Information Security Hacking Defense Intelligence and Misinformation Protection Algorithms

Seong-Kyu Kim[*]

*Assistant Professor (Tenure Track) of Department of Information Security, Joongbu University, Gyeonggi-do Goyang-Si 10279, Republic of Korea.

Department of Public Policy and Information Technology, Seoul National University of Science and Technology, Seoul 01811, Korea.(e-mail : skkim@joongbu.ac.kr or guitara77@gmail.com)(e-mail : skkim@joongbu.ac.kr or guitara77@gmail.com)

Jun-Ho Huh[**]

** Assistant Professor (Tenure Track) of Department of Data Informatics, (National) Korea Maritime and Ocean University, 727, Taejong-ro, Yeongdo-gu, Busan, Republic of Korea (e-mail : 72networks@pukyong.ac.kr or 72networks@kmou.ac.kr)

Abstract

TThis paper has seen a rapid increase in services and users over the Internet. As a result, cyberattacks are increasing, and information leakage and financial damage are occurring. The government, public institutions, and companies use signature-based detection rules to respond to known malicious codes during these rapid cyber attacks, but it takes a long time to generate and verify signature-based detection rules. To address this problem, in this paper, we propose and develop a signature-based detection rule generation and verification system using signature extraction and traffic analysis techniques through latent Dirichlet allocation algorithms. Experiments on the developed system have shown that we generate and validate detection rules much faster and more accurately than existing ones.

I. Introduction

This paper introduces and operates various network security systems such as firewall, intrusion detection system (IDS), intrusion prevention system (IDS), and web firewall (WAF) to cope with cyber attacks through the Internet. Despite the active study of various AI-based and behavior-based malware detection techniques, most security systems use signature-based detection methods that boast high detection performance in known malware detection [1]. Signature-based detection methods inevitably imply an abstract definition of malicious code called detection rules. The effective definition of detection rules requires professional knowledge of networks, security, operating systems, etc. The operation of inappropriate detection rules can cause numerous false positives, causing security system performance degradation, and even paralyzing the entire network with network security systems installed. In this paper, we propose a system that allows not only security experts but also quasi-experts to generate and validate detection rules quickly and accurately to efficiently respond to rapidly increasing malware [2]. The proposed technique is to automatically generate snort detection rules using research from existing signature extraction and traffic analysis results, and to validate detection rules generated by deploying IDS servers in virtual environments [3].

II. Related Research

2.1 Signature-based Detection Rules

Among the signature-based detection rules, we analyze snort detection rules, which are the most commonly used globally and are adopted as TTA standards and are mostly supported by domestic network security equipment. In this paper, we define the requirements needed to design a system that automatically generates and validates snort detection rules based on the adopted TTA standard. Snort detection rules are logically divided into detection rule headers and detection rule options.In the detection rule header, each configuration is separated by a space, the detection rule header and options are separated by square brackets, and the options and options of the detection rule are separated by semicolons [4]. Options consist of optional keys and optional values, separated by colons. And as a result of analyzing the detection rules, the use of many options may increase detection performance depending on the traffic it detects, but vice versa, it may increase the false detection rate. Therefore, we minimize the automatically generated options and implement them so that users can add them arbitrarily.

2.2 Malicious Signature Extraction Study

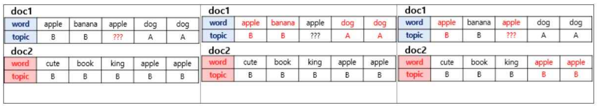
Many studies have been conducted to automatically extract malicious signatures from malicious code or malicious traffic [5]. built two honeypods to extract malicious signatures from the traffic generated by the worms generated using Metasploit, and then used the Longest Common Substring (LCS) algorithm in the traffic sent and received.Recent work has extracted signatures using Voting Experts algorithm and Ranking algorithm using unsupervised learning-based segmentation algorithm from silver malicious traffic and malicious code analysis using Denising Auto-Encoder algorithm from deep trust neural networks (DBNs).

III. Research method

3.1 Study Design

Topic modeling algorithms have been used to extract keywords from social networks to track changing issues, analyze papers published in papers to identify topics that draw attention from time to time, extract topics from newspaper articles, and analyze changes in issues from time to time. As such, topic modeling techniques are algorithms that have been used to analyze trends for a particular field and have proven their performance [6]. Topic modeling algorithms include VectorSpace Model (VSM), Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (pLSA), and Latent Dirichlet Allocation (LDA). In this paper, we propose a system that leverages LARGen techniques developed based on latent Dirichlet allocation algorithms to perform signature extractions necessary for generating detection rules, and uses them to automatically generate snort detection rules and validate generated snort detection rules. The latent Dirichlet allocation (LDA) algorithm is a topic modeling algorithm designed to infer topics inherent in text documents. Recently, it has been widely used as a statistical topic model for finding topics in very large data

(BigData). This section briefly summarizes the LDA [Fig. 1].



doc1

word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2

word	cute	book	king	apple	apple
topic	B	B	B	B	B

doc1

word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2

word	cute	book	king	apple	apple
topic	B	B	B	B	B

doc1

word	apple	banana	apple	dog	dog
topic	B	B	???	A	A

doc2

word	cute	book	king	apple	apple
topic	B	B	B	B	B

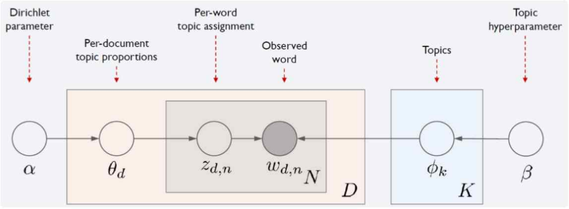
[Fig. 1] How to proceed with the learning

3.2 Research Architecture

The detection rule generation system provides an easy environment to analyze malicious traffic that requires generating detection rules, automatically recommends candidate signatures, and generates detection rules according to snort grammar. In the Recommended Detections string window, select the flow you want to extract the signature from the Malicious Packet File Management window and click the "Extract Malicious Signature" button to ensure that the program has successfully entered the database[7].

We show a list of malicious signatures extracted using . The Edit Detection Rule window also displays the recommended detection string window in the Recommended Detection String window.

When selected, the extracted signature automatically analyzes the corresponding flow information and shows the results of the snort detection rule, which is generated to fit the detection rule format. It can be modified by the user arbitrarily. In addition, the Detection Rule Management window contains a list of detection rules saved after completing the detection rule editing in the Detection Rule Editing window and a list of detection rules applied to real-world information security systems. The initial detection rules generated here can be applied to security equipment after verification is performed [Fig. 2].



[Fig. 2] Dirichlet Distribution Inference Process

IV. Future research and conclusions

In this paper, we propose a system that allows not only experts but also quasi-experts to generate and validate detection rules quickly and accurately, and even experimentally through real-world development. The proposed and developed system can generate detection rules regardless of the type of malicious code if only payload exists in malicious traffic that wants to generate detection rules. System experiments have allowed us to generate and validate detection rules normally, and with this system, there are differences depending on system users and detection rules, but we have been able to generate and validate detection rules within an hour. With the developed detection rule generation and verification system, experts will be able to generate and validate detection rules much faster than existing ones, and semi-professionals who are not yet skilled will be able to generate and validate detection rules more easily than existing ones.

[Reference]

[1] Myung-Hoon Kang, “Completion of information security monitoring system with complete control and security of the IDS pattern matching techniques discussed in Big Data Analytics”, pp 40-41. 5 2013.

[2] Jae Chan Yoo, "A Study on the Protection for Corporation Information Using Scenario Technique," The Graduate of SungKyunkwan University, pp. 14-16, August 2012.

[3] Kelly M, Mark Nicolett, Oliver Rockford, "Magic Quadrant for Security Inofrmation and Event Management", Gartner Group, pp. 2-8, June 2014.

[4] DongSung im, YoungMin Kim, "Enhancement of internal Control by expanding Security Information Event Management System", Korea Society of Computer and Information, pp 36-37. 8 2015.

[5] Zhuo. Zhang, Zhibin Zhang, Patrick P.C.Lee, Yunjie Liu and Gaogang Xie "ProWord: An unsupervised approach to protocol feature word extraction", in INFOCOM, 2014 Proceedings IEEE, pp. 1393-1401, July, 2014.

[6] Omid E. David and Nathan S. Netanahu, "DeepSign: Deep learning for automatic malware signature generation and classification", International Joint Conference on Neural Networks, July, 2015.

[7] Fabrizio Biondi, Francois Dechelle and Axel Legay "MASSE: Modular Automated Syntactic Signature xtraction", IEEE International Symposium on Software REliability Enginerring Workshops, Oct, 2017.