scope, and objective and the as well as chapter 2 reviewed some literatures about intrusion detection system, then Snort and its components. In chapter 3 discusses the methodology to be used in comparing two different pattern matching algorithm, and the research framework as well. Indeed, chapter 4 presents the project initial finding and the discussions. Also, in chapter 5 will present the result and its analysis. Finally, chapter 6 will be the conclusion and further work to be conducted in the project.

REFERENCES

- Mustafa Abdul Sahib Naser, Nur'Aini Abdul Rashid, and Mohammed Faiz Aboalmaaly. (2012). Quick-Skip Search Hybrid Algorithm for the Exact String Matching Problem, International Journal of Computer Theory and Engineering.
- Rafeeq Ur Rehman,(2003). Intrusion Detection Systems with Snort. Library of Congress Cataloging-in-Publication Data. Published in USA.
- P. Garcí'a-Teodoroa, J. Dí'az-Verdejoa, G. Macia'-Ferna'ndeza, E. Va'zquezb. (2008), Anomaly-based network intrusion detection: Techniques, systems and challenges, Department of Signal Theory, Telematics and Communications – Computer Science and Telecommunications Faculty, University of Granada, Granada, Spain
- Moath Hashim Alsafasfeha, Abdel Ilah Noor Alshbatatb. (2008). Configuring Snort as a Firewall on Windows 7 Environment, National university of Malaysia UKM, Selengor, Malaysia. bTafila Technical University, Electrical Engineering Department,Tafila, Jordan, 66110.
- Christian Charras - Thierry Lecroq, <http://www-igm.univ-mlv.fr/~lecroq/string/> Faculté des Sciences et des Techniques, 76821 Mont-Saint-Aignan Cedex, FRANC
- Zhu Yong-qiang, (2011). Two Enhanced BM Algorithm in Pattern Matching, IEEE Computer Society, College of Computer & Information Engineering, Lishui University, China.

- KEN CHEN, FEI YU, CHENG XU, YAN LIU. (2008). Intrusion Detection for High-speed Networks Based on Producing System, IEEE Computer Society, School of Computer & Information Engineering, Hunan Agricultural University, Changsha, china.
- Abuhmed, T., Mohaisen, A., & Nyang, D. (n.d.). Deep Packet Inspection for Intrusion Detection Systems : A Survey. Information Security.
- Alserhani, F., Akhlaq, M., Awan, I. U., Cullen, A. J., & Mellor, J. (n.d.). Snort Performance Evaluation 1. Performance Evaluation.
- Baker, Z. K., Member, S., & Prasanna, V. K. (2005). Flexible Intrusion Detection. October, 13(10), 1179-1189.
- Bansal, K. (2008). The Knuth-Morris-Pratt algorithm.
- Boob, S., & Jadhav, P. (2010). Wireless Intrusion Detection System, International Journal, 5(8), 9-13.
- Brown, D. J., Suckow, B., & Wang, T. (n.d.-a). A Survey of Intrusion Detection Systems 2 Information Sources Analysis Techniques.
- Marcelo Reis, Fabr, cio Paula, Diego Fernandes, and Paulo Geus. (2009). A Hybrid IDS Architecture Based on the Immune System, Computing Institute State University of Campinas Campinas, SP Brazil.
- Brown, D. J., Suckow, B., & Wang, T. (n.d.-b). A Survey of Intrusion Detection Systems 2 Information Sources Analysis Techniques.
- Dhanalakshmi, Y. (2008). Intrusion Detection Using Data Mining Along Fuzzy Logic and Genetic Algorithms, Journal of Computer Science, 8(2), 27-32.
- Di, J. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges, 28, 18-28. doi:10.1016/j.cose.2008.08.003
- EnCase Forensic Version 7 Preview New Features Guide. (n.d.).Managing.

- Hai-sheng, Q. I. N. (2011). Algorithm Based on Intrusion Detection System, 0-3.
- Idika, N. (2007). A Survey of Malware Detection Techniques, Purdue University, 48. Citeseer. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.75.4594&rep=rep1&type=pdf>
- Jyothsna, V., Prasad, V. V. R., & Prasad, K. M. (2011). A Review of Anomaly based Intrusion Detection Systems. *International Journal*, 28(7), 26-35.
- Kumar, S. (2011). Design and Implementation of IDS Using Snort, Entropy and Alert Ranking System. *Source, (Icscn)*, 264-268.
- Mandumula, K. K. (2011). *nu t or ris r at t nu t. History.*
- Nazer, G. M. (2011). Current Intrusion Detection Techniques in Information Technology - A Detailed Analysis, *European Journal of Scientific Research*, 65(4), 611-624.
- Norton, M. (2002). Optimizing Pattern Matching for Intrusion Detection, *System*, 1-11.
- Papadogiannakis, A., Polychronakis, M., & Markatos, E. P. (n.d.). (2010). Improving the Accuracy of Network Intrusion Detection Systems Under Load Using Selective Packet Discarding.
- Rajasekhar, K., Babu, B. S., Prasanna, P. L., Lavanya, D. R., & Krishna, T. V. (2011). An Overview of Intrusion Detection System Strategies and Issues. *Network*, 8491, 127-131.
- Raju, B., & Srinivas, B. (2012). Network Intrusion Detection System Using KMP Pattern Matching Algorithm, *Computer Science and Telecommunications*, 3(1), 1-4. Re, K.-morris-pratt. (n.d.). Pattern Matching.
- Roobahani, A. R. (2009). Service Oriented Approach to Improve the Power of Snort,. doi:10.1109/ICCEE.2009.270

- Awsan Abdulrahman Hasan, and Nur' Aini Abdul Rashid. (2012). Hybrid Exact String Matching Algorithm for Intrusion Detection System, ICCIT,
- Sandhu, U. A., Haider, S., Naseer, S., & Ateeb, O. U. (2011). A Survey of Intrusion Detection & Prevention Techniques. *Management*, 16, 66-71.
- Sandhu, U. A., Haider, S., Naseer, S., & Ateeb, O. U. (2011). A Survey of Intrusion Detection & Prevention Techniques, *Management*, 16, 66-71.
- Sandhu, U. A., Haider, S., Naseer, S., & Ateeb, O. U. (2011). A Survey of Intrusion Detection & Prevention Techniques, *Management*, 16, 66-71.
- Security, C., & Monitoring, T. (2011). Importance of Intrusion Detection System (IDS). *International Journal*, 2(1), 1-4.
- Sedjelmaci, H., & Feham, M. (2011). Novel hybrid intrusion detection system. *Network Security*, 3(4), 1-14.
- Singhrova, A. (2011). A Host Based Intrusion Detection System for DDoS Attack in WLAN. *Engineering*, 433-438.
- Singla, N., & Garg, D. (2012). String Matching Algorithms and their Applicability in various Applications. *Soft Computing*, (6), 218-222. Snort, D. (n.d.). Dissecting Snort Network, Tekniska, K. (n.d.). *Intrusion Detection Systems*.
- The Mendeley Support Team. (2011). Getting Started with Mendeley. Mendeley Desktop, London: Mendeley Ltd. Retrieved from <http://www.mendeley.com>
- Verwoerd, T., & Hunt, R. (2002). Intrusion detection techniques and approaches. *Computer Communications*, 25, 1356-1365.
- Weinsberg, Y., & Dolev, D. (n.d.). High Performance String Matching Algorithm for a Network Intrusion Prevention System (NIPS). P And T.

Wu, S., & Manber, U. (1994). A fast algorithm for multi-pattern searching, *Science*, 1-11.

Zeng, B., Yao, L., & Chen, Z. (2010). A Network Intrusion Detection System with the Snooping Agents, *Source, (Iccasm)*, 232-236.

Salah, K. Ã., & Kahtani, A. (2010). Journal of Network and Computer Applications Performance evaluation comparison of Snort NIDS under Linux and Windows Server, *Journal of Network and Computer Applications*, 33(1), 6-15. Elsevier. doi:10.1016/j.jnca.2009.07.005