신규 파일 시스템에 대한 디지털 포렌식 분석 필요성 연구

이근기*, 이창훈**, 이상진*
*고려대학교 정보보호대학원
**서울과학기술대학교 컴퓨터공학과
e-mail:chlee@seoultech.ac.kr

Digital Forensic Analysis for New File System

Keun-Gi Lee*, Changhoon Lee**, Sangjin Lee*
*Graduated School of Information Security, Korea University
**Dept of Computer Science and Engineering, Seoul National University of Science & Technology

요 약

파일 시스템은 컴퓨터에서 파일이나 자료를 쉽게 발견 및 접근할 수 있도록 보관 또는 조직하는 체제를 가리키는 말이다. 기존에는 Windows에 사용되는 FAT(File Allocation Table) 파일 시스템과 NTFS(New Technology File System), Unix/Linux 등에서 주로 활용되는 ext계열 파일 시스템 등이 주된 분석 대상이었으나 스마트폰과 태블릿 PC, NAS(Network Attached Storage) 서버 등 다양한 IT기기가 보급되면서 이들 기기에서 사용되는 파일시스템을 추가적인 분석이 필요하다. 따라서 본 논문에서는 추가적으로 분석해야할 파일 시스템의 종류를 나열하고 각각의 특성을 서술하여 향후 추가 분석의 지침으로 활용하고자한다.

1. 서론

파일 시스템은 디지털 포렌식 분석에서 가장 중요하고 기본적인 연구 분야이다. 파일 시스템에 대한 이해를 통하여 운영체제와 무관하게 파일 수집의 무결성을 유지할 수도 있으며, 많은 파일이 저장된 디스크를 대상으로 신속한 파일 검색을 수행할 수 있다. 따라서 이러한 파일 시스템에 대한 명확한 이해는 디지털 포렌식 분석의 필요 요소이다.

기존에는 Windows에서 사용 중인 FAT(File Allocation Table) 파일 시스템과 NTFS(New Technology File System)이 주된 분석 대상이었으며, 이를 분석할 수 있는 다양한 디지털 포렌식 조사용 소프트웨어가 출시되어 판매되고 있다. 또한 Windows외에도 UNIX/Linux 계열의 운영체제는 주로 ext(Extended File System)2, ext3, ext4 등이 주된 분석대상이었다.

하지만 최근 스마트폰, 태블릿 PC, NAS 서버 등의 보급으로 인하여 기존에 주로 활용되는 파일 시스템 외에 다양한 파일 시스템이 활용되었으며, 이러한 기기가 조사 대상이 되었을 때, 적절한 대응이 어렵다. 따라서 다양한 파일 시스템에 대한 효율적인 대응을 위하여 체계적인 파일 시스템 포렌식 기술을 확보할 필요가 있다.

따라서 본 논문에서는 최근 많이 활용되고 있는 파일 시스템의 종류를 나열하고 각각의 특징을 고려하여 향후 디지털 포렌식 분석에 활용할 수 있도록 파일 시스템 연구 항목을 제시하고자 한다.

2. 파일 시스템 분류

파일 시스템은 사용 용도와 운영체제에 따라서 다음 [표 1]과 같이 분류될 수 있다.

[표 1]에서 확인할 수 있는 바와 같이 FAT, NTFS, ext 파일 시스템 외에도 최근 많이 활용되고 있는 HFS+ 파일 시스템과 대용량 파일을 처리하기 위한 기능을 내장하고 있는 XFS, Reiser FS등도 많이 활용되고 있다. 또한 다양한 임베디드 장비가 활용되면서 플래시 저장장치용 파일 시스템도 활용되고 있다. 또한 일반적인 파일 시스템 외에도 분산 파일 시스템인 Hadoop Distributed File System(HDFS)는 최근 SNS(Social Network Service)등 실시간으로 처리해야 하는 대용량 데이터를 처리하기 위하여 많이 활용되고 있다.

이렇게 다양한 파일 시스템을 사용하는 장비가 디지털 포렌식 분석대상이 되었을 경우, 파일을 검색하기 어렵고 증거로 사용하기 위하여 추출하기 어렵다. 따라서 이러한 파일 시스템을 정확하게 분석하기 위해서는 정형화된 기능을 중심으로 분석이 가능하도록 해야한다.

3. 파일 시스템 분석 요소

파일 시스템을 정확하게 분석하기 위해서는 다음과 같은 항목을 위주로 분석을 수행해야 한다.

- 파티션 테이블 분석

파티션 테이블은 운영체제가 부팅될 때, 디스크의 구성 요소(Layout)를 확인하기 위한 작업으로 부팅해야할 파티

<표 1> 주요 파일시스템 분류

대분류	중분류	종류	설명
일반 파일시스템	Windows	FAT	- 이동식 저장매체에 주로 활용
		NTFS	- 윈도우에서 주로 활용함
		exFAT	- Windows 7에서 추가된 FAT을 개선한 파일 시스템
	Mac/iOS	HFS+	- 애플 제품(iPod, iPhone, iPad, MAC)등에 활용함
	Unix/Linux	Ext2, Ext3, Ext4	- 리눅스의 기본 파일 시스템 - 최근 안드로이드 스마트폰에서 ext4가 많이 활용됨
		UFS	- 유닉스의 기본 파일 시스템
		ZFS	- 솔라리스 계열의 기본 파일 시스템 - 대용량 저장장치 지원 및 암호화 기능 지원
대용량 파일 시스템	-	XFS	- 대용량 파일 시스템으로 주로 활용 - 최근 NAS장비에서 많이 활용
		ReiserFS	- 대용량 파일 시스템으로 주로 활용
임베디드용 파일 시스템	-	YAFFS	- 플래시 메모리용 파일 시스템
		RFS	- 삼성에서 개발한 임베디드용 파일 시스템
		TFAT	- MS에서 개발한 플래시 메모리용 파일 시스템
분산 처리용 파일 시스템	-	Hadoop Distributed File System	- 최근 많이 활용 중인 분산처리용 파일 시스템
가상 파일 시스템	-	VMDK	- VMWare에서 활용하는 가상 파일 시스템
		VHD	- Virtual PC에서 활용하는 가상 파일시스템

션을 판단하는 작업이다. 파일 시스템의 종류에 따라 GPT, MBR 등으로 구성될 수 있으며 이 영역을 분석하여 디스크의 파티션 수와 크기 등 구성 정보를 확인할 수 있다.

• 메타 데이터 분석

파일 시스템은 통상 저장하고 있는 파일에 대한 정보를 모아서 관리하는 메타 데이터 영역이 존재한다. 이 영역은 Directory Entry, $MFT, inot, Super Block 등의 이름으로 구성되어 있으며, 파일명, 파일 크기, 파일의 위치, 시간 정보, 삭제 여부 등의 정보를 포함하고 있다.

이 영역을 분석하여 디스크의 내의 파일 목록을 신속하게 확보할 수 있다.

• 파일 검색

디스크 내의 많은 파일 중 증거로 활용할 수 있는 파일을 선별할 수 있어야 한다. 파일 검색은 메타 데이터나 파일 내의 내용을 기반으로 수행할 수 있어야 하며 대용량 데이터 중에서 원하는 파일만 신속하게 검색할 수 있는 기술이 필요하다.

• 파일 추출

특정한 파일을 정밀 분석하기 위하여 디스크나 이미지 파일 내의 파일을 별도의 파일로 추출하는 기능이다. 파일

의 메타 데이터가 변경되지 않도록 운영체제의 복사 기능이 아닌 파일 시스템 구조를 분석하여 물리적으로 파일을 추출해야 한다. 이렇게 추출한 파일에 대한 증거 무결성(Integrity)를 유지하기 위한 방안 연구가 반드시 필요하다.

• 삭제 파일 복원

범죄자는 범죄 사실을 은폐하기 위해 실제 조사 수행 전에 파일을 삭제할 가능성이 높다. 또한 일반 사용자도 시스템을 사용하면서 파일 삭제를 빈번하게 수행한다. 따라서 이런 파일을 복원할 수 있는 방안이 필요하다.

• 슬랙(Slack) 영역 분석 및 은닉 데이터 탐지

일반적으로 사용하지 않는 공간인 슬랙(Slack)에 파일을 숨길 가능성이 있기 때문에 이 영역에 대한 분석이 필요하다. 슬랙 영역은 위치에 따라서 디스크 슬랙, 볼륨 슬랙, 파일 슬랙 등이 존재하며 이 영역을 분석할 수 있어야 한다.

• 암호화 파일 시스템 복호화

특정 파일 시스템의 경우 자체적으로 암호화 기능을 제공하는데, 이 암호화가 적용된 디스크를 다른 시스템에 옮기면 정상적으로 분석할 수 없기 때문에 이러한 암호화 기능을 우회하거나 복호화할 수 있는 기법에 대한 연구가 필요하다. NTFS 파일 시스템에서는 EFS(Encrypted File System) 기능, MAC의 FileVault, ZFS의 Encryption 기능 등이 이러한 역할을 수행한다.

• 시간 정보 의미 분석

파일 시스템에서는 파일 생성시간(Created), 파일 수정시간(Modified), 파일 접근시간(Accessed) 등 다양한 시간 정보가 존재하는데 이 시간 정보에 대한 정확한 이해가 필요하다. 특히 이 시간 정보는 운영체제 별로 상이한 관리 정책을 사용하기 때문에 시간 정보의 정확한 의미 분석을 통하여 파일 복사 시점을 확인하거나 시간 역전 현상 등을 통하여 추가적인 분석의 필요성을 확인할 수 있다.

4. 결론

본 논문에서는 최근 많이 활용되고 있는 파일 시스템에 대한 연구 필요성을 제시하고 있다. 현재 파일 시스템에 대한 연구는 기존에 윈도우 환경에 집중되어 있기 때문에 새로운 파일 시스템을 활용하는 장비를 조사해야 할 경우 대응하기 힘들기 때문에 본 논문에서 제시한 분석 요소를 기준으로 분석 기법을 정립해야 한다.

감사의 글

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(No. 2012-0003832)

참고문헌

[1] Brian Carrier. ″File System Forensic Analysis“ Addition Wesley Professional

[2] Harlan Carvey, “Windows Forensic Analysis”, Syngress

[3] 이상진, “디지털 포렌식 개론”, 이룬