악성코드 탐지 시스템 Web-Anti-Malware

Web-Anti-MalWare Malware Detection System

저자
(Authors)

정승일, 김현우
Seung-il Jung, Hyun-Woo Kim

출처
(Source)

한국컴퓨터정보학회 학술발표논문집 22(2), 2014.7, 365-367 (3 pages)
Proceedings of the Korean Society of Computer Information Conference 22(2), 2014.7, 365-367 (3 pages)

발행처
(Publisher)

한국컴퓨터정보학회
The Korean Society Of Computer And Information

URL

http://www.dbpia.co.kr/Article/NODE06603197

APA Style

정승일, 김현우 (2014). 악성코드 탐지 시스템 Web-Anti-Malware. 한국컴퓨터정보학회 학술발표논문집 , 22(2), 365-367.

한국컴퓨터정보학회 하계학술대회 논문집 제22권 제2호 (2014. 7)

악성코드 탐지 시스템 Web-Anti-Malware

정승일°, 김현우*
°*카이스트 사이버보안연구센터
e-mail:sijung@kaist.ac.kr°, babisss@kaist.ac.kr*

Web-Anti-MalWare Malware Detection System

Seung-il Jung°, Hyun-Woo Kim*
°*Cyber Security Research Center, KAIST

● 요 약 ●

최근 웹 서비스의 증가와 악성코드는 그 수를 판단 할 수 없을 정도로 빠르게 늘어나고 있다. 매년 늘어나는 악성코드는 금전적 이윤 추구가 악성코드의 주된 동기가 되고 있으며 이는 공공기관 및 보안 업체에서도 악성코드를 탐지하기 위한 연구가 활발히 진행되고 있다. 본 논문에서는 실시간으로 패킷을 분석할수 있는 필터링과 웹 크롤링을 통해 도메인 및 하위 URL까지 자동적으로 탐지할 수 있는 악성코드 탐지 시스템을 제안한다.

키워드: 악성코드(malware), 탐지(Detection), 웹 크롤링(web crawling)

I. 서 론

최근 악성코드는 갈수록 진화하고 있으며 그 피해는 갈수록 늘어나고 있다. 특히 웹 서비스의 사용이 증가함에 따라 개인정보의 유출 및 금융사기 등 피해사례가 갈수록 늘어나고 있다. 한국 인터넷 진흥원 분석 보고서에 따르면 2013년 한해동안 탐지된 악성코드 은닉사이트(경유지, 유포지)는 17,750건으로 2012년 대비 36%(13,018건) 증가하였다[1]. 갈수록 늘어나는 악성코드를 대비하기 위하여 악성코드 자동 분석 시스템이 주목 받고 있으며, 이에 따라 본 논문에서는 악성코드 뿐만 아니라 사용자가 판단하여 피해를 줄 수 있는 코드는 모두 검사 할 수 있는 자동 분석 시스템을 제안하고자 한다. 시스템은 기존 시스템과 달리 시그니쳐 기반의 악성코드 탐지 뿐만 아니라 기존의 탐지된 악성코드의 특징을 도출하여 악성코드 탐지율 을 높였으며 또한 웹 크롤링을 이용하여 사용자가 의심되는 도메인 뿐만 아니라 도메인의 하위 URL까지 분석하여 미연에 악성코드를 방지할 수 있게 하였다.

본 논문은 구성은 다음과 같다. 2장 악성코드 탐지 관련연구를 살펴본다. 3장 에서는 제안한 시스템의 구성 및 필터링의 기능, 웹 크롤링 기법에 대해 알아보며, 4장 에서는 결론을 맺는다.

II. 관련 연구

악성코드는 O/S 및 IE(Internet Explorer) 등의 취약점을 분석하여 다양하게 진화하고 있기 때문에 신속한 분석이 요구된다. 현재 악성코드 분석을 위해서는 인적자원이 사용되고 있으며 그로 인해 많은 시간이 소요된다. 이러한 약점을 보완하기 위해서는 자동화된 악성코드 분석 시스템이 요구된다. 악성코드 분석을 위해서 다양한 방법이 개발되고 있으며 그 방법들을 조합하여 더욱 진화된 탐지 기법이 만들어지고 있다.

크롤링 엔진을 이용하여 SEED URL에 파생된 다양한 URL을 추출하여 악성코드를 분석하는 방법이 있다. 이 방법은 정확한 탐지율이 다소 떨어질 수 있으나 신속하게 악성코드를 분석하고 탐지 할 수 있기 때문에 악성코드 공격을 사전에 차단할 수 있다.

커널 레벨의 URL 필터링 방법은 빠른 속도로 URL을 비교하여 검사하고 차단 할 수 있다. 대부분의 URL 필터링 기법은 어플리케이션 레벨에서 동작하기 때문에 처리 속도가 느리다.

III. 본 론

Web-Anti-MalWare 시스템은 악성코드의 차단 및 실시간 탐지를 목표로 하는 미러링 장비이다. 본 장비는 실시간으로 네트워크 패킷을 확인. 접속하려는 URL 페이지의 악성유무를 확인하며 크롤링 기반으로 악성코드를 미연에 방지 및 악성코드를 수집하여 사용자가 악성행위를 하는 페이지에 접속하지 않도록 한다.

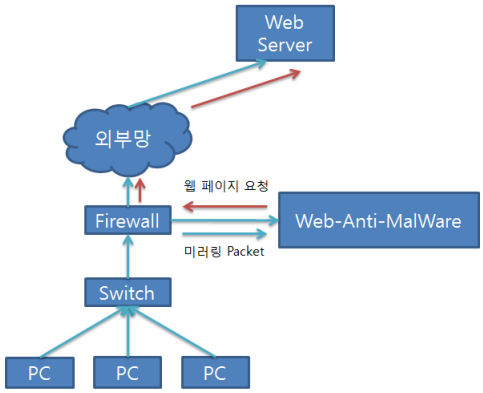

그림 474 system composition

최초 Client PC에서 웹 페이지에 접속했을 경우 Web-Anti-MalWare 장비는 Client PC에서 요청한 패킷을 미러링 받는다. 미러링 받은 패킷을 Web-Anti-MalWare에서 1차적으로 커널 URL 필터링 에서 필터링 하며 2차적으로 응용 실시간 필터링에서 URL의 웹 서버에 접속하여 악성코드가 심어져 있는 URL인지 판단한다.

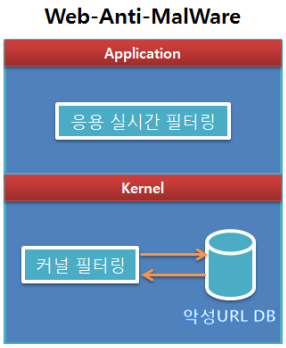

그림 475 filter structure

3.1 URL 필터링(Kernel Level)

커널 URL 필터링은 기존의 응용 어플리케이션에서 탐지하던 시스템들과는 달리 커널에서 악성링크를 탐지한다. 최초 인터페이스에서 들어온 패킷의 URL을 추출하여 사용자가 악성링크로 지정한 URL과 기존의 악성코드로 탐지된 URL의 패턴을 비교하여 악성링크 여부를 판단하여 필터링 한다.

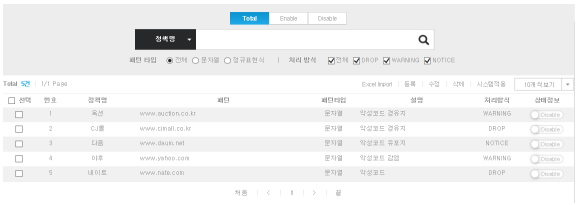

그림 476 URL Filter screen

커널에서 패킷 단위 분석을 통해 HTTP 헤더를 분석하여 URL을 추출하기 때문에 빠른 속도로 처리가 가능하다. URL을 추출 후에 악성링크로 등록된 데이터와 비교하여 매칭하여 처리하는 것도 빠르게 처리할 수 있다.

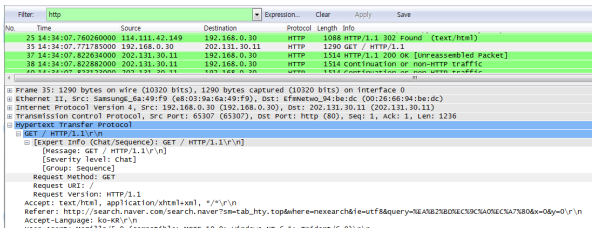

그림 477 URL for HTTP Header information extraction

사용자가 악성링크에 등록된 사이트에 접속시 빠르게 분석 및 비교하여 사용자가 악성링크에 접속했음을 알 수 있다. 미러링을 이용한 시스템이기 때문에 Inline 시스템과 같이 직접적인 차단은 할 수 없으나 방화벽에 악성링크에 접속했음을 알려주거나 RST(Reset) 패킷을 보내 접속을 막을 수 있다.

그림 478 URL filtering results screen

3.2 응용 실시간 필터링

커널 필터링에서 URL에 대한 필터링이 되지 않았다면 분석 할 URL을 응용 실시간 필터링에서 분석한다. 직접 해당 URL 웹 페이지를 요청하여 웹 서버에서 정상적으로 응답이 오면 웹 페이지 악성코드의 유무를 5가지의 필터조건에 의해 판단한다.

- 시그니처 탐지
- 라인당 글자수 초과 탐지
- 비정상 문자열 포맷 스트링 탐지
- 숨김 iframe 탐지
- 굼김 applet 탐지, 숨김 object 탐지

필터 조건에 의해 악성코드라고 판단될 시 실시간으로 연동되어있는 FireWall에 해당 URL을 차단할 수 있도록 정보를 제공해주며, 관리자 알람을 설정하여 관리자가 조치를 취할 수 있도록 한다. 또한 데이터베이스에 저장된 악성코드의 정보(탐지명, 탐지형태, 탐지내용, 출발지 IP / PORT, 목적지 IP / PORT, 국가)를 이용하여 악성코드의 유형 등을 파악. 추후 변종 악성코드를 탐지 및 통계자료로 쓰일 수 있도록 한다.

그림 479 Malware detection screen

3.3 웹 크롤링

크롤링이란 웹 애플리케션의 링크로 해당 애플리케이션에 노출되는 자원을 수집하는 것을 말한다. 해당 시스템은 악성코드를 미연에 방지하기 위하여 크롤링을 사용하였다.

사용자가 지정한 도메인 페이지부터 시작하여 페이지 안에 URL들을 수집 및 해당 페이지가 악성코드가 심어져 있는지 응용 실시간 필터에서 확인한다. 응용 실시간 필터링에서 확인 후 URL을 추출하여 리스트를 만든다. 이 리스트 안에 URL에 접속하여 해당 페이지 안에 링크들을 수집 및 URL리스트에 저장한다. 리스트 저장 과정에서 중복 검사하여 동일한 URL이 저장되지 않도록 한다.

크롤링은 5가지 컴포넌트로 구성되어있다.

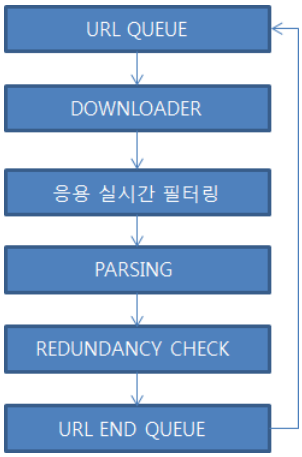

그림 480 Crawl component order

- URL QUEUE : URL 필터링에서 매칭되지 않은 URL, 응용 실시간 필터링에서 크롤링 하기 위한 최초 URL 리스트 이다.
- DOWNLOADER : URL QUEUE에서 URL을 가지고와 해당 웹 서버에 접속하여 정상적으로 응답을 받으면 해당 URL의 웹 페이지를 다운로드 받는다.
- PARSING : 다운로드 받은 웹 페이지의 소스에서 각 Tag(a tag, area tag, frame tag, iframe tag)를 파싱하여 각 tag를 저장 및 type에 따른 URL을 추출한다.
- REDUNDANCY CHECK : Redundancy Check는 추출된 각 tag별로 추출된 URL을 저장 시 URL END QUEUE에 있는 리스트와 매칭하여 이미 분석된 URL은 제외. 분석 및 크롤링 되지 않은 URL만 저장한다.
- URL END QUEUE : 크롤링 하여 분석된 URL 리스트 이다.

그림 481 Web crawling screen

IV. 결 론

본 논문에서는 악성코드에 의한 감염을 방지하기 위해 지속적으로 링크를 검사하여 악성코드를 분석하고 탐지하는 시스템을 제시하였다. 공격을 받은 후 치료하고 차단하는 방어적인 자세에서 능동적으로 검사하고 사전에 차단하여 대응하는 시스템이다. 크롤링 기술을 이용하여 지속적이고 방대한 데이터를 수집 및 분석하고, 탐지된 악성링크를 차단하여 감염에 대한 위험을 예방하는 것이 이 시스템의 목표이다.

악성링크 분석 기술을 더욱 향상 시켜 탐지율을 높이는 연구를 지속적으로 진행하여 현 시스템에 적용한다면 향후 악성코드에 의한 피해를 효율적이고 효과적으로 차단 할 수 있는 악성코드 탐지 시스템이 될 것으로 기대한다.

참고문헌

[1] National Internet Development Agency of Korea.Trend analysis of large-scale spread of malicious code ,2014. p.2

[2] R.tian, L.M. Battern and S.C. Versteeg, Function Le-ngth as a Tool for Malware Classification, Proceedings of the 3rd International Conference on Malicious and Unwanted Software. pp.69-76, October 2008.

[3] Jinkyung Lee. A Study of Malware Detection and Classification by Comparing Extracted Strings. ICUIM 2011.

[4] Michael Weber. Matthew Schmid, Michael Schatz and David Geyer, “A Toolkit for Detection and Analyzing Malicioius Software”, ACSAC’02, IEEE, 2003.

[5] G.Wagener. R.State. and A. Dulaunoy. “Malware behaviour analysis.” Journal in Computer Virology. Nov.2007

[6] P.Vinod. V.Laxmi. and M.Gaur. “Survey on malware detection methods. ”Peedings of the 3rd Hackers’ Workshop on Computer and Internet Security.2009.