홈페이지에 삽입된 악성코드 및 피싱과 파밍 탐지를 위한 웹 로봇의 설계 및 구현

김대유, 강창구*, 김정태

(주)위너다임*, 목원대학교

Analyses of Detection Techniques of Malicious Code in the Homepage

Daeyu Kim, Chang-Ku Kang, Jung-Tae Kim

Mokwon University

E-mail : jtkim3050@mokwon.ac.kr

요 약

본 논문에서 제안하는 웹 서버 취약점 및 악성코드를 탐지하는 웹 로봇의 기술은 인터넷에서 개인정보보호사고의 원인분석을 통해 도출된 요구기능을 통합 구현하는 기술로 인터넷 이용자의 개인정보 피해 원인을 종합적으로 처리한다는 측면에서 효과가 크다. 인터넷에서 개인정보를 유출하는 홈페이지의 악성 코드 및 피싱과 파밍을 종합적으로 탐지기술을 구현함으로써 개인정보를 유출하기 위하여 사용되는 홈페이지의 악성 코드 및 피싱과 파밍 사이트로 유도되는 웹 사이트를 탐지 할 수 있다.

1. 서론

컴퓨팅 환경이 웹으로 전환됨에 따라 다양한 웹 기반 응용 프로그램들이 개발되고 이에 따른 사용자 수 증가로 웹을 통한 트래픽이 증가하고 있다. 또한, 고속 네트워크 기반시설에 힘입어 인터넷 사용자가 전 세계적으로 급증하게 되었고, 유비쿼터스 환경의 도래와 함께 네트워크 및 컴퓨팅 환경은 언제 어디서나 네트워크에 연결되어 컴퓨팅을 수행할 수 있는 웹 컴퓨팅 환경으로 변화해 가고 있다. 오늘날 웹은 게임, 뱅킹, VOD, VoIP 등의 새로운 어플리케이션과 프로토콜의 출현으로 인해 매우 복잡해지고 세분화되는 경향을 보이고 있다. 이에 따라 다양한 웹 기반 응용 프로그램들이 개발되고 사용자 수가 증가함에 따라 웹을 통한 트래픽이 증가하게 되었다. 웹 기반 서비스(WWW, Web mail, VOD, P2P, IM 등)의 증가는 유해트래픽의 다양한 통로로 사용되게 되고, 내부 정보 유출 등 피해가 확산되고 있다. 웹 기반 서비스(WWW, Web mail, VOD, P2P, IM 등)의 증가는 엄청난 기회를 창조하였지만 이와 더불어 유해 트래픽의 통로로 다양하게 사용됨으로써 불안정한 네트워크 환경을 조성하는 등 어려운 과제를 안겨주었다. 이 같은 유해 트래픽의 확산은 네트워크에 직접적인 피해를 유발할 뿐 아니라, 최근에는 내부 정보 유출로까지 이어지고 있어 심각성이 나날이 증가하고 있다. 웹(World Wide Web)은 대부분이 정보제공이나 서비스 제공을 목적으로 하기 때문에 최근 해킹 사고의 대부분이 웹을 통해 이루어지고 있다. 2004년 말부터 급증하고 있는 웹페이지 변조공격과 같이 웹 어플리케이션의 취약점을 노린 해킹 사건들이 다수를 이루고 있으나 다른 인터넷 서비스들과는 달리 접근제어나 방화벽 등으로 보호가 어렵다. 결국 웹기반 서비스를 효과적으로 보호하기 위해서는 웹기반 서비스에 특화된 웹 어플리케이션 보안기술과 도구들이 필요하다. 응용서비스가 웹으로 변화함에 따라 주로 사용자의 개인정보 유출에 그 목적을 두고 있는 스파이웨어를 전파하는 새로운 기법으로 웹사이트를 이용하는 경향이 있다. 최근 홈페이지 변조사고는 단순 홈페이지 초기화면 변조보다 접속자 수가 많은 웹사이트를 해킹한 후 악성코드를 업로드하여 방문자 PC를 감염시킨다. 감염 후 특정 사이트에 접속하여 로그인할 경우 계정 및 비밀번호 정보가 공격자에

게 유출되도록 하는 기법들을 사용한다. 웹 어플리케이션의 보안 솔루션으로 제안되고 있는 웹 방화벽의 경우, 프로그램 개발자의 실수를 룰에 의해 막아주는 기법으로 제안되어, 근본적인 문제 해결이 될 수 없다. 웹 서버상의 잘못된 설정이나, 논리적인 프로그램의 오류로 인한 취약점을 점검하고, 룰에 의한 침입을 차단해 주는 수준이다. 특히, OWASP(웹 어플리케이션 10대 취약점)에 대한 취약한 요소를 차단하는 수준이다.

2. 국내외 연구 동향

2.1. 국내의 경우

안티스파이웨어 제품 활용 외에 스파이웨어를 차단하기 위해 아래와 같은 방법을 활용할 수 있다고 연구되었다. 첫째, 스파이웨어로 의심되는 소프트웨어를 직접 차단하는 방법으로 네트워크 단계에서 네트워크 서비스 차단자가 직접 차단하는 방법이 있다. 그러나 이 경우는 사용자의 권리 침해나 차단 스파이웨어의 범위 등 다양한 문제 해결이 선행되어야 한다. 단, 사내 네트워크와 같이 제한된 범위에서 적용할 경우, 개별 컴퓨터보다 효율적인 수단이 될 수 있다는 점에서 고려되고 있다. 둘째, 스파이웨어의 침입이 가능한 운영체제 기능을 제한함으로써 운영체제 단계에서 스파이웨어의 설치를 제한하는 방법도 논의되고 있으나 이 방법도 운영체제 개발회사에서 기능을 제공해야 하고 사용자의 활용 능력이 요구된다는 점에서 활성화되기 어려운 점이 있다. 셋째, 침입차단시스템(Firewall) 등 기존의 보안제품을 활용하여 내부 네트워크에서 외부로의 불법적인 정보 송신을 통제함으로써 스파이웨어를 차단하는 방법이 있다. 스파이웨어도 바이러스, 웜과 같은 악성 프로그램의 일종으로 스파이웨어의 알려진 악성 기능도 기존의 보안 프로그램을 활용한 기술적 조치로 차단이 가능하기 때문이다.

2.2. 국외의 경우

미국의 NSA(National Security Agency)는 스파이웨어를 비롯한 악성 프로그램의 악성 기능에 대한 계층별 방어 수단을 (그림4)와 같이 구분하고 있다. 1998년 악성 이동코드의 침해 위협에 대응하고자 인터넷 보안 소프트웨어 개발 업체들이 악성 이동코드 컨소시엄을 구성하여 컴퓨터바이러스 업체인 시멘텍, 트랜드 마이크로등이 영역 확장 차원에서 이동코드 탐지 기술개발을 시작하였다. 이 같은 보안 소프트웨어 개발 업체 및 백신 개발 업체들의 이동코드 탐지 기술 개발 시작은 추후 이동코드 생산 제품에 대한 유해성 케스트 및 그에 다른 인증 절차 등의 알고리즘을 개발하고 있다.

<표 1> 악성 기능에 대한 계층별 방어 매트릭스

구분	안티 바이러스	네트워크 단계			OS단계	
		firewall	IDS	IPS	무결성 체크	stack guard
네트워크 취약성 공격(감염)		O	O			O
사용자 속이기(감염)		O				
원격관리상 취약성 공격(감염)		O				
사전설치된 백도어를 이용한 공격(감염)	O	O				
파일시스템 변경				O	O	
시스템 설정 변경				O	O	
일부 프로세스 수정				O		
네트워크 접근	O			O		
시스템 권한 획득				O		
권한된 명령 수행				O		
주요 API 후킹				O		
네트워크 트래픽 폭주 야기		O		O		
운영시스템 속도 저하				O		
시그너처 포함	O	O	O			

3. 웹 유해코드 공격방법과 유형

APWG에 접수된 보고(Phishing Archive)분석을 통해 피싱 사이트의 URL 표현 방법을 크게 세 가지로 분리된다.

<표 2> 악성코드의 형태

구 분	유 형
Type 1	링크 표현은 정상적인 사이트를 나타내고 있으나 실제 접속 시 브라우저의 주소창에 표시되는 내용
Type 2	URL이 http://www.usbank.com 인데 악성 사이트의 URL은 http://www.us-bank.com 로 표현
Type 3	사용자가 악성 사이트로 접속 시 자바스크립트를 내려 받게 하여 브라우저의 URL 주소를 정상적인 URL것처럼 보이게 함

자료출처 : APWG

<표 3> 웹 유해코드의 공통점

공통점	1. Type1,2의 경우는 URL형태가 변하는 경우 2. Type3의 경우는 URL의 경우는 같

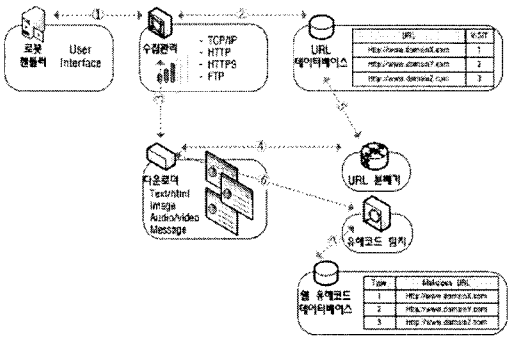
	지만 접속하는 IP가 다른 경우 유해코드가 삽입 된 사이트로 유도하여 악성코드가 삽입된 웹 페이지에 접속하게 하여, 사용자의 정보를 전송하는 방식은 모두 일치

<표 4> 웹 유해코드의 주요 특징

악성코드 종류	주요 특징
HangHack Trojan	- 감염되면 syshlp.dll 파일이 생성되고, iexplore.exe에 인젝션되어 실행 - 온라인게임 계정 및 패스워드를 유출 (smtp.tom.com:25 연결 후 메일전송)
GrayBird.bk Backdoor	- 리버스커넥션하는 백도어 프로그램 - 감염되면 svchost.exe와 랜덤한파일명.dat(ex. yiykbs.dat) 두개의 파일이 생성되고, 랜덤한파일명.dat는 다른 프로세스에 인젝션되어 실행됨 - vip.huigezi.com(61.152.93.13)로부터 특정파일 다운로드하고, 222.181.170.35:21 또는 222.181.169.159:21로 리버스커넥션 요청함 - 은폐형으로 동작하여 프로세스 목록이나 파일보기에서 확인 불가능 (c:\windows\dhcp 폴더 내에 백도어 파일이 존재하지만, 보이지는 않음) - GrayBird Client를 이용하여 시스템 정보 및 파일유출 가능함
phel.q Exploit	- MS04-013, MS05-001 취약점을 이용하는 Exploit 코드 - Psyme Trojan Downloader 파일을 다운로드 후 실행시킴
phel.c Exploit	- MS05-001 취약점Exploit 코드 - HangHack Trojan 파일 다운로드 후 자동설치
Psyme Trojan Downloader	- MS05-001 취약점 Exploit 코드 - GrayBird.bk Backdoor를 시스템에 자동설치하기 위한 도움말 파일

4. 웹 유해코드 탐지방법

웹 수집 로봇의 수집과정에 웹 페이지의 소스코드를 다운로드하는 과정을 거치게 되는데, 해당 과정에서 HTML 소스코드에서 악성 코드를 검출하는 기법을 도입하여 (그림 3. 유해코드를 탐지하는 웹 수집로봇의 구성도와 같이) 웹 유해코드를 검출 할 수 있다.



(그림 1) 유해코드를 탐지 하는 웹 수집로봇의 구성도

웹 유해코드는 페이지를 악성 사이트로 연결시키는 역할을 하고 있다. 이 연결될 수 있는 방법은 HTML 태그로 사용 되거나 자바스크립트를 이용하여 가능하다.

<표 5> 다른 페이지를 불러올 수 있는 태그 및 스크립트

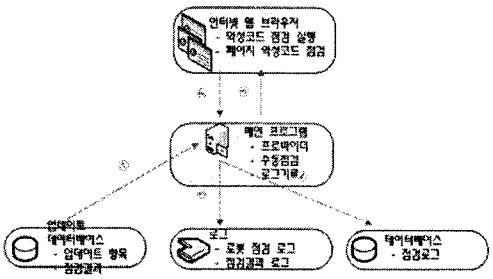
```
<IFRAME SRC= "http://www.link.com" >
<SCRIPT SRC = "http://link.com">
<OBJECT SRC = "http://link.com">
```

또한, 자바스크립트 함수를 이용할 경우에는 다른 현재 나타내있는 페이지에서 수많은 페이지를 불러오도록 할 수도 있다.

우선적으로 이 논문에서 제안하는 악성코드 탐지 규칙은 다음과 같이 5가지형태로 구분하였다.

1. 레퍼런스 웹 주소와 다른 웹페이지로 연결되는 링크들의 국가정보가 틀린 경우
2. 레퍼런스 웹 주소와 다른 웹페이지로 연결되는 링크들의 IP 대역이 다른 경우
3. 성데이터 베이스에 다른 웹페이지로 연결되는 링크가 포함되어 있는 경우
4. 악성 데이터 베이스에 다른 웹페이지로 연결되는 링크의 IP가 포함되어 있는 경우
5. 악성코드에 사용되는 스크립트 함수가 포함되어 있는 경우

위 5가지 조건 중, 1가지 형태가 검출된다면 악성사이트로 의심하고 보고하는 에이전트 형태로 진행을 계속 하였다. 다음 (그림 6) 유해코드를 탐지하는 에이전트의 동작과정의 데이터베이스 점검로그에 해당 기록을 남기도록 한다.



(그림2) 유해코드를 탐지 하는 에이전트의 동작과정

1. 업데이트 서버에는 (표 8) 악성코드 업데이트 항목을 업데이트 한다.

<표 6> 악성코드 업데이트 항목

형태	주소
피싱	http://site.co.kr/page.jsp
파밍	http://site.co.kr/page.asp
악성코드	http://site.co.kr/page.php
성인	http://site.co.kr/page.js
유해	http://site.co.kr/page.html

2. 웹 수집 로봇 핸들러로는 프로바이더(중앙제어) 모듈로 점검 대상사이틀르 전달한다.

3. 메인 프로그램의 프로바이더는 해당 사이트의 점검로그와 점검결과 (표 9)와 같이 나타낸다.

<표 7> 검출결과 내역

시각	형태	검출내역
08/9/28 2:00	<Iframe>	http://host.co.kr
08/9/28 2:00	<Script>	http://host.com
08/9/28 2:00	<Object>	http://host.co.kr

5. 결론

웹 유해코드를 탐지 하는 기법으로 은닉된 변종 악성코드 검출로 변종 악성코드 확산을 원천적으로 차단하는 기술과 웹 환경의 안전성 확대를 통한 사이트의 신뢰성이 증가하여 해당 사업분야의 관련 기술 발전을 촉진할 것이다. 또한 웹 컨텐츠의 안전성을 제공하므로 다양한 형태의 컨텐츠 기술 발전에 긍정적인 영향을 미칠것으로 예상되며, 본 에이전트의 기술은 PC환경의 활용 가능하다. 본 논문에서는 홈페이지에 삽입된 악성코드 연구를 통하여 산업 발전의 기여도 등 국가경제에 미치는 효과로 보아 앞으로 "홈페이지에 삽입된 악성코드 탐지기법 분석" 연구가 더욱 필요할 것으로 예상된다.

참고문헌

[1] http://cafe.naver.com/jmkim9064/778

[2] Li Zhuowei, etcs, "Utilizing Statistical Characteristics of N-grams for Intrusion Detection", Proceedings of the 2003 International Conference of Cyberworlds, pp.212-218.

[3] www.itfind.or.kr, "유비쿼터스 사회의 사이버 공격 기술 동향", 권호: 1259 발행일: 2006.08.16

[4] Joon S. Park and Gautam Jayaprakash, "Component Integrity Check and Recovery Against Malicious Codes" Proceedings of the 20th International Conference on Advanced Information Networking and Applications (AINA"06)

[5] Tony Abou-Assaleh, "N-gram-based Detection of New Malicious Code", Proceedings of the 28th Annual International Computer Software and Applications Conference

[6] Frank Adelstein, Matt Stillerman, "Malicious Code Detection for Open Firmware", Proceedings of the 18th Annual Computer Security Applications Conference (ACSAC.02)

[7] A. Murat Fiskiran, "Runtime Execution Monitoring (REM) to Detect and Prevent Malicious Code Execution, Proceedings of the IEEE International Conference on Computer Design (ICCD'4)