Journal of the Korea Academia-Industrial cooperation Society
Vol. 20, No. 7 pp. 613-621, 2019

https://doi.org/10.5762/KAIS.2019.20.7.613
ISSN 1975-4701 / eISSN 2288-4688

악성 스크립트 패턴 분석을 통한 악성코드 탐지 기법

이용준[1], 이창범[2*]
[1]국방보안연구소, [2]한국교통안전공단

A Malware Detection Method using Analysis of Malicious Script Patterns

Yong-Joon Lee[1], Chang-Beom Lee[2*]
[1]Defense Security Institute, [2]Korea Transportation Safety Authority

요 약 최근 IoT, 클라우드 컴퓨팅 기술이 발전하면서 IoT 디바이스를 감염시키는 악성코드와 클라우드 서버에 랜섬웨어를 유포하는 신종 악성코드가 등장하여 보안 위협이 증가하고 있다. 본 연구에서는 기존의 시그니처 기반의 탐지 방식과 행위기반의 탐지 방식의 단점을 보완할 수 있도록 난독화된 스크립트 패턴을 분석하여 점검하는 탐지 기법을 제안한다. 제안하는 탐지 기법은 웹사이트 통해 유포되는 악성 스크립트 유형을 분석하여 유포패턴을 도출한 후, 도출된 유포패턴을 등록하여 점검함으로써 기존의 탐지룰 기반의 탐지속도를 유지하면서도 제로데이 공격에 대한 탐지가 가능한 악성 스크립트 패턴분석 기반의 악성코드 탐지 기법이다. 제안한 기법의 성능을 검증하기 위해 프로토타입 시스템을 개발하였으며, 이를 통해 총 390개의 악성 웹사이트를 수집, 분석에 의해 도출된 10개의 주요 악성 스크립트 유포패턴을 실험한 결과, 전체 항목 평균 약 86%의 높은 탐지율을 보였으며, 기존의 탐지룰 기반의 점검속도를 유지하면서도 제로데이 공격까지도 탐지가 가능한 것을 실험으로 입증하였다.

Abstract Recently, with the development of the Internet of Things (IoT) and cloud computing technologies, security threats have increased as malicious codes infect IoT devices, and new malware spreads ransomware to cloud servers. In this study, we propose a threat-detection technique that checks obfuscated script patterns to compensate for the shortcomings of conventional signature-based and behavior-based detection methods.

Proposed is a malicious code-detection technique that is based on malicious script-pattern analysis that can detect zero-day attacks while maintaining the existing detection rate by registering and checking derived distribution patterns after analyzing the types of malicious scripts distributed through websites. To verify the performance of the proposed technique, a prototype system was developed to collect a total of 390 malicious websites and experiment with 10 major malicious script-distribution patterns derived from analysis. The technique showed an average detection rate of about 86% of all items, while maintaining the existing detection speed based on the detection rule and also detecting zero-day attacks.

Keywords : Lansomware, Malicious Code, Malware Detection, Malicious Patterns, Secure Web Sites

1. 서론

최근 IT 산업 및 ICT 기술의 발전과 인터넷과 스마트폰의 급속한 확산에 따라, 자동화되고 지능화된 악성코드 및 악성코드 생성 도구가 인터넷 망을 통해 빠르게 유포되고 있다. 또한, IoT(Internet of Things) 및 클라우드

*Corresponding Author : Chang-Beom Lee(Korea Transportation Safety Authority)
email: chblee1225@gmail.com
Received June 3, 2019 Revised July 4, 2019
Accepted July 5, 2019 Published July 31, 2019

컴퓨팅 기술이 등장하면서 IoT 디바이스와 클라우드 서버에 랜섬웨어를 감염시키는 등 새로운 유형의 악성코드들로 인해 보안 위협이 더욱 증가하고 있다[1-4].

유포되고 있는 전체 악성코드 가운데 신종 악성코드는 약 20% 정도로 대부분 기존 악성코드의 변종으로 분류한다[2]. 이러한 악성코드를 이용한 보안 위협은 현재 가장 심각한 보안 위협으로 분류되고 있으며, 악성코드의 탐지를 어렵게 하기 위해, 악성코드 제작은 더욱 지능화되고 정교해 지고 있다. 이에 따라, 약 50% 이상의 신종 악성코드가 안티 바이러스 제품 등 악성코드 탐지 시스템에 의해 탐지되지 않는 등 악성코드에 의한 보안 위협은 더욱 심각한 보안 위협이 되고 있다[5-11].

이러한 악성코드는 다양한 경로를 통해 유포되고 있으며, 유포 경로 중, 웹사이트와 이를 이용하는 이용자 PC의 취약점을 이용하여 악성코드를 유포하는 것이 일반적인 패턴이다. 이용자의 PC를 감염시킬 수 있는 악성코드가 은닉된 웹사이트는 악성코드를 직접 숨기고 있는 유포지와 이러한 유포지로 이용자의 PC를 자동 연결되도록 연결주소가 숨겨진 경유지로 구분된다[12]. 해커는 악성코드 유포지를 개설한 후 경유지 웹사이트를 해킹하여 유포지 URL을 삽입시킴으로써, 경유지에 접속하는 방문자 즉, 이용자가 인지하지 못하도록 한 상태에서 유포지로 이동하도록 유도하여 악성코드에 감염시킨다. 이러한 악성코드 경유지는 포털 사이트, 블로그, 게시판 등 주로 이용자의 방문이 높은 웹사이트가 주된 대상이며, 공격자가 해킹한 경유지 내에 악성코드를 난독화하여 은닉시키기 때문에 쉽게 탐지되지 않는 특징이 있다[13].

웹사이트에 은닉된 악성코드를 탐지하는 방법에는 시그니처 기반 방식과 행위 기반 방식으로 구분된다. 시그니처 기반 탐지는 점검하는 웹사이트 소스코드 내에 악성코드 유포지가 삽입되었는지 검색하는 방식으로, 탐지속도는 빠른 장점이 있지만, 제로데이 공격(zero day attack)에 대한 탐지성능이 낮은 단점이 있다[14]. 행위기반 탐지는 웹사이트에 가상으로 접속시켜 방문자 PC의 파일변조, 악성코드 실행 등의 상태변화를 추적을 통해 탐지하기 때문에 탐지속도가 느린 단점이 있지만, 제로데이 공격에 대한 탐지 성능이 높은 장점이 있다[15].

그동안 다양한 시그니처 기반 탐지기법과 행위기반 탐지기법이 연구되어 왔으나, 약 80% 이상의 신종 악성코드는 난독화, Anti-VM, 실행압축 등 악성코드 탐지 및 분석을 우회하는 기술이 사용되고 있어, 기존의 탐지룰 기반 탐지나 행위기반 탐지 기술로는 효과적인 악성코드 탐지에 한계가 있다[16-18].

따라서 본 연구에서는 기존의 탐지룰 기반의 탐지 방식과 행위기반의 탐지 방식의 단점을 보완할 수 있는 난독화된 스크립트를 패턴분석으로 점검하는 탐지 기법을 제안한다.

제안하는 탐지 기법은 웹사이트를 통해 유포되는 악성코드의 스크립트 유형을 분석하여 유포패턴을 도출한 후, 도출된 유포패턴을 탐지룰에 등록하여 점검함으로써 기존의 시그니처 기반의 탐지속도를 유지하면서도 제로데이 공격 까지 탐지가 가능한 악성 스크립트 유포패턴 분석기반의 악성코드 탐지 기법이다.

2. 관련 연구

2.1 웹 어플리케이션 보안취약점

공격자는 웹 어플리케이션의 보안취약점을 이용하여 경유지 웹사이트 내에 악성코드 유포지를 은닉시켜, 방문자에게 악성코드를 유포시킨다. 해킹의 대상이 웹 어플리케이션의 취약점에 대해 OWASP(Open Web Application Security Project)에서는 Table 1과 같이 정의하고 있다[19].

Table 1. OWASP Top 10 Vulnerabilities

No.	Security vulnerability	Response method
1	Cross Site Scripting	- Standard validation of input values - Encrypting output values
2	Injection Vulnerability	- Parameterized Query Language - Object Relationship Mapping
3	Remote File Execution	- Network Independent Configuration - Design Phase Verification Technique
4	Unstable direct object reference	- Indirect reference map - Allow direct reference after authorisation
5	Cross Site Request Modulation	- Separate authentication mechanism - POST request method
6	Information leakage and improper error handling	- Restricting disclosure of error information - Disabling error information disclosure function
7	Vulnerable Authentication and Session Management	- Single authentication mechanism - Separate session management mechanism
8	Unstable encryption storage	- Proven cryptographic algorithm implementation - Secure encryption key management
9	Unstable communication	- SSL
10	Failure to restrict URL access	- Implementing Access Matrix

2.2 웹사이트를 통한 악성코드 유포 경로

인터넷 사용자가 웹사이트 접속에 의한 악성코드에 감염되는 경로는 크게 두 가지로 분류되며, 이는 웹사이트에 악성코드를 직접적으로 숨기고 있는 유포지와 사용자가 유포지로 자동 연결되어 접속되도록 접속 주소가 숨겨진 경유지이다[20].

2.2.1 최초 경유지(Landing Site)

악성코드의 최초 경유지는 인터넷 사용자들로 하여금 악성코드 유포지로의 접속을 유도하는 첫 번째 웹사이트로 악성코드 감염의 진입로 역할을 한다. 일반적으로 경유지는 상시 접속자수가 많고, 이용자들의 관심이 많은 정보 또는 사회적 이슈를 제공하는 포탈 사이트, 블로그 등이 최초의 경유지로 이용되고 있다. 주로 〈iframe〉 등과 같이 지정된 웹사이트로 연결할 수 있는 코드(HTML 태그)나 키워드의 링크를 통해 최초 경유지로 연결을 한다[21].

2.2.2 중간 경유지(Hopping Site)

악성코드의 중간 경유지는 앞에서 설명한 최초 경유지 이후에 중간 경유지를 통해 접속한 인터넷 사용자를 다음 경유지로의 이동을 중계해주는 역할을 한다. 따라서 최초 경유지와 같이 사회적 관심정보나 이슈를 제공하여 접속을 유도하기 보다는 웹사이트의 보안관리가 취약한 웹사이트가 주로 중간 경유지로 이용된다[22].

2.2.3 최종 경유지(Exploit Site)

악성코드의 최종 경유지로는 중간 경유지와 같이 보안이 취약한 웹사이트가 주로 이용되며, 단순한 중계 역할 외에도 이용자의 PC나 이용자가 PC에서 사용하는 어플리케이션의 보안 취약성을 이용하여 공격하는 exploit 악성 스크립트가 삽입된다[23]. 본 연구에서는 이러한 exploit 악성 스크립트 유포패턴을 분석하여 악성코드 유포하는 웹사이트를 탐지하고자 한다.

2.2.4 유포지(Distribution Site)

악성코드 유포지는 이용자 PC에 직접 악성코드를 감염시키는 웹사이트로 공격자가 직접 운영하기도 하고, 이미 해킹된 웹사이트가 이용되는 경우도 있다. 유포지는 악성코드의 유포를 목적으로 한 특정 코드는 삽입되어 있지 않으며, 사용자가 악성코드 유포지에 접속하면 사용자의 설치에 대한 동의가 없어도 악성코드를 자동으로 설치되는 특징을 가지고 있다[24]. Fig. 1은 웹사이트를 통해 악성코드가 유포되는 경유지와 유포지의 구간을 나타낸 것이다.

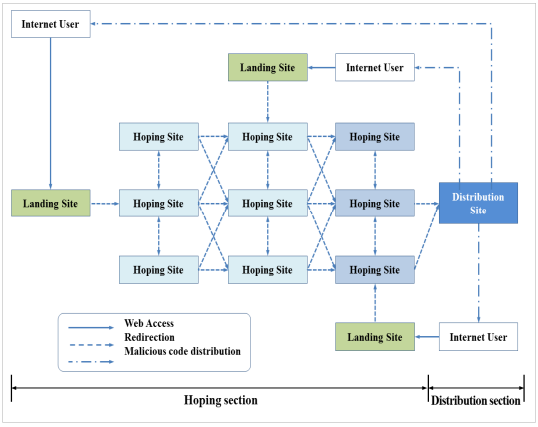


Fig. 1. Path of distribution of malicious code

2.3 웹사이트 은닉 악성코드 탐지방법

웹사이트 내에 은닉된 악성코드를 탐지하는 대표적인 방법은 유포지 목록으로 웹 어플리케이션 내 소스코드를 점검하는 시그니처 기반 탐지 기법(Signature-Based Detection Technique)과 웹사이트에 가상의 사용자 PC를 접속시켜 상태의 변화에 따라서 악성여부를 확인하는 행위 기반 탐지 기법(Behavior-Based Detection Technique)이 있다.

2.3.1 시그니처 기반 탐지방법

시그니처 기반 탐지 방법은 악성코드 유포지를 탐지룰로 등록하고 점검하고자 하는 웹사이트를 크롤링하여 소스코드 내에 유포지가 있는지 검색하는 방식이다. 이 탐지방법을 이용하면 해당 웹사이트의 소스코드에 악성코드 유포지가 삽입된 경우, 악성코드 경유지로 악용되고 있다는 것을 탐지할 수 있다. 시그니처 기반 탐지방법은 유포지 목록으로 검색하기 때문에 탐지 속도가 빠르나, 웹사이트의 악성 스크립트가 난독화되어 변형되는 경우에는 탐지되지 않는 단점이 있다. 따라서 시그너처 기반 탐지를 우회하기 위해 유포지 정보를 난독화 또는 암호화하는 신종 또는 변형 악성코드인 제로데이 공격을 탐지하기에 어렵다[25].

2.3.2 행위 기반 탐지방법

가상의 사용자 PC로 악성코드가 은닉된 웹사이트에

접속시켜 악성 스크립트가 실행될 때 힙 스프레잉 공격 코드의 실행여부를 탐지하는 방법이다. 이 방법을 이용하면 힙 스프레잉 코드가 실행될 때 스크립트의 난독화 및 암호화 여부와 상관없이 최종적으로 정보유출, DDoS공격 등 악성행위를 하기 때문에 제로데이 공격에 대한 탐지가 가능하다. 하지만 웹사이트에 접속한 PC 내의 악성 행위 여부를 계속 모니터링하기 때문에 탐지속도가 느린 단점이 있다[26].

3. 악성 스크립트 유포패턴 분석

3.1 악성 스크립트 수집 및 패턴 분석

악성코드의 유포 패턴을 분석하기 위해 정보보안기관이 공개한 국내외 악성코드 유포지로 확인된 500개의 웹사이트를 크롤링하여 수집한 후, 악성 스크립트 패턴분석을 수행하였다. Fig. 2는 악성코드 스크립트 수집 및 패턴 분석을 위한 프로세스를 나타낸 것이다.

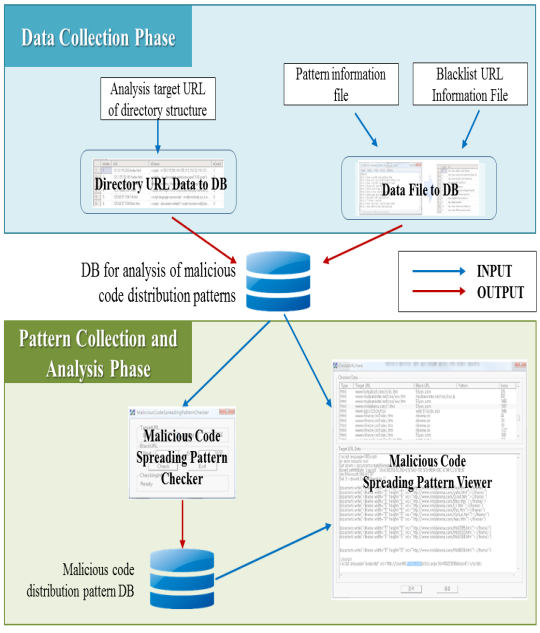


Fig. 2. Process for collecting and analyzing patterns of malicious code distribution

수집한 500개의 악성코드 유포지를 분석한 결과, 전체 악성코드 유포지 중의 95%는 악성 스크립트 형태였으며, 5%는 멀티미디어 파일 내 스크립트를 삽입하는 방식이었다. 분석한 악성 스크립트 유포패턴 분석 결과는 Table 2와 같다.

Table 2. Results of analysis on the distribution patterns of malicious scripts

Item	Use	Not used	Total
Sctipt	470 URLs (94%)	30 URLs (6%)	500 URLs (100%)
Code Obfuscation	315 URLs (63%)	185 URLs (37%)	500 URLs (100%)
document. write()	240 URLs (48%)	235 URLs (52%)	500 URLs (100%)
External link	165 URLs (33%)	335 URLs (67%)	500 URLs (100%)
URL Encoding	70 URLs (14%)	430 URLs (86%)	500 URLs (100%)

Table 2에 타나듯이 스크립트는 일반적으로 악성 스크립트로 은닉하기 위해 사용되며 94%가 스크립트 형태로 분석되었다. 스크립트 코드유형은 javascript 49%, script 28%, vbscript 17% 순위였다.

악성코드 유포지는 시그니처 기반으로 탐지가 가능하기 때문에 공격자는 스크립트 코드를 난독화하거나 문자열로 변환하여 은닉시키며 난독화는 63% 로 분석되었다.

악성 스크립트에 통상 document.write() 함수가 사용되고 있으며, 이는 웹브라우저가 웹페이지를 사용자에게 보여줄 때 내부의 콘텐츠를 구성할 수 있어 악성 스크립트를 은닉시킨 경우, 웹페이지가 보여질 때 악성 스크립트를 조합이 가능하기 때문이다. 패턴 분석을 통해 약 48%의 document.write() 함수가 악용된 것으로 분석되었다.

외부링크는 경유지에서 유포지로 접근하는 가장 중요한 방법으로 외부 도메인이 은닉하는 경우로 33%가 사용되었으며, URL 인코딩은 악성코드 유포지 URL의 탐지를 피하기 위해 유포지 URL을 스크립트 코드에서 인식하기 어려운 형태로 변환하는 방식으로 14%로 분석되었다.

3.2 악성 스크립트 패턴분석을 통한 탐지항목

앞에서 분석한 악성 스크립트 패턴분석을 통해 Table 3과 같이 악성 스크립트 탐지 항목 10개를 도출하였다.

Table 3. Malicious script detection items by distribution Pattern

No.	Malicious Script Distribution Pattern	
	Distribution Pattern	Example
1	User-defined String Processing Function	function
2	Large-scale String Processing Function	unescape, replace, split, fromCharCode
3	Simple String Iteration	+\",&\",+\',&\',\"+,\"&,\'+,\'&
4	Eval function	eval
5	Large-scale Special Character	#, $, %, ^, &
6	Us-ascii, Jscript.Encode Encoding Function	us-ascii, jscript.encode
7	Web Pages containing Multimedia	jpg, gif, swf, wav
8	Executable Webpage Containing Files	exe
9	Shell Script	Shell.Run, ShellExecute
10	img tag Abnormal width, height value	width 〈 4, height 〈 4

Table 3의 1, 2번 유포 패턴은 대량의 문자열을 인코딩시키는 난독화를 탐지하는 것이며, 3번 유포 패턴은 웹페이지 소스코드를 분해 후 재조립 하는 경우로 악성 스크립트의 난독화를 위해 사용한다. 4번 유포 패턴인 eval 함수는 수식형태의 문자열을 숫자로 바꿔주는 함수로써 난독화를 위해 사용한다. 5번 유포 패턴은 인코딩이나 난독화가 된 소스코드의 경우 script 태그 안에서 특수문자와 기호의 사용빈도가 높기 때문에 도출된 항목이며, 6번 유포패턴인 us-ascii, jscript.encode 인코딩 방식은 정상적인 웹페이지에서 사용되지 않으며, 악성스크립트에서 사용빈도가 높기 때문에 도출된 항목이다. 7번 유포패턴은 멀티미디어 파일을 웹페이지에 불러오는 것처럼 악성코드 유포를 위한 파일을 불러오는 경우이며, 8, 9번 유포패턴은 실제 악성행위를 하는 exploit 파일에 사용되는 exe 실행파일과 웹쉘 스크립트를 사용하는 경우를 탐지하기 위한 항목이다. 10번 유포패턴은 그림파일을 보여주는 img 태그를 사용하여 악성코드 유포 파일을 보이지 않게 불러오는 경우를 탐지하기 위해 도출된 항목이다.

3.3 악성 스크립트 패턴분석 기반 탐지 기법

패턴분석에 의해 도출된 악성 스크립트 탐지룰은 정상적인 웹페이지에서도 탐지될 수 있기 때문에 악성 스크립트 유포패턴이 탐지되어도 악성코드 경유지/유포지로 완전히 단정할 수는 없다. 따라서 실험에서는 악성 스크립트 유포패턴이 탐지된 웹페이지는 의심스러운 웹페이지로 1차 분류를 한 후, 정확성을 높이기 위해 중요도를 산정한다.

악성 스크립트 유포패턴의 중요도는 의심스러운 정도를 정량적으로 제시하기 위해 악성 스크립트 유포패턴이 정상적인 웹페이지에서 사용되는 빈도와 관련하여 중요도를 산정하는 것이다. 중요도를 계산하는 식은 다음과 같다.

- $i_{Pattern}(D)$: 중요도
- $N_{Total}(W\_list)$: 전체 화이트리스트의 개수
- $N_{Detected}(Webpages)$: 탐지된 웹페이지의 수

$$i_{Pattern}(D) = 1 - \frac{N_{Detected}(Webpages)}{N_{Total}(W\_list)} \quad (1)$$

대부분의 악성 웹페이지에서 복수의 악성스크립트 유포패턴을 탐지해낼 수 있다. 해당 웹페이지에 대한 최종적인 위험점수는 각 악성 스크립트 유포패턴 항목의 탐지 결과값과 중요도를 이용해서 계산한다. 위험점수를 계산하는 식은 다음과 같다.

- $S_{Danger}(W)$: 해당 웹페이지의 위험점수
- $N_{Detected}(P)$: 탐지된 유포패턴의 수
- i : 중요도 $(0 \le i \le 1)$
- v : 탐지 결과값 $(0 \le v \le 1)$
- k : 악성 스크립트 유포패턴의 개수

$$S_{Danger}(W) = \max_{1 \le k \le n}(i_k \times v_k) \quad (2)$$

이러한 조건들로 위험 점수($S_{Danger}(W)$)는 0에서 100 사이의 값이 계산된다.

4. 구현 및 실험

4.1 유포 패턴 탐지 시스템 프로토타입 개발

제안한 악성코드 유포 패턴 분석 기반 탐지에 대한 성능 검증을 하기 위해 유포 패턴 탐지 시스템을 프로토타입으로 개발하였다.

시스템은 점검대상 웹사이트의 수집정보와 유포패턴

을 매칭하여 악성코드 유포지와 경유지를 탐지하며, 신규 유포패턴은 분석을 통해 악성 스크립트 패턴분석으로 지속적으로 포함시켜서 제로데이 공격에 대응하도록 개발하였다. 시스템 프로토타입에 의한 성능 실험의 프로세스는 Fig. 3과 같다.

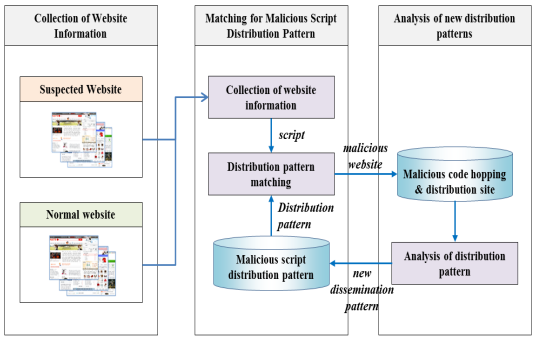


Fig. 3. Process of performance experiment using detection prototype system

웹사이트 정보 수집 단계에서 의심 웹사이트와 정상 웹사이트를 수집한 후 악성 스크립트 유포 패턴 매칭단계에서 악성 스크립트 유포 패턴과 매칭되는 악성 스크립트를 탐지하고, 신규 유포 패턴 분석 단계에서 악성코드 유포지와 경유지를 분석하여 그 분석 결과를 악성 스크립트 탐지룰 DB에 저장하여 관리한다.

Fig. 4는 개발된 프로토타입 시스템의 DLL 정보를 설정하는 화면으로 DLL 파일을 로드하여 각 DLL의 중요도와 가중치 및 임계값을 설정할 수 있다.

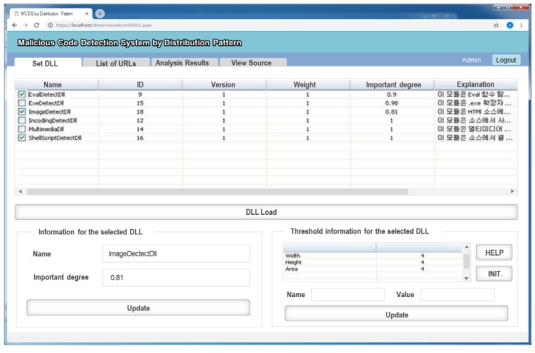


Fig. 4. DLL information setting of prototype system developed for experiment

Fig. 5는 분석결과 화면으로 각 분석 대상 URL별로 악성코드 유포 패턴을 탐지한 결과를 디스플레이한다. 결과로 출력되는 정보는 URL, 위험점수, 위험상태이며, 해당 URL의 위험점수가 임계 위험점수 보다 크면 악성코드 간주하여 붉은 색으로 표시된다.

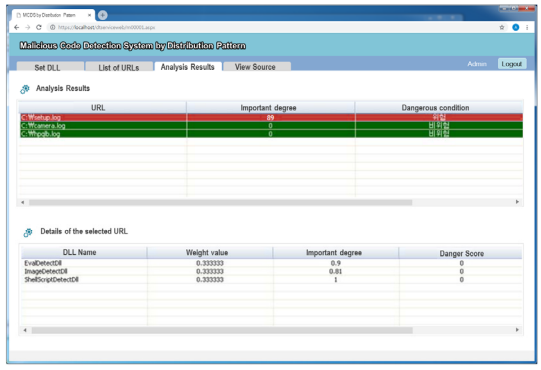


Fig. 5. Analysis result view of prototype system developed for experiment

Fig. 6은 악성 스크립트 소스 보기 화면으로 분석 결과 화면에서 분석 결과에 의해 나타난 URL들을 선택 시, 분석된 URL의 소스코드와 탐지된 패턴들을 확인할 수 있다. 이 때, 탐지된 패턴 들은 붉은색으로 표시된다.

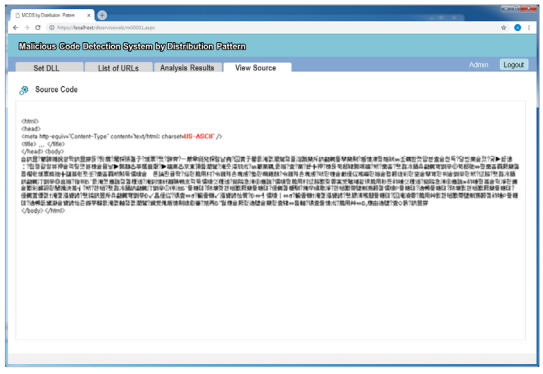


Fig. 6. Search source view of prototype system developed for experiment

4.2 실험 결과

정보보안기관이 공개한 국내외 악성코드 유포지로 확인된 총 390개의 웹사이트를 분석에 의해 도출된 악성 스크립트 유포패턴 탐지 항목으로 모두 점검하였으며, 본 탐지성능 실험에서는 계산된 위험점수가 60점 이상인 경우를 탐지로 설정하였다. 점검을 통해 도출된 각 점검 항목별 탐지 결과는 Table 4와 같다.

Table 4. Detection rate by diffusion pattern detection items

No.	Distribution Pattern	Detection rate	Number of detection	Important degree
1	Eval function	100%	110	89.6
2	Shell Script	100%	53	100
3	Us-ascii, Jscript.Encode Encoding Function	100%	20	0
4	img tag Abnormal width, height value	100%	3	80.6
5	Executable Webpage Containing Files	98.67%	74	98.5
6	Large-scale Special Character	98.00%	196	79.0
7	Web Pages containing Multimedia	97.62%	41	100
8	Large-scale String Processing Function	89.43%	110	96.7
9	User-defined String Processing Function	88.82%	143	72.1
10	Simple String Iteration	88.30%	83	96.3

Table 4에 나타나듯이 점검 항목 중 단순 문자열 반복, 쉘 스크립트, img 태그의 비정상 width, height 값 항목에 있어서 매우 높은 탐지율을 보였고, 전체 항목 평균 약 86%의 높은 탐지율을 보였다. Important degree는 제안하는 위험점수 계산식을 통해 부여된 점수로 패턴 항목에 대한 보안위협 수준을 나타낸다.

5. 결론

본 연구에서는 악성코드 탐지를 위해 기존의 시그니처 기반의 탐지 방식과 행위기반의 탐지 방식의 단점을 보완할 수 있는 악성 스크립트 패턴분석을 이용한 악성코드 탐지 기법을 제안하였다.

우선 악성코드 유포지의 스크립트 패턴을 분석하여 공통적인 특징을 분석과 분류를 통해 유포 패턴을 도출하였으며, 도출된 유포 패턴을 점검 항목으로 하여 웹사이트의 악성 스크립트를 탐지할 수 있도록 악성코드 유포 패턴 탐지 시스템 프로토타입을 개발하였다.

제안한 기법의 성능을 검증하기 위해 개발된 프로토타입 시스템으로 총 390개의 악성 웹사이트를 대상으로 악성 스크립트 탐지 항목을 실험한 결과, 전체 항목 평균 약 86%의 높은 탐지율을 보였다.

제안한 탐지 기법은 웹사이트를 통해 유포되는 악성 스크립트 유형을 분석하여 유포패턴을 도출한 후, 도출된 유포패턴을 탐지룰에 등록하여 점검하는 방식으로 기존의 시그니처 기반의 탐지속도를 유지하면서도 제로데이 공격에 대한 탐지가 가능한 장점이 있다.

기존에는 악성코드 은닉 웹사이트를 분석자에 의해 수동적으로 분석을 수행하였으나 본 연구에서 제안한 기법을 활용하면 보다 자동화된 악성코드 은닉 웹사이트를 탐지 성능이 향상될 것으로 기대한다.

References

[1] S. Y. Min, C. S. Jung, K. H. Lee, E. S. Cho, T. B. Yoon, S. H. You, "Design of Comprehensive Security Vulnerability Analysis System through Efficient Inspection Method according to Necessity of Upgrading System Vulnerability", *Journal of the Korea Academia-Industrial*, Vol.18, No.7, pp.1-8, 2015.
DOI: http://dx.doi.org/10.5762/KAIS.2017.18.7.1

[2] K. S. Jeong, S. Bae, H. Kim, "Evaluation Criteria for Suitable Authentication Method for IoT Service Provider in Industry 4.0 Environment", *Journal of the Society of Korea Industrial and Systems Engineering*, Vol.40, No.3, pp.116-122, 2017.
DOI: https://doi.org/10.11627/jkise.2017.40.3.116

[3] A. Mateen, Q. Zhu, S. Afsar, M. Usman, "IoT and Wireless Sensor Network Monitoring for Campus Security", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.18, No.6, pp.33-41, 2018.
DOI: https://doi.org/10.7236/JIIBC.2018.18.6.33

[4] Y. S. Kim, B. K. Lee, "CoAP/6LoWPAN-based Smart Home Network system using DTLS", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.18, No.6, pp.53-61, 2018.
DOI: https://doi.org/10.7236/JIIBC.2018.18.6.53

[5] S. T. Yu, S. H. Oh, "Malware Analysis Mechanism using the Word Cloud based on API Statistics", *Journal of the Korea Academia-Industrial*, Vol.16, No.10, pp.7211-7218, 2015.
DOI: http://dx.doi.org/10.5762/KAIS.2015.16.10.7211

[6] S. Y. Min, E. S. Cho, B. W. Jin, "A Implement of Integrated Management Systems for User Fraud Protection and Malware Infection Prevention", *Journal of the Korea Academia-Industrial*, Vol.16, No.12, pp.8908-8914, 2015.
DOI: http://dx.doi.org/10.5762/KAIS.2015.16.12.8908

[7] E. S. Lee, S. R. Kim, Y. K. Kim, "A Study on Enhancing Security Management of IT Outsourcing for Information System Establishment and Operation", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.17, No.4, pp.27-34, 2017.
DOI: https://doi.org/10.7236/JIIBC.2017.17.4.27

[8] H. T. Lee, "Analysis of Security Technology for Internet of things", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.17, No.4, pp.43-48, 2017.
DOI: https://doi.org/10.7236/JIIBC.2017.17.4.43

[9] H. H. Jung, H. Y. Kwon, "A Study on the Necessity of the Introduction of Professional Certification System for Financial Security", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.17, No.4, pp.209-218, 2017.
DOI: https://doi.org/10.7236/JIIBC.2017.17.4.209

[10] K. A. Yang, D. W. Shin, J. K. Kim, B. C. Bae, "Trend and Prospect of Security System Technology for Network", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.18, No.5, pp.1-8, 2018.
DOI: https://doi.org/10.7236/JIIBC.2018.18.5.1

[11] S. Y. Lee, J. Y. Kim, "Performance of privacy Amplification in Quantum Key Distribution Systems", *The Journal of The Institute of Internet, Broadcasting and Communication*, Vol.18, No.5, pp.111-116, 2018.
DOI: https://doi.org/10.7236/JIIBC.2018.18.5.111

[12] M. Ahmadi, D. Ulyanov, S. Semenov, M. Trofimov, G. Giacinto, "Novel feature extraction, selection and fusion for effective malware family classification", *Proceedings of the sixth ACM conference on data and application security and privacy*, pp.183-194. March 2016.
DOI: http://dx.doi.org/10.1145/2857705.2857713

[13] S. Hansen, S. T. Larsen, M. T. Stevanovic, J. M. Pedersen, "An approach for detection and family classification of malware based on behavioral analysis", *Proceedings of International Conference, In Computing, Networking and Communications(ICNC)*, IEEE, pp.1-5, Feb. 2016.
DOI: http://dx.doi.org/10.1109/ICCNC.2016.7440587

[14] Y. J. Ki, E. J. Kim, H. K. Kim, "A novel approach to detect malware based on API call sequence analysis", *International Journal of Distributed Sensor Networks*, Vol.2015, No.4, pp. 1-9, 2015.
DOI: https://doi.org/10.1155/2015/659101

[15] K. Rieck, T. Holz, C. Willems, P. Dusse, P. Laskov, "Learning and classification of malware behavior", *Proceedings of International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, Springer, pp.108-125, 2018.
DOI: https://doi.org/10.1007/978-3-540-70542-0_6

[16] Y. Fan, Y. Ye, L. Chen, "Malicious sequential pattern mining for automatic malware detection", *Expert Systems with Applications*, Vol.52, pp.16-25. 2016.
DOI: https://doi.org/10.1016/j.eswa.2016.01.002

[17] J. Saxe, K. Berlin, "Deep neural network based malware detection using two dimensional binary program features", *Proceedings of Malicious and Unwanted Software(MALWARE), 10th International Conference*, IEEE, pp.11-20, Oct. 2015.
DOI: https://doi.org/10.1109/MALWARE.2015.7413680

[18] B. Sun, Q. Li, Y. Guo, Q. Wen, X. Lin, W. Liu, "Malware family classification method based on static feature extraction", *Proceedings of 3rd International Conference, In Computer and Communications (ICCC)*, IEEE, pp.507-513. March 2017.
DOI: https://doi.org/10.1109/CompComm.2017.8322598

[19] S. Acharya, B. Ehrenreich, J. Marciniak, "OWASP inspired mobile security", *Proceedings of International Conference, Bioinformatics and Biomedicine(BIBM)*, IEEE, pp.782-784, 2015.
DOI: https://doi.org/10.1109/BIBM.2015.7359786

[20] P. Royal, M. Halpin, D. Dagon, R. Edmonds, W. Lee, "PolyUnpack: Automating the Hidden-Code Extraction of Unpack-Executing Malware", *Proceedings of 22nd Annual Computer Security Applications Conference (ACSAC'06)*, IEEE, pp.289-300, 2006.
DOI: https://doi.org/10.1109/ACSAC.2006.38

[21] W. Li, C. Li, M. Duan, "Method for detecting the obfuscated malicious code based on behavior connection", *Proceedings of 3rd International Conference on Cloud Computing and Intelligence Systems*, IEEE, Nov. 2014.
DOI: https://doi.org/10.1109/CCIS.2014.7175735

[22] A. Shabtal, R. Moskopvitch, C. Feher, S. Dolev, Y. Elovici, "Detecting unknown malicious code by applying classification techniques on opcode patterns", *Security Informatics*, Vol.1, No.1, 2012.
DOI: https://doi.org/10.1186/2190-8532-1-1

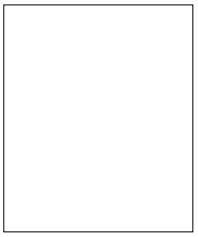
[23] C. She, Y. Ma, J. Wang, L. Jia, "An improved malicious code intrusion detection method based on target tree for space information network", *International Journal of Distributed Sensor Networks*, Vol. 13, No. 12, 2017.
DOI: https://doi.org/10.1177/1550147717747847

[24] D. D. Lille, B. Coppens, D. Raman, B. D. Sutter, "Automatically combining static malware detection techniques", *Proceedings of 10th International Conference on Malicious and Unwanted Software(MALWARE)*, pp.48-55, Oct. 2015.
DOI: https://doi.org/10.1109/MALWARE.2015.7413684

[25] P. Vinod, R. Jaipur, V. Laxmi and M. Gaur, "Survey on malware detection methods", *Proceedings of the 3rd hackers' workshop on computer and internet security*, pp.74-79, 2009.

[26] M. Egele, T. Scholte, E. Kirda, C. Kruegel, "A survey on automated dynamic malware-analysis techniques and tools", *ACM computing surveys (CSUR)*, Vol. 44, No.2, 2012.
DOI: https://doi.org/10.1145/2089125.2089126

이 용 준(Yong-Joon Lee) [종신회원]



- 2005년 2월 : 숭실대학교 컴퓨터학과 박사
- 2010년 2월 ~ 2016년 3월 : 한국인터넷진흥원 수석연구원
- 2016년 4월 ~ 현재 : 국방보안연구소 선임연구원

〈관심분야〉
산업보안, 사이버보안, 기밀유출차단

이 창 범(Chang-Beom Lee) [정회원]

- 2001년 8월 : 전남대학교 전산학과 이학석사
- 2005년 2월 : 전남대학교 전산학과 이학박사
- 2005년 3월 ~ 2006년 8월 : 울산대학교 연구교수
- 2006년 10월 ~ 현재 : 한국교통안전공단

〈관심분야〉
정보보호 정보통신