ISSN : 1976-5304
https://kdfs.jams.or.kr

원격지 디지털증거 수집을 위한 프레임워크

서 강 윤*, 이 상 진**
고려대학교 정보보호대학원 (대학원생)*, (교수)**

A Framework for Acquiring Online Digital Evidence

Kangyoun Seo*, Sangjin Lee**
Graduate School of Information Security, Korea University (Graduate Student)* (Professor)**

요 약

디지털증거는 사이버 범죄뿐 아니라 사이버 환경의 사실관계 입증이 필요한 사안에 결정적인 역할을 한다. 지금까지는 디지털증거가 물리적으로 특정할 수 있는 매체에서 수집해왔으나, 클라우드와 분산 처리 기술이 발전하면서 데이터의 물리적 위치를 특정할 수 없는 경우가 있으며, 이때에는 온라인으로 원격에 있는 데이터를 수집하는 방안이 필요하다. 하지만, 기존의 원격에 있는 데이터의 수집 절차나 규칙은 무결성, 감사 추적 등 원론적인 조건을 강조할 뿐 실제 수집에 적용할 때에는 모호한 부분이 많다. 또한, 원격에 있는 데이터를 수집하며 발생하는 위변조의 위협이 고려되지 않았다. 본 논문에서는 수집 환경의 신뢰성과 네트워크의 신뢰성을 보장하여 위변조되지 않았음을 보이고 동시에 원격에 존재하는 데이터를 원본 그대로 수집하였음을 보일 수 있는 프레임워크를 제안한다. 또한, 이 프레임워크를 가상환경으로 구현하고 사례 연구를 통해 원격에 있는 디지털증거를 수집하여 효용성을 논의한다.

주제어 : 디지털증거, 디지털포렌식 조사

ABSTRACT

Digital evidence is crucial to issues that require not only cybercrime but also to prove the facts of the cyber environment. While digital evidence has traditionally been collected from physically specific media, it is often impossible to specify the physical location of the data as cloud and distributed processing technologies evolve, requiring a ways of collecting data remotely. However, existing procedures or rules for collecting data from remote locations only emphasize fundamental conditions such as integrity and audit tracking, but there are many ambiguity when applied to actual collection. In addition, the threat of falsification that occurs when collecting data from remote locations was not considered. In this paper, we propose a framework that can ensure the reliability of the collection environment and the reliability of the network to show that it has not been falsified and that data that exists remotely has been collected as it was originally. It also implements this framework as a virtual environment and collects digital evidence from remote locations through case studies to discuss its effectiveness.

Key Words : Digital Evidence, Digital Forensic Investigation

※ 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2018-0-01000, 디지털 포렌식 통합 플랫폼 개발)
- Received 07 October 2019, Revised 14 October 2019, Accepted 19 October 2019
- 제1저자(First Author) : Kangyoun Seo (Email : crmsnmercury@gmail.com)
- 교신저자(Corresponding Author) : Sangjin Lee (Email: sangjin@korea.ac.kr)

Ⅰ. 서 론

소셜 미디어의 사용 증가, 클라우드 서비스의 대중화로 인해 사회, 문화, 업무 등 일상생활이 사이버 환경에서 활발히 이루어지고 있다. 이에 따라 사실관계 입증이 필요한 분쟁, 범죄 등의 발생 시에도 사이버 공간에 존재하는 디지털증거 수집이 필요한 사례가 늘어나고 있다. 지금까지는 데이터를 가지고 있는 서버의 물리적 위치를 특정하여 해당 위치에 있는 서버에서 직접 수집해왔으나 분산 처리와 클라우드 기술의 발전으로 데이터가 존재하는 장소를 특정할 수 없는 경우가 많아 데이터의 직접 수집이 곤란한 경우가 종종 발생하고 있다. 따라서, 물리적인 위치를 파악할 필요 없이 원격에 있는 데이터를 신뢰성을 보장하며 수집하는 기술적 방안이 필요하다.

원격에 있는 디지털증거를 수집하는 행위는 유럽 연합, 미국, 중국 등 여러 국가에서 시행되고 있으며 원격 수집에 관련된 규칙과 지침을 제정하거나 모범사례집을 배포하고 있다.[3][4][5][6][7][11] 각각의 규칙이나 지침들은 감사 추적, 무결성 유지, 전문가의 참여 등을 권고하나 원론적인 수준의 제안으로 이를 기반으로 조사자가 원격에 있는 디지털증거를 수집하기에는 어려움이 있다. 따라서, 실제 조사자가 사용할 수 있는 수준의 지침이나 프레임워크가 필요하다.

원격에 있는 디지털 데이터를 수집할 때는 네트워크상이나 수집하는 기기에서 데이터가 변조될 수 있다. 가령 중간에 프록시 서버가 설치되어 있다면 중간에서 오고 가는 네트워크 패킷 데이터를 변조할 수 있고, 수집 기기에 백그라운드 프로그램이 있다면 화면에 나타나지 않으면서 저장 전후에 데이터를 변조할 수도 있다.

이러한 위변조 가능성을 불식시키기 위해서는 증거 수집 기기와 수집 프로그램의 신뢰성과 데이터가 이동하는 네트워크의 신뢰성이 보장되어야 한다. 지금까지는 네트워크나 수집하는 기기는 믿을 수 있다고 가정하고 수집하는 사람의 행위를 지켜봄으로써 원격에 있는 데이터의 수집 과정에 위변조가 없었다고 추정하였다. 하지만, 수집하는 과정을 지켜본다고 하여도 사전에 백그라운드 프로그램이 있거나 프록시 서버가 있다면 이를 탐지할 수 없다. 따라서, 원격에 있는 데이터가 있는 그대로 수집되었음을 사후에 기술적으로 감사할 수 있는 프레임워크가 필요하다.

본 논문에서는 디지털포렌식 아티팩트와 네트워크상의 신뢰할 수 있는 데이터를 활용함으로써 원격 데이터 수집 과정의 신뢰성을 확보하고 실제 조사에 활용할 수 있는 원격 디지털증거 수집 프레임워크를 제안하고자 한다.

Ⅱ. 관련 연구

1. 원격에 있는 데이터 수집 관련 규정

영국의 ACPO(The Association of Chief Police Officers)는 디지털증거에 관한 모범사례 안내[5]를 발간하여 디지털증거를 다룰 때의 원칙을 공표하였다. 원격에 있는 데이터의 수집에 관해서는 섹션 4의 6항과 Appendix A, B에서 수집할 데이터가 속한 환경을 공적인 공간과 사적인 공간으로 분리하여 수집할 때 그 과정을 기록하여 감사 추적할 수 있도록 권고하였다. 하지만, 원격에 있는 디지털증거를 수집할 때 신뢰할 수 있는 전문가의 도움을 받아 수집하도록 규정하였을 뿐 구체적인 방법은 제시하지 않았다.

유럽의 ENISA(The European Union Agency for Network and Information Security)는 전자증거-초동 대응자를 위한 기본 지침[3]을 통해 디지털증거에 관한 기본 원칙을 공표하였다. 이 원칙들은 원격에 있는 데이터를 수집할 때에도 적용된다. 수집하는 방법은 네트워크 포렌식 안내서[4]에서 제시하였다. Wireshark, tcpdump, Argus 등 소프트웨어 네트워크 패킷 수집 도구나 In-Line Network Tap, Hub, Switch 등 물리적 장비를 활용하여 케이블에서 직접 네트워크 트래픽을 수집하는 방법부터 ARP Spoofing, MAC Flooding 등의 해킹 기술을 활용하여 수집하는 방법, 무선 공유기(AP)의 모니터 모드, Promiscuous 모드 등을 활용하여 무선 네트워크 트래픽을 획득하는 방법까지 제시하였다.

중국 공안부의 증거 수집 관련 규정[6]에는 원격에서 디지털증거를 수집할 때 지켜야 할 사항과 수집해야 할 정보를 명시하였다. 원격 압수에 대해서는 제4장 네트워크 온라인 전자 데이터 추출 섹션의 제23조부터 제35조에 걸쳐 명시되어 있다. 제25조에 네트워크를 통해 원격에서 디지털증거를 수집할 때, 변경 가능성이 있어 이후에 같은 데이터를 수집하지 못하는 경우 비디오, 사진 촬영, 컴퓨터 스크린 캡처 등의 방법으로 시스템의 접근 방법, 수집 종료 시각, 사용된 도구 및 방법과 전자 데이터의 네트워크 주소, 저장 경로 또는 데이터 추출 시 절차, 무결성 검사 과정과 결과를 기록하도록 명시되어 있다. 또한, 35조에서는 프록시 서버, 전송 소프트웨어 등의 네트워크 도구를 사용한 경우 사용된 소프트웨어의 이름과 버전을 명시하도록 하였다.

Todd G. Shipley와 Roy Womack는 특허[11]와 저서[7]를 통해서 원격에 있는 데이터를 수집하고 보존하는 방법을 제시하였다. 주요 내용은 앞서 언급한 ACPO의 지침과 미국 내 관련 법을 고려하여 조사자가 지켜야 할 원칙으로 가능한 증거를 변경하지 말 것, 입증 가능한 방법으로 수집할 것, 관리의 연속성을 유지할 것

을 제시하였다. 수집 방법은 특허의 예와 그 설명에 상세히 제시되어 있다. 이 특허에서는 수집 이전에 조사할 대상의 정보를 수집하고 진정성을 위해 IP를 조회하는 것을 제시하고 있다. 수집은 도구를 통해 자동화되어 있으며 네트워크를 통해 원격에 있는 데이터와 조사 과정에서 발생한 데이터를 수집한다.

여러 국가들의 사례집과 지침들은 원격에 있는 디지털증거를 수집할 때도 기존의 디지털증거의 요건에 맞추어 수집하도록 권고한다. 하지만, 원격에 있는 데이터의 수집과 물리적으로 특정된 매체에서의 수집 간의 차이를 고려하여 이때 발생하는 위변조의 위협을 불식시키는 방안은 제시하고 있지 않다. 즉, 조사자가 이 지침과 규정을 따라 수집하고자 해도 신뢰할 수 있는 수집 시스템이 없어 전문가의 도움을 받고 수집 과정을 촬영하는 방법 등으로 수집하고 있으나, 악의적인 수집자가 전문가를 속이고 백그라운드 프로그램을 동작시켜 수집된 데이터를 조작할 수 있는 바 이러한 행위가 발생하지 않았음을 기술적으로 입증할 수 있는 방안이 필요하다.

2. 디지털포렌식 조사 프로세스

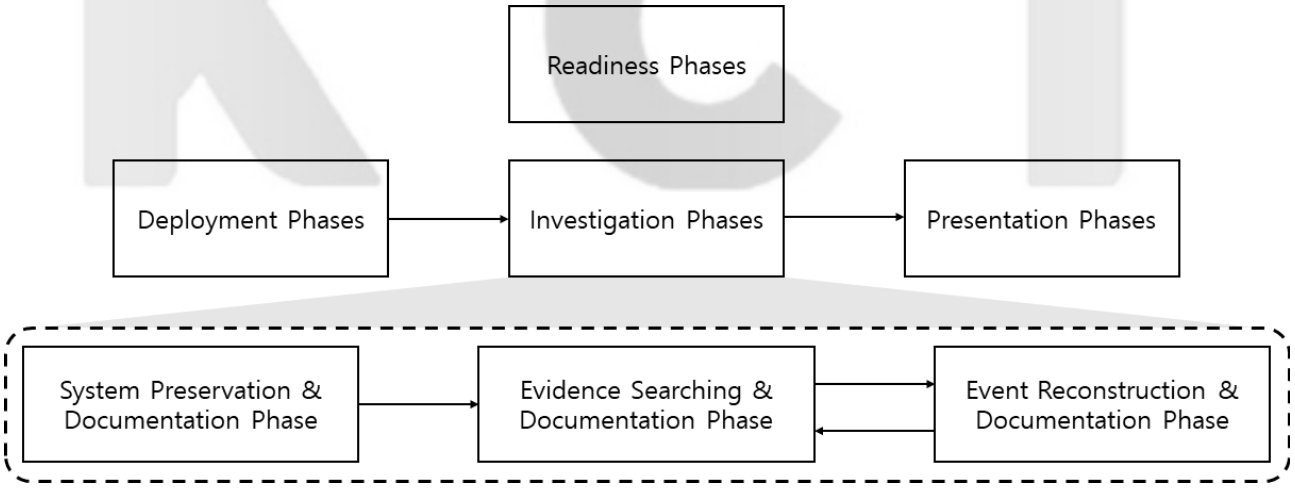
현재의 디지털포렌식 조사 프로세스는 Brian Carrier와 Eugene Spafford가 제안한 모델[1]을 따르고 있으며 전체적인 구성은 [그림 1]과 같다. 이 모델은 물리적 범죄 현장의 조사 모델을 디지털 환경에 차용한 것으로 Mark Reith 등이 제안한 모델[2]과 유사하다.

준비(Readiness) 단계는 전문가 교육, 도구 테스팅 등을 통해 조사에 사용될 인적, 물적 자원을 준비하는 단계이다. 이 단계에서 상황에 따른 적절한 조사 방법, 요구사항, 절차를 수립한다.

전개(Deployment) 단계에서는 조사가 필요함을 인지하고 준비된 절차에 따라 영장이나 소유자의 동의 등을 통해 접근 권한을 확보한다.

앞의 단계에서 대상에 대한 식별과 적절한 조사 절차가 도출되고 접근 권한이 확보되면 실질적인 조사(Investigation)를 시작한다. 이 단계는 다시 보존, 탐색, 재구성의 단계로 구성되어 시스템을 그대로 복사하고 이미지 파일 형태로 데이터를 보존한 뒤 보존된 시스템의 원본 이미지를 탐색하여 사건과 연관된 데이터를 수집하고 가시화하여 재구성한다.

조사가 종료되면 조사된 결과를 적당한 절차에 따라 감사인이나 법원에 제출(Presentation)한다.



[그림 1] 디지털포렌식 조사 프로세스
[Fig. 1] Digital Forensic Investigation Process

이 프로세스는 디지털포렌식 조사의 추상화된 모델로 특정한 조사 상황에 맞추어 각 단계를 구현하여 사용한다. 원격에 있는 디지털증거를 수집할 때는 특히 데이터 수집 과정의 경로상에서 위변조되지 않고 있는 그대로 수집되었음을 보장할 수 있는 시스템을 사용하여 조사 단계를 이행해야 한다.

Ⅲ. 원격에 있는 디지털증거 수집에 대한 요건

원격에 있는 데이터를 수집하여 디지털증거로써 활용할 때는 특별히 주의해야 하는 사항이 있다. 이는 원격으로 디지털증거를 수집하는 환경에 기인한다. 원격에 있는 데이터를 수집할 때에는 〔그림 2〕와 같이 네트워크를 통해 데이터를 받는다. 조사자는 수집 시스템 위에서 동작하는 수집 도구를 이용해 원격에 있는 데이터를 네트워크로 수신받는다. 하지만, 수집 시스템을 제외한 네트워크와 수집 대상은 수집 행위자가 통제할 수 없다. 따라서, 수집 중 외부의 영향을 받아 수집되는 데이터가 위변조될 수 있다. 또한, 수집 시스템에서 발생한 이벤트를 기록하는 장치가 없다면 네트워크로 받은 데이터가 받은 그대로 수집 시스템에 저장되었는지 증명할 수 없다. 그러므로 수집 행위가 일어난 네트워크와 수집 시스템의 신뢰성을 각각 보여야 한다. 수집 시스템을 신뢰할 수 있다면 수집 행위가 일어나는 기기에서 수집되고 저장되는 데이터의 무결성을 보장할 수 있으며, 수집 행위가 일어나는 네트워크를 신뢰할 수 있다면 네트워크로부터 수신받아 수집된 데이터가 원본임을 보장할 수 있다.

따라서, 본 절에서는 수집 시스템과 네트워크라는 두 가지 측면에서 신뢰성의 요건을 충족하는 방법을 제시한다.

〔그림 2〕 원격지의 데이터 수집
〔Fig. 2〕 Data Acquisition Through Network

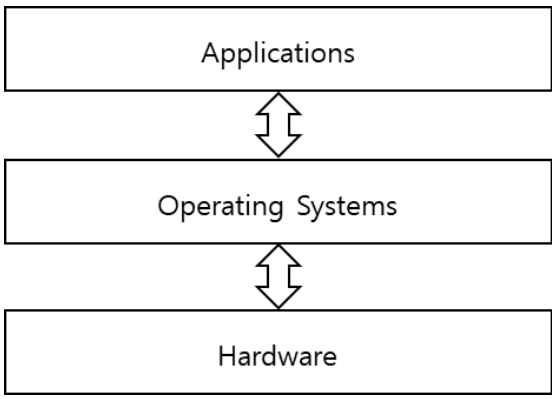
1. 수집 시스템의 신뢰성

수집 시스템에서 수집된 데이터가 수집 기기와 도구에 의해 위변조되지 않고 원본 데이터와 동일함을 보장해야 한다. 컴퓨터 시스템은 〔그림 3〕과 같이 계층적으로 표현할 수 있으며, 각 계층의 요소에 대한 신뢰성을 보여 전체의 신뢰성을 보일 수 있다. 다만, 하드웨어는 그 자체를 신뢰할 수 있다고 가정한다. 본 논문은 운영체제와 응용프로그램 레벨에서 신뢰성을 보장하기 위해 수신받는 패킷을 보존하고, 사용자의 행위를 기록함으로써 사후에 데이터의 위변조를 감사할 수 있도록 할 것을 제안한다. 이런 시스템에서는 보존된 패킷과 최종 수집된 데이터의 일치성을 조사함으로써 데이터의 위변조 여부를 감사할 수 있다. 또한, 사용자가 어떠한 행위를 했는지 기록함으로써 후술하는 네트워크의 신뢰성을 보장할 수 있다. 사후 감사를 위해 수집에 사용했던 시스템을 그대로 보존하는 절차가 필요하다.

먼저, 행위 기록 시스템을 설치하여 수집 환경의 파일 입출력, 네트워크 통신 기록, 커널의 시스템 호출 기록 등을 기록해야 한다. 사후에 이 로그를 근거로 네트워크를 통해 받은 데이터가 저장 매체에 저장되는 과정에 백그라운드 프로그램 등이 동작하지 않았고 그 데이터가 그대로 저장되었음을 입증할 수 있다.

백그라운드 프로그램이 없음을 입증하였기 때문에 이제 수집 도구 그 자체에서 데이터 위변조 여부를 확인해야 한다. 이를 위해서는 보존된 패킷과 수집된 데이터를 비교하면 된다. 또한, 수집 과정에서 발생한 네트워크 패킷에는 콘텐츠를 구성하는 데이터 외에도 재전송 기록, 실제 데이터가 송수신된 IP 주소, 공개키 기반 인증 정보 등 수집 시스템에서 수정할 수 없는 데이터가 포함되어있어 데이터가 위변조되었을 때 이 데이터를 확인하여 탐지할 수도 있다. 네트워크 패킷 수집의 위치를 수집 시스템이 아닌 수집 시스템에서 구동되는 프로그램이 관여할 수 없는 네트워크 인터페이스의 입출력 단에 설치하여 데이터의 독립성을 보장할 수 있다.

최종적으로 수집된 데이터가 위변조 되지 않았음을 보장하기 위해서는 수집시스템 그 자체를 보존하여 사후 감사할 수 있는 체계가 필요하다.



〔그림 3〕 컴퓨터 시스템의 구성
〔Fig. 3〕 Computer-System Structure

2. 네트워크의 신뢰성

수집 시스템과 도구의 신뢰성이 보장되더라도 수집에 사용되는 네트워크를 신뢰할 수 없다면 수집된 데이터가 위변조되지 않았다는 사실을 입증할 수 없다. 그러므로 수집 행위자는 원본 데이터를 보관하고 있는 원격지로부터 수집 환경이 구성된 단말로 데이터가 수신되는 과정에서 위변조가 없었음을 보여야 한다. 하지만, 현재의 인터넷과 같이 수많은 소규모 네트워크들이 연결되는 형태의 네트워크는 전체에 대해 신뢰성을 보이기 매우 어렵다. 따라서, 본 논문에서는 수집 대상과 수집 환경 사이의 일부 네트워크에 대한 신뢰성을 보이는 방안을 제안한다.

네트워크를 통해 접속하여 수신되는 데이터는 송신 측 원본 매체의 변경에 따라 달라지며 이는 예측할 수 없다. 즉, 수집하는 시점에 따라 다른 결과가 나타날 수 있다. 수집 시점을 특정하지 않고 수집하는 경우 수집 이후 수집 대상이 삭제 또는 변경되어 수집된 데이터를 부인할 수 있다. 따라서, 수집하는 시점에 수집 대상이 존재했음을 보이기 위해 수집 시점을 특정해야 한다.

수집 시점을 특정한 다음 수집 대상에 대한 접속의 신뢰성을 보여야 한다. 접속의 신뢰성이 보장되면 현재 접속하려는 대상이 사전에 접속하려고 의도한 그 대상이라는 것을 보임으로써 네트워크로 연결된 대상의 신원이 보장되어 중간자, 탬퍼링, 스푸핑 공격 등으로 접속하는 대상이 위장/변조되지 않았음이 보장된다.

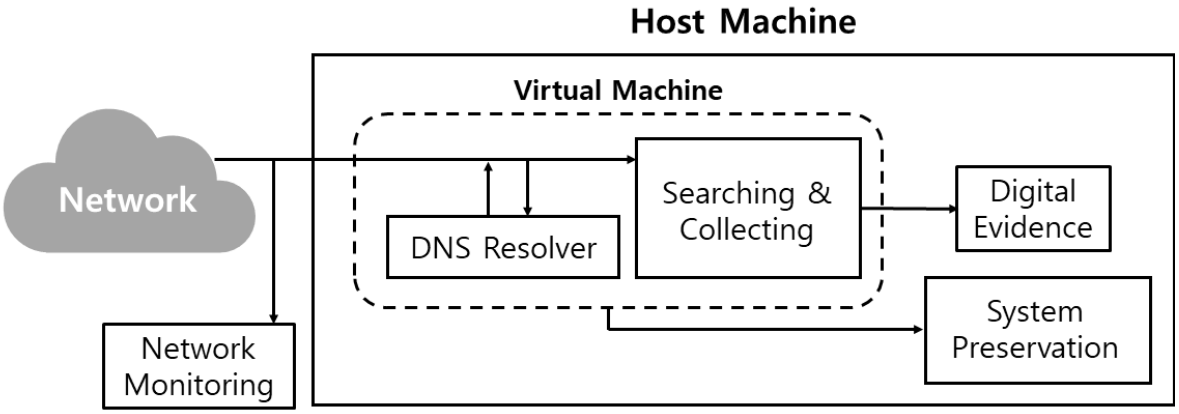
이렇게 수집 시점을 특정하고 수집 대상에 대한 접속의 신뢰성이 확보되면 수집 대상과 수집 시스템 간에 주고받는 데이터가 특정한 시점에 그 대상이 보냈다는 것을 신뢰할 수 있다.

Ⅳ. 원격 데이터 수집 프레임워크

디지털포렌식 조사는 〔그림 1〕과 같이 준비, 전개, 조사, 제출의 단계로 구성된다. 원격에 있는 디지털증거에 대한 디지털포렌식 조사에도 이 프로세스가 적용된다. 본 논문에서 제안하는 프레임워크는 조사 단계에서 원격에 있는 데이터를 수집하고 보존하는 방안을 제시한다. 전체적인 프레임워크는 〔그림 4〕과 같이 구성하였다. 앞에서 언급한 수집 시스템의 신뢰성 확보를 위해 수집 시스템을 보존하는 방법으로 수집 시스템이 구축된 디스크를 보존할 수도 있다. 하지만, 이 방법을 적용하면 수집할 때마다 수집 시스템을 구축하고 이를 검증해야 하며 수집 이후 디스크 자체를 보존해야 해 비효율적이다. 따라서, 제안하는 프레임워크는 가상환경 안에 수집 시스템을 독립적으로 구축하여 재사용 가능하며 수집 이후에는 하나의 파일로 보존하는 방법을 제안한다.

이 프레임워크를 사용한 수집 절차는 가상환경 안에 수집 시스템 구축, 수집 대상의 메타데이터 수집 및 검증, 수집, 보존으로 진행된다. 수집 이후에는 디지털증거를 수집한 가상환경 안의 시스템을 보존하여 신뢰성을 보여야 하며 디지털증거에 대한 분석이 필요한 경우 이 데이터를 활용할 수도 있다. 또한, 앞서 조사한 각국의 원격 디지털증거 수집 시의 지침과 모범사례에서 권고하는 무결성, 감사 추적의 요건도 충족할 수 있다.

앞서 서술한 요건들을 지킬 수 있도록 구축된 원격 디지털증거 수집 시스템이 있다면 전문가의 참여가 수집이 아닌 분석과 감사 추적에만 필요하게 되어 신속한 증거 확보가 필요한 원격 디지털증거 수집 상황에 효과적으로 대처할 수 있다.



〔그림 4〕 원격에 있는 디지털증거 수집 프레임워크
〔Fig. 4〕 Online Digital Evidence Acquisition Framework

1. 신뢰성이 보장되는 수집 시스템 구성 방안

앞에서 언급한 것처럼 수집 시스템은 사후 감사를 위해 보존되어야 한다. 이를 위해 본 논문이 제안하는 프레임워크에서는 가상 머신을 사용한다. 또한, 수집 과정의 신뢰성 담보를 위해 수집 시스템에서 실행된 프로세스의 기록, 네트워크 연결 기록, 파일 입출력이 기록되는 시스템과 네트워크 모니터링 환경을 구축한다. 이때, 행위 기록 시스템을 제일 먼저 설치하고 구동시켜 네트워크 모니터링 환경 구축 과정도 검증할 수 있도록 한다.

먼저, 운영체제 구동을 위한 시스템 파일과 편의를 위해 운영체제와 함께 설치되는 응용프로그램만이 존재하는 가상환경을 구축한다. 이때 설치하는 운영체제 또한 사후에 검증할 수 있도록 설치에 사용된 파일과 버전 정보, 설치 파일의 해시값을 기록하여 보관한다. 가상환경을 구축하여 그 안에서 수집하면 수집 환경 자체를 분리하여 보관할 수 있다. 그렇게 하면 긴 시간을 두고 수집하는 경우에도 수집 환경을 잠시 멈추었다가 수집이 필요한 시점에 구동하여 수집할 수 있다. 또한, 가상환경 구축의 특성상 한 번만 구축해놓으면 언제든 수집이 필요한 시점에 구동하여 긴급히 수집해야 할 경우에 대응할 수 있다.

다음으로, 행위 기록 시스템을 구축한다. 이로써 앞으로 발생하는 프로세스 실행, 프로그램 설치, 파일의 입출력, 네트워크 접속 등의 일련의 활동을 재현할 수 있게 됨으로써 사후 사용자의 행위를 검증할 수 있다.

마지막으로, 수집된 증거의 신뢰성을 보이고 수집 대상과의 네트워크 접속의 신뢰성을 보이기 위해 네트워크 트래픽 모니터링 및 수집 도구를 설치한다. 패킷 수집 도구는 수집 시스템의 영향을 받지 않은 네트워크를 통해 수신된 그대로를 수집하기위해 네트워크 인터페이스에서 직접 수집한다. 이를 위해 수집 상황과 환경을 고려하여 네트워크 탭(TAP:Terminal Access Port), 허브, 스위치 등의 물리적인 장치를 활용한다. 다만, 후술할 신뢰성 확보 방안에서 암호화 통신을 사용하므로 사후에 이를 복호화하기 위해 세션 키 또한 수집한다.

2. 수집 대상과 연결된 네트워크의 신뢰성 확보 방안

원격 디지털증거의 특성상 준비 단계에서 수집 대상에 접근하기 위해 확보한 IP, 도메인 네임, 인증 정보 등이 변경되었거나 위장/위변조되었을 여지가 있다. 따라서, 수집할 데이터의 신뢰성을 위해 수집단계에서 이를 다시 한번 확인한다. 이후, 시점을 특정하고 수집 대상과 수집 시스템 간 네트워크의 신뢰성을 보일 수 있다.

먼저, 수집 시점을 특정하는 방법으로 네트워크상의 시간 동기화를 위한 공용 NTP 서버들과 수집 환경 간의 시간 차이를 조회하는 방법을 제안한다. NTP는 〔그림 5〕와 같이 계층적으로 구성되어 상위 계층으로부터 시간을 동기화한다. NTP의 최상위 계층(Primary Refeence Clock)에는 세슘원자시계 수준의 정확도를 보장하며 시간을 하위 계층에 제공한다. 공공 NTP 서버는 높은 보안 수준을 갖추고 있어 특정 개인이나 기관이 악의적으로 위변조하기 매우 힘들고 설령 하나의 공공 NTP 서버가 위변조되었다고 하더라도 올바르게 동작하는 다른 공공 NTP 서버들에서 조회한 정보와 교차 검증하여 이를 바로 잡을 수 있다. 또한, 암호화 통신하에서 공개키 암호기반 인증서를 교환하여 서버의 신원이 인증되므로 NTP 서버로 위장하여 조작된 시간을 보내는 것은 불가능(infeasible)하다. 따라서, 프레임워크에서는 공공 NTP 서버에 접속하여 암호화 프로토콜의 Hand-Shake 단계에서 교환하는 인증서를 확인하여 신뢰할 수 있는 기관임을 검증한다. 이 방법의 안전성은 전자 서명과 인증서에 적용된 암호 알고리즘의 안전성에 의존한다. 이로써 신뢰할 수 있는 시간 조회 수단을 얻게 되며 이 서버와 수집 환경의 시간 차이를 조회함으로써 수집의 시점을 특정한다.

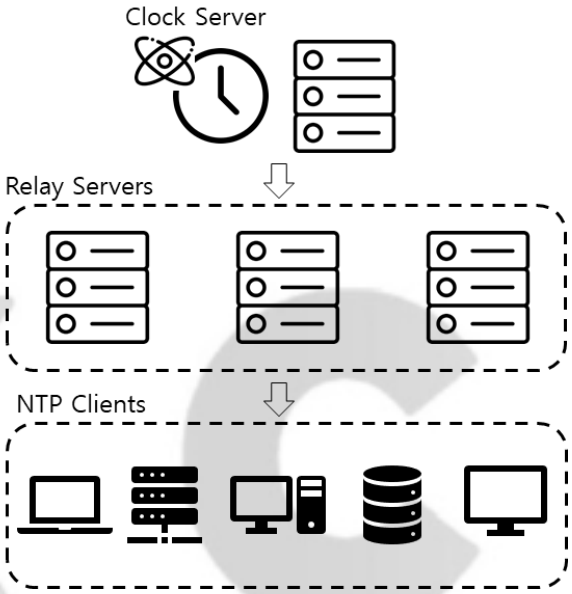
다음으로, 수집 대상과 수집 환경 간 네트워크 통신의 신뢰성을 확보한다. 이 방법으로는 DNS 보안 표준 중 DNSSEC〔16〕, DNS-over-TLS〔17〕를 활용한다.

DNSSEC(Domain Name System Security Extension)〔16〕은 2005년에 공표된 국제표준기술로 기

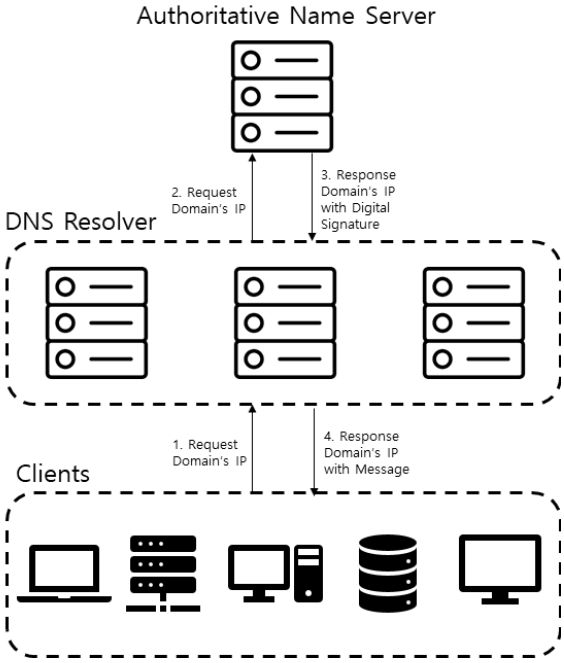
존의 DNS에 전자 서명 기능을 추가함으로써 DNS를 위변조하는 공격을 방지하는 표준이다. DNSSEC이 적용된 시스템은 〔그림 6〕과 같이 도메인에 대한 IP 질의에 대해 도메인 서버는 IP 주소와 전자 서명을 응답하고 이를 DNS 서버가 인증하여 IP 주소가 변조되지 않았음을 보여 DNS에서 응답한 데이터의 무결성을 보장한다. 하지만, DNSSEC 시스템은 사용자 기기, DNS 서버(DNS Resolver), 도메인 서버(Authoritative Name Server) 모두가 DNSSEC을 지원해야 한다. 따라서, 제안하는 프레임워크는 DNSSEC과 함께 DNS-over-TLS로 이를 보완한다.

DNS-over-TLS〔17〕는 2016년 공표된 국제표준기술로 이 기술이 적용된 시스템에서는 DNS 질의/응답 과정에 TLS 통신의 암호화 기술을 적용하여 네트워크상의 도청(eavesdropping)과 경로 변조 위협을 방지한다. 이를 통해, DNS 질의응답 과정에 송수신되는 데이터의 무결성과 기밀성이 보장된다. 다만, 이 기술 또한 DNS 서버(DNS Resolver)가 이를 지원해야 향상된 보안성이 적용된다.

공공 NTP에서 제공한 데이터로 수집 시점을 특정하고 DNSSEC, DoT(DNS-over-TLS)를 통해 받은 데이터로 수집 대상을 특정함으로써 수집 대상과 수집 시스템 간 네트워크의 신뢰성이 확보된다. 네트워크의 신뢰성 확보를 위한 모든 단계는 앞서 구축된 수집 시스템에서 이루어져 모든 접속과 설치의 기록이 남아 사후에 검증할 수 있다.



〔그림 5〕 NTP의 시간 동기화 구조
〔Fig. 5〕 NTP Mechanism



〔그림 6〕 DNSSEC의 IP 질의응답 구조
〔Fig. 6〕 DNSSEC IP Querying Mechanism

3. 구현

앞서 서술한 신뢰성을 보장하는 수집 시스템 구축 방안과 네트워크의 신뢰성을 확보하는 방안을 토대로 실제 원격 디지털증거 조사에 활용할 수 있는 수집 시스템을 구현하였다.

먼저, 수집 시스템을 구현하기 위해 호스트와 독립적으로 분리되어 하드웨어 레벨의 가상화를 지원하는 VMware에 소스코드 및 설치 파일이 공개되어 검증할 수 있는 운영체제인 우분투를 설치하여 가상환경을 구축하였다. 구축된 가상환경에 리눅스 감사 시스템(Linux Audit System)을 설치하여 리눅스 커널 기반 운영체제의 파일 입출력, 시스템 콜 호출, 프로세스 실행 등의 행위를 기록할 수 있도록 하였다. 리눅스 감사 시스템은 사용자가 기록할 행위에 대한 규칙〔15〕을 정하여 데몬의 형태로 백그라운드에서 항상 동작하도록 설정하였다. 수집 및 신뢰성 확보를 위한 도구 설치 이전에 리눅스 감사 시스템을 설치하고 동작시켜 설치 과정과 도구의 동작을 모두 기록할 수 있도록 하였다. 네트워크 트래픽은 수집 환경과 독립된 네트워크 인터페이스 하드웨어에 TAP(Test Access Point)을 설치하여 수집하도록 제안하였지만, 실험실 환경의 한계로 패킷 수집 소프트웨어 WireShark〔18〕를 설치하였다. 이로써, 리눅스 감사 시스템 규칙으로 정한 모든 행위를 기록하고 발생하는 네트워크 트래픽을 수집하는 수집 시스템을 구축하였다.

다음으로 네트워크의 신뢰성을 확보하기 위해 수집 시점을 조회할 수 있는 ntpq를 설치하고 신뢰할 수 있는 공공 NTP〔19〕에서 시간을 조회하도록 "/etc/ntp.conf" 파일을 수정하였다. 수집 대상과 수집 시스템 사이 네트워크의 신뢰성을 보이기 위해 DNSSEC이 적용된 환경에서 수집 대상의 IP 주소를 조회해야 한다. 이를 위해, 로컬 환경에서 데몬 프로그램으로써 DoT(DNS-over-TLS)를 이용하여 DNS stub resolver의 역할을 하는 응용프로그램인 Stubby〔20〕를 설치하였다. 이로써, 수집 시스템에 DoT를 지원하는 DNS Resolver 역할을 하는 프로그램을 통해 DNS를 질의하고 응답받는 환경이 구축되었다.

선별 및 수집을 위한 인터넷 브라우저는 개발자 버전 파이어폭스 브라우저〔21〕를 사용하였다. 우분투 운영체제에 기본으로 탑재된 일반 파이어폭스 브라우저 대신 개발자 버전을 사용하여 암호화 통신의 세션 키를 확보하고 웹 디버거, 프록시 등을 활용한 외부 수집 도구로의 확장성을 고려하였다. 또한, 브라우저를 통한 모든 DNS 질의가 사전에 정한 DNS Resolver로부터 암호화 통신 환경에서 이루어지도록 브라우저의 설정값〔22〕을 변경하였다.

환경 구축에 사용된 운영체제와 소프트웨어의 정보는 〔표 1〕과 같으며 VMware를 제외한 모든 프로그램은 오픈소스 프로그램으로 GitHub〔23〕에 공개돼 있어 누구나 코드 레벨의 기능을 검증할 수 있다.

〔표 1〕 수집 시스템 구현에 사용된 운영체제 및 소프트웨어 정보
〔Table 1〕 Software Info.

이름	설명	버전
VMware	가상머신 관리 도구	12.1.0 build-3272444
Ubuntu	운영체제	18.04.3
Mozilla Firefox Developer Edition	인터넷 브라우저	70.0b10
WireShark	네트워크 패킷 모니터링 도구	2.6.10
Linux Audit System	감사 추적 도구	2.8.2
ntpq	NTP 서버 조회 도구	4.2.8p10
Stubby	DNSSEC, DoT 환경 구축 도구	0.2.2
net-tools	네트워크 관리 도구	2.10-alpha

4. 사례 연구

원격에 있는 디지털증거를 수집하는 경우는 누구에게나 공개된 대상을 수집하는 경우와 대상에 접속하기 위해 인증 과정이 필요한 사적인 대상을 수집하는 경우로 나눌 수 있다. 이 두 경우의 대표적인 예로 구글 뉴스와 구글 이메일이 있다. 따라서, 사례 연구로써 프레임워크의 구현체로 구글 뉴스와 구글 이메일을 수집하여 프레임워크의 효용성을 평가하고자 한다. 각 사례의 수집은 다음의 절차로 수행되었다.

1. auditd와 stubbt의 서비스 상태를 확인한다.
2. 공공 NTP 서버를 조회하여 수집 개시 시점을 특정한다.
   Example Command : ntpq -p
3. WireShark를 구동하고 네트워크 모니터링을 시작한다.
4. 세션 키를 기록하도록 브라우저를 구동한다
   Example Command : SSLKEYLOGFILE=〔Path〕 firefox - no-remote -profile
5. 수집 대상에 접속한다.
   수집 대상의 도메인 네임 : news.google.com, mail.google.com
6. 데이터를 수집한다.
   수집 방법 : 웹 페이지 리소스 저장, 이메일 파일 저장
7. 공공 NTP 서버를 조회하여 수집 종료 시점을 특정한다.
8. 수집한 파일, 패킷과 Linux Audit System의 로그를 저장하고 수집 시스템을 보존한다.
9. 보존한 수집 시스템의 해시값을 계산하여 수집 시스템의 무결성을 유지한다.

구글 뉴스의 수집 결과로 〔표 2〕와 같이 웹 페이지를 구성하는 리소스, 패킷과 복호화를 위한 세션 키, 리눅스 감사 시스템이 생성한 로그가 생성되었고, 구글 이메일의 수집 결과물로는 〔표 3〕과 같이 이메일 파일(.eml), 발생한 네트워크 패킷과 패킷 복호화를 위한 세션 키, 리눅스 감사 시스템이 생성한 로그가 생성되었으며 모든 파일은 보존된 수집 시스템에 저장되어 있다.

리눅스 감사 시스템이 생성한 로그에는 수집 이전에 프레임워크를 구성하는 프로그램의 설치부터 동작 과정 중 발생한 행위까지 모두 기록되었으며 로그 분석 프로그램〔24〕으로 분석하여 수집 과정 중 수집 시스템에서 위변조 행위가 없었음을 보일 수 있다. 또한, 수집된 패킷을 분석하여 수집 시스템이 접속한 대상의 IP와 Hand-Shake 과정에서 받은 전자 서명, 인증서를 확인하여 데이터를 송신한 대상을 식별할 수 있으며, 디지털증거로 활용될 웹 페이지 리소스 파일들과 이메일 파일이 그 대상으로부터 수신된 패킷에서 생성되었음을 보일 수 있었다.

수집 시스템은 가상머신 디스크 파일(.vmdk)로 보존되어 WinHex〔26〕나 EnCase Forensic〔27〕와 같은 분석 도구로 사후에 검증과 감사의 용도로 사용할 수 있다. 그 크기는 6.6GB로 사후 검증과 감사의 용도로 디지털증거와 함께 보관하기에 무리가 없는 크기이다.

〔표 2〕 구글 뉴스 게시물 수집 결과물
〔Table 2〕 Result of Acquiring Google News Contents

파일	설명	크기
Google News.html	웹 페이지 HTML 파일	2.5 MB
Google News_files/*	웹 페이지를 구성하는 리소스 파일들	1.8 MB
audit.log	Linux Audit System이 기록한 행위 기록 파일	3.2 MB
SESSION_KEY	브라우저로부터 추출한 암호화 프로토콜의 세션 키	40.4 KB
Google News Packet.pcapng	수집 과정에서 발생한 패킷	14.3 MB

〔표 3〕 구글 메일 수집 결과물
〔Table 3〕 Result of Acquiring G-Mail Contents

파일	설명	크기
Mail Box.eml	웹 페이지 HTML 파일	141.9 KB
audit.log	Linux Audit System이 기록한 행위 기록 파일	3.2 MB
SESSION_KEY	브라우저로부터 추출한 암호화 프로토콜의 세션 키	48.4 KB
Google Mail Packet.pcapng	수집 과정에서 발생한 패킷	8.3 MB

Ⅴ. 논의

제안한 원격 디지털증거 수집 프레임워크의 구현체로 조사자는 원격에 있는 데이터를 수집 시스템의 신뢰성과 네트워크의 신뢰성을 보장하며 수집할 수 있다. 즉, 수집된 결과물에 포함돼 있는 프로세스 실행 기록, 파일 입출력 기록 등의 행위 기록을 근거로 수집 시스템에서 위변조되지 않았음이 보일 수 있고, 특정된 시간에 전자서명과 인증서로 식별된 수집 대상으로부터 수신된 패킷의 컨텐츠 데이터와 일치함을 보여 네트워크를 거치며 수집 데이터가 위변조되지 않았음을 보일 수 있다. 또한, 수집 행위가 일어난 시스템을 하나의 파일로 보관하여 수집 이후에 검증할 수 있다. 따라서, 제안한 원격 디지털증거 수집 프레임워크를 활용하여 수집하는 것이 기존의 원격 디지털증거 수집 방안들〔3〕〔5〕〔6〕〔7〕보다 더 구체적이며 디지털포렌식 관점에서 건전한 방법이다.

본 논문에서는 원격에 있는 디지털증거를 디지털포렌식 관점에서 건전하게 수집하기 위해 커널 수준의 행위 기록 시스템 도입, 패킷 분석, DNSSEC 환경 구축과 같이 여러 분야의 기술을 복합적으로 활용하였다. 이러한 접근 방식은 수집 시스템이 원격에 있는 대상에 네트워크로 연결되어 데이터를 송수신하는 환경에서 필수적이며 원격 디지털증거 수집에 유의미하게 활용할 수 있을 것이다.

Ⅵ. 결 론

본 논문은 원격에 있는 데이터를 수집할 때 수집 환경과 네트워크의 신뢰성을 보장하는 방법론과 이를 적용한 프레임워크를 제시하였다. 또한, 이 프레임워크를 구현하고 사례 연구에 적용하여 원격에 있는 디지털증거를 수집 시스템과 네트워크 모두 신뢰성이 보장된 환경에서 수집할 수 있으며 모든 과정이 사후 검증 가능함을 보였다.

이 방법론을 적용하여 구축된 수집 시스템으로 누구나 원격에 있는 데이터를 신뢰할 수 있게 수집할 수 있다. 즉, 기존의 전문가의 도움을 받아 해결했던 수집 과정의 신뢰성 보장을 기술적으로 해결할 수 있다. 덧붙여, 원격에 있는 데이터는 조사자의 통제 밖에 있어 긴급한 수집이 요구되는데 이런 수집 환경이 미리 갖추어져 있다면 신속한 대응도 가능하다.

제시한 프레임워크는 수집 환경을 가상화하여 동시에 여러 대상에 대한 수집과 반복 수집이 가능하다. 또한, 이 프레임워크는 디지털포렌식 조사 프로세스 모델의 일부로써 이용되도록 고안되어 기존의 디지털포렌식 조사 솔루션 또는 도구에 하나의 모듈로써 내포시킬 수 있을 것이다.

참 고 문 헌 (References)

〔1〕 Brian Carrier, Eugene Spafford, An Event-Based Digital Forensic Investigation Framework, DFRWS USA. 2004.

〔2〕 Mark Reith, Clint Carr and Gregg Gunsch, An Examination of Digital Forensic Models. International Journal of Digital Evidence. Vol 1, Number 3, Fall 2002.

〔3〕 ENISA(The European Union Agency for Network and Information Security), Electronic evidence - a basic guide for First Responders, 2014.

〔4〕 ENISA(The European Union Agency for Network and Information Security), Introduction to Network Forensics, 2019.

〔5〕 ACPO(The Association of Chief Police Officers), Good Practice Guide for Digital Evidence, 2012.

〔6〕 Ministry of Public Security of the People's Republic of China, 公安机關辦理刑事案件電子數据取証規則, 2019.

〔7〕 Todd G. Shipley와 Roy Womack, Internet Investigation, Syngress. 2014.

〔8〕 Humaira Arshad, Aman Jantan and Esther Omolara, Evidence collection and forensics on social networks: Research challenges and directions. Digital Investigation. Vol 28. pp 126-138, 2019

〔9〕 About ICANN, 〔Online〕 Available at:https://www.icann.org/resources/pages/what-2012-02-25-en

〔10〕 Google, Introducing Google Public DNS, 〔Online〕 Available at:https://googleblog.blogspot.com/2009/12/introducing-google-public-dns.html

〔11〕 Roy Womack, Todd G. Shipley. Online Evicence Collection, US 8,417,776 B2. 2013

〔12〕 "NTP FAQ". The NTP Project.

〔13〕 M.D.Kohn, M.M.Eloff, J.H.P.Eloff, Integrated digital forensic process model, Computer & Security. Vol 38. pp 103-115, 2013.

〔14〕 Ubuntu, 〔Online〕 Available at : https://ubuntu.com/download/desktop

〔15〕 Audit Rules, 〔Online〕 Available at : https://github.com/linux-audit/audit-userspace/tree/master/rules

〔16〕 D. Eastlake "Domain Name System Security Extensions", RFC2535 March 1999.

〔17〕 D. Wessels, and P. Hoffman. "Specification for DNS over Transport Layer Security (TLS)". RFC 7858, May 2016.

〔18〕 WireShark, 〔Online〕 Available at : https://www.wireshark.org/about.html

〔19〕 Public NTP List, 〔Online〕 Available at : http://support.ntp.org/bin/view/Servers/StratumOneTimeServers

〔20〕 Stubby, [Online] Available at : https://dnsprivacy.org/wiki/display/DP/DNS+Privacy+Daemon+-+Stubby

〔21〕 Firefox Developer Edition, [Online] Available at : https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Developer_Edition

〔22〕 Firefox Trusted Recursive Resolver, [Online] Available at : https://wiki.mozilla.org/Trusted_Recursive_Resolver

〔23〕 GitHub, [Online] Available at : https://en.wikipedia.org/wiki/GitHub

[24] aureport Manual, [Online] Available at : https://linux.die.net/man/8/aureport

[25] Virtual Disk Format 5.0, [Online] Available at : https://www.vmware.com/support/developer/vddk/vmdk_50_technote.pdf

[26] WinHex, [Online] Available at : https://www.x-ways.net/winhex/

[27] EnCase Forensic, [Online] Available at : https://www.guidancesoftware.com/encase-forensic

저 자 소 개

서 강 윤 (KangYoun Seo)
학생회원
2018년 2월 : 고려대학교 정보대학 컴퓨터학과 졸업
2018년 3월 ~ 현재 : 고려대학교 정보보호대학원 정보보호학과 석사과정
관심분야 : 디지털 포렌식, 정보보호 등

이 상 진 (Sangjin Lee)

1989년 10월 ~ 1999년 2월 : ETRI 선임 연구원
1999년 3월 ~ 2001년 8월 : 고려대학교 자연과학대학 조교수
2001년 9월 ~ 현재 : 고려대학교 정보보호대학원 교수
2008년 3월 ~ 현재 : 고려대학교 디지털포렌식연구센터 센터장
관심분야 : 디지털 포렌식, 심층암호, 해시함수