Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.24, NO.3, Jun. 2014 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2014.24.3.523

로그 체인을 고려한 디지털증거지도 작성

박 호 진,† 이 상 진‡
고려대학교 디지털포렌식연구실

Build a Digital Evidence Map considered Log-Chain

Hojin Park,† Sangjin Lee‡
Digital Forensic Research Center, Korea University

요 약

컴퓨터 침해사고 대응 시 침입경로를 파악하는데 많은 시간을 소비하므로 피해 범위가 확대되거나 침해사고 발생의 원인을 밝힐 주요 증거를 유실하게 된다. 이로 인해 동일한 원인의 침해사고가 재발하고 있다. 이 논문에서는 침입자를 신속하고 정확하게 찾아낼 수 있도록 침해사고 발생 전, 디지털증거지도 작성을 제안한다. 디지털증거지도는 다양한 IT 장비와 소프트웨어가 만들어 내는 머신 데이터간의 연결 고리가 그물 형태로 만들어진다. 연결 고리는 다양한 외부요인에 민감하기 때문에 지속적인 관리가 필요하다. 침해사고 발생 전, 유효한 디지털증거지도를 숙지함으로써 침해사고 발생 시 신속히 대응하여 피해 범위를 축소하며 침입경로를 제거하여 침해사고 재발을 방지한다. 디지털증거지도는 로그뿐만 아니라 컴퓨터에서 발생하는 다양한 아티팩트들을 채용함으로써 고도화된 APT 공격과 안티-포렌식 기법에 효과적으로 대응한다.

ABSTRACT

It has been spent too much time to figure out the incident route when we are facing computer security incident. The incident often recurs moreover the damage is expanded because critical clues are lost while we are wasting time with hesitation. This paper suggests to build a Digital Evidence Map (DEM) in order to find out the incident cause speedy and accurately. The DEM is consist of the log chain which is a mesh relationship between machine data. And the DEM should be managed constantly because the log chain is vulnerable to various external facts. It could help handle the incident quickly and cost-effectively by acquainting it before incident. Thus we can prevent recurrence of incident by removing the root cause of it. Since the DEM has adopted artifacts in data as well as log, we could make effective response to APT attack and Anti-Forensic.

Keywords: incident response forensic, evidence map, log chain, windows artifacts

I. 서 론

시만텍이 2011년에 비해 2012년 APT공격이 42% 증가했다고 밝히고[1] 있는 것처럼, 최근 사이버 공격은 지속적으로 발생하고 있으며 점차 고도화되고 있다. 사고가 발생하기 전 사고분석을 위한 준비가 되어있지 않으면 최신 공격기법에 적절히 대응할 수 없어, 결국 사고 원인이 제거되지 못해 재발하게 된다. 실제로 한국인터넷진흥원(KISA)에서 2012년도에 대응한 233개 웹하드 업체 중 침해사고 재발률은 73%에 이른다[2].

침해사고의 정확한 원인을 밝혀내기 위해서는 중요 전산 자원이 어떻게 구동되었는지 신속하게 파악할 필요가 있다. Rowlingson R.은 포렌식 준비도(Forensic Readiness)란 조사비용을 최소화 하면서 디지털 증거들을 최대한 사용할 수 있도록 조직의

접수일(2014년 3월 17일), 수정일(2014년 5월 28일), 게재확정일(2014년 6월 2일)
† 주저자, hojinpk@korea.ac.kr
‡ 교신저자, sangjin@korea.ac.kr(Corresponding author)

능력을 최대화 하는 것으로 정의하였다[4]. 이렇듯 사전 준비를 위해 비즈니스 로직을 정확히 이해하고, 잠재적인 증거 파악과 수집 방법을 숙지하는 등 기술적 접근과 함께 보안 절차에 해당하는 포렌식 준비도가 도입되었다. 이 논문에서는 침해사고 사전적(Proactive) 포렌식 준비도 관점[3]에서 디지털증거지도 작성 방법을 제안한다. 디지털증거지도가 제공하는 로그체인과 같은 프레임워크를 사전에 준비함으로써 침해사고 대응 시 분석 속도를 높일 수 있을 뿐만 아니라, 침해사고 경로를 파악할 수 있어 정확한 침해사고의 원인을 제거할 수 있다.

1.1. 침해사고 발생 전 증거 수집의 필요성

최근 발생하고 있는 APT(Advanced Persistent Threat) 공격은 타켓팅한 목표물에 최적화된 공격을 사용하기 때문에 공격 성공률이 높고, 공격 사실을 인지하기 어렵다는 특징이 있다. 2013년도 Verizon의 침해사고 조사 보고서에 따르면 데이터 유출 사례의 2/3가 감지하기까지 짧게는 수개월에서 길게는 수년까지 걸린다고 한다[5]. 이렇듯 안티-포렌식 기법이 사용되어 증거가 인멸되는 것을 사전에 방지해야 할 필요성이 점차 증가되고 있다. 침해사고가 발생한 후 증거를 수집한다면, 분석을 위한 기초자료가 부족하여 침해사고의 침입시점과 피해범위 그리고 사고원인 등 적절한 사고대응을 수행할 수 없다.

이러한 공격에 대응하기 위해서 가장 기본적으로는 방어에 최선을 다해야 하겠지만, 브루스 슈나이어가 보안은 사슬과 같아서 가장 약한 고리만큼만 안전하다[6]고 말한 것처럼 완벽한 보안은 없다고 가정하고 침해사고를 인지한 후 적절한 대응이 신속히 이루어질 수 있도록 체계적인 준비가 필요하다.

1.2. 디지털증거지도 작성의 필요성

침해사고 대응의 신속성과 정확성을 높이기 위하여 잠재적인 증거가 누락 없이 수집되어야 하며, 다양한 장비와 소프트웨어들이 만들어 내는 로그들 간의 상관관계 분석이 가능한 환경이 준비되어 있어야 한다. 또한 기업 내부 침해사고 대응 인력뿐만 아니라 침해사고 대응을 서비스로 제공하는 업체에서도 고객과 업무 협약 후 시스템 구성도를 파악하는 일 못지않게 잠재적인 증거 파악과 머신 데이터(Machine-Generated Data)들 간의 상관관계를 파악하는 일 또한 중요하다[7].

실제로 시스템 구성도는 작성하는 관점에 따라 추상적으로 작성되거나 침해사고 대응을 의뢰한 고객사의 사정으로 부분적인 자료만 제공되어 침해사고 대응 시 부족한 정보로 인해 정확하지 못한 결론이 도출될 수 있다. 그리고 시스템 구성도를 모두 이해하고 있다 하더라도 서로 다른 장비 혹은 소프트웨어간의 로그 형식에 대한 이해는 별개이다. 더욱이 다양한 로그들의 생명주기가 관리되지 않는다면, APT 공격과 같이 침입한지 1년이 넘어 발견되었을 경우 로그가 존재하지 않아 분석이 불가능해질 수 있다.

이 논문에서는 잠재적인 증거의 활용도를 높이고 다양한 로그들 간의 상관관계 분석을 보다 빠르고 정확하게 하기 위해서 디지털증거지도(Digital Evidence Map) 작성을 제안한다. 디지털증거지도란 다양한 머신 데이터들 간의 상관관계를 분석할 수 있도록 시스템을 구축 및 유지할 수 있는 프레임워크이다.

아직까지 국내외에서 발표된 침해사고 대응 절차에서 로그 관리는 디지털증거지도 관점에 대해 고려되지 않고 기본적인 저장에 관해서만 언급되어 있다. 로그 저장은 KISA의 사고 대응 7단계 중 “사고 전 준비”[8]와 NIST의 “Incident Response Life Cycle” 중 첫 번째 단계인 Preparation[9]에 해당한다. 따라서 침해사고 대응 절차상 준비단계, 즉 침해사고가 발생하기 전 단계에서 로그의 저장뿐만 아니라 다양한 로그들 간의 상관관계가 파악될 수 있도록 디지털증거지도가 작성돼야 한다.

II. 디지털 증거의 위치 분석

이 논문에서는 디지털증거를 로그(Log)와 아티팩트(Artifacts)로 구분한다. 로그는 프로그램이 실행되면서 관련된 내용이 기록된 것으로 일정한 형식이 존재한다. 아티팩트는 운영체제 또는 응용프로그램이 구동되면서 남기는 흔적으로 일정한 형식이 존재하지 않는다.

2.1. 로그

로그는 정형화된 규칙 혹은 틀에 맞추어 기록된다. 이러한 데이터는 침입경로 상에서 크게 네트워크 장비, 운영체제 그리고 응용프로그램이 생성하는 로그로 구분된다. 네트워크 장비는 방화벽, IDS/IPS 와 같

은 네트워크 보안장비가 있으며, 윈도우 운영체제에서는 윈도우 이벤트 로그, 개인 방화벽, ETW(Event Tracing for Windows) 등이 있다. 응용프로그램은 IIS와 같은 웹 서비스, 컴퓨터 백신, 데이터베이스 로그 등 다양한 로그가 있다. 이러한 로그는 침해사고의 원인 분석 시 주로 사용된다.

로그는 일정한 양식이 존재하는데, 네트워크의 패킷 데이터를 NetFlow (RFC 3954)나 Syslog (RFC5424) 형태로 정형화된 로그 포맷이 있다. 또한 Fig.1.과 같이 IIS 웹 서비스 프로그램은 W3C ELF, Extended Log Format으로 기록한다. 이와 같이 데이터 내용은 데이터의 종류마다 다르지만 장비나 프로그램에 의해 일정 양식에 맞춰 기록된 것이기 때문에 분석이 용이하다.

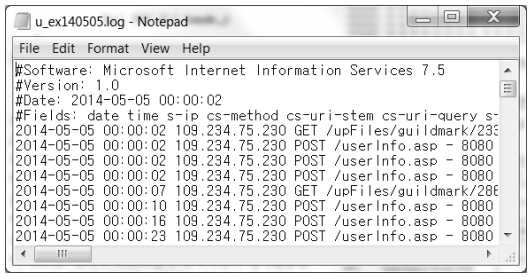

Fig.1. W3C ELF Sample

2.2. 아티팩트(Artifacts)

아티팩트는 컴퓨터 프로그램이 동작하면서 남겨진 흔적을 말한다. 일반적으로 아티팩트는 파일 다운로드, 프로그램 실행, 파일 조작, 하드웨어 사용, 계정 사용, 웹 브라우저 사용으로 구분된다. 아티팩트가 로그와 구별되는 점은 데이터가 기록된 목적에 있다. 로그는 프로그램이 동작하면서 처리한 일련의 데이터를 기록하여 증거를 남기기 위함이지만, 아티팩트는 프로그램이 동작하기 위해 필요한 정보를 기록한 것이다. 예를 들어 로그는 방화벽 프로그램이 동작하면서 받아들였던 데이터와 처리한 결과를 기록하는 것이며, 아티팩트는 파일시스템이 동작하면서 필요한 시간정보, 파일 삭제여부, 소유자 등의 데이터들과 같은 정보를 말한다.

이와 같이 아티팩트는 각 프로그램의 목적에 따라 기록하는 데이터와 형식이 다를 수 있기 때문에 각 데이터 간의 규칙을 찾는데 많은 노력이 필요하다. 특히 아티팩트는 인간이 읽을 수 없는 형태의 바이너리 형태로 기록되는 경우가 대부분이어서 로그로 활용하기 위해서는 읽을 수 있는 문자로 변환하는 과정이 필요하다.

2.2.1 바이너리 데이터

윈도우 운영체제에서 사용하는 대표적인 아티팩트에는 파일 시스템과 레지스트리가 있다. 윈도우 파일 시스템으로 대표적인 NTFS 포맷은 다수의 메타데이터들에 파일시스템 정보가 기록되어 있다. 그 중에 $MFT 파일에는 모든 파일들의 엔트리 정보가 기록되어 있고, 각 파일에 해당하는 다양한 속성들을 가지고 있는데 $STANDARD_INFORMATION 속성에는 파일의 생성, 접근, 수정 시간과 소유자 등의 정보가 기록되어 있다. 해당 구조체가 저장된 데이터를 살펴보면 Fig.2.에서 블록으로 처리된 부분과 같이 사람이 읽을 수 없는 바이너리 데이터이다.

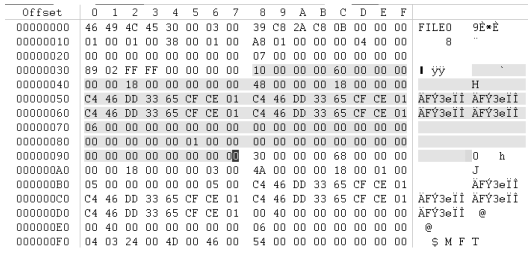

Fig.2. $STANDARD_INFORMATION BINARY Data

$STANDARD_INFORMATION 구조체는 Fig.3.과 같이 정의되어 있다.

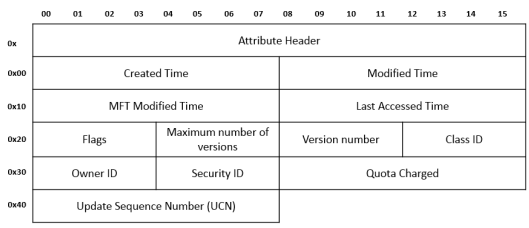

Fig.3. $STANDARD_INFORMATION Structure

윈도우 운영체제는 Fig.3.과 같이 정의된 구조를 기반으로 Fig.2.의 바이너리 데이터를 사람이 읽을 수 있도록 변환하여 Fig.4.와 같이 제공한다. 하지만 구조체에 정의된 모든 속성들이 표기되지는 않는다. 예를 들어 $FILE_NAME 속성의 구조체에도 파일

의 생성, 접근, 수정 시간이 기록되어 있지만 $FILE_NAME 구조체의 정보를 파일 속성 화면에서 사용되지 않으며, 시간을 확인하고 변경할 수 있는 윈도우 API를 이용할 경우에도 $FILE_NAME의 시간정보를 불러오거나 수정하지 않는다.

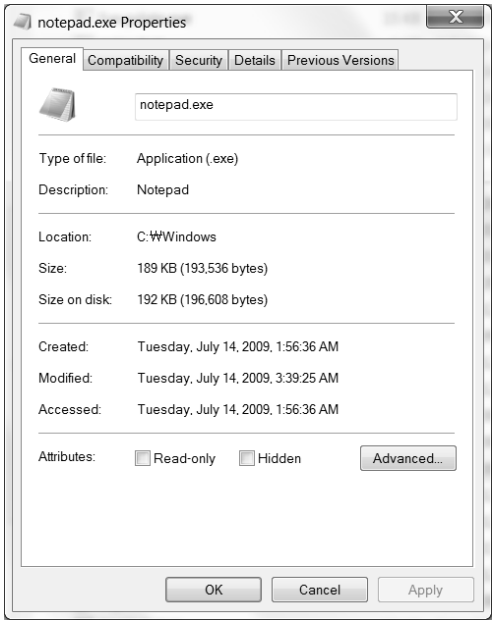

Fig.4. A Windows File Property

파일 시스템 이외에도 윈도우 레지스트리는 지정된 파일들에 HIVE라는 구조의 데이터가 바이너리 형태로 기록되어 있다. HIVE 내부 구조체인 _CM_KEY_NODE 에는 레지스트리 키(Key)가 마지막으로 기록한 시간만을 LastWriteTime 변수에 기록하고 있다. 하지만 기본적으로 윈도우 레지스트리 구조를 보여주는 "레지스트리 편집기"에서는 시간 정보를 확인할 수 없다. 특히 마지막 기록시간만 저장[10]하고 있기 때문에 최초 기록 시간을 알 수 없다는 특징이 있다.

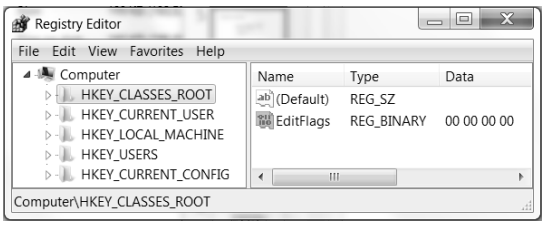

Fig.5. Windows Registry Editor

이렇듯 아티팩트는 저마다의 구조와 특징을 가지고 있어 침해사고 대응 시 활용도를 높이기 위해서는 인간이 인식할 수 있는(Human-Readable) 데이터로 변환 후 기록돼야 한다.

2.2.2 휘발성

로그의 경우 보관을 목적으로 하기 때문에 로그를 기록하는 프로그램에서 로그 보관 기간이 설정된다. 그러나 아티팩트의 경우 휘발성이 높기 때문에 각 아티팩트의 속성에 따라 사라지기 전에 저장하지 않는다면 활용할 수 없다.

NTFS 파일 시스템의 메타 데이터인 $LogFile에는 NTFS 파일 시스템의 안정성을 높이기 위해 모든 트랜잭션이 순환구조로 기록된다[11]. 따라서 악성코드가 로그뿐만 아니라 자신인 실행프로그램을 삭제하였더라도 $LogFile 에 기록된 데이터를 분석하여 악성코드의 흔적을 찾아낼 수 있다. 그러나 일반적인 하드디스크 볼륨에서는 $LogFile 이 64MB 크기만을 가지고 있기 때문에 파일 트랜잭션이 많을 경우 증거가 남아 있을 가능성이 줄어든다. 따라서 평소 아트팩트를 기록하고자 하는 컴퓨터가 얼마나 파일 트랜잭션을 일으키고 있는지 확인할 필요가 있다.

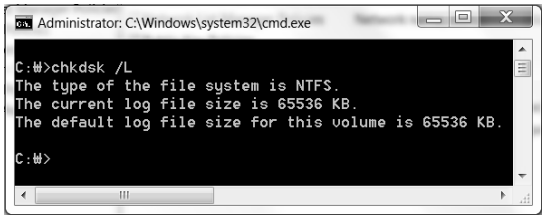

Fig.6. Check $LogFile File Size

실험 결과 $LogFile이 64MB 일 때 일반적인 사무용 컴퓨터는 약 3시간 정도의 데이터가 남아있다. 따라서 적어도 하루 근무시간인 9시간을 기록하고 싶다면 산술적으로 64MB의 3배인 192MB로 지정해야 한다. $LogFile 의 크기를 변경하기 위해서는 chkdsk 의 /L:크기(KB) 옵션을 사용할 수 있다. 예) chkdsk C: /L:196608

2.2.3 안티-포렌식 대응

파일형태로 기록되는 로그는 삭제되기 쉽지만, 앞서 살펴본 파일시스템의 $MFT 파일과 $LogFile 그

리고 레지스트리 HIVE파일인 SAM과 같은 파일들은 운영체제가 동작하는 동안 위변조 되기 어렵다.

침입자가 증거 인멸을 위해 파일삭제, 시간변조 등과 같이 안티-포렌식(Anti-Forensic) 행위를 했을 경우에도 다양한 아티팩트를 이용하여 악의적인 파일이 사용되었다는 증거를 찾아낼 수 있으며, 변조되지 않은 시간을 추출하여 분석에 사용할 수 있다.

따라서 안티-포렌식에 대응하기 위해 안티-포렌식 행위 자체를 차단하는 것도 중요하지만 행위가 일어났을 때를 대비하여 다양한 아티팩트들을 디지털증거지도에 채용해야 한다.

III. 디지털증거지도 작성 방법

디지털증거지도에 증거들을 빠짐없이(Collectively Exhaustive) 반영하기 위해 잠재적인 증거를 파악하는데 체계적인 접근이 필요하며, 이러한 증거들이 서로 연결되어 있어 연관분석을 통해 전체적인 침입경로를 파악할 수 있어야 한다. 디지털증거지도는 이와 같은 목적을 달성하기 위해 "로그체인 방식"과 로그의 신뢰성과 가용성을 높이기 위해 "로그 생명 관리 방식"을 제안한다.

3.1. 증거항목 조사

디지털증거지도를 작성할 때 가장 중요한 것은 잠재적으로 증거가 될 수 있는 모든 디지털 데이터가 고려돼야 한다. 이와 같이 모든 디지털 데이터를 고려하기 위해 Fig.7.과 같은 3단계로 증거목록을 수집할 것을 제안한다.

첫 번째 단계로 "시스템 구성도"를 통해 전체적인 시스템 구조를 파악한다. 이 과정에서 시스템 구성도가 올바르게 작성되었는지 검증하는 것이 중요하다. 시스템 존재 여부를 검증하는 방법은 NMAP과 같은 네트워크 스캐닝 도구로 호스트 탐지(Host Discovery) 작업을 통해 탐지된 서버들과 문서에 나타난 서버들을 비교함으로써 문서에 포함되지 않은 테스트 서버나 방치되어 있는 서버들을 찾을 수 있다. 실제로 이러한 서버들이 공격에 주요 목표가 된다.

두 번째로 모의침투테스트(PT, Penetration Test)와 같은 방법을 통해 침입경로를 찾아내야 한다. 특히 외부에서부터 내부로 들어오는 경로뿐만 아니라 내부에서 외부로 통하는 경로에 대한 파악도 포함돼야 한다. 인바운드 트래픽을 통해서는 침입의 원인을 찾아낼 수 있는 반면 아웃바운드 트래픽을 통해서는 민감한 데이터 유출 여부를 판단할 수 있기 때문이다. 또한 비인가 된 경로뿐만 아니라 정상적인 경로를 통해 유입되는 침입에 대해서도 고려돼야 한다. 내부자 위협뿐만 아니라 관리자의 PC가 탈취되었을 경우 관리자가 접속한 정상적인 경로를 통해 악의적인 사용자는 추가적인 인증절차 없이 접근할 수 있기 때문이다.

마지막으로 네트워크 장비 목록과 소프트웨어 목록을 참고하여 앞서 밝혀진 예상 침입경로에서 어떠한 네트워크 장비를 거치게 되는지 어떤 운영체제와 소프트웨어들이 사용되는지 파악한다. 특히 로그뿐만 아니라 아티팩트들에 대한 파악이 이 단계에서 조사된다. 이렇게 조사된 다양한 로그들이 어떻게 기록, 저장, 활용되는지는 지속적인 관리가 필요하다.

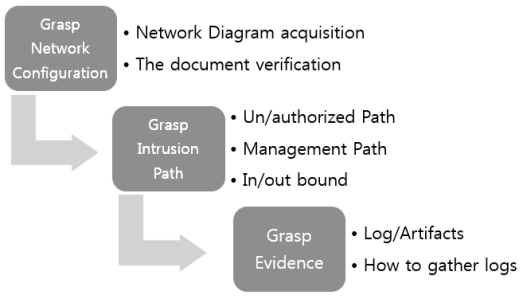

Fig.7. The Extract Process of Potential Digital Evidence

3.2. 로그체인 분석

로그체인(Log-Chain)이란 서로 다른 로그들 간의 연결고리를 일컫는 것으로 서로 다른 시스템 혹은 프로그램에 의해서 생성된 다른 형식의 로그이지만, 침입경로 탐지와 같은 목적을 달성하기 위해 필요한 속성들을 서로 연결시킨 것이다.

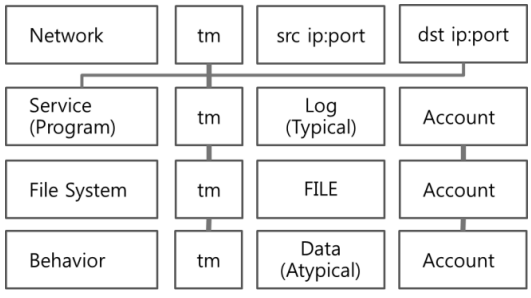

Fig.8. Log-Chain Example

일예로 지금까지도 활발히 활용되고 있는 타임라인 분석 과정에서는 시간 정보를 이용하여 다양한 로그들을 연결한다. 그러나 시간정보는 쉽게 위변조 될 수 있다는 특징이 있어 시간정보만으로 로그를 연결할 경우 안티-포렌식과 같이 공격자가 분석을 방해할 목적으로 시간정보를 변경하였을 경우 상관분석을 할 수 없다. 따라서 이 논문에서 로그와 아티팩트에 기록된 다양한 속성들을 이용하여 그물형태의 로그체인을 구성할 것을 제안한다.

Fig.8.의 tm(시간)이나 Account(계정)과 같이 각 로그에서 동일한 속성들을 연결하여 끊어지지 않고 체인을 구성할 수 있다. 만약 연결고리가 없을 경우 연결할 수 있는 로그를 추가로 기록하거나 침입경로를 물리적 혹은 정책적으로 제한하는 등의 사전 작업이 필요하다.

로그체인은 분석의 신속성과 정확성을 고려하는 가장 중요한 사전 작업이다. 하나의 침해사고 분석 시 다양한 로그들간의 연결을 위해 로그체인이 사용되기도 하지만, 실제로 다수의 침해사고가 발생하였을 경우 각각의 침해사고들 간의 독립성을 위해서도 로그체인은 중요하게 다루어진다. 만약 각각의 침해사고를 구분할 수 없을 경우 침해범위를 정확히 결정할 수 없어 침해사고 대응에 보다 많은 시간이 소요된다.

IV. 디지털증거지도 운영 방법론

디지털 증거는 외부 요인에 의해 끊임없이 변하기 때문에 디지털증거지도를 활용하는데 있어 Fig.9.와 같이 순환구조로 지속적인 관리가 필수적이다. 앞 절에서 언급된 증거항목 조사와 상관분석을 위한 로그체인 분석 이외의 정책수립과 유지보수는 디지털증거지도를 통해 최대의 효과를 얻기 위해 필수이다.

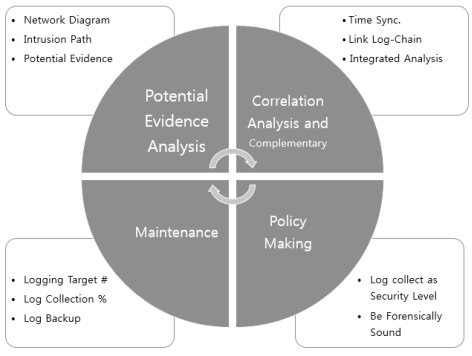

Fig.9. The Digital Evidence Life-Cycle

4.1. 정책 수립

잠재적 증거들을 선별하고 로그체인 분석을 통해 각 데이터들 간의 상관분석이 가능하도록 보완하는 작업이 완료되었다면, 그 다음으로 잠재적인 증거들에게 영향을 줄 수 있는 외부요인들을 사전에 파악하고 관리하여야 한다.

4.1.1 견고한 포렌식 환경

로그의 경우 데이터를 생성하는 시스템이나 소프트웨어를 통해 로깅여부나 로깅 시 기록하는 속성들을 정의함으로써 제어가 가능하다. 그에 비해 아티팩트는 운영체제, 버전 그리고 컴퓨터 사용자의 습관에 이르기 까지 다양한 외부요인으로 인해 데이터 통제가 까다롭다. 이렇듯 로그와 아티팩트를 보다 적극적으로 활용하기 위해서는 견고한 환경(Forensically Sound)이 제공돼야 한다.

시간 동기화

대부분의 디지털 증거에는 시간이 기록된다. 디지털증거지도 작성 시 가장 중요하게 다루는 로그체인에서도 시간에 대한 정보는 다양한 로그들 간의 가장 기본적인 연결 고리이다. 타임라인 분석도구인 log2timeline이 입력으로 받아들이는 로그의 종류는 35개이며[12] 이들 간의 기준이 되는 시간은 동기화되어 있다는 전제를 기본으로 하고 있다.

또한 통합 분석 도구를 사용하여 다양한 로그를 분석할 경우 시간 동기화뿐만 아니라 타임존에 대한 동기화가 필요하다. 예를 들어 일반적인 로그에는 로컬시간이 기록되지만 웹 서비스인 IIS 로그에는 UTC+0 가 기록된다. 따라서 통합분석도구에 IIS 로그와 다른 로그를 입력할 때 각 로그에 맞는 UTC(Coordinated Universal Time)를 설정하여야 정확한 분석을 할 수 있다.

윈도우 이벤트 로그

윈도우 로컬 감사 정책은 기본적으로 윈도우 이벤트 로그에 기록하지 않는다. 윈도우 이벤트 로그는 침해사고 대응 시 결정적인 증거를 제시할 수 있는 유용한 정보가 많기 때문에 침해사고가 발생되기 전에 다양한 감사 이벤트 들이 윈도우 이벤트로그에 기록될 수 있도록 변경 작업이 필요하다.

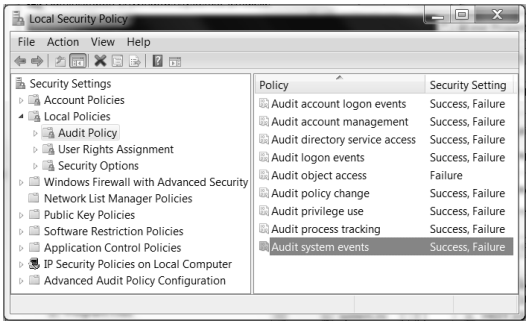

Fig.10. Windows Event Log Audit Settings

윈도우 프리패치

응용프로그램이 실행될 때 윈도우 프리패치 기능이 활성화 되어 있다면 실행되었던 응용프로그램의 구동 흔적이 %SystemRoot%\Prefetch 폴더에 생성된다. 그러나 윈도우 운영체제 버전마다 기본 설정이 다르기 때문에 침해사고 대응 정책에 맞는 설정 작업이 선행돼야 한다. 예를 들어 Prefetch를 결정하는 레지스트리 값인 EnablePrefetcher 가 Windows Workstation 용은 부팅영역과 응용프로그램 실행에 사용할 수 있는 "3"이고 Server 용 버전에서는 부팅영역에서만 사용할 수 있는 "2" 이다.

이렇듯 활용하고자 하는 데이터들의 특성을 사전에 인지하여 상황에 알맞은 정책이 반영돼야 한다.

4.1.2. 보안등급에 따른 로깅 정책

기존에는 로그를 활용함으로써 저장된 로그 파일의 크기를 예측할 수 있었다. 그러나 아티팩트를 분석하기 위해 물리 메모리와 하드디스크를 이미징(Imaging) 할 경우 로그 크기가 급격하게 증가하게 된다. 이러한 점을 보완하기 위해 Guidance Soft의 CyberSecurity 와 MANDIANT의 Intelligent Response 상용제품의 경우 침해사고로 의심되는 이벤트가 발생했을 때에만 물리 메모리를 저장하거나 하드디스크를 이미징한다. 디지털증거지도에서는 각 기업이나 기관의 보안위협 등급에 따라 다양한 로그의 로깅 여부를 결정함으로써 적절한 침해사고 대응을 수행할 수 있다. 또한 보안 등급에 따라 로깅 정책을 수립함으로써 로그를 저장하고 분석하는데 필요한 자원을 기업 규모에 맞출 수 있다.

Table 1. The logging policy as Security Level

	Logging			
	Guarded (Green)	Elevated (Blue)	High (Yellow)	Severe (Red)
Firewall Log	O	O	O	O
Dump Log	X	O	O	O
Tracing Log	X	X	O	O
$MFT File	X	X	X	O

4.2. 유지보수

로그가 생성된 이후 수집되는 로그 관리 시스템에서 일반적으로 수행하는 로그의 효과적인 저장/보호/분석 작업[13] 이외에 디지털증거지도와 관련된 유지보수가 추가로 필요하다. 디지털증거지도를 한번 작성해 봄으로써 로그체인을 연결할 수 있지만, 운영환경에서는 다양한 외부요인으로 시스템이나 아티팩트를 포함한 로그 설정이 계속해서 변하기 때문에 로그체인이 끊어질 수 있다. 따라서 로그체인이 유지될 수 있도록 지속적인 노력이 필요하다.

이와 같이 디지털증거지도의 로그체인 형성을 통한 이론적인 접근 뿐만 아니라, 로그 수집 대상인 모든 곳에서 실제로 로그가 수집되고 있는지 모니터링 되어야 하고, 수집된 로그의 보관일 또한 정책에 따라 동일한 기간 동안 저장해야 하며, 로그의 연계보관성(Chain-of-Custody)과 로그체인 유지를 위해 각 로그의 담당자를 디지털증거지도 생명관리에 포함하여 지정된 관리자에 의해 로그 무결성을 입증할 필요가 있다.

V. 가상환경 실험

가상 공격을 통해 디지털증거지도가 작성되기 이전과 이후의 침해사고 대응을 비교하고, 디지털증거지도를 통해 침해사고 대응 시 어떠한 효과를 얻을 수 있는지 살펴보았다.

네트워크는 가상화 솔루션을 이용하여 NAT 환경을 구축하였으며 방화벽은 리눅스 운영체제에 brctl 명령어를 이용하여 브릿지(Bridge) 방화벽을 구축했다. 웹서버는 윈도우 서버 2008 운영체제에 Apache 웹서버와 MySQL 데이터베이스 그리고 공개용 웹 응

용프로그램 DVWA, Damn Vulnerable Web Application을 사용했다.

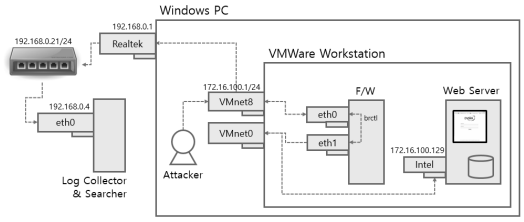

Fig.11. A system structure for Penetration Test

5.1. 디지털증거지도 작성 전 공격

공격은 공개용 SQL 인젝션 도구인 sqlmap을 이용하여 취약점이 있는지 확인하였으며, 발견된 DVWA의 파일 업로드 취약점을 이용해 공개용 b374k 웹쉘을 업로드 했다. b374k 웹쉘의 리버스 쉘 기능을 이용하여 윈도우 쉘 획득 후 mysqldump 명령어로 데이터 추출 후 ftp 명령어를 이용하여 추출된 데이터를 외부로 전송했다. 공격 목적을 달성한 후 분석을 방해하기 위해 timestomp.exe를 이용하여 공격과 관련된 모든 파일을 윈도우 기본 프로그램인 notepad.exe의 시간과 동일하게 했다.

첫 번째 공격에서는 운영체제 및 응용프로그램의 기본 설정값을 그대로 사용했고 아티팩트들도 분석하지 않았다. 첫 번째 공격 후 증거를 수집한 결과 방화벽 로그와 웹 응용프로그램인 Apache 로그 이외에는 얻을 수 있는 분석 데이터가 없어 침해사고가 아닌 침해시도로 분석결과를 결론지을 수밖에 없었다.

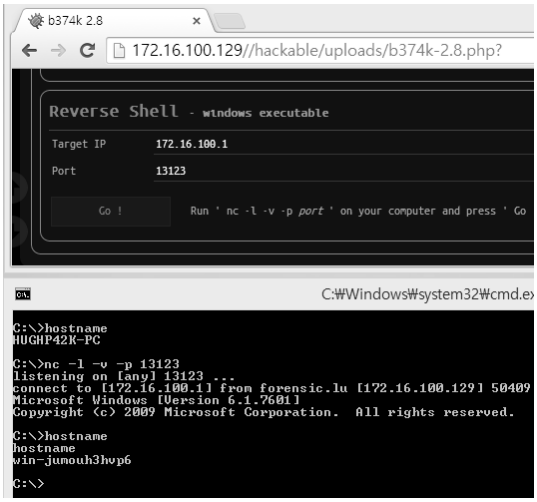

Fig.12. Intrusion success through Reverse Shell

5.2. 디지털증거지도 작성

앞서 발생한 공격 대응 시 부족했던 점을 보완하기 위해 4장에서 언급된 디지털증거지도 생명주기에 따라 수행 내역을 Table 2.와 같이 정리하였다.디지털증거지도 작성 전 발생한 공격 대응 시 분석할 수 있는 정보가 극히 제한적이었던 점을 보완하기 위해 다양한 로그가 기록될 수 있도록 윈도우 이벤트 로그 설정을 변경하였고, 디지털증거지도가 단순한 체인 형태가 아닌 그물형태로 연결될 수 있도록 아티팩트를 적

Table 2. The Details for Digital Evidence Map

Life Cycle	Category	The result of Implement and inspection
Potential Evidence Analysis	Network Diagram	Refer Fig.11. A System Diagram for Penetration Test
	Intrusion Path	1) Web application vulnerability 2) Authorized account
	Potential Evidence	{Path-1} F/W log, Web log, Event log and Artifact log of web server
		{Path-2} Add admin PC Event log and Artifacts
Correlation Analysis and Complementary	Time Sync.	Synchronize local standard date and time through a NTP server
	Link Log-Chain	Link Mash Log-Chain in a Digital Evidence Map
	Integrated Analysis	Build Integrated analysis system by using text search engine, Splunk
Policy Making	Threat Level	Build a logging setting, collecting policy base on threat level, "Guarded, Elevated, High, Severe"
	Log Collection	By using a log collect agent, like Splunk Universal Forwarder
Maintenance	Logging Target #	Maintain the latest Network Diagram and Protection List.
	Log Collection %	Maintain 100%
	Log Backup	Backup Splunk's Index and hashing it

극 활용하고자 하였다. 이를 위해 윈도우 프리패치 설정 값인 PrefetchParameters 을 3 (All)으로 설정하고, 파일의 마지막 접근시간이 기록될 수 있도록 NtfsDisableLastAccessUpdate 값을 0 (Enable)으로 설정했다. 또한 안티포렌식에 대응방안으로 $LogFile 크기를 3배 증가시켰다. 이와 같은 사전준비활동으로 확보할 수 있는 증거들이 증가하였으며, 이를 기반으로 각 로그에서 얻을 수 있는 항목에 ○ 기호를 표기함으로써 Fig.13.와 같이 그물형 로그체인 디지털증거지도를 작성했다.

5.3. 디지털증거지도 작성 후 공격

두 번째 공격도 첫 번째와 동일하게 수행했다. 공격 후 아티팩트를 포함하여 수집된 로그 분석 결과 공격으로 판단되는 로그들을 확인 후 Fig.13.에서와 같이 ● 기호로 표기했다. 침입경로를 확인하기 위해 동일한 시간 혹은 동일한 실행파일 경로를 가지는 검은색 점들끼리 연결하여 침입경로와 침입자를 찾을 수 있었다.

공격자는 시간정보를 손상시켜 윈도우 이벤트로그에 기록된 프로세스 실행 정보와 윈도우 프리패치 파일 그리고 $MFT 파일의 시간정보를 이용해서는 연결고리를 찾을 수 없어 디지털증거지도에는 점선으로 표기했다.

그러나 추가로 수집된 아티팩트 중 $LogFile에는 Fig.12.에서와 같이 리버스 쉘 기능이 있는 b374k.exe 프로세스 실행 기록이 남아 있었다. 또한 $MFT, $LogFile 그리고 윈도우 이벤트 로그에 기록된 프로세스 실행 경로와 네트워크 연결 정보가 기록된 윈도우 이벤트 로그에 기록된 프로그램 경로를 통해 공격자 IP를 확인할 수 있다.

VI. 적용 사례

지속적으로 침해사고가 발생하는 게임업체 A에서 2012년 침해사고가 발생했을 때, 대외적인 대응 등을 제외하고 사고 분석에만 2개월이 소요되었다. 분석결과 몇몇 시스템에서 발견되는 악성코드 역분석 (Reverse Engineering)을 통해 확보한 C&C주소를 기준으로 비인가 된 연결이 있었는지 확인하는 것까지는 분석이 가능했지만, 전체적인 침입경로를 확인하거나 원인을 규명할 수는 없었다. 결론적으로 재발방지를 위한 명확한 대책을 수립하지 못한 체 전제 서버 재설치 후 사건이 마무리 되었다.

이후 A업체를 대상으로 Fig.14.에서 ○표로만 채워진 디지털증거지도를 작성하여 시야확보 및 그물형 로그체인을 확보했다. 또한 Fig.15.에서와 같이 이 논문에서 제시한 디지털증거지도 운영 방법론에 근거하여 지속적인 보안활동을 수행했다. 그 결과 2013년 10월에 발생한 침해사고에서 침입 후 약 4시간 만에 초동대응을 완료했다. 침입을 인지한 직후 피해확산을 막기 위해 침입에 사용된 계정을 비활성화 하였으며 침입경로를 파악한 결과 외부업체를 통한 침입임을 인지하고 외부직원이 사용한 PC를 인터넷으로부터 격리 후 이미징 작업을 통해 상세분석단계로 넘어갔다.

이 사건은 침입자가 A업체 협력사 직원 PC에서 터널링 기법으로, 협력사 직원이 A업체 접속하기 위해 정상적으로 인증 받은 네트워크 경로를 추가적인 인증

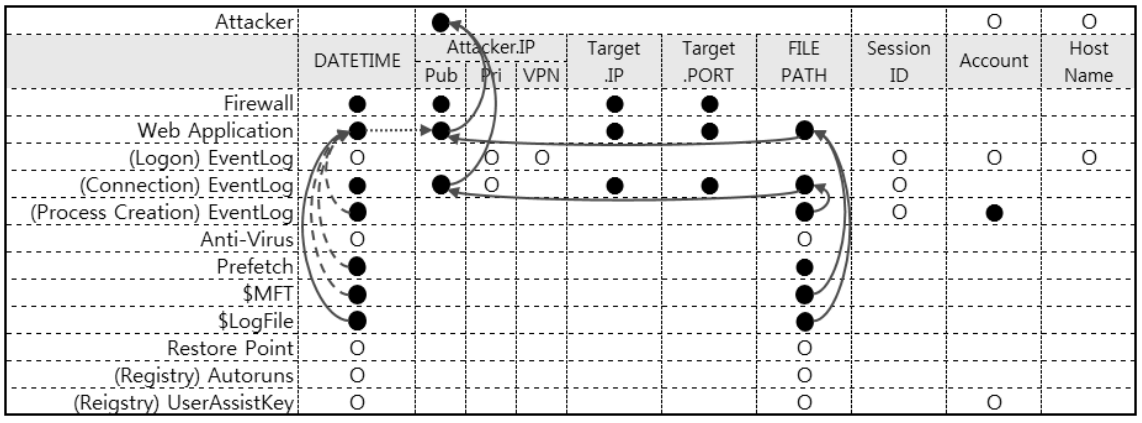

Fig.13. Digital Evidence Map Tracing

절차 없이 A업체의 서버에 접근했다. 침입자는 목적한 것을 얻기 위해 침입한 윈도우 시스템에 로그인되어 있는 다른 사용자의 자격증명(Credential)을 획득하는 방식으로 약 4시간동안 21개의 서버에 접근했다.

침입자가 신뢰된 경로를 이용하였기 때문에 침해사실을 인지하는데 다소 시간이 소요되었지만, 침입자가 다른 사용자의 윈도우 자격을 얻기 위해 실행한 quser.exe, wce.exe 등의 명령어가 윈도우 이벤트 로그에 기록되어 침입사실을 인지할 수 있었다.

최초 침입사실 인지 후 침입경로를 확인하기 위해 사전에 작성된 디지털증거지도 로그체인의 가장 밑 부분부터 하나씩 짚어 올라갔다.

Fig.14.의 디지털증거지도에서 확인할 수 있는 것처럼 quser.exe가 실행된 윈도우 이벤트 Session ID의 로그인 로그에서 접속한 IP를 확인했다. 이 IP는 VPN으로부터 할당 받은 IP임으로 해당 시간에 VPN 로그에서 VPN IP와 일치하는 공인 IP를 확인함으로써 공격자의 IP를 확인할 수 있었다. 또한 Fig.14.에서와 같이 디지털증거지도의 점들을 따라 선을 그어감으로써 전체적인 침입경로를 빠르게 확인할 수 있다.

VII. 결 론

최근 계속되는 침해사고는 점차 고도화된 공격기법으로 침입사실조차 인지하기 어려워졌을 뿐만 아니라 정확한 원인 파악이 되지 않아 침해사고가 끊이지 않고 있다. 이러한 침해사고에 신속하게 대응하여 피해확산을 방지하고, 정확한 침입경로를 파악함으로써 침해사고가 재발되는 것을 방지하기 위해 이 논문에서는 침해사고 발생 이전에 디지털증거지도 작성을 제안했다.

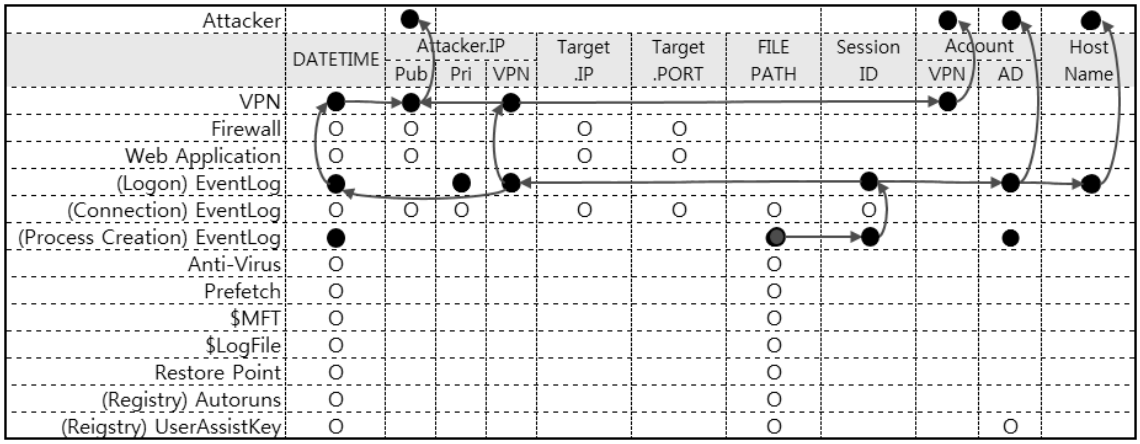

Fig.14. The Log-Chain at Digital Evidence Map

Location	Category	Sub Categories	Manager	Log Collecting		Time Sync.	Log Backup		Threat Level				Log Collection		Tool for log analysis
				%	#		Storage Date	Hashing	green	blue	yellow	red	Tool	Hashing	
Office & 3rd Party (Manager)	Windows Event	Security	* * *	100	128/128	O	365	O	O	O	O	O	Splunk		Splunk
		System	* * *	100	128/128	O	365	O	-	-	O	O	Splunk		Splunk
	Anti-Virus		* * *	100	128/128	O	365	O	O	O	O	O	SEPM		SEPM
	HDD Image		* * *				365		-	-	-	O	Write Blocker	O	EnCase
IDC (Appliance & Server)	VPN		* * *	100	1/1	O	365	O	O	O	O	O	Splunk		Splunk
	Windows Event	Security	* * *	97	338/348	O	365	O	O	O	O	O	Splunk		Splunk
		System	* * *	97	338/348	O	365	O	-	-	O	O	Splunk		Splunk
	Windows Firewall		* * *	2	8/348	O	365	O	O	O	O	O	Splunk		Splunk
	Anti-Virus		* * *	99	347/348	O	365	O	O	O	O	O	SEPM		SEPM
	O.S.	NTFS Metadata	* * *	100	1/1	O	365		-	-	O	O	IR Srcipt	O	log2timeline
		HIVE Files	* * *	100	1/1		365		-	-	O	O	IR Srcipt	O	REGA
		Physical Memory	* * *	100	1/1		365		-	-	-	O	IR Srcipt	O	Volatility
		Volatilty Data	* * *	100	1/1		365		-	-	-	O	IR Srcipt	O	Utilities
	HDD Image		* * *				365		-	-	-	O	Write Blocker	O	EnCase

Fig.15. The Life-Cycle at Digital Evidence Map

모의환경을 통해 기본적인 옵션만으로 설치된 시스템에서는 로그체인이 끊어져 있음을 확인할 수 있었고, 실제 환경에서는 지속적인 보안활동으로 로그체인을 유지할 수 있다는 것을 알 수 있었다.

디지털증거지도를 통해 그물형 로그체인을 구성하고 로그체인이 끊어지지 않도록 효과적인 관리를 수행할 수 있었다. 이로써 디지털증거지도가 작성된 이후 발생한 침해사고 대응 시에는 보다 신속하고 정확하게 침입경로를 파악하여 피해확산을 저지하고 정확한 침해원인을 밝혀 사고가 재발되는 것을 방지할 수 있다.

References

〔1〕 〔1〕 Symantec, Internet Security Threat Report, vol 18, pp. 64. 2013

〔2〕 〔2〕 Suntae Park, "2013 Major Incident Cases and Response," 17th CONCERT, pp. 14, Dec. 2013

〔3〕 〔3〕 Seungjo Baek, Jongin Lim, "A study on the Forensic Readiness as an Effective Measure for Personal Information Protection," Internet and Information Security vol. 3, no. 2, 2012

〔4〕 〔4〕 Rowlingson, R. "A Ten Setp Process for Forensic Readiness, International Journal of Digital Evidence," vol. 2, no. 3, Winter 2004.

〔5〕 〔5〕 "2013 DATA BREACH INVESTIGATIONS REPORT," Verizon, pp. 55, 2013

〔6〕 〔6〕 Bruce Schneier, "Secrets & Lies," WILEY, US, Preface, 2000

〔7〕 〔7〕 Jonghyeon Kim and 4 others "Technical Trends of Cyber Security with Big Data," ETRI Cyber Security Technology Special Issue, pp. 23, 2013

〔8〕 〔8〕 KISA, "An Incident Analysis Process Guide," pp.14, 2010

〔9〕 〔9〕 NIST, SP800-61 Revision 2, "Computer Security Incident Handling Guide," pp. 21, Aug. 2012

〔10〕 〔10〕 Kwonyeop Kim, "A study on the Windows Registry as Digital Forensic," pp. 6, Feb. 2006

〔11〕 〔11〕 Dongeun Lee, "A study on the $LOGFILE of NTFS as Digital Forensic," pp. 37, Feb. 2007

〔12〕 〔12〕 log2timeline, "Current Input Modules," http://log2timeline.net/#input, 2013

〔13〕 〔13〕 NIST, SP800-92, "Guide to Computer Security Log Management," pp. 28. Sept. 2006

〈저자소개〉

박 호 진 (HoJin Park) 정회원
2001년 10월 ~ 2005년 6월: (주)비씨큐어 암호기술연구소
2005년 06월 ~ 2012년 6월: (주)안철수연구소 A-FRIST
2012년 06월 ~ 현재: 넥슨유럽 보안팀
2006년 02월: 한국방송통신대학교 컴퓨터과학과 학사
2008년 03월 ~ 현재: 고려대학교 정보경영공학전문대학원 석사과정
〈관심분야〉 디지털 포렌식, 침해사고 대응, 정보 시각화

이 상 진 (Sangjin Lee) 종신회원
1989년 2월~1999년 2월: 한국전자통신연구원 선임 연구원
1999년 2월~2001년 8월: 고려대학교 자연과학대학 조교수
2001년 9월~현재: 고려대학교 정보경영공학전문대학원 교수
〈관심분야〉 대칭키 암호, 정보은닉이론, 컴퓨터 포렌식