이웃(IoT) : IoT기기 디지털 포렌식 도구

IoT : IoT Device Digital Forensics Tool

요 약

디지털 포렌식은 현대 사회에서 중요한 역할을 수행하고 있는 분야 중 하나로, 디지털 기기 및 시스템에서 증거를 수집하고 분석하여 범죄나 인위적인 활동을 조사하는 핵심 기술이다. 하지만 제조사에 따라 다른 파일시스템으로 인해 많은 어려움이 있다. 본 논문에서는 다양한 파일시스템 내에서 동작이 가능한 효율적인 IoT 디지털 포렌식 방법론을 제시한다. 제안하는 방법은 파일시스템 자동 분석과 딥러닝을 활용한 유해한 파일 추출이다. 이를 통해 기존 포렌식의 복잡성을 해결할 수 있다.

1. 서 론

과거에는 컴퓨터와 스마트폰과 같은 전통적인 디지털 기기가 디지털 포렌식의 주요 대상이었다. 그러나 현재의 사회에서는 IoT 기기가 우리 주변에 널리 보급되어 있다. 스마트 홈 시스템, 스마트 카메라, 스마트 냉장고 및 의료 기기와 같은 IoT 기기는 우리의 일상 생활에 녹아들어 있으며, 기업과 정부 기관에서도 널리 사용되고 있다. 이러한 IoT 기기는 다양한 데이터를 생성하고 저장하며, 이 데이터는 사건 및 사고의 중요한 증거로 사용될 수 있다.

IoT 기기는 기존의 디지털 포렌식 환경[1]과 다른 독특한 특성을 가지고 있어 기존 도구 및 기술로는 충분히 대응하기 어려운 도전 과제를 제기한다. 그 중 하나는 다양한 파일 시스템[2]의 사용이다. IoT 기기는 다양한 플랫폼과 운영 체제로 구성되어 있으며, 일반적인 컴퓨터나 스마트폰과는 다른 특성을 가지고 있다. 또한 메모리가 제한적이기 때문에 디지털 포렌식 도구의 성능 및 용량 제약이 중요한 고려 사항이다.

본 연구의 주요 목표는 다양한 파일 시스템을 가진 IoT 기기에서 동작하는 디지털 포렌식 도구를 개발하는 것이다. IoT기기는 OpenWrt, BusyBox 등과 같은 다양한 플랫폼을 사용하여 동작하는데, 우리는 이러한 기기들에 대한 디지털 포렌식 전문가 및 수사관이 증거 수집 및 분석을 보다 효과적으로 수행할 수 있는 도구를 개발하고자 한다.

IoT 기기는 다양한 센서 및 로그 데이터를 생성하며, 이 데이터는 사건 조사 및 사고 분석에 중요한 역할을 한다. 따라서 이 도구는 파일 시스템, 로그 파일, 데이터베이스 등 다양한 데이터 소스에서 데이터를 추출하고 분석하는데 사용될 것이다.

2. 관련연구

Boztas et al.과 Nemayire et al.의 연구에서는 스마트 TV를 통해 얻을 수 있는 다양한 데이터에 대한 칩오프(Chip-Off) 및 파일 시스템 분석을 수행하여 중요한 정보를 얻어냈다.

Boztas et al.은 MTKII를 사용하여 스마트 TV의 사진, 연결된 장치, 웹사이트 방문 기록과 같은 데이터를 획득하였고 스마트TV의 칩오프 가능성을 보였다.[3]

Nemayire et al.은 칩오프를 수행하여 데이터를 획득할 수 있었다. 파일 시스템 분석을 통해 이미지, 음성, 비디오 파일과 같은 미디어파일, 브라우저 히스토리, 구글 검색기록을 획득하였다.[4]

Jo et al.와 Shin et al.의 연구에서는 AI 스피커를 대상으로 한 디지털 포렌식 분석 방법을 제시하였다.

Jo et al.은 스마트 홈 IoT와 페어링하여 사용되는 AI 스피커와 안드로이드 모바일 앱의 패킷을 분석하였다. 분석을 위해 MitM 기법을 사용하는 Fiddler 도구를 사용하여 패킷 데이터를 획득하였다. 클라우드 서버와의 통신 패킷을 분석하여 사용자이름, 계정 정보, 이메일 등의 사용자 정보와 시간 정보 등을 획득하였다.

Shin et al.은 AI 스피커에 인증서를 주입하여 암호화된 트래픽을 분석할 수 있었고 사용자의 스케줄, 메모, 음성 명령어 등을 획득하였다. 이러한 방법은 AI 스피커뿐만 아니라 칩오프가 가능한 타 스마트 홈 IoT 기기에도 적용할 수 있을 것이다.[5]

3. IoT 기기 포렌식 구조

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학 사업의 연구결과로 수행되었음"
(2017-0-00093)

3.1 라이브 포렌식 (Live Forensic)

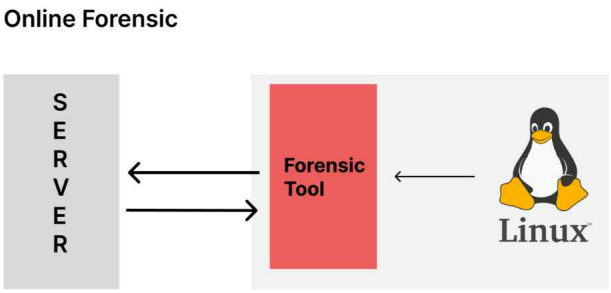

그림 1. 라이브 포렌식 동작 방식

라이브 포렌식은 디지털 포렌식 분야에서 사용되는 기술 중 하나로, 실시간 환경에서 컴퓨터 시스템, 네트워크, 또는 디지털 장치에서 즉시 수집된 데이터를 분석하는 프로세스를 의미한다. 이는 현장에서 사건 조사나 사이버 공격 대응과 같은 긴급한 상황에서 사용된다.

본 연구에서는 실시간으로 파일시스템을 분석하고 파일시스템내에 있는 파일과 디렉토리들을 분석하고, 메타데이터를 수집하여 유해하다고 판단되는 경우 서버로 전달하는 도구를 개발한다.

3.2 오프라인 포렌식 (Offline Forensic)

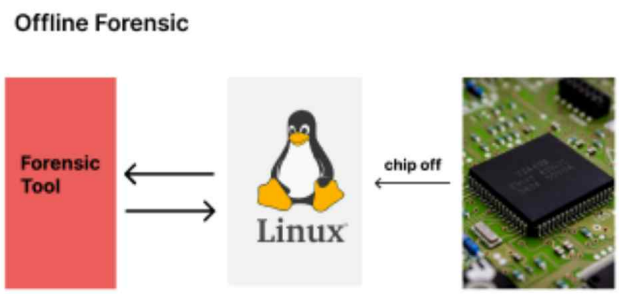

그림 2. 오프라인 포렌식 동작 방식

오프라인 포렌식(Offline Forensics)은 IoT기기에서 칩오프(chip-off)를 수행하여 칩에서 데이터를 수집하고 분석하는 프로세스이다.

본 연구에서는 칩오프를 통해 IoT기기에서 칩오프를 수행한 후에 파일 시스템을 자동으로 분석한다. 그 후 각 파일시스템에 맞는 방식을 통해 파일과 디렉토리들을 분석하고 유해할 것으로 판단되는 파일이나 로그를 추출하는 도구를 개발한다.

4. IoT 기기 포렌식 파일 시스템

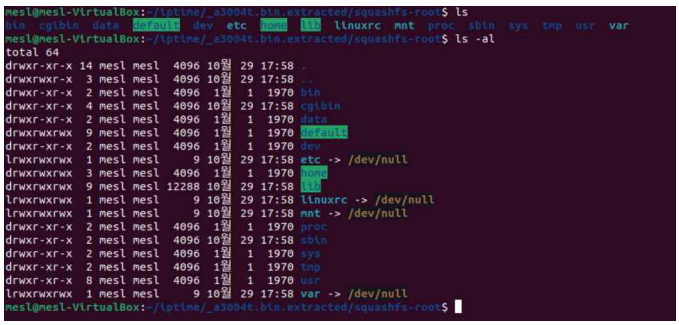

그림 3. iptime A3004T의 파일 시스템 구조

펌웨어 분석 툴인 binwalk를 통해서 펌웨어 파일을 분석하여 파일 시스템 정보와 내부 파일 구조를 확인할 수 있음을 확인하였다. 다음과 같은 방법으로 파일시스템에 대한 정보를 분석한 후 각 파일 시스템에 맞는 분석 방법을 적용한 도구를 개발한다.

파일 시스템별로 파일 및 로그 등을 분석하고 유해한 파일이나 로그 등을 분석한다. 유해한 파일임을 확인하는 것은 IoT에 접속 및 외부프로그램을 실행한 후 남는 로그와 내부에서 동작하였을 때 남는 로그들을 학습시킨 딥러닝 모델을 통해 진행된다. 모델에 관한 자세한 내용은 6절에서 이야기한다.

4.1 ext4 파일 시스템

Ext4 (Fourth Extended Filesystem)은 리눅스 운영 체제의 파일 시스템 중 하나로, Android 운영 체제의 기본 파일 시스템 중 하나이다. 사용되는 IoT기기로는 스마트TV와 스마트워치 등이 있다.

ext4의 전체적인 파일 시스템 정보를 담고 있는 Super block을 먼저 분석 후에 inode의 위치를 확인할 수 있다. inode 테이블에서 실제 파일 위치를 확인하여 Data Block에서 데이터를 얻을 수 있다. 파일이 삭제된 경우, 저널 영역에서 복구를 원하는 파일의 inode를 찾아 파일 카빙을 진행하거나, 기존의 inode에 백업을 진행하는 방식을 통해 삭제된 파일의 복구가 가능하다.

본 연구에서 개발한 도구는 ext4 파일시스템에서 자동으로 Data Block을 통해 inode들을 분석하고 유해한 파일이나 로그를 찾는다. 그 후 라이브 포렌식 모드의 경우 우리의 서버로 전송하며, 오프라인 포렌식 모드의 경우 추출한 파일들을 출력한다.

4.2 플래시 파일 시스템

플래시 파일 시스템(Flash File System)은 플래시 메모리(Flash Memory)를 기반으로 하는 데이터 저장 및 관리 시스

템이다. 예시로는 JFFS(Journaling Flash File System), UBIFS(UBI File System) 등이 있다. IoT기기로는 OpenWrt를 사용하는 공유기 등에서 주로 많이 사용된다.

플래시 메모리 기반 파일 시스템 중 JFFS2에 대하여 분석한 결과, JFFS2는 '0x1985' 라는 매직넘버를 노드 맨 앞에 가지며, 사용하지 않는 영역은 0xFF로 채워져 있다. 여기에서 dirent node의 node type은 '0xE001'이며 node의 버전 정보, 파일명, 수정시간 등에 대한 정보가 저장되어 있고, 해당 노드의 inode number를 검색을 통해 일치하는 inode를 찾을 수 있다.

파일이 삭제된 경우, dirent node에만 존재하고 있으며, 해당 파일의 복구는 inode 번호 및 파일 이름의 검색을 통해 파일이 삭제되기 전의 버전의 inode를 찾아 삭제된 파일을 복구할 수 있다.

본 연구에서 개발한 도구는 플래시 메모리 기반 파일시스템에서 자동으로 dirent node를 통해 inode들을 분석하고 ext4파일시스템과 동일하게 라이브 포렌식과 오프라인 포렌식 모드에 맞게 각각 진행한다.

4.3 읽기 전용 파일 시스템(Read-Only File System)

읽기 전용 파일 시스템(Read-Only File System)은 데이터를 읽을 수만 있고 수정할 수 없는 파일 시스템을 말한다. 읽기 전용 파일 시스템은 데이터의 무결성을 유지하고 변경을 방지한다. 이는 악의적인 변경이나 시스템 오류로부터 데이터를 보호하는 데 도움이 된다. 예시로는 Squashfs, Cromfs등이 있다.

읽기 전용 파일 시스템은 ext4와 파일 구조가 비슷하다. Super block을 통해 inode 테이블에서 실제 파일 위치를 확인하여 Data Block에서 데이터를 얻을 수 있다.

본 연구에서 개발한 도구는 ext4 파일시스템과 동일하게 라이브 포렌식과 오프라인 포렌식 모드에 맞게 각각 진행한다.

5. 아티팩트(artifact) 검출

본 연구에서 파일시스템 내의 유해한 파일과 로그들을 파악하기 위해서 딥러닝 모델을 학습시켰다. 학습시키기 위한 데이터를 수집하기 위해서 IoT기기가 동작할 때 생기는 로그와 직접 접근하여 외부 프로그램을 실행시켰을 때 생기는 로그들을 수집하였다. 파일은 파일명과 파일내용을 특징으로, 로그들은 로그 디렉토리와 로그내용을 특징으로 모델을 학습시켰다. 로그들은 주로 /var/log 경로에 저장되며, 로그들이 저장되는 대표적인 경로는 다음과 같다.

경로	설명
/var/log/messages	리눅스 시스템 전체적인 로그
/var/log/secure	접속한 유저 인증 정보
/var/log/cron	crontab에 등록된 예약 작업 실행 여부
/var/log/lastlog	각 계정의 가장 최근 로그인 기록
/var/log/syslog	syslog가 생성하는 공통 로그

7. 결론 및 기대효과

본 연구는 리눅스 기반의 OpenWrt 또는 BusyBox로 구성된 IoT 기기에 대한 디지털 포렌식 도구를 개발하고자 하는데 주요 목표를 두고 진행된다. 이 도구는 라이브 포렌식을 통해 유해하다고 판단되는 로그 발생시 서버를 통해 관계자에게 전송된다. 서버에 수집된 데이터들과 관계자의 피드백을 통해 딥러닝 모델을 계속해서 학습시킬 수 있다. 또한, 디지털 포렌식 전문가와 수사기관은 이 도구를 활용하여 IoT 기기에서 증거 수집 및 분석을 더욱 효과적으로 수행할 수 있을 것이다. 이를 통해 범죄 조사, 기업 보안 사고 대응, 사회 문제 해결에 대한 능력이 향상될 것이다.

참고문헌

[1] 정익래, 홍도원, 정교일, "디지털 포렌식 기술 및 동향", 전자통신동향분석 제 22권 제 1호

[2] 정규식, 김정길, 곽후근, 장훈, "유무선공유기를 이용한 임베디드 리눅스 시스템 구축 및 응용"

[3] A Boztas*, A.R.J. Riethoven, M. Roeloffs, "Smart TV forensics: Digital traces on television"

[4] Terrence Nemayire, Alex Ogbole, 박성미, 김기철, 정연석, 장윤식,"최신 스마트 TV 데이터 획득 기법 분석: 2018년 삼성 제품의 사례"

[5] Yeonghun Shin, Hyungchan Kim, Sungbum Kim, Dongkyun Yoo, Wooyeon Jo, Taeshik Shon, "Certificate Injection-Based Encrypted Traffic Forensics in AI Speaker Ecosystem"

[6] 이진오, 손태식, "IoT 플랫폼에 탑재되는 안드로이드 및 리눅스 기반 파일시스템 포렌식"