디지털산업정보학회 논문지 제15권 제4호-2019년 12월
http://dx.doi.org/10.17662/ksdim.2019.15.4.111

디지털 이미지 증거에서 사건과 무관한 파일 삭제시 무결성 제공 방안 연구

김 태 경*

A Study on the Providing the Integrity of Digital Evidence while Deleting the irrelevant File

Kim TaeKyung

〈Abstract〉

The digital forensic analysis ensures the integrity of confiscated data by calculating hash values for seizure and search of digital evidence and receiving confirmation and signature from participants. However, evidence that is irrelevant to the alleged offense needs to be deleted even after seizure from the point of view of privacy. But the hash value is altered by deleting the irrelevant data from the image file, one will not be able to prove that the file is in the initial state when it was seized.

Therefore, in this paper, a study was conducted to support the integrity of the digital evidence, even if some of the seized digital evidence was deleted or damaged during the seizure search. The hash value of each data is calculated and hash value of the combination of hash values are also calculated. Even if the unrelated evidence is deleted from the seized evidence regardless of file system such as FAT or NTFS, the suggested method presented a way to provide the integrity that proves there is no change in the evidence file.

Key Words : Digital Forensic, Hash, Integrity, Privacy

Ⅰ. 서론

현재는 디지털포렌식 조사를 위해 디지털증거의 압수·수색시 해시값을 계산한 후 참여권자의 확인 및 서명을 받도록 하여 압수한 데이터의 무결성을 보장하도록 하고 있다[1]. 그러나 범죄 혐의와 관련성 없는 증거 즉, 압수한 전자정보의 목록에 없는 전자정보의 삭제·폐기시에 무결성을 훼손할 수 있으므로 이에 관한 연구가 필요한 실정이다.

'전자정보 압수수색 영장에 관한 새로운 실무 운영 지침(2015. 8. 1. 시행)'에 의하면, 전자정보만을 압수 대상으로 하고 컴퓨터나 외장하드 등 저장매체 자체의 압수를 금지하면서 압수한 전자정보의 목록을 상세히 작성하여 피압수자에게 교부하도록 하는 한편, 위 목록에 없는 전자정보는 범죄 혐의와 관련이 없는 것으로써 삭제·폐기하도록 영장에 명시하도록 되어 있다[2]. 그러나 실제 압수 현장에서는 이미징 복제라는 방식을 사용하고 있다. 이미징은 저장매체의 콘텐츠와 무관하게 각 섹터를 읽어 그 내용을 섹

* 명지전문대학 인터넷응용보안공학과 교수

터 단위로 동일 위치에 그대로 복제하는 것을 말하며 원본과 완전하게 동일하게 만드는 작업이다.

디지털포렌식의 5대 원칙에는 정당성의 원칙, 신속성의 원칙, 재현의 원칙[2], 연계 보관성의 원칙, 무결성의 원칙[3]이 있다. 정당성의 원칙은 획득한 증거 자료가 적법한 절차를 준수해야 하며, 위법한 방법으로 수집된 증거는 법적 효력을 상실한다는 것이고, 신속성의 원칙은 시스템의 휘발성 정보수집 여부는 신속한 조치에 의해 결정되므로 모든 과정은 지체 없이 신속하게 진행되어야 한다는 것이다. 재현의 원칙은 피해 직전과 같은 조건에서 현장 검증을 시행한다면, 피해 당시와 동일한 결과가 나와야 한다는 것이며, 연계 보관성의 원칙은 증거물 획득, 이송, 보관, 법정 제출의 각 단계에서 담당자 및 책임자를 명확하게 해야 한다는 것이다. 마지막으로 무결성의 원칙은 수집한 증거가 위·변조되지 않았음을 증명할 수 있어야 한다는 것이다.

무결성의 원칙과 관련하여 압수수색 시 입수한 이미지 파일에서 정보보호 차원에서 사건과 무관한 정보를 삭제하는 등의 조치를 취하게 되면 파일의 해시[4]값이 달라져서 그 파일이 최초에 압수되었던 상태임을 입증할 수 있는 무결성이 깨지게 된다. 따라서 디지털포렌식 조사기관에서는 향후의 무결성이 보장된 증거의 분석 및 사용을 위해서는 최초에 해시값이 부여된 이미징 파일을 조사기관이 보관하는 경우가 있다. 따라서 디지털 수집 증거 중 사건과 관련 없는 일부 데이터를 삭제할 때 무결성을 보장하는 방안에 관한 연구가 필요하다.

본 논문의 2장에서는 관련 연구에 관해서 기술하였으며, 3장에서는 디지털 이미지 증거에서 사건과 무관한 파일 삭제시 무결성 제공 방안을 제시하였으며, 4장에서는 제안한 방법에 대한 수행평가를 수행하였다. 마지막으로 5장에서는 결론에 대해서 기술하였다.

II. 관련연구

2.1 디지털증거 삭제시 이중 해시를 이용한 무결성 보증

압수수색 한 자료에 대해서 관련성 없는 디지털증거를 삭제할 경우 이중 해시를 이용하는 방안에 대한 연구가 수행되었다[5]. 이중 해시를 활용하여 일부 데이터 블록을 삭제한 후 그 해시값만을 남겨둔 뒤에 나머지 데이터들의 해시값과의 해싱 및 전체 해시값과의 비교를 통하여 무결성 입증이 가능하다는 것이다. 여기서 이중 해시란 같은 메시지에 대하여 해시를 수행한 후에 다시 해시를 취하는 방식이 아닌, 전체 데이터를 A, B 혹은 여러 개의 부분집합으로 나눈 이후에 각 부분집합의 해시값을 구하여 원래의 데이터가 아닌 각 해시값끼리의 해시를 통하여 해시값을 구하는 것을 이중 해시라는 개념으로 정의하였다.

압수수색 시 압수한 디지털증거 중 일부가 삭제되거나 손상되더라도 해당 부분의 해시값을 보존하고 있다면, 나머지 부분의 해시값과 이중 해시를 하여 나온 해시값이 원래 데이터의 해시값과 일치하는지 비교함으로써 무결성의 입증을 할 수 있게 된다는 것이다. 즉, 사건과 무관한 정보의 데이터는 삭제하되 그 해시값은 보존하였다가 추후 무결성 입증시 이를 활용하는 방법이다.

그러나 이 연구는 FAT, NTFS[6] 등 어떠한 파일시스템을 사용하느냐에 따라 트리 구조 형성은 어떠한 방법으로 해야 하는지에 관한 추가적인 연구가 필요하며, 실제로 해시값끼리 해싱하여 원 데이터의 해시값과 일치하는지 여부에 대한 연구가 수행되지 않았다. 따라서 본 논문에서는 파일시스템에 관련 없이 압수된 증거에서 관계없는 증거의 삭제시 무결성을 제시할 수 있는 방안을 제시하고자 한다.

2.2 디지털포렌식 분야에서의 해시 함수

해시함수는 사용 목적에 따라 메시지인증, 키유도, 난수생성용과 단순해시(메시지 압축), 전자서명용으로 나눠진다[3]. 메시지 인증용은 메시지의 위변조를 확인하기 위해 해시함수를 이용하며, 키유도 및 난수생성용은 안전한 키와 랜덤한 난수를 위해 해시함수를 사용하는 것이다. 또한, 단순해시 및 전자서명용은 패스워드의 안전한 저장이나 효율적인 전자서명 생성을 위해 메시지 압축 시 해시함수를 이용하는 것이다.

해시를 이용한 방법은 주로 데이터의 위조나 변조에 대한 검사 혹은 데이터베이스에서 데이터에 대한 빠른 검색, 일방향 암호화를 이용한 비밀번호의 안전한 보관 등에 사용된다. 최근에는 전자서명, 불법 저작물의 차단, 암호화폐의 블록체인, 전자투표 등으로 그 활용범위가 점점 확대되고 있다. 수사기관에서는 해시함수를 1990년대 중반 혹은 후반부터 사이버수사와 디지털포렌식의 무결성 제공을 위해 사용하기 시작하였고, 2007년도에 대법원에서 디지털증거의 증거능력 인정요건으로 동일성 혹은 무결성을 제시하면서 증거법 측면의 주목을 받게 되었다. 2011년도에는 형사소송법에서 디지털증거에 대한 출력 및 복제를 압수의 원칙(제106조 제3항)으로 규정하면서 디지털증거의 복제 사례가 증가하여 해시 함수는 같은 자료임을 입증하는 주요한 기술로 사용하게 되었다[7].

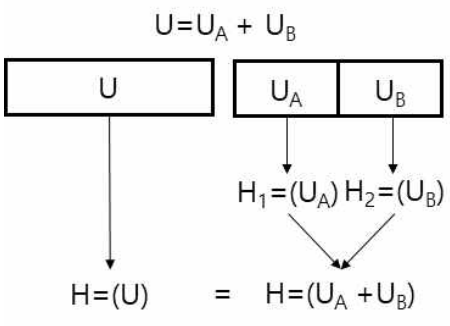
해시 함수는 일방향 암호화 방식이다. 일방향 암호화란 암호화는 가능하지만, 복호화는 불가능한 알고리즘으로 산출된 해시값을 통해 해시 되기 전의 값을 추측하는 것이 불가능하다는 것이다. 또한, 그 크기가 각기 다른 데이터들을 입력값으로 해시 함수에 대입하면 항상 같은 크기의 출력값이 생성되는 특징을 가지고 있다. 이외에도, 데이터 중 아주 작은 부분의 값이 변경되더라도 출력값은 전혀 다른 값이 출력하게 된다.

Ⅲ. 디지털 이미지 증거에서 사건과 무관한 파일 삭제시 무결성 제공 방안

3장에서는 디지털 이미지 증거에서 사건과 무관한 파일 삭제 시 무결성을 제공하는 방안에 대해 제시하였다.

3.1 기존 제안 모델

이전 연구에서 디지털포렌식에서 압수한 증거에서 관련성 없는 데이터의 삭제시 사용한 이중 해시는 다음의 <그림 1>과 같다[5].

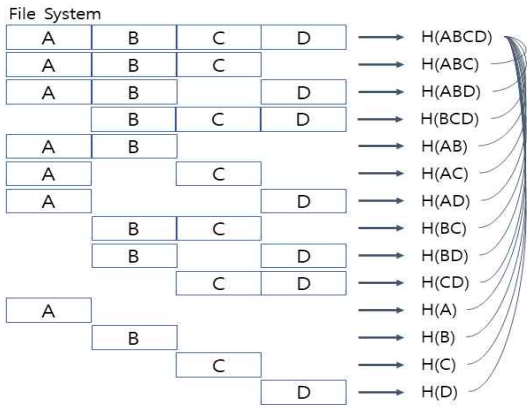


<그림 1> 이중 해시

<그림 1>은 하나의 데이터를 반으로 나눠 각각 해시값을 구하고 그 해시값들에 대해서 해시값을 구하면 원래의 데이터에 대한 해시값과 같다는 방법을 이중 해시로 제안한 것이다. 그러나 이 방법은 하나의 파일을 어떠한 크기의 파일들로 나눌지 그 기준도 불명확하며, 실제 압수수색 후 압수한 파일들에서 사건과 관련 없는 파일을 선택할 때도 파일별로 포함 여부를 결정하게 된다.

3.2 제안 모델

제안하는 모델은 다음의 <그림 2>와 같다.



<그림 2> 제안하는 모델

<그림 2>에서 디스크에 있는 파일시스템에 A, B, C, D 네 개의 파일이 있다고 가정하자. 먼저 전체 이미지 파일에 대해 해시값(H(ABCD))을 구한다. 그 후에는 파일을 하나씩 빼고 해시값(H(ABC), H(ABD), H(BCD)) 을 구하게 된다. 이후에는 계속 빼가면서 해시값(H(AB), H(AC), H(AD), H(BC), H(BD), H(CD), H(A), H(B), H(C), H(D))을 구하게 된다. 해시값을 구한 이후에는 전체 이미지 파일에 대한 해시값과 나머지 해시값들을 합하여 각각의 해시값을 구하면 된다.

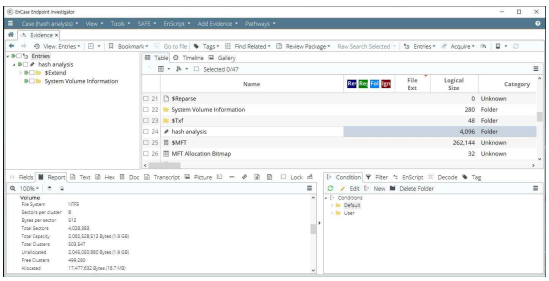
해시값은 파일의 크기와 상관없이 일정한 크기로 생성되므로 다양한 조합으로 인해 생성되는 해시값들이 차지하는 영역은 큰 저장공간이 필요하지는 않게 된다. 또한, 사건과 관련 없는 어느 파일을 삭제하더라도 무결성을 입증할 수 있게 된다. 즉 모든 상황에 해당하는 해시값들을 미리 계산하여 생성함으로 개인정보를 보호하며, 효과적으로 디지털포렌식 업무를 수행할 수 있다.

IV. 성능평가

4장에서는 제안한 모델에 대한 성능평가를 수행하였다. 해시값을 계산하기 위해서 사용한 프로그램으로는 Encase 8.08버전[8]과 Hashcalc 프로그램을 이용하였다. USB(1.9G) 메모리를 NTFS로 포맷하고 4개의 파일(IOT1.pdf, IOT2.pdf, IOT3.pdf, IOT4.pdf)을 생성하고 Encase 프로그램을 이용하여 이미지 파일을 생성하였으며, 생성한 이미지 파일은 다음의 <그림 3>과 같다.

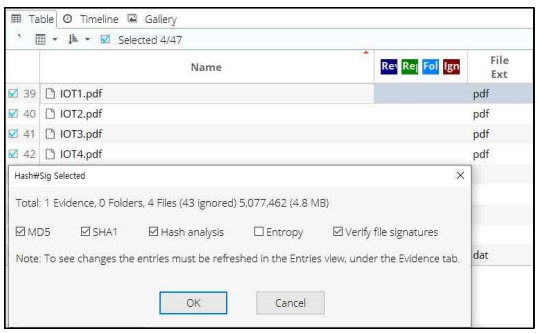
이미지 파일에 대한 정보는 다음의 <그림 4>와 같다. Sector per cluster는 8로 Bytes per sector는 512Byte로 설정하였다.

이미징한 USB 메모리에서 4개의 파일에 대한 해시값을 생성하는 과정은 다음의 <그림 5>와 같다.



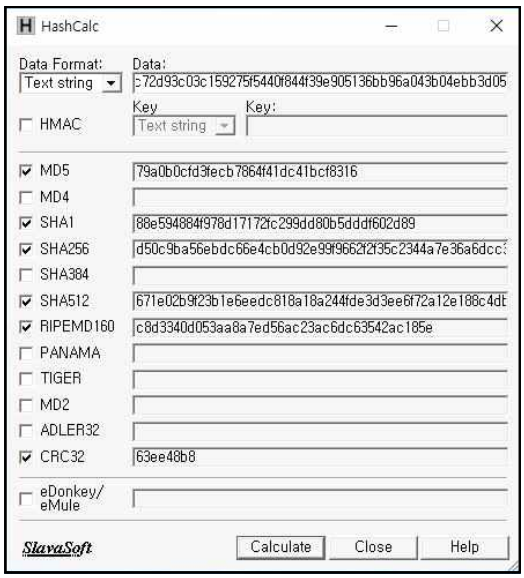
<그림 3> Encase 이미지 파일

<그림 4> Encase 이미지 파일 Report 정보



<그림 5> 파일에 대한 해시값 생성

또한, 3장에서 제안한 모델에 대한 다양한 해시값들에 대해 조합을 통한 해시값을 계산하기 위해 다음의 <그림 6>과 같이 Hashcalc 프로그램을 이용하였다.



<그림 6> 다양한 해시값들의 조합을 통한 해시값 생성

Encase와 Hashcalc 프로그램을 이용하여 생성한 해시값은 다음의 <표 1>과 같다.

<표 1> 무결성 보장을 위한 해시값 계산

항목	해시값
USB	c6da17afe369526150b52fd4e5a3a5a6
IOT1 파일	32e5afe63335fe351577a3f10a4fc976
IOT2 파일	d1154e637881652d3dc50a31e668be2b
IOT3 파일	5b0b86540fac72d93c03c159275f5440
IOT4 파일	f844f39e905136bb96a043b04ebb3d05

<표 1>에서는 USB 전체에 대한 해시값 및 USB 메모리 안에 있는 IOT1.pdf, IOT2.pdf, IOT3.pdf, IOT4.pdf의 4개 파일에 대한 해시값을 생성하였다. 다음의 <표 2>는 다양한 해시값의 조합을 통한 해시값의 생성을 계산한 것이다.

<표 2> 다양한 해시값 조합의 해시값 계산

항목	해시값
USB and IOT1,2,3,4	79a0b0cfd3fecb7864f41dc41bcf8316
USB and IOT1	79c53aea671ea8ddb94b2bb57ad6bbeb
USB and IOT2	ddf8129d830d1e83c1e2ddbc6cf4d42e
USB and IOT3	63e7b06601eed5b7e10fa59feabf51f8
USB and IOT4	184bf0dc53ba719b0d83c69ea9e4946d
USB and IOT1,2	b09ff136e15289fc872689c2b5d9f94d
USB and IOT1,3	79e1abb9e16c6ef821cd52bc79515cce
USB and IOT1,4	6d0482cddde11cea5dd086bfcf999f96
USB and IOT2,3	efb7c929b536db859a601a4612900900
USB and IOT2,4	9e48505eb84ded5c242d8e1fb8162a3d
USB and IOT3,4	4a167591d068a0e3e4b4ec9b1b47daaa
USB and IOT1,2,3	a4cb977f739eff7b19d35f8bb4d9ac86
USB and IOT1,2,4	866c4ea30d7315f7da71545631d49756
USB and IOT1,3,4	f21c81e9d324ac4bef3407a76eac3618
USB and IOT2,3,4	cb73b8ef75f57af75a92359fde63fb25

<표 2>에서 제시한 것처럼 다양한 해시값들의 조합을 이용한 해시값들을 계산하고, 압수한 증거파일에서 사건과 무관한 파일들은 삭제하고, 삭제한 파일들의 해시값들을 보관한다면, 무결성의 원칙과 관련하여 디지털 수집 증거 중 사건과 관련 없는 일부 데이터를 삭제할 때에도 무결성을 제공할 수 있다. 또한, 전체 이미지 파일에 대한 해시값 및 각 파일에 대한 해시값을 이용하므로 파일시스템 종류에 상관없이 적용이 가능하다.

Ⅴ. 결론

디지털포렌식 조사에서는 디지털증거의 압수·수색시 해시값을 계산한 후 참여권자의 확인 및 서명을 받도록 하여 압수한 데이터의 무결성을 보장하도록 하고 있다. 그러나 범죄 혐의와 관련성 없는 증거는

개인정보보호 차원에서 압수한 이후에도 삭제하는 것이 필요하다. 따라서 본 논문에서는 압수수색 시 압수한 디지털증거 중 일부가 삭제되거나 손상되더라도 해당 부분의 해시값을 보존하여 추후 사건 분석 및 관련 기관 제출 시에도 무결성을 입증하는 방안에 대한 연구를 수행하였다.

본 논문에서는 FAT, NTFS 등 파일시스템에 관련 없이 압수된 증거에서 관계없는 증거의 삭제시 각 데이터의 해시값을 계산하고, 해시값들의 조합을 통해 새로운 해시값들을 생성함으로 압수한 데이터와 현재 데이터에 변화가 없다는 것을 증명하는 무결성을 제공할 수 있는 방안을 제시하였다.

참고문헌

[1] 이완규, “디지털 증거 압수 절차상 피압수자 참여 방식과 관련성 범위 밖의 별건 증거 압수 방법,” 형사법의 신동향, 제48호, 2015.

[2] 정교일, “디지털 증거의 압수와 공판정에서의 제출방안”, 형사법 신동향, 2010.

[3] 이인수, “디지털증거 확보체계,” 정보보호학회지, 26권 5호, 2016.

[4] 김기범, “해시함수의 형사법적 고찰,” 한국형사정책연구원 형사정책연구, 29권 2호, 2018년.

[5] 이상미, “관련성 없는 디지털증거 삭제시 이중해쉬를 이용한 무결성 입증 방안,” 서울대학교 융합과학기술대학원 학위논문(석사), 2016.

[6] 박송이 · 허지민 · 이상진, “파티션 복구 도구 검증용 데이터 세트 개발 및 도구 평가,” 정보보호학회논문지, 27권 6호, 2017.

[7] 암호 알고리즘 및 키 길이 이용 안내서, KISA-GD-2018, 한국인터넷진흥원, 2018.

[8] 이보만 · 박대우, “기업회계장부 압수수색과 DB파일 포렌식 기술 적용방법 연구,” 한국컴퓨터정보학회 하계학술대회 논문집, 19권 2호, 2011.

■ 저자소개 ■

김 태 경
Tae Kyung Kim

2017년 9월~현재
명지전문대학
인터넷응용보안공학과 수
2008년 3월~2017년 8월
서울신학대학교 교수
2006년 3월~2008년 2월
서일대학 정보전자과 교수
2005년 8월 성균관대학교 전기전자 및
컴퓨터공학과(공학박사)

관심분야 : 네트워크보안, IoT 보안,
개인정보보호, 디지털포렌식
E-mail : tkkim@mjc.ac.kr

논문접수일 : 2019년 12월 3일
게재확정일 : 2019년 12월 12일