韓國 컴퓨터情報學會 2006동계학술발표논문집&학회지
第14卷 第2號, 2006. 12.

로그 히스토리 분석을 사용한 웹 포렌식 알고리즘 연구

정정기*, 박대우*

A Study of Web Forensics Algorithm that used Log History Analysis

Jeung-Ki Jeung*, Dea-Woo Park*

요 약

수많은 로그 히스토리의 자료에서 컴퓨터 사이버범죄에 대한 증거자료로 채택되기 위한 기술적인 웹 포렌식 자료의 추출에 사용되는 웹 포렌식 알고리즘은 필수적인 요소이다. 본 논문에서는 웹 포렌식 알고리즘을 제안하고 설계하여, 실제 기업의 웹 서버 시스템에 제안한 알고리즘을 구현해 본다. 그리고 웹 로그 히스토리 정보에 대한 무결성이나 정보출처에 대한 인증을 적용한 웹 발신 로깅 시스템 구성실험을 한다. 회사의 이메일, 웹메일, HTTP(웹게시판, 블로그 등), FTP, Telnet 및 메신저(MSN, NateOn, Yahoo, DaumTouch, BuddyBuddy, MsLee, AOL, SoftMe)의 서버에서 웹 로그 히스토리 분석을 위해 사용한 웹 포렌식 알고리즘과 플로우를 설계하고 코딩을 통한 구현을 한다. 구현 결과 웹 포렌식을 통한 컴퓨터 사이버범죄에 대한 학문적 기술적 발전에 기여하고자 하는데 본 논문의 목적이 있다.

Abstract

Web Forensics algorithm used to an extraction of technical Web Forensics data to be adopted to proof data regarding a crime cyber a computer at data of a great number of log History is an essential element. Propose Web Forensics algorithm, and design at these papers, and try to implement in a Web server system of an actual company. And make the Web dispatch Loging system configuration experiment that applied integrity regarding Web log History information or authentication regarding an information source. Design Web Forensics algorithm and the Flow which used for Web log History analyses at server of e-mail, webmail, HTTP (Web BBS, Blog etc.), FTP, Telnet and messengers (MSN, NateOn, Yahoo, DaumTouch, BuddyBuddy, MsLee, AOL, SoftMe) of a company, and implement through coding. Therefore have a purpose of these paper to will contribute in scientific technical development regarding a crime cyber a computer through Web Forensics.

▸ Keyword : Computer Forensics, Forensics Algorithm, Integrity, Ubiquitous Security, Web History

• 제1저자 : 정정기, 교신저자 : 박대우(prof1@paran.com)
* 숭실대학교 정보과학대학원 정보보안학과

Ⅰ. 서 론

지식 정보화 사회에 진입하여 시간과 장소와 기기에 제한이 없는 유비쿼터스(Ubiquitous) 컴퓨팅과 네트워크를 이용한 업무 전환이 급속하게 이루어지고 있다. U-IT839전략 추진과 전자정부의 출범으로 인터넷의 업무에도 많은 변화를 겪고 있다. DMB, WiBro, Home Network, W-CDMA, RF-ID, VoIP 등의 신규 서비스와 BcN, USN, IPv6의 유비쿼터스 네트워크가 발달 할수록, 사용의 편리성을 악용한 사회적, 경제적 폐해와 개인정보의 도용과 같은 반정보화의 현상도 더욱 심화되고 있다.

정보화와 더불어 개인과 기업, 공공기관의 업무에서도 인터넷을 사용한 업무가 급격히 증가하고, 저장기술과 사용기술도 같이 발전하였다. 또한 정보의 보관방법도 종이, 테이프와 같은 기존의 저장장치에서 컴퓨터, PDA, 이동식 USB 저장장치 등으로 다양화되고 있다.

현재 컴퓨터 범죄에서 디지털 증거의 압수·수색 및 분석을 위한 다양한 방법이 연구되어 활용되고 있다. 그림 1. 과 같이 경찰청 사이버테러대응센터의 사이버방지 및 해킹 기술동향자료[1]에 의하면 1997년 126건에서 2003년 51,722건으로 급격한 증가 및 검거 현황을 보여주고 있다.

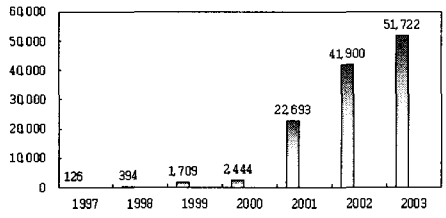

그림 1. 사이버방지 및 해킹기술동향
Fig. 1 cyber prevention and hacking technology tendency.

따라서 사이버 범죄의 급속한 증가에 따라, 증거 수집의 기술 발전의 필요성도 증가한다. 컴퓨터를 이용한 사이버범죄를 수사하고 재판할 때 디지털 자료에 대한 증거로써 포렌식(Forensic)의 필요성과 중요성이 증대되고 있다. 이러한 사이버 범죄 유형으로는 개인 정보의 도용으로 사용자의 정보를 제 3자가 이용하거나 데이터베이스에 불법적으로 접근해서 정보를 빼내는 행위가 있으며, 악성 프로그램과 스파이웨어를 유포하여 불특정 컴퓨터의 정보를 획득하거나 악용하는 사례도 있다. 2006년 2월에는 인터넷 게임의 주민등록번호를 도용한 사건이 발생하여 사회적인 이슈를 제기하며 재판을 하였다. 또한 게임 아이템을 이용한 불법적인 사기나 절도가 빈번하게 발생하면서, 불법 전자상거래 사이트를 구성하여 상품의 금액만을 챙기는 범죄가 발생하고 있다. 이외에 컴퓨터를 이용한 이중장부를 저장하거나 다른 사람의 약점으로 이용할 수 있는 음란 동영상, 문서 등을 불법적으로 가지고 있을 경우도 있다.

이러한 컴퓨터 관련 사이버범죄가 일어났을 후, 침입자의 흔적을 찾고자 할 때, 우리가 가장 먼저 취하는 행동은 침입자의 흔적(Digital Evidence)을 찾는 행위이다[2]. 이러한 행위에 가장 잘 사용되는 정보가 컴퓨터 서버 내에 남아 있는 로그(Log) 히스토리(History)에 관한 정보라 하겠다. 이러한 이유로 로그 정보는 불법적인 범죄자를 수사하기 위한 최소한의 흔적이 될 수 있고, 재판에서는 범죄자를 구속하기 위한 법적인 증거자료가 될 수 있다.

특히 초고속 인터넷에서의 웹(Web)을 이용한 중요한 정보와 금융 업무환경에서 로그 기록이 존재하고, 로그 기록에서 시간이 적용된 로그 히스토리가 반드시 존재한다. 하지만 수많은 웹의 접속자와 제공된 서비스에 대한 로그 히스토리에 관한 자료의 양은 너무 많다.

따라서 수많은 로그 히스토리의 자료에서 컴퓨터 사이버 범죄에 대한 증거자료로 채택되기 위한 기술적인 웹 포렌식 자료의 추출에 사용되는 웹 포렌식 알고리즘은 필수적인 요소이다.

하지만 이러한 로그 히스토리 정보는 분석 작업을 수작업 또는 비 전문지식으로 하는데서 기인하는 수사의 비효율성 문제가 존재한다. 또한 위조나 변조 가능성 혹은 그 정보의 출처에 대한 인증의 부재로 인하여 수사 시나 법정에서 참고 자료로서의 효력 밖에 가지지 못하는 문제점[3]이 있다.

따라서 본 논문에서는 여러 곳에 위치한 로그 히스토리 정보를 자동으로 한 곳에 모아 줄 수 있고, 로그 히스토리를 자동으로 분석해 줄 수 있다면, 사이버 범죄에 대한 수사의 편의성을 도모 할 수 있을 것이다.

그리고 로그 히스토리 정보에 대한 무결성(Integrity)이나 정보출처에 대한 인증을 추가할 수 있다면, 재판과정에서나 혹은 수사과정에서 지금 보다 더 나은 효력을 발휘할 수 있을 것이다.

이 문제점에 대한 해결책을 위해 본 논문에서는 웹 포렌식 알고리즘을 제안하고 디자인하여, 코딩으로 구현한 후, 실제 기업의 웹 서버 시스템에 제안한 알고리즘을 적용해본다. 그리고 웹 로그 히스토리 정보에 대한 무결성이나 정보출처에 대한 인증을 적용한 실험을 통해 로그 히스토리 분

석을 사용한 웹 포렌식 알고리즘 연구을 하여, 웹 포렌식을 통한 컴퓨터 사이버범죄에 대한 학문적 기술적 발전에 기여하고자 하는데 본 논문의 목적이 있다.

II. 관련 연구

초고속 인터넷과 웹을 이용한 컴퓨터 사이버범죄에 대한 증거 자료 수집 및 보존에 대한 일련의 과정들을 웹 포렌식이라고 가정한다. 웹 포렌식을 위한 웹 히스토리 자료의 의미와 포렌식을 수행하기 위한 도구인 툴과 웹 브라우징 및 웹 히스토리 자료 및 메일에 대한 분석 자료에 대한 법률적 포렌식의 의미와 검토를 통한 관련 연구를 한다.

2.1. 포렌식

포렌식에서 수집된 데이터는 법정에서의 증거자료로써의 효력을 발휘하기 위해서는 데이터의 특성을 잘 알아 안전하게 다루어 져야 한다.

포렌식의 분류로는 그림 2.와 같이 나눌 수 있다. 데이터 포렌식과 시스템 포렌식을 나누는 보다 상세한 기준은, 휘발성 관련된 데이터 쪽은 데이터 포렌식으로, 시스템에서 확인할 수 있는 증거는 시스템 포렌식으로 분류한다.

본 논문에서는 추적성 및 프로세싱 확보 차원에서 네트워크 포렌식 및 시스템 포렌식에 대하여 연구하고자 한다.

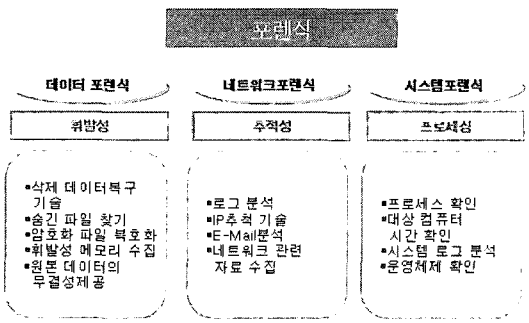

그림 2. 포렌식의 분류 및 기술
Fig. 2 Classifications of Forensics and Technology.

2.2. 로그 히스토리

컴퓨터 시스템에 불법으로 침입한 공격자[4]는 흔적을 남기게 되는데 이러한 흔적이 저장되어 지는 곳을 로그 히스토리 파일이라 할 수 있다. 이러한 로그 히스토리 파일에는 시스템에 대한 스캔 행위, exploit 툴을 이용한 공격, 특정 사용자 계정으로의 접속, root 권한의 획득, 트로이 목마 설치, 자료 유출 및 삭제 등 공격자의 행위[5]들이 기록되어 진다.

이미 시스템에는 이러한 로그가 다량 존재하며, 이를 분석하고 조합하고 추리하여 공격자의 행동을 추적하는 것이 포렌식 관점에서의 로그 히스토리의 의미라 할 수 있을 것이다.

로그 히스토리에 대한 분석 작업은 실제 범죄에 대한 분석 작업과 비슷한 형태를 지닌다. 다음은 실제 범죄가 일어났을 때 분석 방식으로 다음 두 가지를 생각할 수 있다.[6]

1) 다의성 포렌식 분석(Equivocal Forensic Analysis)

수집된 정보 최대한 객관적으로 검토하고 모든 것에 의문을 가지고, 이용 가능한 증거자체의 출처와 의미를 입증하여 조사관의 가설과 견해를 발전시켜 나가는 방법이다.

2) 행동 증거 분석(Behavioral Evidence Analysis)

수사의 초점, 용의자 범위 제한, 범죄자 선택 이해, 용의자들의 인터뷰 등을 통해 도움을 줄 수 있는 형태로 증거를 압축시키는 방법이다.

2.3. 포렌식 자료 분석

1) 포렌식 수사 도구 확보

컴퓨터 사이버범죄의 수사를 위한 포렌식 수사 툴[7]도구는 그림 3.의 포렌식 툴 테스팅 프로젝트 사이트[8]에서 다운받거나 정보를 찾을 수 있다.

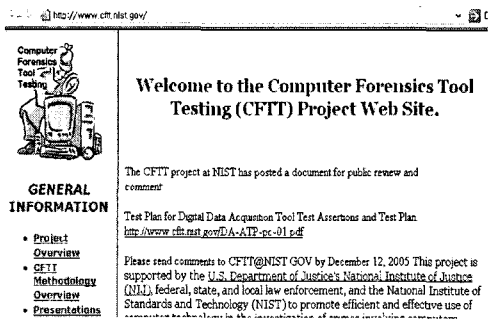

그림 3. 포렌식 수사 툴 제공 사이트
Fig. 3 Site Forensic Invest Tool

2) 포렌식 수사 툴

포렌식 수사를 위해 사용되는 툴은 표 1.과 같다.

Guidance Software의 EnCase[9]는 1980년대에서

개발되어, 증거 보존 및 분석 기능을 모두 갖고 있으며, 미국 법원에서도 EnCase를 통해 얻은 내용을 증거로 채택한 판례가 있다. NTFS, FAT 12(Floppy)/16/32를 모두 지원하며, Linux의 EXT2/3, Reiser, Sun Solaris 의 UFS, AIX의 JFS, Macintosh의 HFS, HFS+, FFS(OpenBSD, NetBSD, and FreeBSD), Palm, CDFS, ISO 9660, DVD 파일시스템 분석이 가능하며, 자료수집 후 강력한 리포팅 기능도 제공 된다.

또한 FastBloc 및 각종 I/O 자원을 통한 이미지 추출, 자료 분석 기능이 있고, 이미지에서 바로 분석 기능을 수행할 수 있으며, 타임라인 분석 기능인 파일의 생성, 편집, 최종 접근시간 등을 볼 수 있는 기능을 통해 하드디스크에 남겨진 흔적을 특정 파일에 대해 시간대 별로 어떤 작업이 수행되었는지 까지 추적할 수 있다.

표 1. 포렌식 수사에 사용되는 툴
Table. 1 The Tool which is Used to a Forensics Criminal Investigation.

구 분	포렌식 수사에 사용되는 툴 종류
디스크 쓰기 방지	A-Card, FastBlock, NoWrite
이미징	DD, Safe Back, SnapBack, DatArrest
검색	Grep, dtSearch, Text Search Plus, A/H/Sfind, , Sfind
문서/파일 분석	Quick View Plus, ThumbsPlus, Winhex. Ultra Edit
복구, 분석	Hash Keeper, TCT, Final Data, FileRepair
통합 툴	EnCase, iLook, Autopsy, Hellix, DEAS

미국 FBI에서 엔론사에 대한 조사에서 대규모의 파기 문서, 대용량의 삭제 이메일을 분석하기 위하여 사용되기도 하였다.

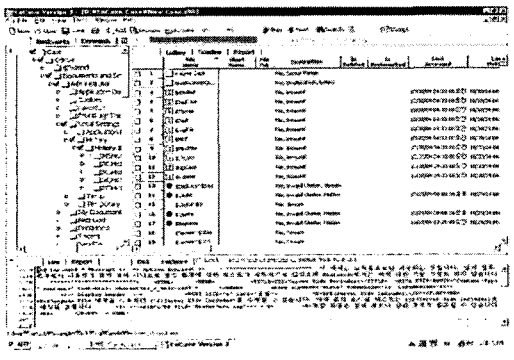

그림 4. iLook 포렌식 툴
Fig. 4 iLook Forensics Tool

iLook[10] 포렌식 툴은 그림 4. 처럼 FBI에서 컴퓨터 포렌식 분석에 수행되며, iLook은 미국 국세청 조사국(Internal Revenue Service Criminal Investigation, IRS-CI)에서 개발하여 IRS, FBI, NASA, 연방 및 지방 수사기관에 배포되어 사용되었다. iLook을 사용하여 하드디스크 분석과정을 자동화하고, 컴퓨터에서 찾아낸 수십 기가 바이트의 하드디스크를 시기적절하게 조사할 수 있는 2개의 기능으로 구성되어 있다. iLOOK Imager는 Computer Investigative Specialist들이 찾아낸 PC기반의 하드디스크를 이미지를 추출할 수 있는 Linux 기반의 제품이다. iLOOK Investigator는 제품의 분석 부분이다. 대상 이미지 파일로부터 삭제된 파일을 복구하고 특정한 내용을 검색할 수 있다.

2.3. 웹 브라이징 분석 툴

1) Pasco[11]

Pasco는 Index.dat 파일을 받아들이고, 데이터를 재구성하고, 정해진 텍스트 파일 구성에 정보를 출력한다. Microsoft Excel와 같은 스프레드시트로 데이터를 내포하여야 할 때 유용하며, Pasco는 웹사이트 방문으로부터 아래와 같은 정보를 보여준다.

가. 레코드 타입- 읽혀졌거나 또는 다른 위치에 사용자의 브라우저 URL활동.
나. URL- 사용자가 방문하였던 현실 웹사이트.
다. 수정된 시간- 마지막 순간에 수정된 웹사이트..
라. 접근 시간- 사용자가 웹사이트를 탐색하였던 순간.
마. 파일 이름- URL의 사본을 포함하여 목록에 나열된 로컬 파일 이름.
바. 자료방- 파일 이름을 발견할 수 있는 장소, 주소록.
사. HTTP Headers- HTTP URL을 탐색하였을 때 사용자가 제공받는 부분.

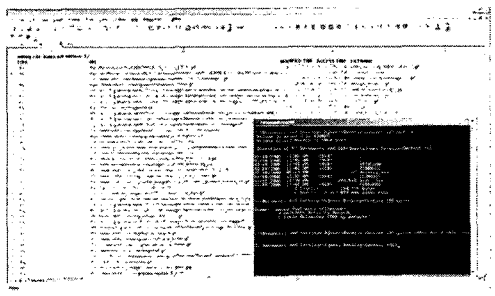

그림 5. Pasco 브라이징 분석 툴
Fig. 5 Pasco Browsing Analysis Tool

2) 웹 Historian[12]

웹 Historian은 디렉토리 구조 및 웹 브라우저의 모든 것 즉 인터넷에서 활동하는 다음 내용들을 분석한다.

가. 인터넷 익스플로어

나. Firefox

다. 넷스케이프

라. 애플

바. Opera

위 사항들이 의미하는 것은 웹 브라우저를 통해 활동하는 데이터를 엑셀이나 스프레드쉬트, HTML, Text file로 분석한다.

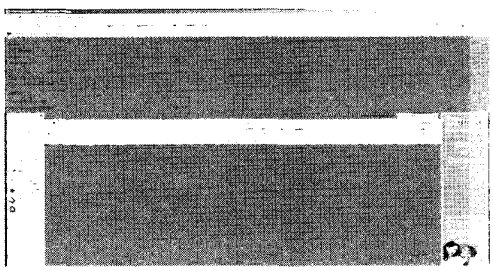

그림 6. 웹 Historian
Fig. 6 Web Historian

그림 6.처럼 특정인이 Hotmail을 방문하여 웹서핑을 하거나 업무를 수행하여 웹 Historian로 분석한 결과는 계정 및 모든 로그 정보를 분석 가능케 한다.

3) Tool Kit(FTK)[13]

저장된 웹페이지를 탐색할 수 있고, 웹 브라우저와 같은 인터페이스를 통해 볼 수 있다.

단점은 FTK를 사용할 때 Index.dat을 신청하여 탐색할 수 있으나, 선택된 Index.dat FTK내의 파일, 데이터는 탐색할 수 없다.

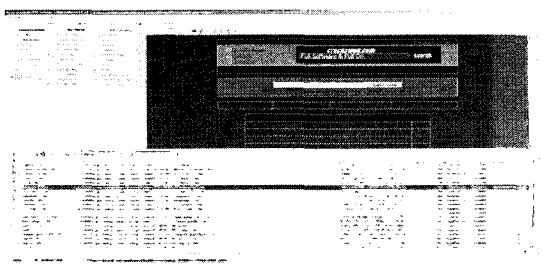

그림 7. Tool Kit
Fig. 7 Tool Kit

2.4. 로그 히스토리 자료 분석의 법률적 검토

웹 포렌식 알고리즘분석 전에 웹에 대한 로그 히스토리 분석 자체가 법률적으로 위법인지 합법인지에 대한 검토를 선행한다.

1) 국내 관련 법 조항

현재 국내에서는 이메일 모니터링과 관련된 법령이 정비되어 있지 않고, 이메일 모니터링의 가부 및 적법성을 갖추기 위한 요건과 절차 등에 관한 유권적인 판례가 없는 상태로서 관련 법 규정의 해석에 의존하고 있는 상태이다.

이메일 모니터링과 관련된 형사 법규는 정보통신망이용촉진및정보보호에관한법률, 통신비밀보호법, 전기통신사업법, 형법 등이 있다.

기업 및 공공 기관의 경우 이메일 모니터링 제도를 운영하고 있는바, 관련 법 규정의 해석에 입각하여 입사시나 재직 중에 징구하는 비밀유지서약서 상에 발송한 이메일 모니터링에 관한 동의 조항을 두고, 사내 제 규정에 이메일 모니터링에 대한 근거 조항을 두어 적법성을 담보하기 위한 중첩적인 장치를 마련하고 있다.

정보통신망이용촉진 및 정보보호에 관한법률 제49조(비밀 등의 보호)에서 누구든지 정보통신망에 의하여 처리/보관 또는 전송되는 타인의 정보를 훼손하거나 타인의 비밀을 침해/도용 또는 누설하여서는 아니된다. 라는 조항에 2가지 해석이 가능하다.

하나는 "타인의 비밀"과 관련하여 통신과 관련된 비밀의 보호주체는 수신인 및 발신인 양 자를 포함하므로 단순히 발신인의 사전 동의만을 얻는 것은 위법(통신비밀보호법상의 당사자에 전기통신의 송신인 및 수신인 모두를 포함하고 있으므로)이다.

다른 하나는 "비밀"이 되는 정보는 정보의 소유자 또는 관리자가 다른 사람에게 공개되는 것을 거부한다는 의사가 현실적으로 또는 잠재적으로 내재되어 있어야 할 것이므로 이메일(정보)의 소유자 또는 관리자로서 그에 대한 처분권이 있다고 볼 수 있는지는 오로지 발신인에 한정된다.

이메일 모니터링으로 인한 수신인의 비밀노출은, 송신인의 수신인에 대한 프라이버시 침해행위에 해당할 뿐, 송신인의 동의를 받아 모니터링을 한 회사 측의 책임은 인정하기 어려고, 통신비밀보호법 제3조(통신 및 대화비밀의 보호)누구든지 이 법과 형사소송법 또는 군사법원의 규정에 의하지 아니하고는 우편물의 검열/전기통신의 감청 또는 통신사실 확인 자료의 제공을 하거나 공개되지 아니한 타인간의 대화를 녹음 또는 청취하지 못한다.)에 대한 하나의 의

견은 동 법의 감청의 정의를 보면, "당사자의 동의 없이 전자장치, 기계장치 등을 사용하여 통신의 음향, 문언, 부호, 영상을 청취/공독하여 그 내용을 지득 또는 채록하거나 전기통신의 송수신을 방해하는 것"으로 되어 있으므로 당사자의 동의가 없는 경우에는 감청으로 간주된다.

또한 동 법상의 "당사자"의 정의에 따르면 전기통신의 발신인 및 수신인 모두를 포함하고 있으므로 발신인의 동의만으로는 당사자의 동의에 해당하지 않으므로 감청에 해당되어 위법이다.

또 하나의 의견은 동 법에서 말하는 "전기통신"이란 "전화, 전자우편. 등과 같이 유선, 무선 등의 전자적 방식에 의하여 음향, 문자 등을 송신하거나 수신하는 것"을 의미하는 바, 이미 의사전달행위가 완료된 이후의 이메일은 여기서 말하는 전기통신에 해당한다고 보기 어렵다. 또한 "감청"의 의미는 동시성을 부여하기 때문에 발신된 이메일을 모니터링 하는 것은 이에 적용되지 않는 것으로 볼 수 있으므로 (판례 2003도 3344판결의 내용), "이메일 모니터링은 동 법의 적용 대상에서 제외된다고 해석하는 것이 타당하다"라고 명기되어 있어, 웹 포렌식에 대한 명확한 위법과 적법에 대한 분명한 기준 척도의 마련이 시급하다.

3) 미국

미국의 경우 다수의 판례는 종업원이 회사가 정기적으로 이메일 모니터링을 하는 것을 인지한 이상 사적용도 금지에 관한 서약서를 제출했는지 여부를 불문하고 개인정보보호(Privacy)권을 주장할 수 없다고 한다.

III. 웹 로그 히스토리 분석

3.1. 웹 메일 분석의 필요성

최근 대기업 및 공공기관을 중심으로 발신로깅시스템[11]을 통한 내부보안에 대한 대응책을 마련하고 성공적인 운영을 통해 기업 및 기관의 안정성을 강화하고 있다. 이에 대한 웹 메일 분석의 필요성은 아래와 같다.

1) 내부 정보 유출 방지
   가) 직원들의 인터넷 사용증가에 따른 정보유출 채널의 다양화.
   나) 통신매체의 광대역화로 인한 대량정보의 단기 유출 가능성 증대.
   다) 아웃소싱, 파견 근무의 보편화로 인한 인적 통제의 어려움 증가.

2) 기업 자산 보호
   가) e-Data(견적서, 기술정보, 도면, 사업계획서 등)의 보편화.
   나) 인터넷이 핵심전송 수단임에도 불구, 관리 방법의 모호함으로 인한 관리정책의 부재.

3) 법률적 요건 강화
   가) 외부감독기관의 내부정보 유출통제에 대한 요건 강화.
   나) 고객정보 유출로 인한 사고발생시 기업의 기본 의무사항(Due Care) 불이행으로 인한 패널티의 증가.
   다) 보안사고시 대응책 수립 및 법적증거 자료의 확보.

4) 경쟁 우위 확보
   가) 대다수 기업들이 내부정보발신통제를 구축하고 있음.
   나) 증권/금융권에서도 최근 들어 내부정보 발신통제를 활발하게 추진하고 있음.
   다) 경쟁자보다 한발 앞선 대응으로 대고객 신뢰도 및 대외 신인도 향상이 필요함.

3.2. 웹 발신 로깅 시스템 구성

웹 네트워크에 발신 로깅 시스템을 구성하고, 이 시스템을 통한 웹 로그 히스토리 분석을 위한 포렌식의 필요 기능을 요약하면 다음과 같다.

1) 네트워크를 통하여 외부로 전송되는 컨텐츠를 송수신자와 동일한 형태로 모니터링.
2) 이메일, 웹 메일, 웹 게시판, FTP, Telnet 을 통한 내용과 첨부/압축파일을 모니터링.
3) 메신저의 경우, MSN, NateOn, DaumTouch, Yahoo, MsLee, SoftMessneger, BuddyBuddy, AOL을 모니터링.
4) 송수신자/IP, 날짜, 시간, 제목, 내용, 메일크기, 첨부/압축파일명, 파일크기, 실수취인, 참조, 숨은 참조를 모니터링.
5) 저장된 로그데이터에 대한 항목별 검색 및 내용과 첨부파일의 키워드 검색.
6) 각 항목별 and, or, not 의 불린식 검색기능.
7) 주민번호, 전화번호 등 일정패턴에 대한 검색기능.
8) 사용자가 원하는 다양한 조건을 조합하여 검색.
9) 다수의 관리자에 차등 작업권한을 부여.
10) 장기간의 로그에 대한 스케줄링 작업 (리포팅, 검색)기능을 제공.

11) 관리자가 원하는 폴더를 별도로 관리하는 기능.
12) 원격에서 서버시스템의 운영현황을 파악 가능.

3.3. 로그 히스토리 분석 방법

분석방법은 사용자가 웹서버에 신호를 보내고 Relay Server를 통하여 Web Server서버에 저장되는 모든 데이터의 조회, 검색을 통하여 로그를 분석한다.

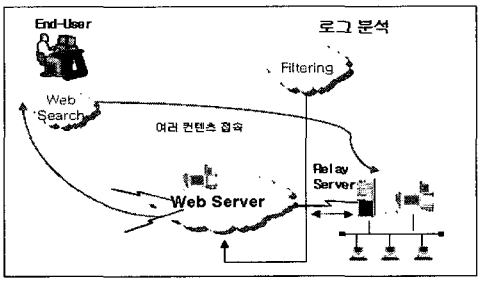

그림 8. 로그 히스토리 분석 프로세스
Fig. 8 log HiSeuToRi analysis process

1) 키워드 분류 관리
2) 분류명 별 키워드 조회
3) 일자, 아이디, 분류 명, 키워드, 작업내용(추가, 수정, 삭제)등의 조건으로 검색
4) URL 필터링
5) 메시지 자동 체크
6) Protocol 별 조회
7) 조회 조건 차별화

3.4. 로그 히스토리 분석 알고리즘

1) 로그 히스토리 자료 분석의 알고리즘 디자인

로그 히스토리 자료 분석의 알고리즘의 요건을 정의한다.

가) 검색은 "전수검사"를 원칙으로 하되 특성을 고려하여 Key-Word 중심의 필터링 검색을 실시한다.
나) 메일/웹 크기별 필터링을 통해 송수신 사항을 검색해 낸다.
다) 메일/게시물 내용 중 특정 단어(발신자가 자주 쓰는 단어 및 특정 단어 등)위주로 검색한다.
라) 첨부 화일 중 특정 첨부파일(ppt, doc, hwp , zip, pdf, mp3 등) 위주로 검색한다.
마) 메일/게시물의 목적지중 특정 목적지 위주로 검색한다.
바) 특히, 퇴직자 및 예정자에 대해서는 3개월 이상의 메일/웹 전수검사 실시한다.
사) 이메일, 웹메일, 웹게시판, TTP, Telnet, 메신져 등 네트워크를 통하여 전송되는 모든 로그를 중심으로 검색한다.
아) 특정인, 특정 IP를 알고 있다고 가정하였을 경우, 특정부분에 대하여 검색한다.

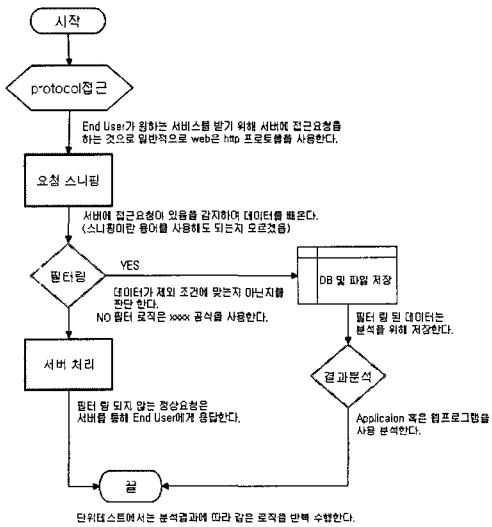

그림 9. 웹 히스토리 알고리즘 플로우
Fig. 9 Web History Algorithm Flow

2) 로그 히스토리 분석 알고리즘 플로우

로그 히스토리 자료 분석의 알고리즘의 플로우는 그림 9. 와 같다

```
public class FilterInsert {
        1. 필터 링 및 저장 단계
        public FilterInsert() {
                String strValue = get(request);

                if(!filter(strValue)) {
                        return 0;
                } else {
                        sqlInsert(strValue);
                }
        }

        2. 분석 단계
        public void analyze() {
                int estimation_cnt = 100;
                String[] arrayResult = sqlSelect();

                if(arrayResult == null) {
                        return 0;
                } else {
                        if(arrayResult.length == estimation_cnt) {
                                return 0;
                        } else {
                                return new FilterInsert();
                        }
                }
        }

        public String get(String value) {
                return strResult;
        }

        public boolean filter(string value) {
                return true;
        }

        public void sqlInsert(string value) {
        }
}
함수 설명
get(string value) : 요청 데이터 스니핑 함수
filter(string value) : 필터 로직
sqlInsert(string value) : 필터 링된 데이터 저장 함수
```

그림 10. 웹 히스토리 알고리즘 코딩
Fig. 9 Web History Algorithm Coding

3) 로그 히스토리 분석 알고리즘 코딩

로그 히스토리 자료 분석 알고리즘의 코딩은 그림 10.과 같으며, C^{++} 로 하고 컴파일해서 프로그램을 실행 시킨다.

IV. 포렌식 웹 히스토리 증거수집 분석

본 논문에서 제안한 웹 발신 로깅 시스템을 웹 네트워크에 구축하고, 웹 히스토리 자료를 포렌식 증거로 채택하기 위한 실험을 실시한다. 실험은 기업의 실제 현장에서 포렌식 툴을 사용하고 제안한 알고리즘을 적용한 시스템을 구축하여 실험한다.

4.1. 웹에 대한 불법 공격 히스토리 분석

컴퓨터상에서 보낼 수 있는 모든 메일과 회사 이메일, 웹메일, HTTP(웹 게시판, 블로그 등), FTP, Telnet 및 메신저(MSN, NateOn, Yahoo, AOL, DaumTouch, BuddyBuddy, MsLee, SoftMe)의 서버를 대상으로 실시한다.

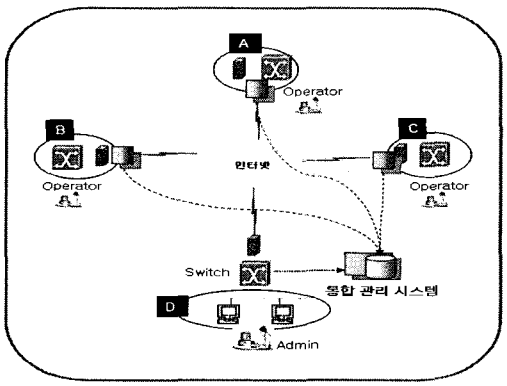

그림 11. 실험을 통한 웹로그 분석
Fig. 11 Analyze WepRoGeu Through Experiment

기업 현장의 웹 서버에서 실험 시스템을 그림 11.처럼 구축하고 공격에 대한 웹 로그 히스토리의 자료 분석을 실시한다. 실험에 사용된 웹 서버의 상세 사양을 표 2와 같이 구성하고 분석 테스트를 실시한다.

4.2. 웹 메일의 웹 히스토리 자료 수집

본 논문에서 제안한 웹 발신 로깅 시스템에서의 웹 메일 분석 시스템을 포렌식 사용 시에 기본적인 시스템의 상세 내역은 다음과 같다.

표 2. 서버 스펙
Table. 2 Server Spec

구분		서버 스펙
	cpu	Xeon 3.4G~3.6G * 2 ea
	Memory	4G
	HDD	73G * 2 ea(OS mirror구성), 146 * 3 ea(조그)
	NIC	수집, 통신 각각
	OS	MS Windows 2003 Server
	백업	
	MS SQL	1copy
	기타	Tap
	cpu	Xeon 3.0 * 1 ea
	Memory	1.5 G
	HDD	73G * 1 ea
	OS	MS Windows 2003 Server
	PC 급	Pentium 4
		512 M
		2 G HDD
		OS

실험을 위한 서비스를 웹 시스템에 대한 사양과 분석도구의 사양은 표 3.와 같다

표 3 웹 서비스 분석도구의 사양
Table. 3 Refusal of a Web Service Analysis Tool.

부문	내 용
S/W	Mail-i 800 Users (A지역 300 / B지역 500, C지역 50)
	Messenger-i 1500 Users (300 / 500, 700, 50)
	Packet Receiver
	MS SQL 2000 Server, standard
H/W	A지역과 B지역에서 수집과 병행한, 통합관리 서버 C지역과 D지역은 현 장비 이용
	Tap - 패킷수집용, 스위치 미러링이 불가능할 경우 사용

웹 메일 분석 시스템의 포렌식 자료로서의 상세 내용은 다음과 같다.

1) 보낸 사람/받은 사람.
2) 보낸 IP/받은 IP.
3) 날짜, 시간, 제목, 내용, 메일크기.
4) 첨부파일명(압축파일 포함), 파일크기.
5) 실수취인, 참조, 숨은 참조, 읽음 여부.

포렌식을 위해 웹 메일을 파악할 수 있도록 본 논문에서 제안된 시스템을 사용하였고, 실험을 통해 즉시 웹 로그를 분석할 수 있었다.

그림 12.~14.의 데이터는 컴퓨터의 시간을 2006년 11월 특정일로 분석하였고, 기업의 업무에서 일어나는 일반적인 내용은 블록처리 하였다.

그림 12.는 웹메일에서 상대방에게 발송한 내역을 검색한 데이터이다. 그림 13.은 MSN 메신저를 통해 대화한 증

거를 수집한 것이다. 그림 14.는 FTP를 통해 송수한 내역을 검색한 데이터이다.

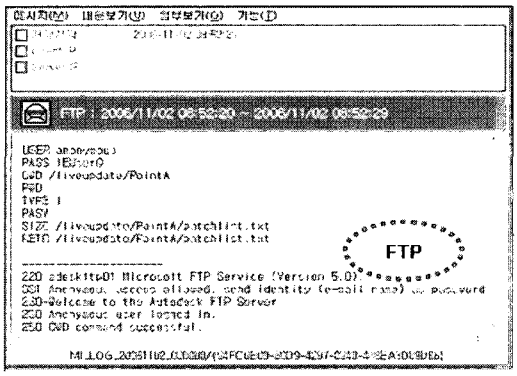

그림 12. FTP를 통해 송수한 내역 히스토리
Fig. 12 Receive/Send Details History Through FTP

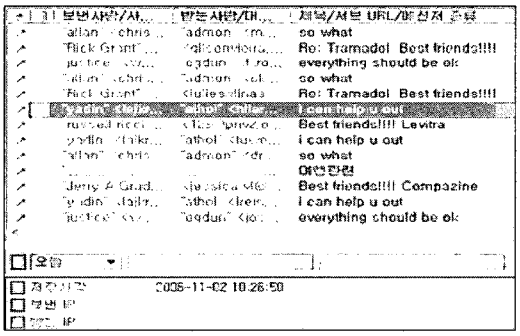

그림 13. 웹메일 히스토리 분석
Fig. 13 Webmail History Analysis

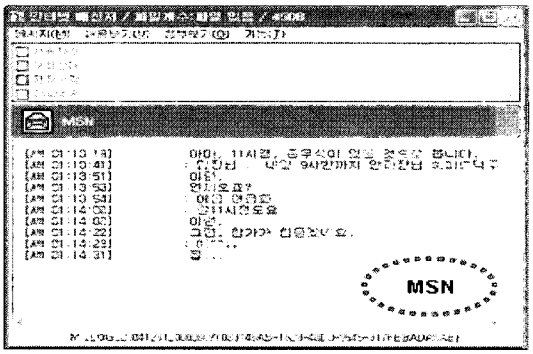

그림 14. MSN 메신저 히스토리 분석
Fig. 11 MSN Messenger History Analysis

4.3. 웹 히스토리 자료 분석 결과

웹 메일 히스토리 자료 분석 결과 거의 모든 이메일, 웹메일, HTTP(웹게시판, 블로그 등), FTP, Telnet 및 메신저(MSN, NateOn, Yahoo, DaumTouch, 로그 히스토리 자료 분석의 알고리즘의 BuddyBuddy, MsLee, AOL, SoftMe)등에서 발신된 내역은 완벽하게 검색할 수 있었다. 실시간 조회도 가능했으며, 웹 서버에 저장 및 백업만 제대로 되어 있다면 수개월 또는 수년 된 자료도 분석 가능할 것이다.

이런 데이터들은 무결성과 신뢰성을 확보할 수 있어 곧바로 포렌식 수사 자료로서 재판에서 증거자료로 채택될 수 있다.

V. 결론

본 논문은 로그 히스토리분석을 사용하여 웹 서버를 통해 발송된 데이터들을 검색하였다. 포렌식 기법이 계속 연구되고, 기술이 발전하고 있지만, 법과 제도의 뒷받침으로 명확한 유권해석이 선행되고, 또한 증빙 자료의 분석 및 검증 데이터들의 무결성을 위해서 로그 히스토리 분석과 분석을 위한 알고리즘의 연구를 통한 웹 포렌식은 더욱 발전 할 것이다.

하지만 현재 국내에서는 이메일 모니터링과 관련된 법령이 정비되어 있지 않고, 이메일 모니터링의 적법성 및 적정성을 갖추기 위한 요건과 법적 절차 등에 관한 유권적인 판례가 없는 상태이다. 현재에는 관련 법 규정의 해석에 의존하고 있는 상태이기 때문에 전자증거 수집절차도 due process를 준수해야 하며 전자증거는 위·변조 및 증거인멸이 용이하기 때문에 수사절차에서 새롭게 나타나는 문제해결을 위한 구체적인 법규정립이 요구된다.

이러한 상황에서 본 논문에서는 웹 포렌식 알고리즘을 제안하고 설계하여, 실제 기업의 웹 서버 시스템에서 제안한 알고리즘을 구현해 보았다. 그리고 웹 로그 히스토리 정보에 대한 무결성이나 정보출처에 대한 인증을 적용한 웹 발신 로깅 시스템 구성 실험을 하였다. 회사의 이메일, 웹메일, HTTP (웹 게시판, 블로그 등), FTP, Telnet 및 메신저(MSN, NateOn, MsLee, Yahoo, DaumTouch, BuddyBuddy, AOL, SoftMe)의 서버에서 웹 로그 히스토리 분석을 위해 사용한 웹 포렌식 알고리즘과 플로우를 설계하고 코딩을 통한 구현을 하였다.

구현 결과 웹 포렌식 알고리즘을 통한 웹 히스토리 자료의 추출 및 분석이 가능 하였다. 이 논문의 결과는 웹 포렌식을 통한 컴퓨터 사이버범죄에 대한 학문적 기술적 발전에 기여하게 될 것이다.

향후 연구에서는 명확한 법적 근거를 통대로 한 웹 포렌식 알고리즘의 발전 방향성이 제시되어야 하며, 추가적으

로 웹 포렌식과 함께 고려해야 할 사항으로는

1) 컴퓨터 통신 감청을 통한 증거수집 절차.
2) ISP가 운영하는 서버에 대한 수사절차.
3) 수사를 위한 ISP의 로그기록 보관 의무.
4) 암호화된 전자증거 수집 시 대응방안.
5) 유비쿼터스 환경의 도래에 대한 대응.
6) 각종 서버에 대한 백업 스케쥴링 절차 등이 요구된다.

또한 증거 수집의 무결성 확보 절차는 적법성을 인정하는 법적 근거 또는 사회적 합의가 중요하므로, 웹 포렌식의 지속적인 연구가 필요하다.

참고문헌

[1] 포렌식 법적 문제. http://www.cftt.nist.gov. 2006.8.

[2] Luoma, V. Forensics and electronic discovery: The new management challenge. Computers & Security, 25(2), 91-96. 2006.

[3] Bhaskar, R. State and local law enforcement is not ready for a cyber Katrina. Communications of the ACM, 49(2), 81-83. 2006.

[4] 박대우, 서정만. "TCP/IP 공격에 대한 보안 방법 연구." 한국컴퓨터정보학회논문지, 제10권 제5호, pp217-226, 2005. 11. 30.

[5] 박대우, 임승린. "해커의 공격에 대한 지능적 연계 침입 방지시스템의 연구." 한국컴퓨터정보학회논문지, 제11권 제2호, pp44-50, 2006. 5. 31.

[6] Kanellis, P., Kiountouzis, E., Kolokotronis, N.,& Martakos, D. (Eds.). "Digital crime and forensic science in cyberspace". Journal of digital forensic practice. Hershey: Idea Group. 2006.

[7] Busing, M. E., Null, J., & Forcht, K. Computer forensics: The modern crime fighting tool. Journal of Computer Information Systems, 46(2), 115-119. 2005.

[8] Computer Forensics Tool Testing Program. http://www.cftt.nist.gov/ 2006.11.

[9] EnCase Enterprise V5. http://www.guidance-software.com/support/downloads.aspx. Oct.6, 2006.

[10] ILook v8. http://www.ilook-forensics.org/. 2006. 11.

[11] Pasco v1.0. http://www.foundstone.com/index.htm?subnav=resources/navigation.htm&subcontent=/resources/forensics.htm. 2006. 10.

[12] Web Historian 1.3. http://www.softpedia.com/get /Security/Security-Related/Web-Historian.shtml. 2006. 11.

[13] Forensic Toolkit®(FTKTM). http://www. access-data.com/products/. 2006. 11.

저 자 소 개

정 정 기

1996년 조선대학교 물리학과 졸업 (이학석사)
2006년 숭실대학교 정보과학대학원 정보보안학과 (석사과정)
2002년 BS-7799 Leader Audit
2003년 LG전자 사내 보안 강사
2005년 LG전자 정보보안 과장 근무
2006년 LG이노텍 정보보안 과장 근무
〈관심분야〉 컴퓨터 포렌식, 네트워크 보안 시스템, 유비쿼터스 보안, 웹 포렌식

박 대 우

1998년 숭실대학교 컴퓨터학과 졸업 (공학석사)
2004년 숭실대학교 컴퓨터학과 졸업 (공학박사)
2000년 매직캐슬정보통신 연구소 소장, 부사장
2004년 숭실대학원 정보과학대학원 정보보안학과 겸임조교수
2006년 정보보호진흥원 선임연구원
〈관심분야〉 유비쿼터스 보안, 네트워크 보안 시스템, VoIP 보안, 이동통신 및 WiBro 보안, Cyber Reality