목적행위를 기반으로 한 악성코드 분류 방식에 관한 연구

김호연*, 박민우*, 서상욱**, 정태명***
*성균관대학교 전자전기컴퓨터공학과
**한국인터넷진흥원
***성균관대학교 정보통신공학부
e-mail : *{hykim, mwpark,}@imtl.skku.ac.kr, **swseo@kisa.or.kr
***tmchung@ece.skku.ac.kr

A Study on Classification of Malware Based on Purpose of Behavioral

Ho-Yeon Kim*, Min-Woo Park*, Sangwook Seo**, Tai-Myoung Chung***
*Dept of Electrical and Computer Engineering, Sungkyunkwan Univ.
**Korea Internet & Security Agency
***School of Information Communication Engineering, Sungkyunkwan Univ.

요 약

악성코드 개체 수의 급격한 증가와 정형화되지 않은 악성코드 분류 기준 때문에 업체별, 연구기관별 악성코드 분류 방식이 서로 상이하다. 이 때문에 악성코드를 분석하는 분석가들은 모호한 악성코드 분류 방식 때문에 업무에 불필요한 시간이 소요되고 있다. 또한 안티 바이러스 제품을 사용하는 최종 사용자로 하여금 혼란을 유발하고, 악성코드에 대응하기 위해 진행되는 연구에서 악성코드에 대한 정확한 분류 지표가 없어, 연구에 혼선을 빚고 있다. 본 논문에서는 악성코드의 정확한 분류와 새로운 악성코드가 발견되고, 새로운 매체가 출현하여도 이에 유기적으로 대응할 수 있도록 악성코드의 목적행위에 따라서 총 7개 그룹으로 나누었다. 제안 분류 방식을 사용할 경우 분류된 악성코드에 대하여 보다 정확한 정보를 얻을 수 있을 것으로 기대한다.

1. 서론

최근 악성코드의 전파되는 개체 수가 급격히 증가되고 있다. 안티 바이러스 툴을 테스트하는 독일의 AV-test는 자사의 2010년 12월 31일자 보고서를 통하여 1984년부터 2010년까지의 발견된 악성코드 개수가 약 4,400만 개라고 발표한 바 있다. 또한 2010년 4월에는 누적 악성코드가 3천만 개를 돌파하였다고 발표하였다. AV-test의 보고서에 따르면 악성코드의 누적 수치는 2010년 4월에서 2010년 12월 31일 까지 약 1,400만 개가 증가하였다. 일단위로 환산하면 매일 새로운 악성코드가 약 4만 1천 개 가량이 되는 것을 알 수 있다[1]. 근래에 들어, 악성코드들은 하나의 악성행위만을 수행하는 것이 아니라 다양한 악성행위를 복합적으로 수행하여 단기간에 많은 악성행위를 수행하도록 개발되어 유포되고 있다. 이러한 악성코드의 증가와 복합적인 악성행위로 인하여 악성코드 분석가들은 업무량이 증가하고 악성코드에 대응하기 위한 다양한 연구에서는 악성코드 분류의 모호성으로 인하여 업무량이 가중되고 있다.

이와 같은 문제를 해결하기 위하여 본 논문에서는 악성코드의 명확한 분류와 새로운 악성코드와 새로운 매체에 대해 유기적인 대응을 위하여 악성코드의 형태별, 특징별 분류가 아닌 목적행위에 따른 분류 방식을 사용한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 안티 바이러스 제품을 제공하는 업체들의 분류 방식을 살펴보고 3장에서는 연구되고 있는 악성코드 분류 방식에 대하여 살펴본다. 4장에서는 기존 악성코드 분류 방식에 문제점에 대하여 설명하고 5장에서는 제안하는 악성코드 분류 방식에 대하여 설명한다. 마지막으로 6장에서 본 논문의 결론을 맺는다.

2. 관련연구

악성코드의 분류 기준은 정형화되지 않아 안티 바이러스 솔루션을 제공하는 업체들 또한 자사에서 악성코드 분류 기준을 새롭게 정의하여 사용하고 있다. 따라서 새로운 악성코드 분류 기준을 위해, 안티 바이러스 업체들의 분류 기준에 대한 선행 연구가 필요하다.

2.1. 카스퍼스키랩(Kaspersky Lab.)

카스퍼스키랩은 1997년 설립된 안티 바이러스 소프트웨어 회사로 러시아 모스크바에 본사를 두고 있다. 설립된 년도에 비하여 많은 고객들을 확보하고 있는데, 29개국의 지사와 전 세계 3억 명이 카스퍼스키랩의 안티 바이러스 솔루션을 사용하고 있다[2].

<표 1> 카스퍼스키랩의 악성코드 분류 구분

대분류	소분류	표시형태
Network Worm	E-mail Worm	E-mail-worm
	Messenger Worm	IM-Worm
	Internet Worm	Net-Worm
	P2P Worm	P2P-Worm
Classic Virus	File Virus	Virus
	Boot sector Virus	BWME
Trojan Program	Backdoor	Backdoor
	General	Trojan
	PSW	Trojan-PSW
	Clicker	Trojan-Clicker
	Downloader	Trojan-Downloader
	Dropper	Trojan-Dropper
	Proxy	Trojan-Proxy
	Spy	Trojan-Spy

카스퍼스키랩은 악성코드를 'Malware'로 정의하고 있으며 악성코드의 특징에 따라 네트워크 웜, 클래식 바이러스, 트로이목마 프로그램, 기타 악성코드 네 분류로 나뉘며 각 분류마다 전파방법, 형태에 따라서 세분화 하고 있다. <표 1>은 카스퍼키랩에서 정의하고 있는 악성코드의 분류 그룹과 해당 그룹에서의 소그룹을 나타낸다.

2.2. 안철수 연구소

안철수 연구소는 무료 안티 바이러스 제품인 V3 lite를 비롯한 다양한 안티 바이러스 솔루션을 제공하고 있으며, ASEC(시큐리티대응센터)라는 긴급 대응 조직과 CERT (컴퓨터침해사고대응센터)를 운영하며 악성코드를 사전에 예방/차단하고 있다[3].

안철수 연구소는 악성코드를 특징별로 분류하고 있으며 악성코드의 분류를 위하여 다음 <표 2>와 같이 분류하고 있다[4].

<표 2> 안철수 연구소의 악성코드 분류 기준

악성코드 그룹	세부 내용
Adware	특정 프로그램과 동반 설치되며 설치된 후 사용자의 정보를 감시하거나 광고 창을 팝업
Appcare	자체적으로 악성행위를 수행하지는 않지만 악성코드가 악용할 수 있도록 보안위협을 높이는 프로그램
Clicker	허위광고 등의 메시지를 출력하여 클릭 유도
Downloader	특정 서버에 접속하여 악성코드를 다운로드
Dropper	자체적으로 다른 악성코드를 생성하거나 설치하는 악성코드
Script	자바 스크립트, 비주얼 스크립트와 같이 스크립트로 만들어진 악성코드
Spyware	사용자 정보를 감시하거나 감시하여 정제된 악성코드를 외부로 유출
Trojan	정상적인 프로그램을 가장하여 악성행위를 수행하는 악성코드
Virus	악성코드의 실행을 위해선, 악성코드가 감염된 숙주 파일이 실행되어야 하는 악성코드
Worm	자체적인 전파 기능이 있는 악성코드
etc.	다른 분류에 속하지 않는 악성코드가 해당

3. 악성코드 분류 기법 동향

안티 바이러스 업체들이 사용하는 분류 기준이 상이함에 따라 이를 해결하기 위하여 악성코드의 명명 형식 통일을 위한 시도가 있었다. CARO(Computer Anti-virus Research Organization)[6]는 199년에 악성코드의 명명 형식 통일을 위하여 명명기준을 제안하였지만, 새로 발견되는 악성코드들에 대하여 명명할 수 없다는 단점 때문에 2002년 Nick Fitzgerald가 새로운 진단명을 제안하였다[7]. Nick Fitzgerald의 명명 형식은 몇몇 업체들이 사용하고 있지만, 이미 구축해놓은 악성코드 명들을 바꾸기가 쉽지 않고, 동일 악성코드라도 분석가마다 분석하는 방식에 따라서 서로 다르게 명명되는 등의 문제로 안티 바이러스 업체들은 기존에 명명 형식을 고수하고 있다. 하지만 악성코드 분류 방식 을 통합하고, 악성코드에 대하여 유연하게 대처하기 위해 새로운 분류 방식에 대한 연구가 끊임없이 진행되고 있다.

3.1 Malicious Codes in Depth

M. Heidari는 'Malicious Codes in Depth'를 통하여 컴퓨터 바이러스의 분류 Depth를 호스트 프로그램이 필요한 악성코드들과 독립적으로 수행 가능한 악성코드로 나누고 있다[8]. M.Heidari는 악성코드가 악성행위를 수행하기 위하여 호스트 프로그램이 필요한 그룹과 독립적으로 악성행위를 수행할 수 있는 악성코드로 분류하고 있다. 호스트 프로그램이 필요한 악성코드의 종류로는 'Trap Doors', 'Logic Bombs', 'Trojan Horses', 'Viruses'가 있으며 독립적으로 수행할 수 있는 악성코드는 'Worms'과 'Zombie'로 구분하고 있다.

3.2 On malicious software Classification

멀웨어는 악성행위를 수행하는 소프트웨어를 총칭하지만 J Lin의 방식과 같이 바이러스와 웜을 포함하지 않고 분류할 수 있다. J Lin은 웜과 바이러스를 멀웨어로 정의하지 않고 다음과 같이 악성코드를 분류한다[9].

- Adware : 사용자의 동의 없이 다운로드 되고 설치되어 금전적인 목적을 수행하는 악성코드
- Spyware : 사용자의 동의 없이 설치되고 백도어를 설치하여 사용자의 정보를 수집하는 악성코드
- Browser Hijacking : BHO를 통하여 브라우저 창을 특정 웹사이트로 고정시키는 악성코드
- TrackWare : 사용자의 개인정보를 분석하여 사용자의 습성(검색어, 사용자의 취미 등)을 분석하는 악성코드
- Malicious shareware : 셰어웨어의 형태로 제공되는 악성코드

3.3 윈도우 악성코드 분류 방법론의 설계

날로 급증하는 악성코드의 출현으로 분석가가 하나하나 분석하고 분류하는 방식을 개선하기 위하여 악성코드를 자동으로 분류하는 방식 또한 연구되고 있다. 서희석 등은 '윈도우 악성코드 분류 방법론의 설계'를 통하여 악성코드의 그룹을 총 9개로 분류하고 각 그룹에서 클러스터별로 2차 분류를 수행하고 있다[10]. 다음 <표 3>은 9개의 그룹과 해당 그룹에 세부 클러스터를 나타낸 표이다.

<표 3> 악성코드 그룹과 클러스터

악성코드 그룹	클러스터
트로이목마	F_CREATE, P_CREATE_OTHER R_MODIFY_YN, N_USE
백도어	Packet 전송 유/무
다운로더	TROJAN, DROPPER FILE VIRUS, KEYLOGGER BOT, WORM BACKDOOR ADWARE/SPYWARE
파일 바이러스	F_CREATE, P_HOOKING_API_YN R_MODIFY_YN
웜	F_CREATE, P_CREATE_OTHER R_MODIFY_YN
드롭퍼	F_CREATE, P_HOOKING_API_YN R_MODIFY_YN
키로거	F_CREATE, N_USE

<표 3>에서 F는 파일을, P는 프로세스, N은 네트워크를 뜻한다. 그룹별 클러스터는 같은 그룹 내에서도 악성코드의 구분을 위함으로, 봇과 스파이웨어/애드웨어는 세부 분류가 나타나지 않으므로 그룹별 클러스터를 만들지 않는다.

4. 기존 악성코드 분류의 문제점

최근에 악성코드들은 하나의 악성코드에 다양한 악성행위를 적용함으로써 악성코드 분류에 대한 모호성이 존재하게 되었다. 또한 악성코드를 업체마다 서로 다른 분류방식을 사용함으로써, 동일한 악성코드를 분석하여도 업체마다 서로 다르게 분류하게 되어 사용자에게는 혼란을 유발하고 악성코드에 효율적으로 대응하고자 하는 정책적, 학술적 연구에 불편함을 초래하게 된다. 다음 <표 4>은 하나의 악성코드를 업체마다 다르게 분류하는 것을 나타내고 있다[5]. 안티 바이러스 업체들은 <표 4>와 같이 하나의 악성코드를 EMSI Software와 AVG는 백도어로 분류하고, Avast와 Dr.Web은 다운로더로 분류하고 있다.

<표 4> 동일 악성코드의 서로 다른 분류 방식

업체	분류 / 탐지 명
EMI Software	Backdoor.Win32.Shiz.kmx!A2
Avast	Win32:Downloader-JRO[Trj]
AVG(GriSoft)	BackDoor.Generic14.XDL
Avira	TR/Dldr.JRO
Kaspersky	Trojan.Win32.Jorik.Shiz.cge
BitDefender	Gen:Variant.Kazy.35202
Dr.Web	Trojan.DownLoader4.41721
Microsoft	VirTool:Win32/Injector.gen!BQ
McAfee	Generic.bfr!cp!659679D9B0C1

5. 악성코드 분류 방식 제안

5.1 악성코드 그룹

위에서 에서 살펴본 바와 같이 기존 악성코드의 분류 방식에 문제점을 개선하기 위하여 악성코드의 형태, 형식, 특징 별 구분이 아닌 목적행위에 따른 분류가 필요하다.

본 논문에서는 악성코드가 만들어진 목적행위에 따라 총 7개로 분류하였다. 악성코드 분류 그룹은 각 악성코드들이 하나에 그룹으로 종속되고, 다양한 매체들을 수용할 수 있도록 정보탈취형, 과금유발형, 시스템파괴형, 모듈형, 원격제어형, 유해가능형, 혼란야기형으로 분류 하였다. 다음 <표 5>는 각 그룹에 대한 분류 기준을 나타내고 있다.

<표 5> 악성코드 분류 방법 모델

그룹	분류 기준
정보 탈취형	-사용자의 정보를 외부로 유출시키는 악성코드 -키 스트로크, 특정 게임 계정, 사용자 정보
과금 유발형	-금전적 이익을 위하여 개발된 악성코드 -광고 팝업 / 많은 금액을 청구하는 서비스 연결 / 결제 유도
시스템 파괴형	-시스템 자원을 고갈시키는데 목적을 두는 악성코드
모듈형	-자체적인 악성행위를 수행하지 않는 악성코드 -다른 악성코드를 다운로드 및 생성
원격 제어형	-외부로부터의 침입 및 시스템 권한을 획득하여 공격자가 시스템을 제어하는 악성코드
유해 가능형	-보안 위협을 증가시키는 세어웨어/ 정상 프로그램 -악용 가능한 행위를 수행하는 프로그램
혼란 야기형	-사용자로 하여금 불편함을 느끼게 하는 악성코드 -공포사진 팝업 / 스크린 세이버 변경

- 정보탈취형 : 사용자의 개인정보를 외부로 유출 시키는 것에 목적이 있는 악성코드이다. 기존 분류 방식으로는 스파이웨어, 트로이목마로 분류되던 악성코드가 해당 그룹으로 분류될 수 있다.

- 과금유발형 : 악성코드 제작자가 금전적 이익을 목적으로 하여 제작된 악성코드가 해당 분류에 속하게 된다. 기존 분류방식으로 허위 안티 바이러스, 애드웨어, 크라임웨어, 다이얼러로 분류되는 악성코드가 해당 분류에 속할 수 있다.

- 시스템파괴형 : 시스템의 자원을 고갈시켜 사용자가 원하는 서비스를 제공받지 못하게 하는 악성코드가 해당 악성코드에 속한다. DoS와 같은 서비스 거부 공격을 수행하는 악성코드, 시스템 파일을 복구 불가능 하도록 변조 및 삭제하는 악성코드가 해당 분류에 속한다.

- 모듈형 : 시스템에 직접적인 악성행위를 수행하는 대신에 다른 악성코드를 생성하거나 다운로드 하는 악성코드 그룹으로 드롭퍼, 다운로더가 해당 분류에 속한다.

- 원격제어형 : 악성코드가 외부에 있는 공격자에게 시스템 권한을 제공하는 악성코드로 공격자는 해당 분류의 악성코드를 통하여 시스템 내부로 침입, 제 2, 제 3의 악성행위를 수행할 수 있다. 악성 봇, 백도어가 해당 분류에 속할 수 있다.

- 유해가능형 : 개발자가 악성행위를 목표로 하여 개발하지 않았지만, 악성코드가 해당 프로그램을 악용하여 악성행위를 수행할 수 있는 프로그램들이 해당 그룹에 속한다. 기존의 악성코드 분류로 휴리스틱(Heuristic), 콘텐츠 차단형 프로그램이 해당 분류에 속한다.

- 혼란야기형 : 단순 장난으로 사용자들로 하여금 불편함을 느끼게 하는 악성코드로 공포 사진을 팝업하거나 사용자 배경화면을 고정하는 조크 프로그램 또는 혹스(Hoax)가 해당 그룹으로 분류될 수 있다.

5.2 제안 방식의 이점

목적행위를 기반으로 한 분류 방식의 이점은 다음과 같다.

- 악성코드 분류에 대한 업무 노드 및 오류 감소 : 악성코드를 특정 카테고리만으로 분류가 가능하여 분석 작업 후 명명 및 분류작업이 용이
- 확장의 용이성 : 새로운 매체가 출현하여도 악성코드의 목적행위는 일반적이므로, 변경 없이 사용 가능하며, 새로운 목적행위가 나타날 시 기존 분류 방식에 수정 없이 새로운 악성코드의 행위만을 추가하여 사용 가능
- 자동분석과의 융합 : 시그니쳐 기반의 탐지 방식이 아닌 행위 기반의 자동화된 분석방식 사용 시 악성코드의 분류 기준이 목적행위 이므로 자동화된 악성코드 탐지와 연계하여 분류가 가능

6. 결론

본 논문에서 제안한 악성코드 분류 방식은 악성코드를 분류할 때 모호성을 해결하기 위해 목적행위에 따른 분류 방식을 사용하였다. 제안하는 악성코드 분류 방식은 새로운 악성코드가 발견되고, 새로운 매체가 출현하여도 이에 유기적으로 대응할 수 있으며, 자동화된 악성코드 탐지 기술인 행위 분석을 사용할 경우 자동화된 악성코드 분석과 연계하여 효율적으로 분석할 수 있다. 따라서 제안 분류 방식을 활용할 경우 분석가에 대하여 업무노드가 감소와 향후 악성코드 대응을 위해 다양한 연구가 진행될 시 악성코드에 대한 정확한 정보를 얻을 수 있을 것으로 기대한다.

Acknowledgements

본 연구는 2011년 한국인터넷진흥원의 “악성코드 유사 및 변종 유형 예측방법 연구” 위탁과제의 지원을 받아 수행된 연구임

참고문헌

[1] AVtest, http://www.av-test.org, 2011.
[2] Kaspersky Lab, http://www.kaspersky.com, 2011.
[3] AhnLab, http://www.ahnlab.com, 2011.
[4] ASEC Team, ASEC Report, 안철수 연구소, 2011.
[5] Symantec, http://www.symantec.com/index.jsp, 2011.
[6] CARO, http://www.caro.org, 2011.
[7] N. FitzGerald, "A virus by any other name: The revised CARO naming convention," , 2002.
[8] M. Heidari, "Malicious codes in depth," Securitydocs.com, 2004.
[9] J. Lin, "On malicious software classification," , pp.368-371, 2008.
[10] 서희석, 최중섭, 주필환, "윈도우 악성코드 분류 방법론의 설계," 정보보호학회논문지, Vol.19, No.2, pp.83-92, 2009.