효율적인 정규식 표현을 이용한 XSS 공격 특징점 추출 연구1)

허승표* 이대성** 김귀남*
*경기대학교 산업보안학과
**경기대학교 산업기술보호특화센터
e-mail:huhspunk@gmail.com

A Study on XSS Attacks Characters, Sample of Using Efficient the Regular Expressions

Seung-Pyo Huh* Dae-Sung Lee** Gui-Nam Kim*
*Dept of Industry Security, Kyonggi University
**Center for Industry Security, Kyonggi University

요 약

OWASP에서 발표한 2007년 웹 애플리케이션 취약점 중 하나인 XSS 공격이 사용자 브라우저에서 스크립트를 실행하게 함으로써 사용자의 세션을 가로채거나 웜을 업로드하여 악성코드를 삽입하는 공격이다[2]. 하지만 많은 XSS 방어 기법에서는 단순 스크립트 우회기법과 강제적인 스크립트 차단 방법을 채택하고 있다. 또한 강제적인 XSS 필터 적용으로 과탐지로 인한 정상적인 웹 페이지가 출력 되지 않는 사례가 나타나고 있다. 따라서 본 연구는 효율적인 정규식을 이용하여 XSS 공격 특징을 분석하여 특징점들을 추출하고 이 특징점들을 기반으로 특정한 규칙을 가진 문자열들을 모든 문자가 유효한지 확인할 수 있는 정규식 표현 방법을 이용하여 다양한 응용프로그램에 적용할 수 있는 기술을 연구하고자 한다. 또한 이를 기반으로 포털 사이트와 브라우저에서 제공하는 XSS 필터들과 비교하여 과탐지율 및 오탐지율 서로 비교하여 본 연구가 효율성 면에서 효과가 있는지 우위를 둘 것이며, 브라우저 벤더, 포털 사이트, 개인 PC 등 충분한 시험 평가와 수정을 통해서 응용할 수 있는 계기를 마련할 것이다.

1. 서론

최근 대두되고 있는 보안사고와 관련하여 악성코드를 이용한 사용자PC 보안 위협사례가 증가되고 있다. 이러한 심각한 보안 위협으로 2009년 7월 7일 발생한 대대적인 DDoS 공격은 국가, 기업, 개인에게 엄청난 손실을 입히게 되었다. 특히 웹을 겨냥한 공격이 늘어나고 있고 공격의 형태로는 웜, 악성코드, 봇넷 등 해커의 별도 명령 없이 몇 시 몇 분에 어느 목표로 지정하면 자동적으로 위협을 가하는 지능적인 공격 형태를 띠는 위협이 증가하고 있다.

웹 애플리케이션 취약점 해킹으로는 SQL-injection, XSS(크로스 사이트 스크립팅), 인증우회, 디렉터리 리스팅, 디렉터리 탐색, 파일 업로드가 있으며, 현재 XSS를 이용한 공격이 급증하고 있다[1]. 웹 애플리케이션 취약점을 이용한 해킹 중 XSS(Cross Site Scripting) 공격은 웹 애플리케이션에 클라이언트로부터 악의적인 데이터를 받아들일 때 제 3자의 사용자PC로 악성코드를 다운로드 받게 함으로써 악성코드 삽입뿐 만 아니라 다양한 공격을 추가하여 공격을 할 수 있다.

1) 본 연구는 지식경제부 지역혁신센터사업인 산업기술보호특화센터 지원으로 수행되었음.

2. 관련연구

가. 파라미터 매치 방법을 이용한 필터 방식

기존의 연구방법에는 게시물 입력, 댓글 작성 등에 입력값 검증 및 출력 값 검증을 실시하고 XSS 공격이 용이한 주요 메뉴에 대해서 프로그램 소스 내 입력 값 검증 필터를 삽입하여 악성 스크립트를 입력하지 못하게 하고 있으며 신규 악성스크립트 패턴을 이용해 이를 우회한 경우에도 출력 값을 검증 필터를 즉시 적용하여 XSS 공격을 대비하고 있다[2]. 하지만 XSS 필터를 자칫 잘 못 적용할 경우 과탐지로 인한 올바른 스크립트도 차단하는 경우가 생길 수도 있다. 기존에 있는 브라우저 XSS 필터 기능에는 100%차단을 적용하고 있어 사용자로서는 올바른 웹 페이지를 보기위해 해당 기능을 해제해야 하는 번거로움과 필터로서의 제 역할을 못하는 사례도 있다.

나. HTML 코드 변환 방법을 이용한 필터 방식

어플리케이션이 모든 header, cookie, query string, form field, hidden field 등 모든 parameter들에 대해 유효성 검사를 실시하고 있다. 사용자의 입력이 화면에 출력되어 나

오는 경우, 사용자의 입력 값을 HTML Entity로 변환시켜야 한다. 또한 위의 태그들을 HTML으로 인식시키지 않기 위해 스크립트에서는 함수를 지원한다. 단순히 특별 문자를 식별하여 인코딩하는 기법은 매우 위험할 수도 있다. 공격자는 변환된 HTML코드를 쉽게 디코딩하여 우회할 수 있기 때문이다.

다. 유효성 검사 컨트롤 및 클래스를 이용한 필터 방식

마이크로소프트사에서 사용하고 있는 유효성 검사를 살펴보자면, 위에 사항과 같이 단순히 필터를 적용하지 않았다. 허용하는 HTML태그를 정의해 놓고 RegularExpressionValidator 컨트롤 또는 Regex 클래스를 통해 정규식을 사용하였다. 하지만 정규식이 복잡해지면 오류가 높아지고 이에 따라 공격에도 취약해 진다는 것을 지적하였다. 구체적인 사항을 언급하지 않아 분석하지는 못했지만 복잡한 필터를 적용하면 할 수록 처리하는 양도 많아지기 때문에 그만큼 위험 노출되고 오히려 비효율적인 처리가 될 수도 있다.

3. 제안하는 방법

본 연구에서 효율적인 정규식을 이용한 XSS 공격 특징점을 추출하기 위해서 웹 애플리케이션 취약점 공격 중 하나인 XSS 공격에 대해 이해하고 여러 스크립트들의 형태를 비교 분석 할 것이며 효율적인 정규식을 이용하여 XSS 공격의 특징점을 추출 할 것이다. 본 연구를 성공적으로 완수하기 위해 다음과 같은 전략을 수립하여 추진하기로 했다.

□ XSS 공격 분석 가이드 라인 설정 : 현 웹 애플리케이션 취약점 분석을 바탕으로 XSS 공격의 형태를 비교 분석한다.

□ 효율적인 정규식 표준화 정립 : XSS 공격 특징점을 주어진 문자, 문자열, 문자열 집합과 패턴들을 가지고 각각의 생성규칙에 해당하는 값을 구하고 정의된 정규문법을 이용하여 결과물을 추출할 수 있는 효율적인 표준화를 정립한다. 정규표현을 정립하기 위해선 다음과 같은 방식으로 기술 할 수 있다.[3]

- 역참조 조건으로 조건 처리
- 오토마톤이 받아들이는 문자열. 오토마톤에는 튜링기계, 유한상태기계 등이 있다.

[그림1]은 (<[Aa]\s+[^>]+>\s*)?는 <A>나 <a> 시작태그가 일치한다 이때 속성이 있다면 속성도 함께 일치하며, 물음표는 패턴이 일치하면 있다는 것을 의미한다. <[Ii][Mm][Gg]\s+[^>]+>는 <IMG> 태그를 대소문자 구별없이, <IMG> 태그에 속한 속성도 모두 포함해 일치한다. (?(1)\s*</[Aa]>)는 조건으로 시작하는, ?(1)은 역참조 1(<A> 시작 태그)이 있을 때만 수행하라는 의미이다. 이와 같이 같은 맥락의 <IMG SRC=javascript:alert!('XSS')> 와 같은 XSS 샘플 코드도 역참조 정규표현식을 이용한다면 모든 스크립트들에 대해서 특징점을 추출할 수 있다.[4]

```
일반적인 태그
<!-- Nav Bar -->
<TD>
<A HREF='/home">
<IMG SRC="/images/home.gif"></A>
<IMG SRC="/images/spacer.gif">
<A HREF='/search">
<IMG SRC="/images/search.gif"></A>
<IMG SRC="/images/spacer.gif">
<A HREF='/help">
<IMG SRC="/images/help.gif"></A>
</TD>
정규표현식
(<[Aa]\s+[^>]+>\s*)?<[Ii][Mm][Gg]\s+[^>]+>(?(
1)\s*</[Aa]>)
```

[그림1] 역참조 정규표현식

4. 결론

본 연구에서는 XSS 공격 특징점들의 추출을 정규식으로 적용하여 함수와 클래스 형태가 아닌 실제 공격대상인 Java Script, VBScript, ActiveX, HTML, Flash 스크립트들의 비교 및 분석하여 비교적 가볍고 처리속도를 향상시킬 수 있는 정규식을 정립하여 적용할 것이다. 기존의 공격 방어는 패턴 매칭 방법을 이용해 마치 시그니처 방식과 유사하다고 볼 수 있다. 해당 시그니처가 있으면 방어를 하고 없으면 공격에 노출 될 수 있는 허점이 있다. 본 연구는 단순히 시그니처 방식이 아닌 XSS 공격 특징을 분석하여 XSS 공격 특징점들을 추출하고 정규식 표현 방법을 이용하기 때문에 다양한 응용프로그램에 적용하기 쉽다는 장점을 이용하여 XSS 공격으로부터 방어 할 수 있는 기술을 연구한다.

5. 참고문헌

[1] The Open Web Application Security Project, "OWASP TOP 10 Project", http://www.owasp.org/.

[2] Philipp Vogt, Philipp Vogt, Nenad Jovanovic, Engin Kirda, Christopher Kruegel, and Giovanni Vigna, "Cross-Site Scripting Prevention with Dynamic Data Tainting and Static Analysi", in Proceeding of the Network and Distributed System, 2007

[3] 정규표현식, http://ko.wikipedia.org

[4] Ben Forta, "Regular Expressions in 10 Minutes", pp. 99-100, 2004.