情 報 保 護 學 會 誌
第 18 券 第 1 號, 2008. 2

안티 포렌식 기술과 대응 방향

이 석 희*, 박 보 라*, 이 상 진*, 홍 석 희*

요 약

최근 디지털 포렌식 기술이 발전함에 따라 안티 포렌식 기술 역시 그 대항력을 발전시키고 있다. 안티 포렌식 기술은 크게 데이터 삭제 기술, 데이터 은닉 기법, 데이터 변조 기술로 구분할 수 있다. 현존하는 다양한 형태의 안티 포렌식 기술을 설명하고, 현재 안티 포렌식 기술의 동향과 안티 포렌식 대응 기술을 간략히 소개하고자 한다. 이를 바탕으로 디지털 포렌식 수사능력을 증대시키고 디지털 포렌식 기술 발전 방향을 수립하는데 도움을 주고자 한다.

Ⅰ. 서 론

최근 컴퓨터에 저장되어 있는 데이터가 법정에서 다루어지는 경우가 많은데 이와 관련된 분야를 컴퓨터 포렌식이라고 하고, 정보처리 기기를 통하여 이루어지는 각종 행위에 대한 사실관계를 확증하거나 증명하기 위해 행하는 각종 절차와 방법으로 정의할 수 있다[12].

일반적으로 '사이버 범죄'란 컴퓨터를 이용한 모든 범죄를 의미한다. 좀 더 실무적인 측면에서의 사이버 범죄는 컴퓨터를 포함한 일련의 디지털 기기나 정보통신 기술이 어떤 행위의 수단이나 목적인 모든 범죄를 뜻한다. 사이버 범죄의 등장에 따라 그에 대응할 수 있는 보안 기술과 관련 정책들이 요구되었다.

그 결과로 다양한 디지털 포렌식 기술이 개발되었고, 실제 사이버 범죄에 대하여 대응하는 사례가 증가함에 따라 일반적인 사이버 범죄 기술과 더불어 디지털 포렌식 기술에 대응할 수 있는 '안티 포렌식 기술' 또한 발전하고 있는데, 디지털 포렌식 수사에 의해 증거가 발견되지 않도록 하기 위한 기술로 정의할 수 있다[1]. 안티 포렌식 기술은 컴퓨터와 같은 디지털 기기에 대한 일정 수준 이상의 지식과 기술력을 가지고 있어야 이를 개발하거나 활용할 수 있다. 그리고 안티 포렌식 기술은 범죄에 대한 결정적인 증거를 은닉하거나 훼손하는 형태로 발전하고 있는데 대표적인 기술로는 파일 시스템의 특성에 따른 정보의 은닉 및 삭제, 스테가노그래피, 메타데이터를 이용한 정보의 은닉 그리고 물리적으로 정보를 파괴하는 것 등이 있다.

이에 대항하여, 증거의 은닉, 삭제 및 파괴 기술에 대항하여 증거를 찾아내는 안티 포렌식 대응 기술 역시 디지털 기기에 대한 전문적인 지식을 바탕으로 하며, 범죄자가 기존에 습득한 기술과 범죄의 의도를 제대로 파악해야만 해당 증거를 찾아낼 수 있다는 점에서 매우 높은 기술력을 요한다.

본 논문에서는 현존하는 다양한 형태의 안티 포렌식 기술을 소개하고 이에 대한 대응 기술을 간략히 언급하고자 한다. 안티 포렌식 기술에 대한 이해는 해당 기술을 이용한 범죄의 동기를 줄일 수 있을 뿐만 아니라 디지털 포렌식 기술의 발전 그리고 디지털 포렌식 수사의 질을 향상시키는 데 큰 도움이 될 것으로 판단된다.

Ⅱ. 데이터 삭제 기술

2.1. 데이터 영구 삭제

일반적인 디지털 데이터는 자기적 매체에 저장되는데, 이 때 자기적 매체란 하드 디스크, 플로피 디스크, 자기 테이프 등을 의미한다. 이러한 자기적 매체는 마그네틱 도메인(Magnetic Domain)이라 불리는 매우 작은

본 연구는 정보통신부 및 정보통신연구진흥원의 IT신성장동력핵심기술개발사업의 일환으로 수행하였음.
[2007-S019-01, 정보투명성 보장형 디지털 포렌식 시스템 개발]

* 고려대학교 정보경영공학전문대학원 (gosky7@korea.ac.kr, danver123@korea.ac.kr, sangjin@korea.ac.kr, hsh@cist.korea.ac.kr)

영역으로써 구성되는데 이것들은 자기적 성질을 각 상황에 따라 변화시키는 구성요소이다. 이러한 자기 성질을 없애거나 변형을 시도하면 그에 담겨진 데이터 역시 삭제할 수 있다.

데이터를 영구적으로 삭제하기 위하여 디가우저를 사용하는 경우가 많다. 디가우저(Degaussers, 소자장비)는 자기 테이프와 디스크 미디어에 있는 데이터나 신호들을 완전히 제거하기 위한 장치(Sanitizing Equipment)이다. 디가우징(Degaussing)은 자기 장치가 다른 강력한 자기장에 노출 되는 일련의 과정을 의미한다.

자기적 매체를 디가우징 한다는 것은 이러한 장치가 띄는 자성을 초기 상태로 만들거나 자기적 랜덤(Random)한 상태로 남겨두면서, 이전에 쓰였던 모든 데이터를 제거한다. 따라서 디가우징 후에 이전의 데이터는 거의 복구가 불가능 하다. 그러나 디가우징 후에도 자성의 정렬이 랜덤화 되지 않는 영역이 있을 수 있다. 이러한 영역을 '잔류 자기 영역'이라고 한다. 디가우징을 올바르게 수행했다면 이러한 잔류자기로부터 원본 데이터를 복구하는 것이 불가능해야 한다.

디가우징을 통한 데이터의 삭제는 2가지 방법으로 이루어질 수 있다.

첫 번째, AC Erasure 방법은 원본 데이터가 그 것보다 높은 전기나 자성으로 여러 번 자극을 받는 것이다. 즉 자성이 훨씬 강한 신호에 장치를 여러 번 노출시키는 방법이다.

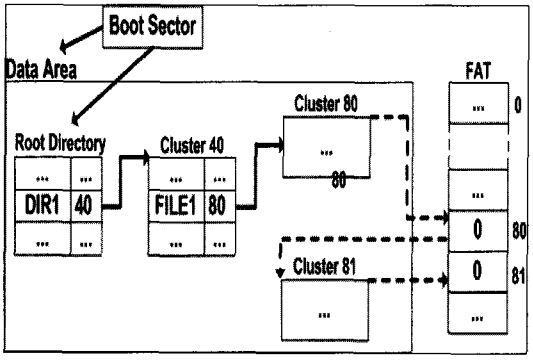
두 번째, DC Erasure 방법은 영구 자석으로 자기적 매체를 천천히 초기화하는 것이다. 즉 자성을 띄는 장치에 자기적 매체를 오랜 시간 지속적으로 노출시키는 것이다. 기록된 데이터들을 지우기 위해서는 자기적 매체가 가지는 자력의 크기보다 그 장치에 가하는 자기력이 더 강해야 한다. 특히, 보자력이란 강자성체의 자력인데, 이 보자력이 기존 장치가 가진 자력을 변화시킬 수 있을 정도로 커야 한다. 자기적 성질을 띤 매체에 기록된 보자력보다 강력한 자력을 자기 매체 면에 직접 노출시켜야 완벽하게 자기매체 상의 기록된 데이터를 삭제할 수 있다.

이러한 디가우저 장비는 현재 다양한 제품들이 상품화되어 판매되고 있으며, 또한 누구나 쉽게 입수할 수 있다. 디가우저 장비에 대한 이해와 그 대응책이 필요한 실정이지만, 디가우저로 삭제된 데이터를 복구하기란 현실적으로 상당히 어렵다.

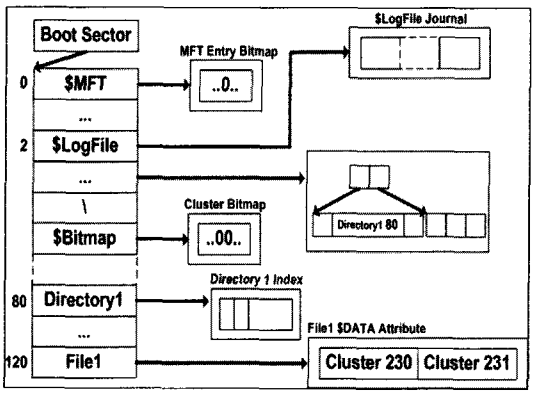
2.2. 하드 디스크 및 파일 와이핑(Wiping)

와이핑 기술은 개인 정보보호 및 중요 국가 기밀 자료의 외부 노출을 방지하기 위한 보호 수단으로 사용되는 방법이지만, 이러한 보호 수단 역시 관점을 달리하여, 증거를 삭제할 경우 에도 사용될 수 있다. 먼저 일반적인 파일 시스템의 파일 삭제 방법에 대한 이해가 필요하다.

파일 시스템이 FAT인 경우 파일의 디렉토리 엔트리(Directory Entry)에서 삭제하고자하는 파일의 Name[0]값을 0xE5로 표기한다. 이는 해당 디렉토리 엔트리가 담고 있는 데이터는 삭제된 데이터라는 것을 의미한다. FAT파일 시스템에서는 파일이나 디렉토리가 삭제되면 해당 디렉토리 엔트리를 초기화하는 것이 아니라 단순히 Name항목의 첫 번째 바이트를 0xE5로 바꾼다. 그러므로 파일이 삭제되어도 파일의 내용은 그대로 존재하고 있다[2].



〔그림 1〕 FAT 파일 시스템



〔그림 2〕 NTFS 파일 시스템

〔표 1〕 Hard disk 관련 보안 규정

국가	적용 기준	특징
미국 (국방성 안보국 기준)	Overwrite 도구: 7회 이상 실시, 분해 장비: 1/4 inch 이하로 분해 Eraser 장비: 대상 디스크의 보자력 보다 강한 장치에 통과	많은 시간 소요 (1회:2~3시간) 별도의 분쇄장비 필요
캐나다	Overwrite 도구: 3회 이상 실시, 분해 장비: 1/4 inch 이하로 분해 Eraser 장비: 대상 디스크의 보자력 보다 강한 장치에 통과	많은 시간 소요 (1회:2~3시간) 별도의 분쇄장비 필요
한국(국정원 정보시스템 불용 처리 지침 기준)	Overwrite 도구: 3회 이상 실시, 분해 장비: 0.25mm 이하로 분해 Eraser 장비: 대상 디스크의 보자력 보다 강한 장치에 통과	많은 시간 소요 (1회:2~3시간) 별도의 분쇄장비 필요(CD, FDD)

NTFS의 경우, MFT 엔트리의 $INDEX_ROOT 속성으로 지우고자 하는 파일의 엔트리를 찾는다. $INDEX_ROOT 속성은 인덱스의 루트에 해당하는 인덱스 노드를 담고있는 속성이다. NTFS는 인덱스의 최상위 노드를 빠르게 찾기 위해 별도로 속성을 만들었다. 찾은 파일엔트리의 헤더 중, 플래그 항목을 비 할당으로 설정하고 엔트리 비트맵에서 해당 영역을 0으로 설정하는 것으로 삭제한다[2].

이와 같이 일반적인 파일 시스템에서는 엔트리의 연결을 끊는 방식으로 파일을 삭제하고 있기 때문에 엔트리 정보를 가지고 있는 영역과 해당 클러스터가 덮여쓰여지지 않았다면 파일의 이름과 시간정보를 완전히 복구할 수 있다. 이를 방지하기 위하여 해당 디스크 영역에 '난수' 혹은 '0'으로 중복 덮어쓰는 기법을 와이핑(wiping)이라 한다[3].

데이터 복구를 막기 위한 방법 중 데이터 부분을 덮어쓰는 과정을 반복하는 방법이 있다. 미 국방성(DoD)의 경우 기밀 자료를 삭제하기 위한 표준(DoD5220, 22-M)을 다음과 같이 제시한다[4].

〔표 2〕 증거 삭제 프로그램 목록

증거 삭제 프로그램	삭제 대상
Evidence Elimibator (http://www.evidence-eliminator.com)	웹 페이지, 그림, 동영상, 음성 파일, E-mail
CyberScrub (http://www.cyberscrub.com/)	그림, 동영상, 히스토리, 웹사이트 주소록, 윈도우 임시파일, 메신저 데이터, E-mail, P2P 어플리케이션 데이터
Window Washer (http://www.webroot.com/)	휴지통, 파일 검색 목록, 최근 접근 문서 목록, 윈도우 임시파일, 쿠키, 히스토리, 저장된 웹사이트 주소록, 저장된 사용자 패스워드, 설치된 Active-X 데이터
Max Pc Privacy (http://www.maxpcprivacy.com/)	윈도우 캐시 데이터, 쿠키, index.dat, 휴지통, 윈도우 임시 파일, 최근 문서 파일, 윈도우 플러그 인, 메신저 대화 기록
Windows Internet Cleaner (http://www.neoimagic.com/)	웹 브라우저 캐시 데이터, 히스토리, 쿠키, 저장된 주소목록, index.dat, 웹 브라우저 자동완성 데이터, 최근 문서 목록, 휴지통, 윈도우 임시파일, 웹 브라우저 플러그 인, 윈도우 클립보드, 윈도우 미디어 플레이어 재생 목록
Privacy Guardian (http://www.pctools.com/)	웹 브라우저 캐시 데이터, 쿠키, index.dat, 히스토리, 휴지통, 윈도우 임시파일, 최근 문서(이미지)목록, 윈도우 플러그 인, P2P 데이터, 메신저 대화 기록
Privacy Eraser (http://www.privacyeraser.com/)	웹 브라우저 캐시 데이터, 쿠키, index.dat, 히스토리, 저장된 주소목록, 웹 브라우저 자동완성 데이터, 최근 문서 목록, 휴지통, 윈도우 임시파일, 웹 브라우저 플러그 인, 윈도우 미디어 플레이어 재생목록
Privacy Protector (http://www.zj-fountain.com/)	저장된 웹 사이트 주소목록, 쿠키, 인터넷 캐시 데이터, 히스토리, 웹 브라우저 자동완성 데이터, 웹 브라우저 플러그 인, index.dat, 최근 문서 목록, 윈도우 검색어 목록, 윈도우 시작 프로그램 목록, 윈도우 임시 파일, 휴지통, 윈도우 미디어 플레이어 재생 목록
Tracks Eraser Pro (http://www.acesoft.net/)	저장된 웹 사이트 주소목록, 쿠키, 인터넷 캐시 데이터, 히스토리, 인터넷 검색 디렉토리, 검색 히스토리, 웹 브라우저 자동완성 데이터, index.dat, 웹 브라우저 플러그 인, 윈도우 임시 파일, 휴지통
Internet Washer (http://www.internetwasher.net/)	최근 문서 목록, 윈도우 검색어 목록, 윈도우 임시 파일, 윈도우 클립보드, 쿠키, 메신저 대화 기록, 방문한 웹 사이트

1. 임의의 문자로 데이터를 덮어씀.
2. 첫 번째 문자의 보수로 덮어씀.
3. 다시 임의의 문자로 데이터를 덮어씀.
4. 이 과정을 7회 반복.

국내의 경우 정보시스템 저장매체 불용처리지침에 따르면 저장매체 전체의 자료 저장 위치에 새로운 자료(0 또는 1)를 중복하여 저장하는 과정을 3번 반복하도록 규정하고 있다.

2.2. 증거 데이터 자동 삭제

증거 데이터를 삭제한다는 것은, 컴퓨터 운영체제에 의해 생성되는 사용자의 개인 정보 또는 증거물이 될만한 모든 데이터를 삭제하는 행위이다. 웹페이지, 문서, 그림, 동영상, 음성파일, E-mail, 레지스트리, 쿠키 그리고 히스토리 등이 주로 삭제 대상이 된다.

웹 페이지는 사용자가 방문한 사이트의 웹페이지로서 이미지나 웹 스크립트 등을 저장하고 있다. 쿠키 값은 사용자가 방문한 웹 사이트 정보를 가지고 있으며 index.dat에는 사용자의 인터넷 사용 내역을 기록하고 있다. 인터넷 히스토리에는 사용자가 접속했던 인터넷 사이트가 기록되어 있다. 이와 같은 인터넷 사용 내역에 대한 기록은 범죄 해위에 대한 단서나 그 정황 증거에 대한 정보를 제공하기 때문에 범죄자들이 삭제하고자 한다.

또한 시스템에서의 최근 문서 목록은 사용자가 최근 열람하거나 수정한 문서 목록이 저장되어 있고, 휴지통에는 사용자가 삭제한 파일을 복원할 수 있도록 저장하고 있다. 윈도우 시작 프로그램에는 사용자가 지정한 윈도우 시작 프로그램이 지정되어 있으며 저장된 인터넷 주소 목록에는 사용자가 방문한 웹사이트의 주소 목록이 저장되어 있다. 또한 P2P 프로그램 정보에는 P2P 프로그램으로 주로 사용되는 프로그램의 설정 파일을 저장하고 있고, 윈도우 미디어 플레이어 재생 목록에는 사용자가 재생한 미디어 목록이 저장되어 있다. 또한 윈도우 검색어 목록에는 사용자가 검색한 파일의 목록이 저장되어 있다. 앞서 설명한 것과 같은 맥락으로 이러한 정보를 삭제하고자 하는데, 인터넷 사용정보 및 관련 시스템 정보를 자동으로 삭제해주는 프로그램들이 현재 존재하며, 인터넷이나 P2P 사이트에서 쉽게 입수할 수 있다.

위와 같은 증거 데이터를 자동으로 삭제할 수 있는 도구에는 [표 2]와 같다.

Ⅲ. 데이터 은닉과 변조 기술

3.1. 스테가노그래피

스테가노그래피기법은 문서의 저작권을 보호하는 데에 사용되기도 한다. 하지만 이러한 기술이 범죄 사실을 은닉하거나 비밀 메시지를 전송하는데 쓰여지곤 한다.

스테가노그래피(steganography)는 메시지(massage)가 전송되고 있다는 사실 자체를 은닉함으로써 데이터를 숨기는 정보 은닉 기술이다. 이는 데이터 암호화와 약간 대조적인 기술인데 암호화는 데이터를 획득할 수는 있지만 스테가노그래피는 데이터 자체를 획득할 수 없다는 점이 다르다. 주로 스테가노그래피 기술은 이미지 데이터에 정보를 숨기는 데에 사용된다.

일반적으로 스테가노그래피 기법을 사용하여 은닉된 메시지는 다른 형태의 데이터에 내재되어 있다. 즉 그림, 기사, 쇼핑 목록 혹은 또 다른 메시지들 속에 포함되어 있다. 이런 식으로 숨겨진 메시지를 나타내고 있는 메시지를 'Cover Text'라고 한다. 평범한 문서에 나타난 선들 사이에 보이지 않는 링크(link)를 설정해두고 메시지를 숨겼다면 그 문서가 Cover Text가 된다. 은닉될 메시지를 평문(plaintext)이라 하면, 이 평문은 종종 우선 암호화되고(chiphertext) 그 다음에 Cover Text가 이 평문을 포함하도록 한다. 이렇게 조작된 메시지를 스테고텍스트(stegotext)라 명명한다.

이러한 스테가노그래피 기술을 이용한 파일은 실제 탐지도 어려우며, 정보를 은닉한 이미지 파일을 획득하였더라도 정보를 추출하고 확인하기 위해서는 사용자가 설정한 비밀번호를 알고 있어야 한다. 때문에 실제 정보를 확인하기가 쉽지 않다. 이러한 스테가노그래피 기술의 특성 때문에 범죄자들이 안티 포렌식 기술로 이용할 가능성이 높다[5].

3.2. 디스크내 데이터 은닉

포맷된 하드 드라이브는 물리적 장치에 대응된 논리적 구조라고 생각할 수 있다. 이러한 논리적 구조는 파티션(partition), 파일 시스템(file system), 파일(file), 레코드(record), 그리고 여러 field로 구성되어 있다. 물리

적 장치는 디스크(disk), 실린더(cylinder), 트랙(track), 클러스터(cluster) 그리고 섹터(sector) 등으로 구성되어 있다. 이러한 논리적 영역과 물리적 영역 사이의 1:1 대응을 하였을 때 발생하는 차이는 데이터를 숨기기 위한 은닉 영역이 되기 충분하다[14].

현재 존재하는 포렌식 수사 툴들은 이러한 디지털 잉여 공간을 밝혀내지 않고 있고, 데이터 영역의 내용에만 탐지를 하고 있다. 포렌식적으로 의미가 있는 접근은 이미 알려진 데이터 영역에 대한 접근이 아니라 알려진 데이터 영역이라 하더라도 그 곳의 어디에 데이터를 은닉했는가에 중점을 두는 것이다.

예를 들면, 새로운 컴퓨터의 하드디스크를 포맷할 때 2개의 예약된 영역을 생성한다. Host Protected Area(HPA)와 메타 데이터를 위한 Device Configuration Overlay (DCO)영역이다. 일반적으로 쓰여지지 않는 영역이며 하드웨어 제조사에서 특별한 용도로 사용하기 위해 하드디스크에 남겨두는 공간이다.

운영체제를 통한 이러한 영역으로의 접근은 디스크 컨트롤러에 의해서 제한되는데, 이 영역에 접근하기 위해서는 디스크 컨트롤러를 직접 제어하는 low level에서 작업이 필요하다. 따라서 디스크 컨트롤러에 접근할 수 있다면, HPA나 DCO영역에 데이터를 숨길 수 있다.

이는 하드 디스크이 구조와 물리적인 경계선을 알고 있고, live CD[15] 같은 또 다른 OS로 부팅할 수 있다면 그다지 어려운 작업이 아니다.

그리고 파티션(partition)영역을 생각할 수 있다. 최근의 운영체제는 관리자가 디스크를 임의의 크기와 여러 개의 파티션으로 분할 가능하게 한다. 따라서 각각의 파티션으로 분할된 영역에 다른 운영체제를 설치하고 여러 어플리케이션들을 설치할 수 있다. 여기서 정보 은닉의 가능성을 찾을 수 있다. 왜냐하면 논리적 파티션은 디스크의 실제 물리적 파티션 영역과 완전히 일치하지 않기 때문이다. 파티션으로 인한 슬랙은 논리적 파티션의 끝 부분과 물리적 블록의 파티션의 끝 부분의 차이에서 발생한다.

슬랙 공간에는 어떤 어플리케이션을 설치한다던가, 어떤 운영체제를 설치한다던가 하는 일은 할 수 없지만, 데이터는 숨길 수 있다. 또한 확장 파티션은 임베디드(embedded) 논리적 파티션들을 가능하게 함으로써 더욱 데이터를 숨기기가 용이하다. 왜냐하면 이때의 각각의 논리적 파티션들은 데이터를 여분으로 숨길 수 있는 공간을 62섹터씩 더 확보하기 때문이다[2].

또한 디스크 상의 슬랙 공간을 생각할 수 있다. 디스크 슬랙 공간이란 파일의 크기가 데이터 단위 크기의 배수가 되지 않을 때, 저장 매체에서 파일이 저장되고 남은 잉여 공간을 지칭한다. 예를 들어, 파일이 2,816byte이고, 클러스터 크기가 4,096byte 이면 나머지 1,235byte는 슬랙 공간이 된다. 슬랙 공간은 두 부분으로 나뉘는데, 슬랙공간의 처음부터 섹터가 끝나는 부분은 운영체제가 파일의 끝임을 표시하기 위해 0값으로 채우는 램 슬랙이라고 부르는 공간이며, 나머지 부분은 운영체제에서 아무런 조작을 실시하지 않은 공간인 일반 슬랙 공간이다[7].

3.3. 암호화

안티 포렌식 기술이 포렌식 수사를 더디게 하는 기술임을 감안할 때, 가장 대표적인 안티 포렌식 기술은 '암호화'라고 할 수 있다. 이 역시 정보를 보호하는 수단으로 일반적으로 사용되는 기술이지만 안티 포렌식 기술로도 사용될 수 있다. 파일이나 디렉토리를 암호화하여 아무나 접근할 수 없도록 하는데 목적이 있다. 해당 파일이나 디렉토리를 암호화 하는 것은 그 암호를 어떤 방법으로든 풀어야 하는 과정을 거쳐야 원하는 파일을 얻을 수 있기 때문이다. 즉 데이터 암호화는 의미 있는 평문파일을 어떤 암호화키를 이용하여 엔트로피가 높은 다른 형식의 파일로 변환시키는 기술이다.

3.4. 데이터 변조 기술

데이터 변조 (Obfuscation)는 데이터를 포렌식 분석하는데 어려움을 주기 위한 작업이다. 가장 기초적인 것은 파일의 확장자를 변경하거나 파일의 헤더 정보 등을 조작하여, 실제 원본 파일의 내용과 시스템에서 인식되는 파일 형식을 다르게 할 수 있다. 따라서 수사관이 특정 파일을 검색할 때 회피 수단으로 사용할 수 있다.

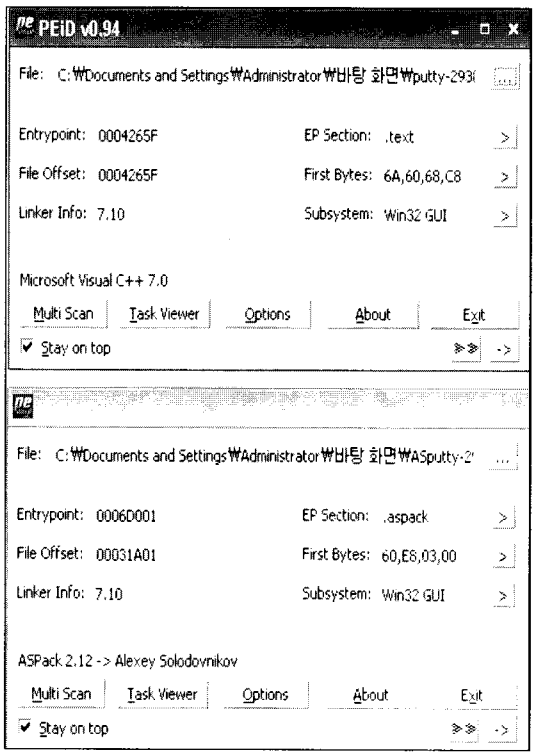
실행파일의 경우에는 기존 코드의 기능은 그대로 유지할 수 있고, Reverse Engineering이 어렵도록 소스코드를 재구성하여 분석에 필요한 비용과 시간을 증가시키는 기술이 사용된다. 하지만 이 역시 악성코드를 작성하여 이를 탐지하지 못하도록 하는데 사용되고 있다. 대표적으로 실행압축(Packing) 기술이 있다.

3.4.1. 실행압축(Packing) 원리

일반적인 압축과 같이 여러 파일을 하나로 묶어 압축을 수행하는 것이 아니라, 각각의 실행 파일 형태를 그대로 유지하면서 크기를 줄여 주는 압축방식을 사용한다. 이에 관한 패킹 툴의 종류와 해당 버전이 다양하게 존재한다.

Packing을 수행하는 프로그램이 압축을 수행할 프로그램의 실제 Code 및 Data를 다른 영역에 압축, 저장한다. 그러고 나서 프로그램의 Entry Point를 실행하여 압축 해제루틴을 먼저 가리키게 한 후, 실제로 실행압축해제(Unpacking)가 먼저 이루어진 후에 프로그램이 동작하는 방식이다[16]. 실제 data packing을 수행한 방법은 아래 [그림 3]과 같다. 패킹을 수행한 후, PE explorer를 이용하여 원본 파일과 패킹된 파일의 헤더 섹션의 차이점을 분석하였다.

[표 2]에 나타나듯이 변경된 Field 값으로 패킹된 이후 새로운 섹션이 2개 추가되었고, EntryPoint의 위치가 변경되었으며, 파일의 사이즈와 헤더의 크기가 변경되었다. 그리고 File Alignment 값으로 파일 상에서의 섹션의 배치간격이 줄어들었다.



〔그림 3〕 putty-2938-rm2rob.exe 실행

〔표 3〕 변경된 필드 값

필드네임	패킹된 putty	원본 putty
NumberOfSection	0006h	0004h
AddressOfEntryPoint	0046D001h	0044265Fh
FileAlignment	00000200h	00001000h
SizeOfImage	00071000h	0006D000h
SizeOfHeaders	00000600h	00001000h

Ⅳ. 안티 포렌식 대응 기술

4.1. 스테가노그래피 탐지

스테가노그래피를 이용하여 인코딩 된 메시지를 찾는 것을 'Steganalysis'라고 한다. 이러한 파일을 찾는 가장 간단한 방법은 이를 원본 파일과 비교해보는 것이다. 만약 단 하나의 정보라도 다른 것이 있다면 둘 중 하나는 조작된 파일이라고 예상할 수 있다.

일반적으로 압축률이 높은 압축 알고리즘으로 압축된 이미지는 데이터 은닉이 불가능하지는 않지만 어려워진다. 극단적인 경우에 높은 압축률로 압축을 하면 숨겨진 데이터를 찾기가 훨씬 쉬워진다.

Steganalysis는 이미지 속의 메시지를 찾는 것이 목적이라기보다는 이미지 속에 숨겨진 파일이 존재한다는 것을 탐지하는 것이 목적이다.

특별한 경우에 시각적인 탐지방법의 Steganalysis를 시도할 수도 있다. 즉, 눈으로 이미지의 변화를 감지하는 방법이다. 통계적인 기술을 이용하여 이미지의 픽셀 간의 점진적인 색상 변화를 비교하는 방법으로 시각적인 탐지를 한다. 간단히 말하면, 이미지 내에서 특정 필셀 근방(neighborhood)에서 색상의 순서쌍을 측정하여 서로 색상 간의 차이가 얼마나 나는 지를 계산하는 것이다. 정보가 은닉되지 않은 일반적인 이미지는 이 차이가 매우 작은 데 비해 정보가 은닉된 이미지는 일부 영역에서 이 차이가 원본 이미지와는 크게 차이가 난다.

이와 같이 다양한 스테가노그래피 탐지 기법들이 연구되고 있지만, 현실적으로 상품화하여 적용하기에는 많은 무리가 있다. 앞서 말한 바와 같이 이미지 속에 숨겨진 데이터가 있다는 것을 탐지만 한다고 하더라도, 수사에 도움을 줄 수 있으며, 그것을 바탕으로 수사를 진행할 수가 있다.

4.2. 데이터 검색 및 탐지

많은 경우 하드 디스크에 숨겨진 데이터는 사이버 범죄를 해결하는데 매우 중요한 역할을 한다. 이들 데이터 중 일부는 파일이 삭제되거나 디스크가 다시 파티션 될 때 남아있던 '주변' 데이터 일수도 있고 사용자가 일부러 은닉해 둔 데이터일 수도 있다.

하드 디스크를 조사하는 방법은 파일 위주로 조사하는 방법, 슬랙 공간을 위주로 조사하는 방법 그리고 비할당영역을 위주로 조사하는 방법 등이 있다. 주로 EnCase와 같은 포렌식 수사 도구를 이용하여 조사하게 되는데 이는 키워드를 선정하여 탐색하는 형태로 이루어진다. 증거를 탐색하는 방법을 명확히 해야 하는데, 이는 주로 2가지로 구분된다. Index에 기반한 탐색 방법과 Bitwise 방법이 있다.

Index에 기반한 탐색 방법은 모든 파일들을 훑으면서 키워드를 찾아나간다. 이는 수사관이 키워드를 어떻게 만드느냐에 매우 의존적인 방법으로 EnCase와 같은 포렌식 수사 도구를 이용하여 일반 드라이브, 이미지, 파티션등의 모든 영역을 모두 검색해야 한다. 이러한 검색 방법은 몇 시간에서 몇일까지 소요될 수 있지만, Index 탐색은 2가지의 장점이 있다. 우선, 이 방법은 파일에 기반한 탐색 방법이기 때문에 이는 특정 파일의 특정 포맷에 관계없이 관련된 모든 파일을 검색할 수 있다. 즉, PDF 파일이든, XLS파일이든, 심지어는 압축된 파일일지라도 파일의 특성에 관계없이, 그리고 각 파일들이 생성된 시간에 관계없이 양적으로 많은 파일들을 확보할 수 있다는 장점이 있다. 두 번째로는 첫 번째 색인을 만든 후에는 그 속도가 매우 빠르다는 것이다. 이는 증거 탐색에서의 효율성이 매우 높음을 의미한다[9].

Bitwise 방법은 디스크 내의 섹터나 슬랙 공간에서 찾을 수 있는 비 할당 영역에 존재하는 간단한 텍스트나 특정 표현들을 찾는 방법이다. 이 방법은 수사관으로 하여금 일반적인 텍스트가 아닌 형식의 표현들을 찾을 수 있게 한다. 즉 파일 헤더와 같은 이진수 표현을 검색할 수 있다. 파일 헤더를 이용할 수 있는 이러한 검색은 해당 개체뿐 아니라 다른 파일 속에 삽입된 해당 포맷의 파일까지 찾아낼 수 있는 장점이 있다. 그러나 방법의 특성상 index 탐색 방법보다는 그 속도가 매우 느리다. 또한 파일의 경계(boundary)를 찾아낼 수는 없는데 이는 주로 심하게 단편화 된 디스크에서 나타나는 현상이다. 단편화가 심한 디스크는 많은 파일이 불연속적인 섹터에 저장되는 경우가 많기 때문이다. 따라서 파일의 경계를 인식하지 않은 키워드 선정은 결정적인 증거를 놓칠 수 있다.

4.2.1. 파일 포맷 분석 및 해쉬 검증

알려진 파일을 찾을 때는 MD5나 SHA-1과 같은 파일의 해쉬 값을 비교해서 찾을 수도 있다. 알려진 파일의 해쉬값은 National Software Reference Library에 테이블 형태로 제공되고 있다[6].

해쉬 값을 분석함으로써 해당 파일을 찾는 경우는 2가지 유형으로 나뉠 수 있다. 우선 앞서 언급했듯이 기존에 존재하는 해쉬 테이블을 기반으로 하는 Positive 해쉬 분석과 알려진 해쉬 값이 없이 분석에 임해야 하는 Negative 해쉬 분석이 있다.

Positive 해쉬 분석의 경우, 해쉬 값은 유일하고 그 컨텐츠를 나타내는 값이므로 해당 파일의 이름, 속성, 날짜 등의 메타데이터가 변경되더라도 그 파일을 찾을 수 있다.

Negative 해쉬 분석의 경우, 이미 알려진 해쉬 값을 가진 파일을 모두 제외하고 알려지지 않은 해쉬 값을 가진 파일을 우선 수집한다. EnCase의 경우 이러한 파일들을 배제하거나 혹은 따로 수집할 수 있는 기능을 가지고 있다. 이러한 파일들의 분석 후 해쉬 테이블에 추가함으로써 차후의 분석에 기여할 수 있다[9].

또한, 파일의 포맷을 분석함으로써 원하는 파일을 찾을 수도 있다. 대개의 경우 각 파일의 시그니쳐(signature)에 의존하는 경우가 많은데 이 역시 파일 포맷을 분석한 결과 얻은 결과이다. 이는 누군가 해당 파일의 이름을 바꾸거나 확장자를 바꾸어도 그 파일을 찾을 수 있는 가장 좋은 방법이다.

4.3. 데이터 복구

일반적으로 파일이 삭제되었다 함은 디스크 상에서 해당 파일이 사라져 버린 것이 아니다. 파일이 삭제되었다고 느끼는 이유는 그 파일로의 링크(link)가 삭제 된 상태이기 때문이다. 따라서 삭제한 파일은 새로 덮여 쓰여지지만 않았다면 디스크의 어느 부분에 남아있을 수 있으며 이는 삭제된 파일을 복구할 수 있는 근거가 된다. 대부분의 포렌식 도구들은 이러한 데이터 복구 기능을 제공하고 있다.

4.3.1. 슬랙공간 및 미할당 영역에서의 복구

앞에서 설명한 바와 같이, '슬랙공간'이란 파일의 크기가 데이터 단위 크기의 배수가 되지 않을 때 즉, 저장매체에서 파일이 저장되고 공간이 남았을 때 남은 잉여공간을 지칭한다. 슬랙공간은 원본 파일의 첫 부분의 정보가 완전히 사라져, 삭제 파일 복구로도 원본 파일을 복원할 수 없다. 그러나 파일의 가독성 있는 일부 정보가 남아있을 수 있으며, 파일의 유실 정도가 낮다면 사라진 파일 헤더를 적절히 재구성하여 본래의 파일로 복구할 여지가 있다. 때문에 각종 디지털 포렌식 도구에서는 슬랙공간 검색 기능을 제공한다[7].

4.3.2. 메모리 및 스왑(Swap) 영역 검색

메모리에는 캐쉬 데이터, 라우팅 정보, 프로세스 정보, 응용 프로그램이 사용했던 데이터들이 남아있다. 또한 메모리에는 패스워드나 암호문의 평문이 존재할 가능성이 있는데 따라서 메모리에 남아있는 데이터를 분석하는 것은 매우 의미가 있다. Swap 파일은 가상메모리 시스템에서 사용되는 메모리 공간이며 하드디스크의 일부 공간을 메모리처럼 사용한다. 윈도우 시스템에서는 Pagefile.sys로 존재하는 데, 하드디스크의 섹터를 직접 읽어 파일 접근 시간 등을 변경하지 않고 파일을 획득할 수 있다[10].

메모리에는 운영체제 및 각종 응용프로그램들이 사용하고 있던 정보들이 담겨져 있다. 이러한 정보들은 시스템이 동작하고 있는 동안에만 유지되는 정보들이기 때문에 반드시 획득해야 한다. 프로세스 정보와 네트워크 상태 등의 정보를 휘발성이 높은 순으로 판별하고 획득하는 방법을 연구한다. RAM에 존재하는 Memory data는 운영체제 영역과 사용자 응용프로그램 영역으로 분리되어 있다. 운영체제 영역의 메모리는 현재까지 획득할 수 있는 도구가 존재하지 않는다. 따라서 운영체제 영역의 메모리까지 획득할 수 있는 방안을 연구하고, 응용프로그램의 메모리는 사용자가 선택하여 수집할 수 있는 도구가 필요하다.

Swap 파일은 가상메모리 시스템의 일부분으로써 보통 시스템 RAM 크기의 1.5배의 크기로 하드디스크에 할당된다. 시스템이 동작하고 있는 활성 시스템에서는 운영체제가 파일을 점유하고 있어 일반적인 파일 복사 방법으로는 획득이 불가능하다. 따라서 하드디스크의 파일 시스템을 분석하고, Swap 파일을 하드디스크로부터 직접 물리적으로 읽어 들이는 방식을 사용한다.

NTFS 파일 시스템에는 시스템의 모든 파일에 대한 정보를 저장하고 있는 $MFT라는 특수한 파일이 존재한다. $MFT 파일에는 Pagefile이 존재하는 하드디스크의 위치정보를 담고 있으며, 이를 분석하여 하드 디스크의 Pagefile을 직접 획득할 수 있는 Swap 파일 수집 소프트웨어가 필요하다[2][11].

4.4. 패스워드 검색

패스워드를 알아내는 가장 쉬운 방법은 무작위 공격(Brute Force Attack)이다. 무작위 공격은 패스워드를 추측하기 위해 시도하게 되는 모든 가능한 값을 체계적으로 생성하여 패스워드를 찾아나간다. 하지만 시간이 오래 걸린다는 단점이 있다. 사전 공격(Dictionary Attack)은 시스템 내에서 높은 가능성을 지니고 사용되는 단어(이름, 장소 등)들을 사용하여 가능한 모든 패스워드를 체계적으로 찾아낸다. 단어 사전은 패스워드를 사전에 모든 단어들을 조사하여 공격자에게 알려준다. 2가지 방법 중에서 상황에 맞게 적절하게 사용되어야 하지만, 범죄 수사는 신속하게 진행되어야 하기 때문에 사전공격을 더욱 개량하고 정확성이 높은 사전을 구성하는 방법에 대한 연구가 진행되어야 한다.

V. 결 론

안티 포렌식 기술을 알고 이에 대응 방안들을 숙지하는 것은 현재 디지털 포렌식 수사에서는 꼭 필요한 사항이다. 대부분의 디지털 포렌식 툴들은 은닉된 데이터나 변조된 데이터들을 자동으로 탐지할 수 없다. 따라서 디지털 포렌식 수사관의 능력이 매우 중요하며, 관련 전문 지식을 습득할 필요가 있다. 또한 범죄자들이 습득한 안티 포렌식 기술과 범죄 의도들을 제대로 파악할 수가 있어야 한다.

본 논문에서는 안티 포렌식 기술들과 그에 대한 대응 방안을 간략하게 기술하였다. 하지만 안티 포렌식 기술은 점점 발전하고 있기 때문에 향후에도 지속적으로 동향을 파악해야 하며, 그에 대응할 수 있는 기술 연구가 필요하다.

참고문헌

[1] Ryan Harris, "Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem", Digital Forensic Research Workshop, Digital Investigation, Elsevier, 2006

[2] Brian Carrier. "File System Forensic Analysis", Addison Wesley, 2005.

[3] File wipe Definition: TechEncyclopedia from TechWebhttp://www.techweb.com/encyclopedia/defineterm.jhtml?term=Filewipe

[4] Peter Gutmann, "Secure Deletion of Data from Magnetic and Solid-State Memory", the Sixth USENIX Security Symposium Proceedings, July 22-25, 1996

[5] Husrev T. Sencar, Mahalingam Ramkumar, Ali N. Akansu, "Data Hiding Fundamentals and Applications", Elsevier Academy press, 2004

[6] NSRL, http://www.nsrl.nist.gov

[7] Eoghan Casey, "Digital Evidence and Computer Crime", p. 62~p.63, ACADEMIC PRESS 2003

[8] 김현상, 박상헌, 이상진, 임종인, "디지털 포렌식을 위한 최적화된 슬랙 공간 검색 기법", 한국정보보호학회 동계학술대회, 2005

[9] Richard P. Salgado, "FOURTH AMENDMENT SEARCH AND THE POWER OF THE HASH", Searches and Seizures in a Digital World, 119 HARV. L. REV. 531, 2005

[10] C. Kevin. Incident Response and Computer Forensics. McGraw-Hill, 2003.

[11] Seokhee Lee, A. Savoldi, Sangjin Lee, and Jongin Lim, "Password recovery using an evidence collection tool and countermeasures", Intelligent Information Hiding and Multimedia Signal Processing, Proc. IEEE, 2007.

[12] 이형우, 이상진, 임종인, "컴퓨터 포렌식스 기술", 한국정보보호학회지, 2002년 10월

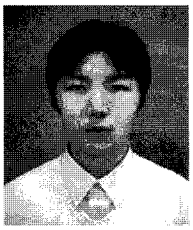
[13] Department of Defense 5220-22-M Standards

[14] 이홍재, "하드 디스크 이해", 전자신문사 , 2003

[15] Kyle Rankin, "KNOPPIX HACKS", O'RELLY, 2005

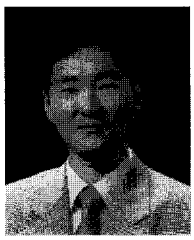
[16] "실행 파일 압축", 위키백과, http://ko.wikipedia.org/wiki

〈著者紹介〉

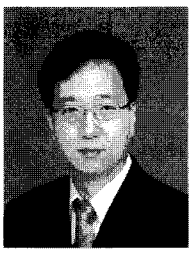


이 석 희 (Seokhee Lee)
학생회원
2003년 : 부경대학교 컴퓨터공학과 졸업(학사)
2004년~2006년 : 고려대학교 정보보호 대학원 졸업(석사)
2006년 3월~ : 고려대학교 정보경영공학 대학원 박사과정
<관심분야> 디지털 포렌식, 수사자동화

박 보 라 (Bora Park)
학생회원
2007년 2월 : 부산대학교 수학교육과 졸업(학사)
2007년 3월 ~ : 고려대학교 정보경영공학전문대학원 석사과정
<관심분야> 디지털 포렌식



이 상 진 (Sangjin Lee)
정회원
1987년 2월 : 고려대학교 수학과 학사
1989년 2월 : 고려대학교 수학과 석사
1994년 2월 : 고려대학교 수학과 박사
1989년 2월~1999년 2월 : 한국전자통신연구원 선임 연구원,
1999년 2월~2001년 8월 : 고려대학교 자연과학대학 조교수,
2001년 9월~현재 : 고려대학교 정보경영공학 대학원 교수
<관심분야> 대칭키 암호, 정보은닉이론, 컴퓨터 포렌식



홍 석 희 (Seokhie Hong)
정회원
1995년 2월 : 고려대학교 수학과 학사
1997년 2월 : 고려대학교 수학과 석사
2001년 2월 : 고려대학교 수학과 박사
2005년 3월~현재 : 고려대학교 정보경영공학전문대학원 조교수
<관심분야> 암호알고리즘, 컴퓨터 포렌식