논문 2012-49CI-4-5

디지털 증거 선별 조사의 효율성을 위한 Digital Evidence Container 설계 및 구현

(A New Design and Implementation of Digital Evidence Container for Triage and Effective Investigation)

임 경 수*,***, 이 창 훈**, 이 상 진*

(Kyung-Soo Lim, Changhoon Lee, and Sangin Lee)

요 약

최근 국내외 수사 기관은 초동 수사 시 현장에 컴퓨터가 있을 경우, 사이버 범죄 수사가 아닌 경우에도 시스템을 압수 또는 확보하는 것이 필수적인 단계로 자리 잡고 있다. 이렇게 확보한 시스템으로부터 용의자의 범죄 사실 입증을 위한 정황 증거를 확보하여 수사에 활용되고 있다. 하지만 사이버 범죄가 아닌 일반 범죄 사건에서 확보한 시스템에서 디스크 이미지를 확보한 후, 면밀히 조사하는 것은 시간이 많이 소요되며 신속한 사건 대응이 필요한 납치, 살인사건과 같은 범죄 유형에는 더욱 어려움이 따른다. 또한 기업 수사에서도 대용량 데이터베이스나 파일 서버 조사에서 나아가 클라우드 환경에서는 디스크 단위의 복제는 불가능하므로 선별 수집한 디지털 증거를 분석에 이용해야 한다. 하지만 다양한 종류의 디지털 증거를 선별 수집하더라도 이를 저장 및 보관하기 위한 표준화된 데이터 포맷이 존재하지 않아 법정에서 증거력을 증명하기 어려운 것이 현실이다. 따라서 본 논문에서는 선별 수집된 다양한 디지털 증거를 보관하기 위한 새로운 증거 보관 포맷을 제시한다. 본문에서 제시하는 디지털 증거 포맷은 다양한 디지털 증거 자료에 활용할 수 있도록 범용성과 확장성에 중심을 두었으며, 일반적인 XML 기술과 압축 파일 포맷을 이용하여 기존 시스템에 적용하기 쉽도록 설계하여 기존 연구들보다 활용하기 쉬운 장점이 있다.

Abstract

The law enforcement agencies in the worldwide are confiscating or retaining computer systems involved in a crime/civil case, if there are any, at the preliminary investigation stage, even though the case does not involve a cyber-crime. They are collecting digital evidences from the suspects's systems and using them in the essential investigation procedure. It requires much time, though, to collect, duplicate and analyze disk images in general crime cases, especially in cases in which rapid response must be taken such as kidnapping and murder cases. The enterprise forensics, moreover, it is impossible to acquire and duplicate hard disk drives in mass storage server, database server and cloud environments. Therefore, it is efficient and effective to selectively collect only traces of the behavior of the user activities on operating systems or particular files in focus of triage investigation. On the other hand, if we acquire essential digital evidences from target computer, it is not forensically sound to collect just files. We need to use standard digital evidence container from various sources to prove integrity and probative of evidence. In this article, we describe a new digital evidence container, we called Xebeg, which is easily able to preserve collected digital evidences selectively for using general technology such as XML and PKZIP compression technology, which is satisfied with generality, integrity, unification, scalability and security.

Keywords : Digital Forensics. Digital Evidence Container, Evidence Management, Triage Investigation

* 정회원, 한국전자통신연구원 사이버융합보안연구단
(Cyber Security-Convergence Research Laboratory, Electronics and Telecommunication Research Institute)
** 정회원-교신저자, 서울과학기술대학교 컴퓨터공학과
(Dept. of Computer Engineering, Seoul National University of Science and Technology)
*** 정회원, 고려대학교 정보보호대학원 (Graduate School of Information Security, Korea University)
※ 본 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행되었습니다. (No. 2012-0003832)
접수일자: 2012년6월18일, 수정완료일: 2012년7월2일

Ⅰ. 서 론

기존의 포렌식 패러다임은 사건 현장에서 확보한 컴퓨터에서 하드 디스크 드라이브를 획득하고 이를 디스크 복제나 디스크 이미징 기술을 이용해 사본을 생성한 뒤, 이를 이용하여 조사를 진행하였다. 디스크 이미징을 통해 생성된 이미지 파일은 원본 하드 디스크와 동일한 증거로 입증하기 위해 암호학적 해쉬 함수를 이용해 무결성을 증명하게 된다.

한편 최근 컴퓨터 포렌식 패러다임은 기존의 디스크 이미지 기반의 조사 방법론에서 사건 조사에 필요한 증거 데이터 중심으로 변화하고 있다. 이는 사이버 범죄 수사에 국한되어 있던 디지털 포렌식 기술이 일반 민/형사 사건에도 보편화되고 있기 때문이며, 또한 일반 사용자 컴퓨터의 하드 디스크 용량이 1 TB 이상일 정도로 급증함에 따라 조사 시간을 줄일 수 있는 효율적인 조사 방법론이 필요하기 때문이다. 이에 따라 현장에서 필요한 증거를 선별 수집 및 분석하는 접근 방법이 대두되고 있다.[1, 4~5, 9~10]

이처럼 확보한 시스템에서 디스크 이미지를 획득한 후 면밀히 조사하는 기존 방법론은 시간이 많이 소요되는 단점이 있으므로, 신속한 사건 대응이 필요한 사건의 경우는 사건과 관련성이 있는 디지털 증거 선별 조사 기법이 필요하다. 따라서 사용자 행위를 조사할 수 있는 윈도우 사용 흔적 정보나 문서 파일만을 선별하여 수집하는 것이 효과적이다. 특히 국내 디지털 포렌식 워크샵 및 컨퍼런스에 참가한 현장 수사관들은 현장의 수많은 시스템에 대한 조사를 하기에는 인력이 부족한 경우가 대부분이며, 현장의 모든 컴퓨터를 압수 수색하는 것은 불가능한 경우가 많다고 언급되고 있다. 이러한 환경에서는 조사가 필요한 증거 자료만을 선별하여 수집할 수 있는 다양한 방안이 필요하다. 또한 클라우드 서비스, 스마트워크와 같이 대용량 스토리지에 데이터가 저장되는 미래 IT 환경에서는 디스크 이미징이 불가능하므로, 활성 상태의 서버나 원격지의 클라이언트 시스템에서 필요한 디지털 증거 자료만을 선별 수집하는 방법이 필요하다.

기존 국외 현장 대응 포렌식 솔루션은 선별 수집은 가능하나 증거 보관 데이터 포맷이 공개되지 않아 개인 컴퓨터 조사 외에는 적용하기 어렵다. 예를 들어 기업 사건에서의 데이터베이스 레코드 선별 수집, 네트워크 보안 장비의 로그 파일, 모바일 장비에서 디지털 증거, 내비게이션과 같은 임베디드 장비와 같이 사용자 관련 증거 자료 선별 수집 등에 적용할 수 없다. 따라서 이러한 다양한 디지털 장비에서 수집한 증거 데이터를 선별 수집하여 보관 및 분석이 가능한 표준화된 형태의 디지털 증거 보관 포맷이 필요하다.

현재는 이렇게 다양한 디지털 기기에서 선별 수집한 디지털 증거를 보관하기 위한 표준화된 데이터 포맷이 존재하지 않아 이를 법정에서 표현하기 어려운 것이 현실이다. 디지털 증거 보관 포맷과 관련한 국제 학회의 연구는 디스크 이미지에 대한 개선 연구가 대부분으로, 국내 환경에 적합하고 디지털 증거 선별 조사 기법에 적합한 증거 보관 포맷이 필요하다. 따라서 본 논문에서는 선별 수집된 디지털 증거를 보관하기 위한 새로운 데이터 포맷을 제시한다.

Ⅱ. 관련 연구

1. Digital Evidence Container

디지털 증거 포맷에 대한 연구는 디스크 원본과 동일한 크기와 형태를 지니도록 비트스트림 복제를 통해 생성된 DD 이미지에서 출발하였으며, 현재는 여러 디스크 분석 솔루션에서 사용되는 각각의 디스크 이미지 파일 형태들로 확대 되었다. 대표적인 예로 EnCase 이미지 파일이 있다. EnCase 이미지 파일은 전체 디스크를 특정 용량으로 분할하여 압축하고, CRC나 해쉬 함수를 이용해 무결성을 증명하도록 설계되었다.

학계에서는 EnCase 이미지 파일 포맷처럼 원본 디스크 이미지에 대하여 부가적인 기능이나 압축 기능이 추가된 디스크 이미지 포맷을 디지털 증거 보관 용기(Digital Evidence Container, 이하 DEC 지칭) 라고 정의하였다.[9]

이는 곧 디지털 증거 보관 용기를 디지털 기반의 증거 보관 포맷이라는 개념으로 변화하고 있다. 즉 그림 1과 같이 현재 물리적 증거를 보관하기 위한 증거물 봉투나 용기를, 디지털 증거에도 적용하여 사건 관련 정보, 시간 정보 등의 메타 정보나 무결성 증명을 위한 해쉬 값 등을 기록한 데이터를 수집한 증거 자료를 함께 보관할 수 있는 포맷으로 발전하고 있다. 이러한 개념을 처음 제시한 것이 2005년 DFRWS 에 발표된 Philip Tuner 의 논문으로 “Digital Evidence Bag” 이라는 개

그림 1. 물리 증거에 대한 보관 포맷 (Evidence Bag)
Fig. 1. Various evidence bag samples of physical evidence.

념을 제시하였다.[6~8]

"Digital Evidence Bag"은 디지털 증거를 저장할 포괄적인 개념의 보관 포맷으로, 다수의 디지털 증거를 포함하는 계층적 구조의 증거 객체 집합이라고 볼 수 있다. 이러한 설계 사상은 컴퓨터, 모바일, 네트워크 등의 다양한 기기에서 수집한 디지털 증거를 보관하기 위한 표준화된 포맷을 개발하는데 목적이 있다. 이러한 표준화된 포맷은 단일화된 포맷에 저장하여 일관성 있게 처리할 수 있어 데이터 처리 과정을 간소화할 수 있다. 또한 분산 환경이나 병렬 처리에 효과적으로 이용 가능하다.

또한 수집한 증거 데이터에 대한 감사 정보를 비롯한, 조사 과정에서 발생한 의미 있는 메타 데이터가 포함되도록 설계하였다. 모든 메타 정보는 평문으로 저장되며, 모든 데이터는 raw 바이너리 포맷으로 기록하여 범용성을 높이는데 초점을 맞추었다.

단순히 이미지 포맷의 개념이던 DEC 는 이와 같은 연구 발표에 의해 단순히 디스크 이미지의 용량을 줄이거나 무결성을 보장하기 위한 장치에서 다양한 기능을 가진 포맷으로 발전하고 있다.

2. 현장 상황을 고려한 증거 선별 수집

현장에서 신속한 사건 대응을 목적으로 개발된 방법론으로, The Cyber Forensic Field Triage Process Model (CFFTPM)이 있다.[5] CFFTPM 은 납치 사건, 인질 사건, 살인 사건과 같이 신속한 시간을 요구하는 경우에 최소한의 시간으로 조사가 가능하도록 개발된 모델이다. 현장에서 사건 조사에 필요한 파일을 선별하여 수집하고 조사하는데 초점을 맞추고 있으며, 디스크 이미지 기반이 아닌 파일 기반 조사를 수행한다. 결과적으로 데이터의 중요성 및 우선 순위 중심으로 필요한 증거 자료를 선별 수집하여 현장에서 이를 조사하는 프로세스 모델이다.[5]

그림 2의 CFFTPM 모델의 첫 단계는 발생한 사건에 대하여 조사 계획 (Planning) 을 수립한다. 다음은 사건의 유형에서 필요한 증거자료를 우선 순위 별로 선별하여 수집하는 증거 선별 (Triage) 단계이다. 사용자 프로파일 분석(User Usage Profiles) 은 Microsoft Windows 운영체제의 "내 문서", "바탕화면" 디렉토리와 같이 주로 사용자의 데이터가 저장되는 디렉토리를 조사하며, 파일의 생성, 수정, 접근 시간 정보나 파일 종류와 같이 파일 속성 정보도 조사 과정에 포함된다. 또한 레지스트리에 저장되어 있는 시스템, 네트워크, 사용자 흔적 조사를 통해 사용자의 시스템 흔적을 파악하게 된다.

타임라인 분석 (Chronology Timeline)은 앞서 수집한 정보들을 특정 시간 순서로 정렬하여 시간대 별로 사용자의 행위와 패턴을 분석하며, 인터넷 사용 기록 분석 (Internet)은 웹브라우저 사용 기록이나 전자 메일, 메신저와 같이 인터넷을 사용하는 개인 정보들을 수집하여 분석한 뒤, 마지막으로 분석한 정보에 대한 사건 명세 (Case Specific) 과정을 거치는 단계로 진행된다.

CFFTPM 방법론을 실제 포렌식 도구에 적용한 것이 GuidanceSoft 사의 EnCase Portalbe Kit이다.[10] EnCase Portable Kit은 EnCase Portable 소프트웨어가 저장되어 있는 4GB USB 드라이브 외에 추가 저장을 위한 16GB USB 드라이브, USB 허브, 동글 키 등을 하나의 솔루션으로 출시한 제품이다.

EnCase Portable의 주요 기능은 대상 컴퓨터에서 수사에 필요한 데이터들을 사용자가 선별하여 자동으로 수집하는 것이다. EnCase Portable Kit의 사용 방법은 현장의 수사 대상 시스템을 EnCase Portable 소프트웨어을 이용하여 부팅한 뒤, 수사관이 조사에 필요한 데이터 종류를 선택하면 이를 자동으로 수집하여 설치한 USB 저장장치에 저장한다. 수집 대상 데이터로는 문서 파일, 인터넷 사용 기록, 이미지 파일 등이 있으며, 전체 디스크에 대한 이미징 기능을 지원한다. 수집한 데이터는 EnCase 제품 군에 사용되고 있는 파일 포맷 (EnCase Logical Evidence Format) 을 이용하여 데이터를 저장한다.이처럼 EnCase Portable을 계기로 디지털 증거의 선별 수집 기법이 현실화되고 있다. 하지만 EnCase Portable에서 사용하는 DEC 포맷은 EnCase

제품군에서만 사용할 수 있어 다른 포렌식 도구에 적용하기 어렵다.

3. XML과 디지털 증거 포맷

XML을 활용한 디지털 포렌식 기술은 다양한 분야에서 현재 활용되고 있다. 디지털 증거에 대한 XML의 활용은 본 논문의 초점과 유사한 DEC, 보안 사고의 사전 대응을 위한 정보 교환 포맷, 디지털 증거 시스템에서의 데이터 처리 포맷 등이 있다.

먼저 DEC 분야는 최근 디스크 이미지에 대한 공개 표준을 위해 연구 중인 AFF(Advanced File Format) 포맷에서 활발하게 이용되고 있다. AFF는 기존 디스크 이미지 (DD 이미지)에 대해 별도의 후처리 과정을 거쳐 새로운 AFF 이미지 포맷을 생성하는데, 이는 조사 과정의 시간적 효율성과 신속성을 개선하기 위함이다.

AFF는 기존 디스크 이미지에서 AFF 포맷으로 변환하기 위해 fiwalk 라는 도구를 사용한다. fiwalk 는 입력받은 디스크 이미지에 대하여 파일시스템을 인식하여 자동으로 메타데이터를 생성하며, 출력 결과는 XML 문서로 출력된다. 여기서 생성된 XML 문서는 결과적으로 디스크 이미지를 인덱싱하여 특정 파일을 찾고 조사하는 과정에 이용되며 조사 시간의 효율성을 개선할 수 있다.

그림 3은 fiwalk 엔진으로 생성된 XML 문서의 예이

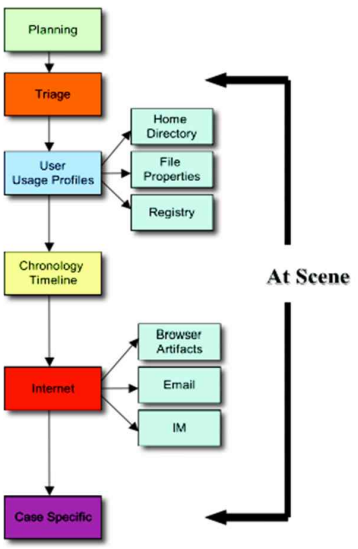

그림 2. CFFTPM 모델 절차
Fig. 2. The Cyber Forensic Field Triage Process Model.

```
<?xml version='1.0' encoding='ISO-8859-1'?>
<fiwalk>
 <metadata>
    <dc:type>Disk Image</dc:type>
 </metadata>
 <Imagefile>/corp/ubnist1.gen0.raw</Imagefile>
 <fiwalk_version>0.5.1</fiwalk_version>
 <Start_time>Sun Mar 8 22:13:10 2009</Start_time>
 <tsk_version>3.0.0</tsk_version>
 <aff_version>3.3.4</aff_version>
 <volume offset='32256'>
  <Partition_Offset>32256</Partition_Offset>
  <block_size>512</block_size>
  <ftype>8</ftype>
  <ftype_str>fat32</ftype_str>
  <block_count>4114340</block_count>
  <first_block>0</first_block>
  <last_block>4114339</last_block>
  <fileobject>
   <id>1</id>
   <filesize>14607</filesize>
   <partition>1</partition>
   <flags>5</flags>
   <ALLOC>1</ALLOC>
   <USED>1</USED>
   <inode>4</inode>
   <type>1</type>
   <mode>73</mode>
   <nlink>1</nlink>
   <uid>0</uid>
   <gid>0</gid>
   <mtime>1230525210</mtime>
   <atime>1230451200</atime>
   <crtime>1230525210</crtime>
   <filename>ldlinux.sys</filename>
   <byte_runs>
     <run file_offset='0'
     fs_offset='4127744'
     img_offset='4160000' len='14607'/>
   </byte_runs>
   <md5>a40ba2f7239bdae2193dfd1089856f38</md5>
  </fileobject>
```

그림 3. AFF 포맷에서 fiwalk XML 문서
Fig. 3. An example of fiwalk XML document in AFF.

```
<EventData>
  <DetectTime>2006-05-31T07:28:03-04:00</DetectTime>
  <AdditionalData dtype="xml">
    <phish:PhraudReport   FraudType="phishemail">
      <phish:FraudParameter>Urgent Sercurity Warning
      </phish:FraudParameter>
      <phish:FraudedBrandName>Bank of AMerica</phish:FraudedBrandName>
      <phish:LureSource>
        <System category="source">
          <Node>
            <Address>216.118.97.160</Address>
          </Node>
        </System>
      </phish:LureSource>
```

그림 4. IODEF 문서 예
Fig. 4. An example of IODEF

다. 출력된 XML 문서를 살펴보면 파일시스템, 볼륨 정보, 파티션 정보과 같은 조사 대상 이미지 파일에 대한 기본적인 내용을 파악할 수 있으며, 특정 파일에 대해서는 접근 권한, 시간 정보, 파일 크기, 해쉬 정보 등이 생성되는 것을 확인할 수 있다.

보안 사고의 사전 대응을 위한 정보 교환 포맷으로 이용되는 경우는 IETF에서 표준으로 진행 중인 IODEF (Incident Object Description Exchage Format) 포맷이

있다. IODEF 포맷은 현재 피싱 사이트 발견 시 해당되는 정보를 CERT 팀이나, DNS 서버 등에 전파하여 악성 서버를 빠른 시간 내에 차단하기 위한 수단으로 이용되고 있다. 그림 4는 IODEF 포맷의 한 예로, “Bank of America” 사이트를 위장하고 피싱 사이트에 대한 URL 과 IP 주소를 나타내고 있다.

마지막으로 디지털 증거 처리 시스템의 데이터 처리를 위한 포맷으로 XFRAME에서 사용하는 XML 문서 포맷이 있다. XFRAME 은 활성 데이터와 같이 정형화된 디지털 증거 포맷이 없는 데이터를 XML 문서 형태로 저장하여 관리하는 데서 출발하였다.

활성 시스템 조사의 경우 물리 메모리에 상주 하고 있는 프로세스 정보나 네트워크 연결 정보 등은 운영체제에서 제공하는 커맨드라인 명령어를 사용하여 콘솔창에서 데이터를 확인하거나, 또는 공개 버전의 포렌식 툴을 이용하여 이러한 콘솔 명령의 결과를 그대로 텍스트 형태로 수집한다. 먼저 콘솔 명령어를 사용할 경우는 별도로 데이터를 수집할 방법이 없어 직접 문서에 기록하거나 스크린 샷을 이용하였다. 한편 공개된 포렌식 도구를 사용하더라도 출력 포맷이 각각 달라 이를 상관 분석에 활용하기에 번거로운 점이 있었다.

이러한 문제점들을 해결하고자 개발된 것이 XFRAME 프레임워크이다. XFRAME은 사건 유형에 맞게 조사에 필요한 데이터를 선별하여 하나의 수사 프로파일로 생성하고, 이를 통해 자동으로 데이터를 수집하여 XML 형태로 데이터를 저장한다.

```
<?xml version="1.0"?>
<MessengerProfiles xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:noNamespaceSchemaLocation="C:\MessengerProfiles.xsd" investigator="임경수"
 beginningTime="2007년 09월 27일, 23시 40분 19초" completedTime="2007년 09월 27일,
23시 50분 50초">
  <Messenger appName="MSN_Messenger" defaultPath="C:\\Program Files\\MSN
Messenger\\">
   <Account_Data isRegData="true" dataType="REG_SZ" hKey="HKEY_CURRENT_USER"
   regPath="Software\\Microsoft\\MSNMessenger\\PerPassportSettings" key="
MessageLogPath" ID="">
    <Account regSubKey="1072943641" regKey="MessageLogPath">lloydlim80</Account>
   </Account_Data>
   <Password_Data isRegData="true">Test_Password</Password_Data>
   <ChatLog_Data regSubKey="1072943641" regKey="MessageLogPath" filePath=
"C:\\Documents and Settings\\임경수\\My Documents\\받은 파일\\lloydlim1072943641">
     <ChatLog isAcquired="true">
       <FileName>lovenorang1429267126.xml</FileName>
     </ChatLog>
     <ChatLog isAcquired="true">
       <FileName>marymoore20062893569322.xml</FileName>
     </ChatLog>
   </ChatLog_Data>
   <DownFile_Data isRegData="false" path="">
   C:\Documents and Settings\임경수\My Documents\받은 파일\
   </DownFile_Data>
   <Cookie_Data isRegData="false" path="C:\\Documents and Settings\\임경수\\Cookies\\">
   파일 수집 완료 </Cookie_Data>
   <Extra_Data name="Default_User" isRegData="true" dataType="REG_BINARY" hKey="
HKEY_CURRENT_USER"
    regPath="Software\\Microsoft\\MSNMessenger\\PerPassportSettings" key="
DefaultMemberName" ID="">
    <Extra name="Defuault User">lloydlim80@hotmail.com</Extra>
   </Extra_Data>
  </Messenger>
</MessengerProfiles>
```


그림 5. XFRAME의 XML 문서 예
Fig. 5. An example of XML document in XFRAME.

Ⅲ. 디지털 증거 선별 조사를 위한 DEC 설계 및 구현

논문에서 제안할 DEC를 소개하기에 앞서 새로운 증거 포맷 개발에 필요한 기능들에 대해 먼저 살펴본다. 본 요구사항은 디지털포렌식워크샵을 통해 일선 수사관 및 포렌식 분야 종사자들의 의견을 바탕으로 정의하였다.[1] 새로운 DEC 설계가 만족해야 할 요구사항을 정의하면 다음과 같다.

- 보존성 (Preservation)
  원본 데이터를 비롯한 관련 메타데이터를 수집할 수 있어야 하고 사후 감사를 위한 분석 로그 기록도 보존할 수 있어야 한다.
- 무결성 (Integrity)
  증거 데이터 원본의 무결성을 보장할 수 있는 장치가 있어야 한다.
- 범용성 (Generality)
  다양한 데이터 자원에 대하여 범용적으로 사용될 수 있도록 공개적인 기술에 기반을 두어야 한다.
- 단일성 (Unification)
  다양한 증거 데이터 종류에 대해서도 단일화된 포맷으로 적용할 수 있어야 한다.
- 확장성 (Scalability)
  데이터 량의 증가에도 유연하게 확장할 수 있어야 하며, 제안된 증거 보관 포맷을 적용하더라도 데이터의 전체 크기가 현저히 증가하지 않아야 한다.
- 압축성 (Compressibility)
  원본 데이터의 크기를 줄일 수 있는 기능을 제공해야 한다.
- 보안성 (Security)
  증거 보관 포맷을 보호할 수 있는 보안 매커니즘을 제공해야 한다.

논문에서 제안하고자 하는 새로운 DEC 은 일반 파일, 데이터베이스 레코드, 모바일 장비(임베디드 시스템) 사용 이력, 웹브라우저 사용 기록과 같은 다양한 증거자료에 대해 표준화된 형태의 데이터 포맷을 기반으로 다른 디지털 포렌식 도구와도 상호운용이 가능한 형태를 제안하는 것이 목적이다. 나아가 급변하는 IT 환경 속에서 기존의 저장 장치 복제 기반이 아닌 컨텐

츠 기반의 디지털 증거 자료를 수집하여 보관하고, 이를 분석한 이후에도 수행 로그 기록과 분석 결과를 함께 보관하여 단일화된 형태로 유지 관리될 수 있어야 한다.

새로운 디지털 증거 보관 포맷은 일반적인 디지털 증거의 효력을 보장할 수 있는 보존성, 무결성을 비롯하여 다양한 자원에서도 활용이 가능한 범용성 그리고 여러 디지털 증거 자료를 수집하더라도 동일한 포맷에 보관이 가능한 단일성이 필요로 한다. 또한 데이터 크기나 수집한 증거 자료가 많더라도 확장 가능해야 하며, 압축성을 보장하여 효과적인 저장 성능과 접근제어에 필요한 보안성을 고려해야 한다.

1. 디지털 증거 보관 포맷 XeBag 소개

본 논문에서는 이러한 요구사항을 반영하여 새로운 디지털 증거 보관 포맷인 XeBag (XML, PKZip-based digital evidence Bag) 을 소개한다. XeBag 은 새로운 디지털 증거 보관 포맷에 적합하도록 현장 대응 단계에서의 선별 수집에 용이하게 설계되었으며, 다양한 증거 자료에 적용 가능하도록 표준화된 형태로 설계하였다. XeBag 포맷의 기본적인 파일 포맷은 PKZip이나 WinRAR 과 같은 압축파일 포맷을 기반으로 하여 수집한 증거 자료를 압축하여 저장하며, 파일에 대한 접근제어가 필요한 경우 패스워드를 적용하여 저장한다. 또한 포렌식 관점에서 표현에 필요한 사건 관련 정보나, 메타데이터, 로그 기록과 같은 데이터는 XML 문서를 활용하여 이를 함께 압축 파일에 저장한다. 또한 파일 단위로 저장이 되지 않는 증거 자료, 예를 들어 데이터베이스 레코드, 임베디드 장비의 사용자 기록, 네트워크 보안 장비의 로그 기록 등은 XML 문서 형태로 정의하거나 변환하여 보관하는 것이 가능하다.

XML 과 PKZip 압축 포맷의 활용은 Microsoft Office 2007, 2010 제품군에서 사용하는 OpenXML 에서도 사용되고 있다. PKZip은 압축 기술로 널리 사용되므로, 증거 포맷의 사용으로 인해 메타데이터 추가 시 발생하는 용량의 증가를 압축 기술로 대응할 수 있다. 또한 다양한 디지털 증거를 하나의 포맷에 통합하여 제공할 수 있으므로 단일성을 만족한다.

한편 XML 기술의 적용은 디지털 증거에 대한 기술 내용이 증가함에도 XML 기본 개념인 확장성을 이용하여 이에 대응할 수 있다. 또한 증거 포맷이 향후 개선되

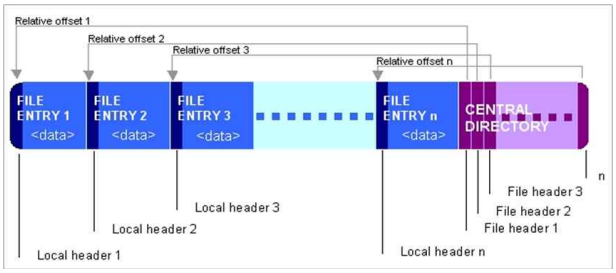

그림 7. PKZip 파일 포맷의 구조
Fig. 7. he structure of PKZip file format.

더라도 기존 포맷에 그대로 확장하여 사용가능하므로 유연성이 높으며, XML 관련 기술에 대한 인프라가 보급되어 있어 다양한 분야에도 활용할 수 있는 범용성을 만족한다. 예를 들어 일반적인 DBMS 는 특정 레코드나 테이블을 XML로 내보내어 저장하는 기능을 지원한다. 또한 수집한 정보에 대한 개인 정보보호 문제나 수집 정보에 대한 보안성도 기존의 PKZip 암호화 및 XML 보안 솔루션을 바로 적용할 수 있는 장점이 있다. 또한 디지털 증거의 해쉬값이나 메타데이터를 XML 문서 안에 같이 정의함으로써 무결성을 확보하는데 이용할 수 있다.

2. PKZip 파일 포맷 구조의 활용

그림 7의 Zip 파일 포맷의 구조처럼, 압축의 대상이 되는 파일들은 File Entry 항목에 추가되는 형태를 지니며, 파일 각각의 메타 정보는 Central Directory 에 저장된다.

XeBag 포맷은 사건 관련 정보와 수집한 증거 자료에 대한 포렌식 데이터를 XML 문서 파일로 생성한다. 이 XML 문서는 수집한 증거 자료와 같이 압축하거나 ZIP 파일 구조 상에서 이용할 수 있는 에누리 영역들을 활용하는 형태로 구성할 수 있다. 전자의 경우는 가능한 일반적인 압축 소프트웨어나 라이브러리로 증거 객체의 열람과 추출이 가능하게 하여 범용성을 최대한 보장할 수 있으며, 후자의 경우는 수집한 증거 객체 리스트에는 DEC 에 의해 생성되는 파일을 추가하지 않게 하고 일반적인 압축 소프트웨어로는 포렌식 관련 데이터는 별도의 추출 도구를 이용하여 열람하게 하여 보안성을 강화할 수도 있다.

후자의 방식을 사용하기 위한 방법으로 PKZip 구조를 예로 들면, 사용자가 정의하여 별도로 사용할 수 있는 에누리 영역은 크게 3가지로 Local File Header의

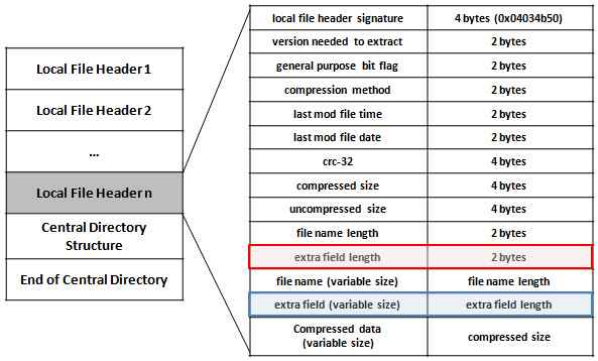

그림 8. PKZip 구조에서 Local File Header
Fig. 8 Local File Header of PKZip file format.

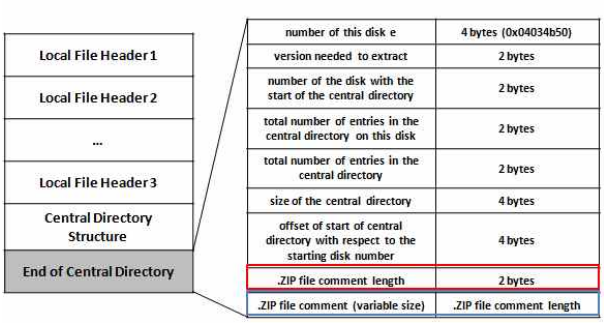

그림 9. PKZip 구조에서 End of Central Directory
Fig. 9. End of Central Directory of PKZip file format.

extra field, Central Directory Structure의 extra field, End of Central Directory의 zip file comment 영역이다.

그림 8은 PKZip 파일 포맷에서 Local File Header의 구조를 나타낸다. 이 항목 중에서 extra field를 사용자가 별도로 지정할 수 있는데, 2 바이트 길이까지 사용 가능한 가변 영역이므로 65536 바이트 만큼 이용할 수 있다. 이와 같이 유사한 형태로 End of Central Directory의 zip file comment 영역도 동일한 크기로 사용이 가능하다. 이러한 구조를 활용할 경우 포렌식 관점에서 중요한 기록이 저장되어 있는 XML 문서는 위와 같은 별도의 공간에 저장할 수 있다.

3. XeBag 포맷 구조

XeBag 구조는, 수집된 증거 파일이나 데이터 셋의 저장은 PKZip을 기반으로 저장하게 되고, 각 증거 자료에 대한 주요 포렌식 데이터는 XML 로 표현이 된다.

기본적인 XeBag 파일 포맷의 구조는 PKZip 파일 포맷의 구조를 따르며, 그림 7의 PKZip 파일 포맷 구조와 동일하다. PKZip 파일 포맷 구조에서 각 파일 엔트리들

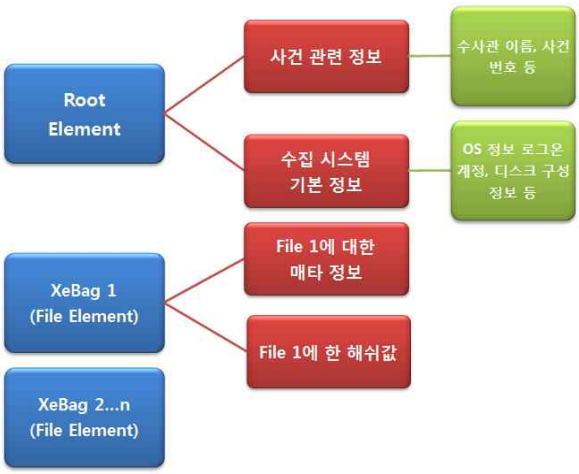

그림 10. XeBag 파일 포맷의 구조
Fig. 10. he structure of XeBag file format.

이 일반적으로 압축 소프트웨어를 사용할 때 압축 대상이 되는 파일들이다. XeBag에서는 디지털 증거 선별 기법을 통해 수집된 주요 증거 자료나 파일이, 파일 엔트리로 하나씩 추가되는 구조를 가지게 된다. XeBag에서는 이러한 파일 엔트리를 증거 객체 (Evidence Object) 라는 이름으로 정의해서 사용한다.

XeBag에서 사용되는 XML 문서는 증거 객체를 포렌식 관점에서 기술할 수 있도록 각 객체마다 생성되는 File element 와 사건에 대한 기본적인 정보가 저장되는 Root element 로 구성된 XML 문서가 생성된다. XeBag에 저장되는 파일 요소를 관리하고 조사 대상 컴퓨터의 기본 정보를 저장하기 위한 Root 요소가 있으며, 파일 요소는 수집한 파일의 MAC 시간 정보나 해쉬값 등이 저장된다. 이처럼 XML 문서는 각 증거 객체를 서술하는 파일 요소와 이를 관리하기 위한 루트 요소로 나뉘는 계층 구조를 지닌다. 이는 Digital Evidence Bag 에서 제시하는 계층적 구조 집합과 유사하며, 이러한 설계 방식은 수십 개의 파일이 저장된 여러 XeBag 파일을 분석할 때 XeBag 파일을 각각압축을 해제하지 않더라도, XML 문서만을 이용해도 증거 객체의 선별이나 간단한 열람이 가능하도록 하여 효과적인 사전 조사가 가능하다.

XeBag 포맷을 이용하여 10개의 문서 파일을 저장한다면, 각각에 해당되는 10개의 파일 요소와 사건 정보, 수집된 파일 리스트 등이 저장되는 1개의 루트 요소가 저장된 하나의 XML 문서로 구성된다. 이러한 XeBag의 XML 문서에 대한 예는 부록 1과 같다.

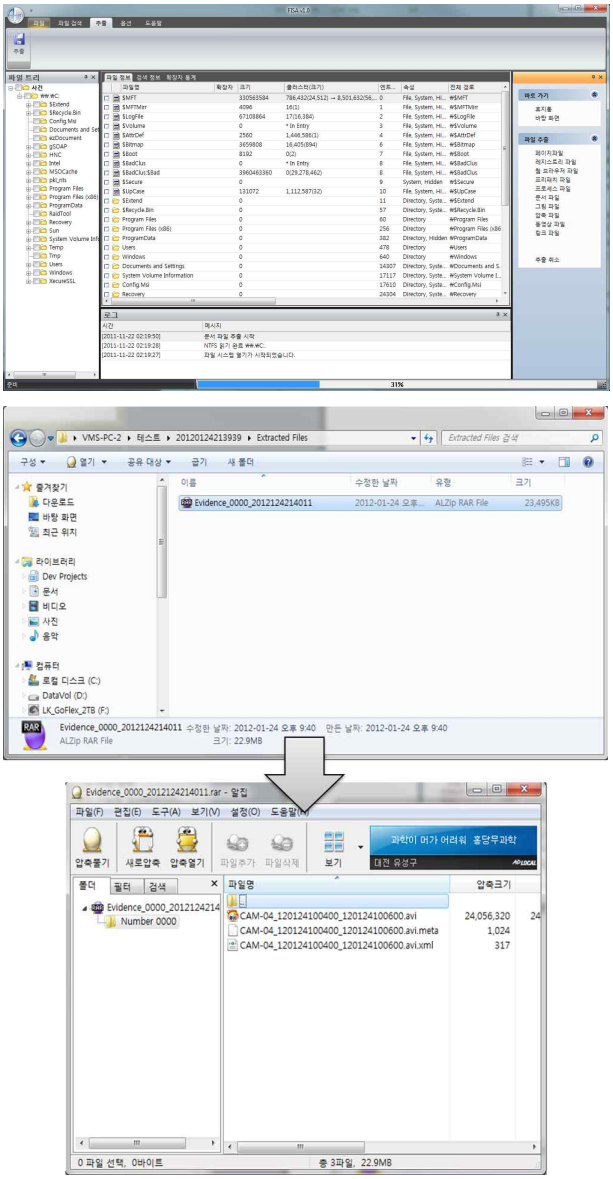

그림 11 파일시스템 분석도구 FISA 메인화면과 WinRAR 포맷기반의 XeBag 증거 파일로 저장한 결과
Fig. 11. A main window of File System Analyzer FISA.

4. 구현

XeBag 파일 포맷을 적용하여 구현하기 위해서는 기본적으로 수사관이 조사 대상 시스템에서 필요한 디지털 증거 자료를 선별 수집하여 저장할 수 있는 포렌식 소프트웨어가 필요하다. 따라서 상용 파일시스템 분석 도구인 포앤식스테크의 FISA[11] 적용하여 수집한 자료를 XeBag 에 보관하여 저장하도록 구현하였다. FISA는 현장 중심의 디지털 포렌식 도구 중의 하나로 활성 컴퓨터의 파일시스템을 분석하여 조사가 필요한 파일을 선별하여 추출할 수 있는 도구이다. 휘발성 데이터 수집 및 분석, 파일시스템 정보 수집 및 출력, 파일 검색 및 추출, 파일 확장자 별 통계 분석 등을 지원한다. 그림 11은 FISA의 메인화면에서 수집할 파일을 선택한 후, WinRAR 기반 XeBag 포맷으로 저장한 결과를 보여준다.

표 1. EnCase Portable 과 XeBag 비교
Table 1. The comparion of EnCase Portalble and XeBag.

요구사항	EnCase Portable	XeBag
범용성	X	O
보존성	O	O
무결성	O	O
단일성	X	O
확장성	△	O
압축성	O	O
보안성	X	O

FISA는 활성시스템에 대한 포렌식 조사가 가능하고 파일 단위의 수집이 가능하기 때문에 EnCase Portable 과 유사한 기능을 가지지만 EnCase Portable은 문서나 이미지 파일과 같은 특정 종류에 대한 모든 파일을 수집하므로 사용자가 선택하여 수집하는 FISA와는 큰 차이점이 있다. 따라서 XeBag 포맷이 적용된 FISA와 EnCase Portalbe 간의 구현 관점에서 객관적인 비교는 어렵다.

또한 본 연구에서 제안한 XeBag 과 기존 연구와의 비교는 Tuner의 Digital Evidence Bag 이나 AFF 포맷은 기본적으로 디스크 이미지 기반에 적용되므로 차별성을 비교하기는 어렵다. 한편 EnCase Portable 의 경우는 EnCase Logical Evidence Format 은 비공개이므로 객관적인 비교는 불가능하지만 앞서 제안한 새로운 DEC 필요 요구사항과 기능적인 관점에서의 비교는 가능하다.

표 1은 EnCase Portable 과 XeBag에 대해 앞서 제안한 DEC 필요 요구사항과 비교한 것이다. XeBag은 기본적인 파일 포맷은 PKZip 파일과 동일하므로 다른 포렌식 도구를 이용하여 분석이 가능하기 때문에 범용성을 만족한다. 또한 수집한 증거자료에 대한 포렌식 정보는 별도의 XML 문서로 보존하고 XeBag 자체는 원본과 동일한 파일을 보관하므로 보존성을 만족한다. 또한 무결성 보장을 위해 수집한 파일들에 대한 해쉬값을 계산하여 보관하므로 이를 만족한다. 단일성의 경우,

다양한 파일 또는 로그나 데이터베이스 레코드 등을 하나의 XeBag 파일에 압축하여 보관하므로 이를 만족한다. 또한 선별 수집한 디지털 증거들은 계속해서 XeBag의 기본 파일 포맷인 압축파일의 엔트리에 추가하므로 확정이 가능하며, 추가적으로 수집할 경우에도 새로운 XeBag 파일에 보관이 가능하므로 확장성을 보장하며, 압축파일 포맷을 이용하기 때문에 압축성도 보장한다. 마지막으로 PKZip 자체의 지원기능인 패스워드 보안을 이용하여 압축할 수 있으므로 보안성을 만족한다.

EnCase Portable의 경우, EnCase 포렌식 도구만을 이용하여 수집한 증거 자료의 열람 및 분석이 가능하므로 다른 포렌식 도구와 연동되는 것이 불가능하기 때문에 범용성을 만족하지 않는다. EnCase Portable에서 지원하는 문서, 이미지와 같이 일반적인 개인 컴퓨터에 존재하는 파일만 수집하므로, 기업 수사 시에 발생하는 데이터베이스 레코드나 서버의 로그 파일과 같은 다양한 증거 데이터 수집에는 한계가 있어 단일성을 만족하지는 않는다. 확장성의 경우, 수집한 데이터의 용량이 수집 완료 후 증가하지는 않으나, 증거 포맷 대해 별도의 메타데이터나 사용자가 작성한 기술 내용을 추가할 수 없으며 수집이 완료된 후 특정 파일을 선별하여 다시 수집하는 것이 불가능하므로 완전히 만족한다고 볼 수는 없다. 또한 수집한 DEC 파일에 대한 별도의 접근 제어 기능이나 암호화 기능이 없으므로 보안성을 만족하지 않는다.

Ⅳ. 결론 및 향후 연구

본 논문에서는 범용성, 단일성, 확장성 등의 DEC 의 요구 사항을 만족하는 새로운 증거 보관 포맷을 제안하였다. XeBag 은 PKZip과 XML을 사용하여 범용성을 최대한 고려하였으며 다양한 데이터 자원에서도 이를 활용하여 포렌식 조사가 가능하도록 하였다. XeBag 의 활용성은 단순히 기존의 압축 해제 도구만 사용하면 기존 파일 단위의 포렌식 도구에 바로 적용이 가능하다는 장점이 있다. 일반 압축도구를 이용해 압축 해제 후 일반적인 열람도구에 입력으로 이용하더라도, XML에 원본 데이터의 시간 정보가 포함되어 있고 XeBag 자체는 압축해제 후, 별도로 보관 가능함으로써 안전성이 보장된다고 볼 수 있다. 단지 증거 데이터 대한 XML 문서를 열람하기 위해서는 이를 추출하여 보여줄 수 있는 열람 도구만 별도로 활용하면 되므로 바로 현장에서 사용 가능하다. 향후 XeBag 포맷의 활용을 넓히기 위해 XeBag의 XML 문서를 열람할 수 있는 뷰어 및 수집 도구의 평가판을 공개할 계획에 있으며, 국내 수사기관과 협조할 계획이다.

향후 연구로는 각각의 디지털 기기 자원에서 XeBag을 활용하기 위한 구체적인 방안에 대하여 분석하고, XeBag 포맷을 활용하기 위한 포맷 개선 연구가 필요하다. 또한 수십 개 이상의 파일을 XeBag 에 담을 경우 발생할 수 있는 데이터 처리량의 증가 등에 대한 성능 개선 연구를 수행할 예정이다. 또한 일반 파일단위가 아닌 다양한 데이터 자원에서도 적용할 수 있도록 체계적인 연구를 수행할 계획이다. XeBag 은 기본적으로 디지털 증거 선별 조사에 적합하도록 설계되었지만, 디스크 이미지에 적용이 가능하므로, 다양한 디지털 증거에 모두 적용할 수 있도록 확장할 수 있다. 이 경우, Zip 포맷에서 기본적으로 제공하는 용량제한으로 인해 XeBag을 분할하는 등의 방안을 고려해야 한다.

참 고 문 헌

[1] 임경수, 이상진, “신속한 사건 대응을 위한 휴대용 포렌식 도구 설계 및 구현,” 2009 디지털 포렌식워크샵, 2009년 8월
[2] 임경수, “디지털 증거 수집을 위한 XML 기반 프레임워크의 설계 및 구현”, 고려대학교 정보경영공학전문대학원, 석사 학위 논문 2008.
[3] Kyungsoo Lim, Seokhee Lee, Jong Hyuk Park, Sangiin Lee “XFRAME: XML-baed framework for efficient acquiring digital evidence on Windows live system”, Proceedings of 4th Annual IFIP WG11.9 International Conference on Digital Forensics, Kyoto, Japan, 2008.
[4] Kyung-soo Lim, SeungBong Lee nd Sangjin Lee,“Applying a Stepwise Forensic Approach to Incident Response and Computer Usage Analysis”,2nd International Conference on Computer Science and its Application,(CSA 2009)
[5] Marcus K. Rogers, James Goldman, Rick Mislan, Timothy Wedge,Steve Debrot, “Computer Forensics Field Triage Process Model”, Conference on Digital Forensics, Security and Law, 2006.
[6] Philip Turner, “Unification of Digital Evidence

from Disparate Sources(Digital Evidence Bags)",Digital Forensic Research Workshop (DFRWS), New Orleans, 2005.

[7] Philip Turner, "Selective and intelligent imaging using digital evidence bags ", Digital Investigation, Volume 3, Supplement 1, September 2006, Pages 59-64

[8] Philip Turner, "Applying a forensic approach to incident response, network investigation and system administration using Digital Evidence Bags", Digital Investigation, Volume 4, Issue 1, March 2007, Pages 30-35

[9] Golden G. Richard III, Vassil Roussev, Lodovico Marziale. "Forensic discovery auditing of digital evidence containers", Digital Investigation, Volume 4, Issue 1, March 2007, Pages 88-97

[10] EnCase Portable, Gudiance Soft http://www.guidancesoftware.com/encase-portable.htm

[11] FISA-File System Analyzer, ㈜포앤식스테크 http://www.4n6tech.com/pro_kr/info/info.php?pn=1&sn=1&dn=1

부록. XeBag XML 문서 예


```
<?xml version="1.0" encoding="UTF-8" ?>
- <XeBag xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <CaseInfo beganTime="2010-09-09 17:21:30" completedTime="2010-09-09 17:40:12" investigator="임경수"
    caseNum="100909-TEST-CASE" comName="LUKE369-VAIO" />
- <BasicInfo appName="System_Basic_Configuration" systemTime="">
  - <IPConfigs appName="ip_list">
      <IP_Info NICName="Intel(R) Centrino(R) Ultimate-N 6300 AGN" HardwareID="{98A8CAD2-11DE-435C-
        ABF1-E8892A44A7DE}" IPAddress="10.13.12.6" MACAddress="00-24-D7-0D-77-6C" ID="2" />
    </IPConfigs>
  - <DiskInfo appName="physical_disk">
      <Disk DriveID="PhysicalDrive0" DiskSize="128034708480(Bytes)=119(Gb)" Cylinders="15566"
        TrackPerCylinders="255" SectorPerTrack="63" BytesPerSector="512" ID="0" />
    </DiskInfo>
  - <PartitionInfo appName="partition">
      <Partition DriveID="" PartitionID="" PartitionType="NTFS" Bootable="YES" Length="16056240" PartitionNumber=""
        PhysicalNumber="" StartingOffset="80" HiddenSectors="" Recognized="" Rewrite="" ID="2" />
      <Partition DriveID="" PartitionID="" PartitionType="NTFS" Bootable="NO" Length="222412800" PartitionNumber=""
        PhysicalNumber="" StartingOffset="27666432" HiddenSectors="" Recognized="" Rewrite="" ID="1" />
      <Partition DriveID="" PartitionID="" PartitionType="NTFS" Bootable="YES" Length="204800" PartitionNumber=""
        PhysicalNumber="" StartingOffset="27461632" HiddenSectors="" Recognized="" Rewrite="" ID="0" />
    </PartitionInfo>
  - <LoggedOnAccountList appName="logon_user">
      <LoggedOnAccount UserName="Luke369" LogOnDomain="Luke369-VAIO" LogOnServer="LUKE369-VAIO"
        Administrator="" ID="0" />
    </LoggedOnAccountList>
  - <Win32AccountList appName="accounts">
      <Win32Account UserName="Luke369" Group="[HomeUsers] [Administrators]" ID="3" />
      <Win32Account UserName="HomeGroupUser$" Group="[HomeUsers]" ID="2" />
      <Win32Account UserName="Guest" Group="[Guests]" ID="1" />
      <Win32Account UserName="Administrator" Group="[HomeUsers] [Administrators]" ID="0" />
    </Win32AccountList>
  </BasicInfo>
- <XeBagFile FileName="EvidenceBag_진행보고.ppt" IndexNumber="1" FilesystemType="NTFS">
  - <FileInfo>
      <Name>EvidenceBag_진행보고.ppt</Name>
      <FileSize>2,479,616</FileSize>
      <ClusterSize>1,063,606</ClusterSize>
      <EntryNumber>2197</EntryNumber>
      <Property>File</Property>
      <FullPath>C:\Users\Luke369\Desktop\XeBag 프로젝트\발표자료\EvidenceBag_진행보고.ppt</FullPath>
    </FileInfo>
  - <MetaData>
    - <StandardInfo>
        <Created>2010-08-30 17:21:30</Created>
        <LastAccessed>2010-08-30 17:21:30</LastAccessed>
        <FileModified>2010-07-08 17:40:30</FileModified>
        <EntryModfied>2010-08-30 17:21:30</EntryModfied>
      </StandardInfo>
    - <FileNameInfo>
        <Name />
        <Created>2010-08-30 17:21:30</Created>
        <LastAccessed>2010-08-30 17:21:30</LastAccessed>
        <FileModified>2010-08-30 17:21:30</FileModified>
        <EntryModfied>2010-08-30 17:21:30</EntryModfied>
      </FileNameInfo>
    </MetaData>
  - <HashValue>
      <MD5>17F5D1369CA224AD88E761385226CF</MD5>
      <SHA1>EDB365A9FB6149746318CCEDC0D711472E3ACCB6</SHA1>
    </HashValue>
  </XeBagFile>
</XeBag>
```

저 자 소 개

임 경 수(정회원)
2006년 부경대학교 컴퓨터멀티미디어공학과 학사 졸업.
2008년 고려대학교 정보보호대학원 석사 졸업.
2010년～현재 고려대학교 정보보호대학원 박사과정 재학 중
2010년～현재 한국전자통신연구원 연구원
<주관심분야 : 디지털 포렌식, 디지털 증거 처리, 정보보호>

이 상 진(정회원)
1987년 고려대학교 수학과 학사 졸업.
1989년 고려대학교 수학과 석사 졸업.
1994년 고려대학교 수학과 박사 졸업.
1989년～1999년 한국전자통신연구원 선임연구원
1999년～2001년 고려대학교 자연과학대학 조교수
2001년～현재 고려대학교 정보보호대학원 교수
<주관심분야 : 디지털 포렌식, 암호학, 정보보호>

이 창 훈(정회원)-교신저자
2001년 한양대학교 수학과 학사 졸업
2003년 고려대학교 정보보호대학원 석사 졸업.
2008년 고려대학교 정보보호대학원 박사 졸업.
2009년～2010년 한신대학교 컴퓨터공학부 전임강사
2010년～2011년 한신대학교 컴퓨터공학부 조교수
2012년～현재 서울과학기술대 컴퓨터공학과 조교수
<주관심분야 : 암호학, 정보보호, 디지털포렌식>