사이버 범죄수사를 위한 디지털 포렌식 표준화 동향

한재혁 _ 고려대학교 정보보호대학원
이상진 _ 고려대학교 정보보호대학원

1. 머리말

디지털 포렌식은 정보저장매체에 저장되어 있는 데이터를 수집하여 이미 발생한 사실 관계를 규명하고 범죄의 단서와 증거를 찾아내는 과학수사 기법을 말한다. 사이버 공간에서 발생하는 범죄의 경우에는 범죄의 단서가 저장되는 영역이 특정 매체에만 국한되지 않고 원격지에 존재하는 경우(예: 클라우드 저장소)가 많아 최근에는 데이터의 수집 범위가 확대되고 있으며 디지털 포렌식의 역할이 갈수록 중요해지고 있다.

디지털 포렌식에서 다루는 대상인 디지털 데이터는 눈으로 직접 내용을 확인하기 어려우므로 이를 다룰 수 있는 도구(HW, SW)를 사용해야 하고, 삭제, 변경 등에 취약하며 복제되기가 쉽다. 이러한 특징들로 인해 법적으로 증거능력을 가지는 디지털 증거를 추출하기 위해서는 전문적인 기술과 논리적인 절차와 방법이 요구되는데 이와 관련한 과정이나 원칙 등을 포괄하는 내용은 단체별 규정[1][2]이나 표준문서로 작성된다.

본고에서는 디지털 포렌식과 관련된 국내외 표준화 동향을 살펴보고, 디지털 포렌식 조사 과정에서 요구되는 효율적인 정보 처리와 통합 관리를 위해 개발된 규격에 대해 소개한다.

2. 디지털 포렌식 표준화 동향

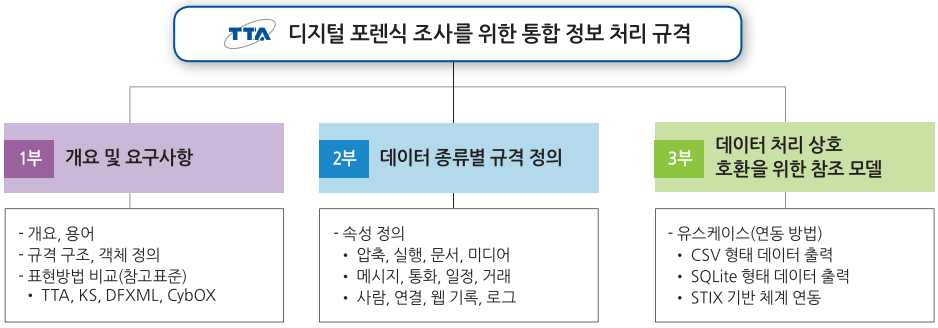
디지털 포렌식과 관련하여 국내 표준화 활동으로 디지털 증거의 증거능력 확보를 위한 절차와 도구의 신뢰성 검증에 대한 표준화가 2007년부터 2009년에 TTA에서 추진되었으며[3][4] 최근에는 개정된 형사소송법과 각종 판례의 요구사항을 반영한 가이드라인[7] 개정과 네트워크 패킷을 처리하는 절차를 규정하는 내용의 지침[8] 제정이 진행되었다.

국제 표준화 활동을 살펴보면, ISO와 IEC에서 추진하는 디지털 증거에 대한 가이드라인[9]과 침해사고(Incident response) 대응 등의 방법이나 절차를 주요 내용으로 하는 정보기술표준이 있다. 이러한 경향은 나라별로 법률이나 제도가 상이하므로 규정으로 통합하는 것이 어렵고 기술 변화의 속도에 맞춰 국제표준을 추진하는 과정이 제한되기 때문에 절차상 원칙을 바로 세우는 데 초점을 맞춘 것으로 보

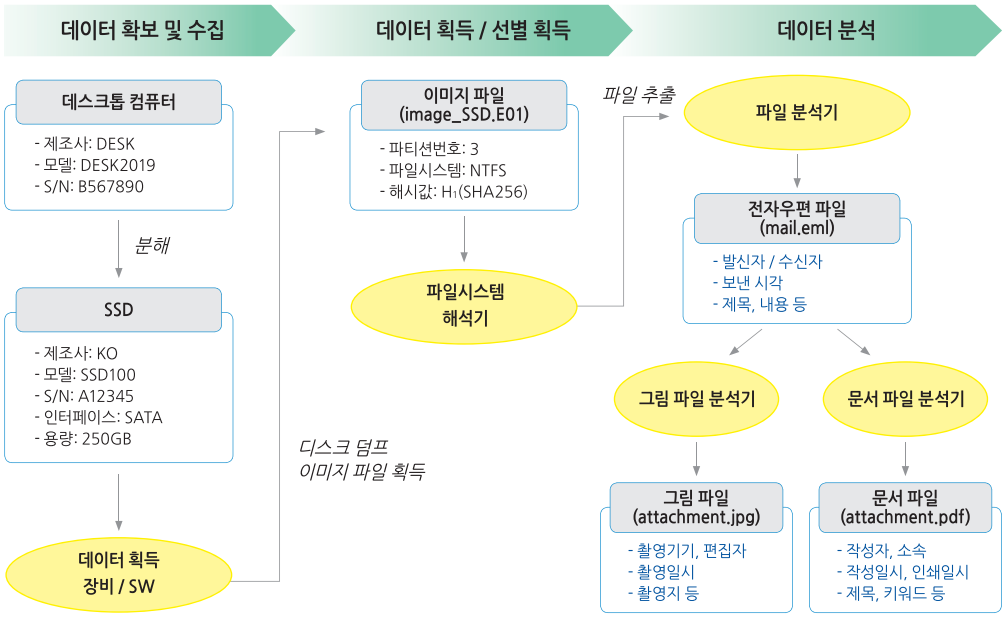
인다. 이에 디지털 포렌식 분야에서는 사실상 표준(de facto standard)의 형태로 참고하는 문서가 다양한데, 미국 국립표준기술연구소(NIST)에서 출판하는 문서(예: SP, Special Publication)나 제시되는 방법론(예: CFTT, Computer Forensics Tool Testing Program)이 대표적이다.

이와 같이 디지털 포렌식 조사 과정에서의 원칙을 다루는 표준 외에도, 디지털 증거에 취해진 조사자의 행위, 보관에 대한 기록 등을 구조적이고 통합적으로 관리할 수 있는 정보를 포함하는 포맷을 기술적으로 정의하는 국내표준[5][6]이 있다. 이 포맷은 디스크 덤프 이미지 파일과 같이 비교적 크기가 큰 파일을 교환하거나 무결성 검증이 반드시 필요할 때 많이 사용되며, 체크섬(checksum)과 해시값, 그리고 관리 연속성(Chain of Custody)을 용이하게 하는 정보(예: 사건 정보)를 저장할 수 있다. 하지만 현재 가장 많이 쓰이는 포맷은 EWF(Expert Witness Compression Format)[10]라고 할 수 있으며, EnCase의 E01, Ex01이나 FTK의 SMART는 모두 EWF를 기반으로 개발되었다. 공통적으로 이러한 포맷은 안전한 데이터 교환을 보장하고 있다는 점에서 앞으로도 꾸준히 사용될 것이나, 파일 단위로 생성되는 포맷이므로 조사 과정에서 불필요한 데이터가 다수 포함될 가능성이 높아 정보를 효과적으로 공유하기 위한 CybOX(Cyber Observable eXpression)가 개발되었다.

CybOX는 현재 STIX 2.0[11]에 통합되었으며 사이버 공간에서 관찰되는 모든 이벤트나 대상을 구조화하여 일관된 분석과 자동화된 해석이 가능하게 하는 정보 표현 규격이다. STIX는 시스템이나 네트워크 상에서 발생된 침해사고나 악성코드 감염을 식별하고 사이버 위협 정보를 공유하는 것을 목적으로 하므로, 표현할 수 있는 대상이 주로 동작 중인 시스템에서만 수집이 가능한 정보(예: 프로세스, 네트워크, 메모리)로 한정적이다. 또한 탐지를 위한 메타데이터를 표현하는 데 필요한 규격만 정의되어 있다. 예를 들어, 전자우편 파일을 살펴보면 STIX에는 송수신자, 메시지ID와 같은 헤더 정보, 첨부된 파일, 본문에 포함된 URL만 포함된다. 만약 조사관이 이 정보를 제공받았다면 전자우편에서 본문의 내용은 확인하지 못하는 상태에서 조사가 진행되는 것이므로 상당한 제약이 따를 것이다. 이는 디지털 포렌식 조사에서 다뤄지는 데이터들이 처리하는 주체나 방법에 따라 표현 방식이 다양하기 때문이며 범죄수사, 감사 등 디지털 포렌식 조사를 목적으로 정보의 교환과 통합적인 관리에 대한 요구사항을 만족시킬 수 있는 규격이 필요해졌다.



[그림 1] 디지털 포렌식 조사를 위한 통합 정보 처리 규격 표준의 시리즈 구성



[그림 2] 디지털 포렌식 조사에서 데이터가 분석되는 과정(예: 전자우편 파일)

3. 디지털 포렌식 조사를 위한 통합 정보 처리 규격 개발

이 표준은 디지털 포렌식 조사에서 사용되는 정보를 표현하기 위해 개발된 규격이며, 사이버 위협 정보 공유 등 사용범위를 포괄적으로 수용할 수 있고 효율적인 정보 교환 및 공유를 지원할 수 있다. 그래서 사이버 보안을 강화시킬 수 있으며, 정형화된 데이터 기반의 디지털 포렌식 분석 도구 개발을 촉진시킬 수 있다. [그림 1]과 같이 총 3부로 구성되어 있으며, 1부는 규격의 구조와 객체를 정의하며 관련한 참고표준들과의 비교를 통해 사용범위를 명확하게 한다. 2부는 데이터 종류별로 규격에서 표현되는 속성(파일, 위치, 문자열, 사람, 일정, 통화 등 20종)을 정의하고 3부는 이 규격이 활용되기 위한 사용법과 유스케이스, 호환성을 고려한 참조모델을 제시한다.

디지털 포렌식 조사 과정을 단계별로 도식화하면 [그림 2]의 예시와 같다. 먼저 확보된 정보저장매체로부터 데이터 획득 장비나 소프트웨어를 이용하여 시스템 동작이나 사용자 행위에 의해 생성된 기록과 관련된 데이터를 획득한다. 획득한 이미지 파일에서 파일 단위로 접근이 가능하도록 파일시스템을 해석하고, 특정 파일의 구조를 분석함으로써 사건의 단서나 사람의 행위, 데이터의 의미를 해석하여 사건의 실체를 파악할 수 있는 정보를 도출한다. [그림 2]에서는 전자우편 파일에 대해 헤더 정보와 본문 내용을 파악한 후 첨부된 파일인 그림 파일과 문서 파일에 대해 추가적으로 정보를 도출하는 예시이다. 그림 파일의 구조 분석을 통해 사진을 촬영한 기기, 편집한 사람, 사진이 촬영된 시각이나 촬영된 장소와 같은 정보를 파악할 수 있고, 문서 파일에서는 작성자, 작성시각 등의 정보를 알 수 있다. 이외에도 디지털 포렌식 조사에서 다루는 파일의 종류는 그 범위가 넓고 다양하지만 정보를 처리하는 방법이 제한적

[그림 3] 통합 정보 처리 규격을 이용한 표현(정보 유출 조사 시나리오)

이므로, 이 표준은 파일 단위로 분석한 결과를 통합된 형태로 관리할 수 있도록 정보를 표현하는 방법을 정의한다.

모든 정보는 객체(object) 단위로 표현되는데, 정보의 속성에 따라 상태나 설정을 표현하는 '기록(trace)'과 시스템 동작이나 사용자 행위를 표현하는 '이벤트(event)'로 구분된다. 이벤트 속성은 육하원칙에 기반을 둔 표현방법을 사용하므로 여러 출처로부터 파악된 정보를 정확하고 명료하며 간결한 형태로 표현할 수 있게 한다. 육하원칙은 '누가(who)', '언제(when)', '어디서(where)', '무엇을(what)', '어떻게(how)', '그 외(etc)'로 구성되며, 조사를 통해 '왜(why)'를 밝히는 과정이므로 '그 외'가 '왜'를 대체한다. [그림 3]은 정보 유출 조사 시나리오에서 조사관이 문자를 주고받은 상황을 파악하는데 필요한 정보를 표현한 결과(제3부, 그림 5-2 참고)이다.

이러한 정보를 처리하는 규격은 먼저 정보를 효율적으로 공유하는 데 활용할 수 있다. 수집되는 데이터의 출처가 다양하고, 현실적으로 조사관들이 사용하는 도구와 분석방법이 가지각색이므로 정보를 표현하는 형태가 상이하다. 주로 데이터 모델에 의거하여 스프레드시트(spreadsheet)나 데이터베이스(database) 형태의 출력을 많이 사용하나, 분석 결과를 통합하여 관리하는 데 발생되는 자원(예: 인력, 시간) 소비를 최소화할 수 있다. 두 번째로 분석 도구의 검증에 활용할 수 있다[12]. 디지털 포렌식 도구는 분석 결과가 서로 상이하므로 신뢰성 검증을 위한 분석결과 비교가 어렵다. 이 규격은 기존 레거시 시스템과의 호환성을 고려하여 개발되었으므로, 여러 도구로부터 출력된 결과를 하나의 표현방식으로 변

환함으로써 도구의 성능을 객관적으로 검증하는 데 활용할 수 있다.

4. 맺음말

IT기술은 꾸준히 발전하고 있으며 디지털 포렌식의 적용범위가 정보저장매체를 포함하여 사이버 공간으로 점차 확대되고 있다. 또한 조사 과정에서 있어서도 증거가 디지털화되어 가고 있는 경향을 보면 디지털 포렌식의 중요성은 점차 커지고 있다.

본고에서는 디지털 포렌식과 관련하여 국내외 표준화 동향을 살펴보았고, 디지털 포렌식 조사 과정에서 효율적인 정보 처리와 통합 관리를 위한 규격을 소개하였다. 이 규격은 정보 공유와 디지털 포렌식 도구 검증을 위해 유용하게 활용될 수 있을 것으로 기대된다.

디지털 포렌식은 특정 국가에 종속되는 기술이나 절차가 아니므로 최근에는 커뮤니티나 포럼을 통한 표준화와 피드백 활동이 활발하다. 이에 유관기관 및 관련 전문가는 TTA에서 추진되고 있는 단체표준에 대한 관심과 더불어 국제 표준화 활동에 적극 참여하는 노력과 후원이 필요할 것이다.

※ 본 연구는 2019년도 과학기술정보통신부의 재원으로 정보통신기획평가원의 지원을 받아 수행됨[No.2018-0-01000, 디지털 포렌식 통합 플랫폼 개발].

[참고문헌]

[1] 경찰청(디지털포렌식센터), '디지털 증거 수집 및 처리 등에 관한 규칙', 경찰청훈령 제845호, 2017. 8. 28. 일부개정
[2] 대검찰청(디지털수사과), '디지털 증거의 수집·분석 및 관리 규정', 대검찰청예규 제991호 2019. 5. 20. 전부개정
[3] TTAS.KO-12.0059, '이동전화포렌식가이드라인', 2007. 12. 26.
[4] TTAK.KO-12.0113, '디지털 증거 조사 모델', 2009. 12. 22.
[5] TTAK.KO-12.0080, '디지털 증거 파일 교환 포맷', 2008. 12. 19.
[6] KS X 1220, '정보기술—보안기술—디지털증거데이터패키지', 2014. 11. 28.
[7] TTAK.KO-12.0058/R1, '디지털 증거 수집 보존 가이드라인', 2017. 12. 13.
[8] TTAK.KO-12.0339, '네트워크 포렌식을 위한 패킷 처리 지침', 2018. 12. 19.
[9] ISO/IEC 27037:2012, 'Information technology—Security techniques—Guidelines for identification, collection, acquisition and preservation of digital evidence', 2012. 10.
[10] Joachim Metz, 'EWF specification', https://github.com/libyal/libewf
[11] TTAE.OT-12.0019, '구조화된 위협 정보 표현 규격(STIX) 버전 2.0', 2018. 12. 19.
[12] 윤우성, 한재혁, 이상진, '효율적인 디지털 포렌식 조사를 위한 육하원칙 중심의 정보 처리 규격', 디지털포렌식연구, 제13권 2호, pp. 125~134, 2019.