Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.33, NO.5, Oct. 2023 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2023.33.5.737

구글 아카이빙 데이터 기반 멀티 행위 분석*

김 예 은,[1†] 홍 사 라,[1] 김 성 민[2‡]
[1,2]성신여자대학교 (대학원생, 교수)

Multi-Behavior Analysis Based on Google Archiving Data*

Yeeun Kim,[1†] Sara Hong,[1] Seongmin Kim[2‡]
[1,2]Sungshin Women's University (Graduate Student, Professor)

요 약

기업 및 개인의 데이터가 온프레미스 환경에서 클라우드로 이동하면서 클라우드 포렌식 중요성이 증가하고 있다. 클라우드 데이터는 모바일 기기부터 데스크톱 등 여러 장치에 저장될 수 있으며, 연동된 계정과 클라우드 서비스로부터 생성되는 정보 등 다양한 행위 아티팩트가 존재한다. 그러나 데이터의 분산 저장과 아티팩트 연계성 부족 등 클라우드의 환경적 제약으로 인해 디지털 증거를 확보하고 분석하는 것에 한계점이 있다. 이를 해결할 수 있는 수단 중 하나로 아카이빙 서비스가 있으며, 대표적으로 구글의 Takeout이 있다.

본 논문에서는 아카이빙 데이터 기반의 클라우드 포렌식을 위해 사용자 행위 데이터를 분석하고 수사 관점에서 필요한 항목을 선별한다. 또한 아티팩트의 연관성과 멀티 행위를 유의미하게 판단하기 위해 선별한 데이터를 시간 정보 기준으로 분석하는 과정과 웹 기반 시각화를 제안한다. 이를 통해 클라우드 데이터 증거수집의 중요성이 증가함에 따른 아카이빙 데이터의 활용 가치를 보여주고자 한다.

ABSTRACT

The importance of digital forensics in the cloud environment is increasing as businesses and individuals move their data from On-premise to the cloud. Cloud data can be stored on various devices, including mobile devices and desktops, and encompasses a variety of user behavior artifacts, such as information generated from linked accounts and cloud services. However, there are limitations in securing and analyzing digital evidence due to environmental constraints of the cloud, such as distributed storage of data and lack of artifact linkage. One solution to address this is archiving services, and Google's Takeout is prime example.

In this paper, user behavior data is analyzed for cloud forensics based on archiving data and necessary items are selected from an investigation perspective. Additionally, we propose the process of analyzing selectively collected data based on time information and utilizing web-based visualization to meaningfully assess artifact associations and multi-behaviors. Through this, we aim to demonstrate the value of utilizing archiving data in response to the increasing significance of evidence collection for cloud data.

Keywords: Cloud, Google, Archiving Service, Digital Forensic, Multi-Behavior Analysis

Received(07. 11. 2023), Modified(08. 07. 2023), Accepted(08. 07. 2023)

* 본 연구는 2023년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2021R1G1A100632611), 산업통상자원부의 재원으로 한국산업기술진흥원의 지원(P0008703, 2023년 산업혁신인재성장지원사업), 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0 사업(IITP-2022-RS-2022-00156310)의 연구결과로 수행되었음.

† 주저자, 220224005@sungshin.ac.kr
‡ 교신저자, sm.kim@sungshin.ac.kr(Corresponding author)

I. 서 론

클라우드를 활용한 디지털 인프라 혁신이 시작되면서 사용자들의 데이터는 온프레미스 환경에서 클라우드로 확장되었다. 저장 공간 확보 및 백업을 위한 자동 동기화 기능도 더해지면서 PC뿐만 아니라 모바일 기기에 저장된 데이터도 클라우드에 보관된다. 이로 인해 모바일 기기로부터 생성된 사건 관련 데이터 등 사실 증명에 필요한 중요 정보가 클라우드에 저장되고 있다[1]. 실제 2021년 광주에서는 남성이 여성의 신체를 불법 촬영한 영상이 자동 동기화되어 적발되면서 불법 촬영 및 성적 촬영물 소지에 따른 성폭력 범죄 처벌 등에 관한 특례법 위반 혐의로 입건된 사례가 있으며[1], 같은 해 미국에서는 약 150개의 아동 성 착취 이미지 및 비디오 파일을 저장한 41세 남성의 클라우드 스토리지 계정이 적발된 사례가 있다[2]. 이에 따라 디지털 포렌식에서 클라우드 데이터는 중요 수집 대상이 되고 있지만, 환경 특성상 데이터 수집 시 어려움이 있다. 첫 번째로, 사용자 행위 데이터 대부분이 원격지 서버에 분산 저장된다[3]. 클라우드의 특징으로 증거는 여러 위치에 흩어져 물리적 접근이 불가하다. 두 번째로, 제공 서비스에 대한 아티팩트 연계 부족이다. 클라우드 서비스 공급자들은 시장 점유율을 높이기 위해 스토리지 기능 외 메일, 원격 화상 통화 등 다양한 기능을 제공하지만, 각 서비스에 대한 아티팩트 연계는 미비한 상태이다. 이러한 클라우드 포렌식의 문제점을 보완할 수 있는 수단으로 아카이빙 데이터가 있다. 아카이빙 서비스는 클라우드의 데이터를 추출해 타 솔루션과 연계할 수 있도록 해주는 것이며, 대표적으로 Google에서 개발한 Takeout이 있다. 이를 통해 유튜브, 크롬 등 구글에 저장된 사용자 활동 데이터를 압축된 파일 형태로 백업하고, 목적에 따라 응용해 활용할 수 있다. 예시로 검색 기록 및 유튜브 활동 데이터를 통한 우울증 환자 분류 및 탐지 방법[4]과 불안 장애 탐지 방법 등 특정 사용자의 정신 건강 문제 징후를 드러내는 연구가 수행된 바 있다[5]. 그러나 앞선 연구와 같이 아카이빙 데이터를 분석해 건강 통계 정보를 생성하는 연구 및 개발은 이루어지고 있지만, 수사적 관점의 연구는 상대적으로 부족하다.

구글 서비스 활동 로그는 연동된 계정에 자동 동기화되어 클라우드에 저장되기 때문에 아카이빙 데이터를 추출하면 사용자 행위를 파악하는 데 유용하게 활용할 수 있다[6]. 또한 구글은 지속적으로 새로운 기능과 서비스를 개발하고 업데이트하기 때문에 새롭게 생성된 데이터와 이를 포함한 방대한 데이터를 사건 유형에 따라 어떻게 효율적으로 활용하고, 목적에 맞는 의미를 도출할지에 대한 데이터 처리 분석 연구가 필요하다[7]. 이에 수사 관점에서 필수적이며 선행 연구에서는 분석되지 않은 정보를 확인하고, 아티팩트의 연계성을 한눈에 이해하기 쉽도록 시간 정보 기준으로 데이터를 처리하는 방식을 제시한다. 시간은 사건의 흐름을 파악할 수 있는 중요 정보로 이를 통해 타임라인을 구성해 사건의 선후 관계를 파악할 수 있다. 본 연구에서 시간은 다른 이벤트의 분석 시점 범위를 정할 수 있는 주요 연계 데이터가 된다. 그러나 각 이벤트 데이터 속 시간 형식은 다르기 때문에 행위 분석 시 어려움이 있다. 이에 따라 사용자 행위를 특정할 수 있는 정보를 추출하고 시간 데이터를 통일화해 정확한 분석이 이루어질 수 있도록 한다. 마지막으로 수사 관점에서 아카이빙 데이터를 유의미하게 활용하기 위해 분석 결과와 연결 가능한 웹 기반 시각화 방법을 제안하고 가상의 기밀 유출 시나리오를 통해 분석 프레임워크를 검증한다. 이를 통해 아카이빙 데이터 기반의 효율적인 수사 방법을 파악하고, 구글 클라우드 포렌식에서의 아카이빙 데이터 활용 가치를 높이는 데 기여하고자 한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구를 서술한다. 3장에서는 구글 클라우드 데이터를 분석 대상으로 선정하게 된 배경을 제시하고, 4장과 5장에서는 구글 아카이빙 데이터의 종류와 분석할 수 있는 아티팩트 정보 확인 후 행위 분석을 위한 프레임워크 구조 및 동작 흐름을 설명한다. 또한 각 이벤트에서 확인할 수 있는 시간 정보를 기준으로 시각화하는 과정을 보인다. 6장에서는 성능 평가를 위해 가상의 시나리오를 설정하고 분석 과정을 확인한다. 마지막으로 7장에서는 결론 및 향후 연구 방향에 관해 서술한다.

II. 관련 연구

김도현 외 2는 Takeout을 활용해 약 11개의 구글 클라우드 서비스 데이터를 수집한 후 파싱, 통합, 정규화를 통해 사용자 행위를 분석하는 도구 'gtForensics'를 개발하였다. 해당 도구를 이용해 사용자 행위를 시간과 이벤트 기준으로 재구성하고, 검색 및 시청 기록, 결제 내역을 통한 사용자의 관심

사를 파악하였다. 또한 시간 정보, 위치 정보, IP 주소를 통해 특정 시각에 사용자가 머물렀던 장소에 대한 정보 수집이 가능함을 보였다[6].

김현우 외 1은 구글에서 보관하는 다양한 데이터로부터 수사에 활용할 수 있는 정보를 추출하고, 이를 시각화해 수사에 유의미한 아티팩트를 확보할 수 있는 수사 기법을 제시하였고[7], 김동호 외 1은 스마트폰 포렌식을 보완하기 위해 구글 클라우드 데이터를 활용한 아티팩트들이 실제 수사에서 어떻게 활용될 수 있는지 분석하는 방법을 제안했다[8].

선행 연구에서 언급된 아카이빙 데이터는 지도, 드라이브 등 기본 계정에 주어지는 11개(중복 제외)의 데이터를 수사 관점으로 활용하는 방법을 다루었다. 구글은 시장 점유율을 높이기 위해 다양한 서비스와 제품을 수시로 개발하고 있어 데이터의 분석 및 처리에 관한 연구가 지속적으로 요구된다. 이에 본 논문에서는 수사 관점의 데이터 구조분석과 처리 과정에 중점을 두고, 기존 선행 연구에서 언급된 데이터의 구조 분석과 언급되지 않은 Google Chat, Google Pay 데이터를 추가로 확인하였다. 또한 각 이벤트 데이터는 형식이 다른 시간 정보를 가지고 있어 타임스탬프를 해석하는 데 어려움이 있기 때문에 이벤트 시간 정보를 확인하고, 이를 형식화해 서비스 간 연계 분석을 진행하고자 한다.

III. Google Cloud Data

구글은 계정 연동된 기기의 서비스 데이터를 동기화해 관리한다. 가장 유명한 Chrome 검색엔진 외에도 동영상 서비스 YouTube, 커뮤니케이션을 위한 Gmail 및 Google Chat, 애플리케이션 배포를 위한 Google Play Store 등 다양한 서비스 및 제품으로 구성되어 있다.

실제 트래픽 분석 사이트 스탯카운터에서도 2023년 세계 브라우저 순위는 크롬(74.25%)이 1위로 압도적인 점유율을 보였으며, 2위인 사파리(10.65%)와의 격차가 상당한 것을 볼 수 있다[9]. 또한 모바일 인덱스 분석 결과 2022년 국내 유튜브 사용자 수는 4,183만 명으로 대한민국 인구 81%가 사용하는 것으로 나타났다[10]. 이러한 통계자료를 통해 디지털 세계에서 정보를 검색하고 콘텐츠를 소비하는 과정에 구글 서비스가 차지하고 있는 비중이 크다는 것을 알 수 있다. 일상적으로 사용되는 서비스가 다양

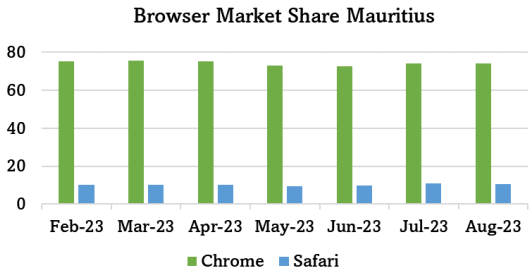

Fig. 1. Browser Market Share Mauritius

한 만큼 구글 클라우드에는 사용자의 행위 관련 데이터가 다량 보관되고 있으며, 이를 분석한다면 구글 클라우드 포렌식 수사에 유용하게 활용될 수 있다.

구글은 클라우드 스토리지 내 데이터와 자사 타 솔루션들을 연계해 주는 아카이빙 서비스를 배포하고 있어 서비스 관련 행위 데이터를 쉽게 수집할 수 있다. 특히 인터넷 검색, 유튜브 등 일상적으로 많이 사용되는 서비스 로그를 수집해 사용자의 관심 분야를 파악하면 수사 방향을 신속히 설정하고 용의자 범위를 좁힐 수 있다. 이러한 아카이빙 데이터를 통해 구글 클라우드 포렌식을 수행할 경우 분산되어 있는 아티팩트 문제를 해결하고, 시간 정보로 사건의 흐름 구성 및 이벤트 연계를 통한 필요 정보 습득 등 수사에 필요한 증거를 확인할 수 있다. 이에 본 연구에서는 구글 아카이빙 서비스를 통한 클라우드 데이터를 분석하고자 한다.

IV. Google Archiving Service Data

구글은 Data Liberation Front 프로젝트 일환으로 출시된 아카이빙 서비스 Takeout을 공개했다[11]. 이는 사용자 계정을 기반으로 보관된 구글 서비스 데이터를 추출해 외부에서도 활용할 수 있도록 하는 것이다. 분석 가능한 데이터로는 유튜브 시청·검색 기록, 인터넷 검색, 지메일 등 구글 서비스 사용 정보와 구글 포토, 드라이브 등 스토리지 관련 정보이다. 다운로드 가능한 사용자 데이터는 각 계정의 서비스 접근 권한마다 다르며, 데이터 종류에 따라 JSON, HTML, JPG 등 분석에 용이한 형식으로 추출이 가능하다. Table 1.은 구글 아카이빙 서비스로 확인 가능한 데이터 중 기본 계정에서 추출이 가능하며 시간 정보, 위치, 결제, 검색 등 사용자 행위 분석에 중요한 주요 아티팩트를 정리한 것이다.

Table 1. Google Archiving Service Data

Contents	Folder	Artifacts	Data Type
Mail	Gmail	Mail Account, Content, Attached File, Deleted Mail	Mbox
Drive	Drive	Upload File (Photo, Document, etc.), Access Info, Deleted File Name	JSON
Service Usage Information	You Tube	Subscription Info, History	JSON / HTML
	Chrome	Bookmark, Browser History, Extensions, Search Keywords, OS Setting, Device Info	JSON / HTML
	Google Account	IP, Login/ Logout Info, Access Device Info	HTML
Location	Location History	Latitude, Longitude	JSON
Android Platform	Google Play Store	Installation App list, Installation Device Info, IP, Credit card Info(used for payment)	JSON
	Android Device Configuration Service	Device Info (HW, SW, network, etc), Access Time	HTML
Photo	Google Photos	Photo, Video File, Metadata, Location, Longtitude/Latitude	JSON ,JPG/ PNG
Google Chat	Google Chat	User Info, Group Info, Messages, Attached File	JSON
Google Pay	Google Pay	Google Payment History (Transaction ID, Payment method, Payment Status, Payment amount)	CSV

4.1 Gmail & Google Drive

지메일 데이터는 사용자의 메일 송수신 내용, 첨부파일 확인이 가능하며 mbox 형식으로 다운로드 된다. mbox는 전자메일 메시지 확인이 가능한 기본 파일 형식으로 썬더버드(Thunderbird)와 같은 프로그램을 통해 원하는 정보를 필터링하여 얻을 수 있다. 메일 내용, 송수신 날짜 및 시간, 참조, 첨부파일 확인이 가능하며 휴지통에 보관된 이메일도 확인 가능해 용의자가 삭제했다고 판단한 데이터를 수집할 수 있다.

구글 드라이브는 스프레드시트 등 구글 문서와 사용자가 업로드한 파일에 대한 정보가 보관된다. 추출된 아카이빙 데이터를 분석하면 파일 이름, 확장자, 접근 권한, 다운로드 권한, 크기 정보 확인이 가능하다. 해당 정보를 확인하기 위해서는 데이터 추출 시 파일에 대한 고급 설정(업로드 버전 포함, 파일 정보 포함)을 추가로 선택해야 한다. 드라이브에서는 업로드된 폴더 및 파일뿐만 아니라 삭제된 파일 정보도 확인할 수 있다. Fig. 2.는 구글 드라이브에 저장된 파일을 삭제한 후 아카이빙 데이터를 추출한 결과이다. 휴지통에 버려진 파일(Fig. 2. 윗부분)이 아카이빙 결과로 추출된 것(Fig. 2. 아랫부분)을 확인할 수 있다. 이를 통해 삭제되었다고 생각한 파일이 의도치 않게 발견될 수 있다. 드라이브에서 확인할 수 있는 시간 정보는 마지막 타인에 의해 수정된 일시 last_modified_by_any_user, 사용자에 의해 수정된 일시 last_modified_by_me, 수정 일시 content_last_modified, 생성 일시 created가 있으며, 사용자가 드라이브에서 파일 이름을 수정할 시 last_modified_by_any_user, last_modified_by_me,

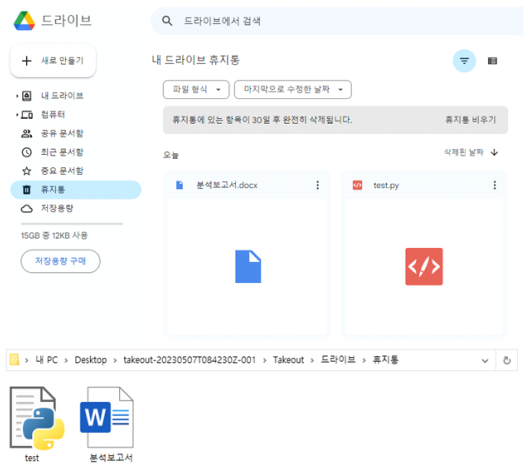

Fig. 2. Google Drive

content_last_modified가 수정된다.

중요 파일 전송 및 저장에 활용되는 지메일, 구글 드라이브 데이터는 기업의 정보유출과 같이 파일 및 송수신 메일이 주요 수단인 사건에서 용의자의 행위 정보를 획득하는 주요 데이터가 된다.

4.2 Service Usage Information

유튜브는 사용자가 검색하고 시청했던 콘텐츠에 대한 정보가 저장된다. 시청 및 검색한 영상의 채널, 제목, URL, 접속 시간 정보를 획득할 수 있으며 검색 기록, 구독 정보 확인도 가능하다. 크롬은 사용자가 검색하고 접속한 사이트의 URL과 검색 키워드, 검색 시간을 확인할 수 있다.

Table 2. Chrome Data

File Name	Description
Autofill.json	Words generated by AutoComplete
Bookmarks.html	List of bookmarks
BrowserHistory.json	Visited browser Info (URL, Time)
Device Information.json	Connected Device Info (Device Info, OS, Manufacturer, Chrome Version)
Extensions.json	Google Chrome Extension
Omnibox.json	Keywords in Google Chrome Address
SearchEngines.json	SearchEngine Info (URL, Time, Keyword)

크롬 데이터를 추출하면 Autofill, Bookmarks, BrowserHistory, Device Information, Extensions, Omnibox, OS Settings, SearchEngines 정보를 확인할 수 있으며 Table 2.는 각 데이터에 대한 설명이다. 유튜브와 크롬 로그는 용의자의 관심사와 검색 키워드를 확인하는 데 유용하다. 구글의 높은 점유율을 차지하는 서비스 데이터를 통해 용의자의 관심사를 파악한다면 증거 수집을 위해 어떤 정보나 자료를 추가적으로 수집해야 하는지 결정하는 데 도움이 될 수 있다. 또한 특정 관심사와 연관된 사건이나 장소를 예측해 용의자의 행동 패턴을 파악하는 정보가 될 수 있다.

4.3 Google Photo & Location

구글 포토는 구글이 제공하는 클라우드 기반의 미디어 파일 스토리지 서비스이다. 사용자가 촬영한 사진 및 비디오가 업로드되거나 자동 동기화된 파일이 백업되어 기기에 저장된 파일이 손실 및 삭제되어도 원본을 보존할 수 있다. 파일에 대한 메타 데이터 확인도 가능해 로그를 분석하면 파일 이름, 크기, 시간 정보를 확인할 수 있다. 이때 사진이 찍힌 시간은 photoTakenTime, 클라우드에 동기화된 시간은 creationTime으로 저장된다. 만일 사진 파일에 EXIF(Exchangeable Image File)가 포함된 경우 촬영 시 위치 정보도 확인할 수 있다[6].

구글 서비스 및 애플리케이션 이용 시 위치 정보 제공에 동의하면 기기의 위치 정보가 수집되어 Location 데이터에 저장된다. 이를 통해 사용자가 이동한 경로의 위도/경도와 시간 정보를 알 수 있고, 방문 장소 파악도 가능하다. 위치 정보는 사건 수사 시 이동 경로를 파악하는 데 중요한 역할을 한다.

또한 사건 발생 시간과 장소를 고려해 용의자의 행위 추적과 사건의 흐름 파악이 가능하고, 피해자 또는 증거에 대한 상대적 위치를 빠르게 파악할 수 있다.

4.4 Android Platform Service

구글 플레이스토어는 사용자가 원하는 영화, 게임 등 콘텐츠를 쉽게 다운로드해 즐길 수 있도록 하는 온라인 앱스토어이다. 2021년 전 세계 구글 플레이

스토어를 통해 다운로드 된 모바일 애플리케이션 수는 1,113억 건으로 2018년의 760억 건보다 증가했다고 나타났다[13].

구글은 사용자가 플레이스토어에서 다운 및 구매한 앱과 결제 정보를 보관한다. 이를 통해 사용자가 구매한 애플리케이션 정보인 설치 날짜, 기기 정보, IP, 결제 카드 정보를 확인할 수 있다. 아카이빙 데이터를 추출하면 다음 Table 3.의 정보를 확인할 수 있다. 이를 통해 안드로이드 환경에서 애플리케이션 사용으로 인해 발생하는 로그를 수집하고, 사용자 행위를 식별하는 데 활용할 수 있다. 이러한 정보는 애플리케이션과 디바이스를 통한 범죄가 증가하는 시점에 용의자의 행위를 분석하기 위한 중요 증거 데이터로 활용될 수 있다.

Android Device Configuration Service 데이터에는 구글 계정으로 접속된 디바이스 정보와 해당 디바이스가 연결된 시간 정보가 저장된다. Android ID, 단말기 고유 일련번호인 IMEI(International Mobile Equipment Identity), MEID(Mobile Equipment Identifier), 펌웨어 버전, 기기별 접속 시간 확인이 가능하다. 이를 통해 수사관은 모바일 포렌식 수행 시 필요한 기기 정보를 확인함으로써 분석 대상 디바이스에 관련된 사전 정보를 획득할 수 있다.

만일 분석 대상 디바이스 정보와 구글 플레이스토어 데이터에서 획득한 설치 애플리케이션 정보를 함께 분석한다면, 사용자의 기기에 설치된 애플리케이션을 디바이스 모델별로 확인하는 멀티 행위 분석이 가능하다.

Table 3. Google Play Store Data

File Name	Description
Devices.json	Play Store Access Device Information
Installs.json	Application Installation List
Library.json	Application Download List (Basic Application Information)
Purchase History.json	Application Purchase List (Purchasing Application Information)

4.5 Google Chat

구글 챗은 구글이 제공하는 실시간 메시지 전송 및 협업 도구이다. 팀원 간 채팅, 파일 공유, 등 협업 기능을 제공한다. 1:1 또는 1:N 채팅 기능을 통해 실시간으로 대화하고, 특정 그룹의 업무 진행을 위한 채팅방인 스페이스를 만들어 사용한다. 이를 통해 팀원과 소통하며 파일과 문서를 공유할 수 있다. 또한 Google Workspace의 다른 앱과 통합되어 Google Meet을 통한 화상 회의, 드라이브를 통한 파일 공유, 구글 캘린더를 통한 일정 관리 등 기능을 쉽게 이용할 수 있다[14].

Table 4. Google Chat Data

File Path	File	Description
Google Chat\ Groups	DM [id]	1:1 Chat Info(Group_info, messages, tasks)
	Space [id]	Space Info (Group_info, Messages, tasks)
Google Chat\ Users\ User [id]	user_info	1:1 chat room and Space info

아카이빙 데이터로 구글 챗 데이터를 확인하면 다음 Table 4.와 같은 데이터를 확인할 수 있다. 크게 Group과 Users 폴더로 나뉘며, Group에서는 해당 계정의 1:1 채팅(DM) 정보와 그룹 채팅(Space) 관련 정보를 수집할 수 있다. 이때 함께 채팅에 참여한 다른 사용자가 아카이빙 데이터를 추출할 경우 사용자와 같은 id로 DM 데이터가 생성된다. 이를 통해 사용자의 채팅방 대화 로그, 생성된 스페이스 목록 리스트, 공유된 파일 정보 확인이 가능하다. 구글 챗은 기업 내 팀 협업과 프로젝트 관리, 효율적인 업무 처리를 위해 사용되기 때문에 기업의 산업기밀 유출 수사에 유용한 데이터가 될 수 있다.

V. 사용자 행위 분석을 위한 구글 아카이빙 데이터 처리 과정

구글 아카이빙 데이터를 활용한 증거수집을 위해서는 수사 대상의 계정을 사전에 확보해야 한다. 이에 본 연구에서는 기업 내 기밀문서 유출자로 의심되는 사원의 계정을 확보하여 데이터를 수집한 상황을

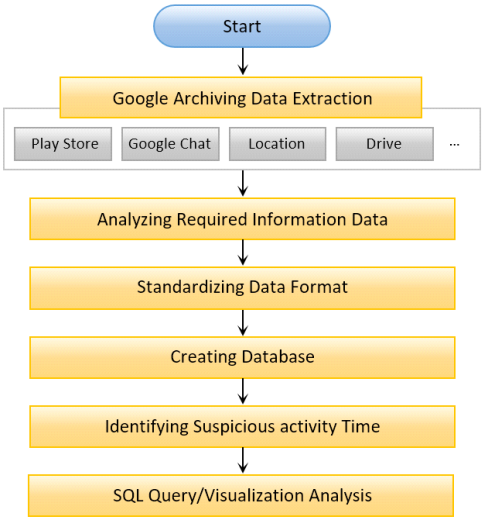

Fig. 3. Archiving Data Analysis

가정한다. Fig. 3.은 제안하는 구글 아카이빙 데이터 기반의 클라우드 포렌식 분석 과정이다. 먼저 Takeout을 통해 구글 클라우드 데이터를 추출한다. 추출된 데이터에는 디지털 포렌식 관점에서 불필요한 데이터가 포함되어 있으므로, 수사 목적에 필요한 정보를 추출하는 작업이 요구된다.

이후, 직관적인 판단을 위해 추출된 데이터의 형식을 통일하기 위한 데이터 변환 작업을 거친다. 특히 아카이빙 데이터의 시간 정보는 각 이벤트 파일마다 다른 형식을 가지고 있어 분석 시 어려움이 있기 때문에 시간 형식 변환이 필요하다. 수사에 필요한 행위 데이터 수집과 표준화 작업이 완료되면 각 행위를 이벤트 별로 구분해 DB를 구성한다. 이를 통해 수집된 데이터를 효율적으로 관리하고 분석할 수 있다.

5.1 Google Archiving Data Extraction

구글 아카이빙 서비스인 Takeout은 구글 계정을 소유하면 누구든 접근할 수 있으며, 추출 데이터는 서비스 항목에 따라 정해진 형식으로 제공된다. 데이터 추출 시 전송 방법(이메일을 통해 다운로드 링크 전송, 드라이브에 추가), 추출 실행 빈도(한 번, 1년간 2개월마다), 파일 형식(.zip, .tgz), 파일 크기를 설정한 후 선택된 전송 방식으로 다운로드 가능하다.

일부 데이터는 JSON, HTML 등 데이터 형식을 선택할 수 있다. 특히 활용도가 높은 JSON의 경우 이름과 값의 쌍으로 구성되어 있으며, 테이블 변환 시 이름은 각 테이블의 속성으로, 값은 속성 값으로 저장한다. 형식화된 데이터는 데이터베이스로의 변환이 가능하며, 이를 통해 구글에서 추출한 다량의 데이터 중 필요 정보를 빠르게 추출해 분석할 수 있다.

이렇게 구글 아카이빙 데이터를 분석하고 활용함으로써, 구글 클라우드 포렌식의 효율적인 수사 프레임워크를 구축할 수 있다.

5.2 Analyzing Required Information Data

Fig. 4.는 아카이빙 데이터에서 추출된 유튜브 시청 기록의 일부이다. 그림 윗부분은 아카이빙 데이터의 원본이며, 이를 JSON Crack으로 구조 분석하면 Fig. 4.의 아랫부분과 같다. 수사 관점에서 사용자 행위를 분석하기 위해 최소한으로 필요한 정보로는 영상 제목(title), 시청 시간(time), 영상 링크(url)로 판단된다. 이때 영상 제목의 경우 이모지가 포함될 수 있어 UTF8MB4 인코딩을 사용해 4바이트의 이모지가 저장될 수 있게 해야 한다. 구글 아카이빙으로 추출되는 각 이벤트 데이터의 필요 정보는 수사 목적에 따라 다를 수 있으며, 같은 이벤트라도 계정마다 데이터의 구조가 다를 수 있다. 따라서 데이터 분석 및 구조화 시 이를 고려해야 한다.

Fig. 5.는 구글 플레이스토어 설치 데이터로 애플리케이션 이름 (title), 설치/업데이트 시간 (firstInstallationTime, lastUpdateTime), 기기 정보(model)가 최소 필요 정보라고 판단된다. 이렇게 유튜브, 크롬, 위치, Android Device

Fig. 4. YouTube

```
{
  "install": {
    "doc": {
      "documentType": "Android Apps",
      "title": "Chrome: 빠르고 안전한 브라우저"
    },
    "firstInstallationTime": "2022-12-29T07:35:16.694017Z",
    "deviceAttribute": {
      "model": "SM-G986N",
      "carrier": "No carrier",
      "manufacturer": "samsung",
      "deviceDisplayName": "samsung SM-G986N"
    },
    "lastUpdateTime": "2022-12-29T07:35:16.694017Z"
  }
}
```


Fig. 5. Google Play Store_install

Configuration on Service, Google Chat 등 아카이빙 데이터 구조를 분석한 후 확보 가능한 데이터 유형을 비교해 중복 정보를 제거한다면 원본 데이터보다 경량화된 데이터 분석이 가능하고, 신속한 수사를 진행할 수 있다.

Fig. 6.은 구글 챗의 사용자 정보를 담은 user_info 데이터이다. 구조를 확인하면 계정 정보와 해당 계정 사용자가 참여한 DM, Space 정보를 확인할 수 있다.

1:1 채팅 또는 그룹 채팅이 생길 때마다 하나의 group_id가 생성되며 이를 통해 사용자가 참여한 채팅 목록을 확인할 수 있다. 각 채팅에 따른 메시지 내용과 공유된 파일 등 자세한 정보는 Fig. 7. 과같이 Groups의 해당 id 값 데이터에서 확인이 가능하다.


```
{
  "user": {
    "name": "[redacted]",
    "email": "[redacted]@gmail.com",
    "user_type": "Human"
  },
  "membership_info": [
    {
      "group_id": "DM gdp76UAAAAE",
      "membership_state": "MEMBER_JOINED"
    },
    {
      "group_id": "DM ppN4GUAAAAE",
      "membership_state": "MEMBER_JOINED"
    },
    {
      "group_name": "space1",
      "group_id": "Space AAAAsH69zuE",
      "membership_state": "MEMBER_JOINED"
    }
  ]
}
```


Fig. 6. Google Chat User_Info

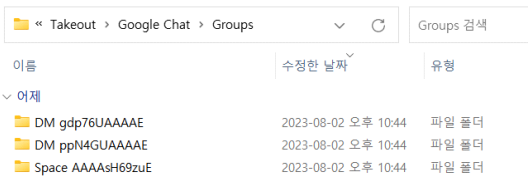

Fig. 7. Google Chat_Groups

5.3 Standardizing Data Format

구글은 다양한 서비스를 제공하기 때문에 보관되는 사용자 데이터도 방대하다. 이러한 정보에는 사용자의 위치 정보인 위도/경도, 타임스탬프와 같이 직관적으로 이해하기 어려운 데이터도 포함된다. 따라서 데이터를 직관적으로 이해하고 활용하기 위해서는 변환이 필요하다.

아카이빙 데이터에는 각 이벤트에 따른 사용자 행위 시간 정보가 포함되어 있다. 디지털 포렌식 수사에서 시간 정보는 사건의 전체적인 흐름을 빠르게 파악할 수 있는 중요한 요소이며, 시간 데이터를 통해 사용자의 의심 행위를 분석하고 기준 시간을 파악해 보다 신속히 수사를 진행할 수 있다. 또한 아카이빙 데이터에서 다른 이벤트와의 멀티 분석을 위한 연계 역할을 지원한다. 따라서 시간 정보를 포함하는 로그 데이터를 분석하고, 시간 형식을 표준화해 일관된 방식으로 활용하고자 한다. 예시로 Fig. 8.은 BrowserHistory 데이터 중 일부이며 Fig. 4.와 5의 시간 정보와는 형식이 다른 것을 확인할 수 있다. 이와 같은 시간 표현 방식을 Unix Time 또는 Epoch Time이라 부르며, 1970년 1월 1일 00:00:00 협정 세계시(UTC)로부터 경과된 시간을 초로 환산한 값을 나타낸 것이다[15]. 이는 유닉스 계열 운영체제나 파일 타입에서 사용되는 시간 표현 방식으로 초 단위로 표현되어 사건 발생 시 무엇이 먼저 일어났는지를 판단할 수 있는 지표가 될 수 있다. 그러나 년, 월, 시, 분 초와 같은 정보를 단번에 알 수 없어 가독성이 떨어지기 때문에 제안 프레임워크에서는 단번에 파악 가능한 시간 형식으로 분석한다. Table 5.는 각 이벤트 별 파악 가능한 시간 정보를 정리한 것이다.


```
{
    "favicon_url": "https://www.youtube.com/s/desktop/ea447fad/img/favicon_32x32.png",
    "page_transition": "LINK",
    "title": "구글 계정 여러개 깔끔하게 삭제 및 복원하기!! - YouTube",
    "url": "https://www.youtube.com/watch?v=M9mmCkEqDws",
    "client_id": "ROMBvO9MsxUF3vxusEdLIg==",
    "time_usec": 16702964659060111
}
```


Fig. 8. BrowserHistory_Time Data

5.4 Creating DataBase

필요 정보로 판단된 데이터를 정리해 데이터베이스로 구성하는 단계이다. Takeout 데이터를 추출하면 이벤트 데이터가 각 폴더에 저장되기 때문에 분석 시 어려움이 있다. 이에 추출한 Takeout 데이터를 DB로 재구성해 수집된 데이터를 효율적으로 관리하고 분석하고자 한다. 또한 쿼리(SQL Query)를 통해 사용자의 행위 및 타임라인을 파악할 수 있다. Fig. 9.는 DB 스키마의 일부이다. MySQL 기준으로 작성되었으며 주요 정보는 다음과 같다. 사용자 계정과 이름을 저장한 user, 사용자가 로그인한 디바이스 정보를 담은 user_device가 있다.

구글은 하나의 계정으로 여러 기기에서 접근이 가능하므로 user 테이블과 user_deivce는 1:N 관계로 연결된다. 따라서 user_device 테이블에는 기기 정보와 사용자 계정 및 이름이 담긴 user 테이블의 id가 함께 저장된다. 사용자 정보를 담은 user와 user_device 테이블이 생성되면 수사 목적에 따라 주요 데이터로 판단된 이벤트 테이블을 생성한다.

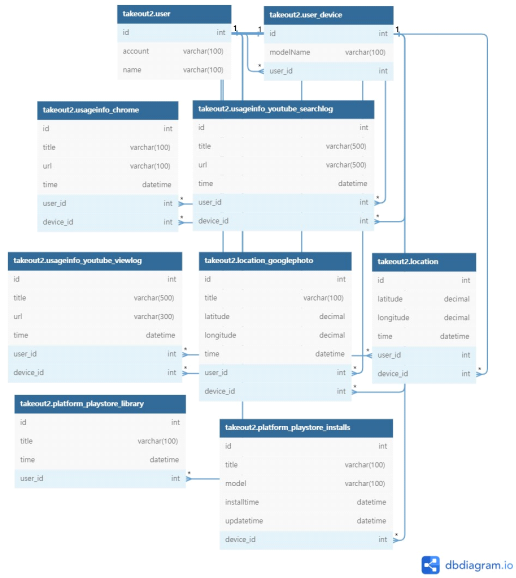

Fig. 9. DB Schema

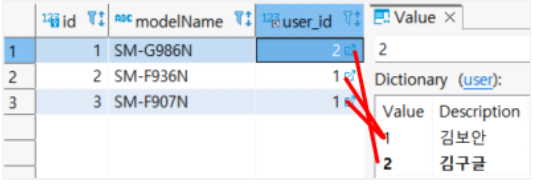

Fig. 10. User_device Table

5.5 Identifying Suspicious activity Time & SQL Query Analysis

Takeout으로 추출된 데이터를 모두 분석하려면, 오버헤드가 크고 확인해야 할 양이 방대해 비효율적이다. 이에 필요 데이터를 분석한 후 DB로 구성되도록 한다.

이를 통해 분석하고자 하는 이벤트 테이블을 선정해 의심시간과 행위 정보를 확인할 수 있으며, 이벤트 단위로 흩어진 테이블에서 분석 기준 시간으로 판단된 시점을 필터링해 시간 복잡도를 줄일 수 있다. 인터넷 검색 기록을 저장한 테이블에서 '계정 삭제', '검색 기록 삭제' 등 이상 키워드가 보인 시간 전·후를 기준으로 필터링해 동일 시간대 다른 이벤트에서의 용의자 행위를 빠르게 판단한다. 예를 들어 유튜브 시청 기록 테이블(Fig. 11.)에서 의심 활동이 발견된 시간을 기준으로 다른 이벤트(Fig. 12.)에서 필터링을 수행함으로써 수사 범위를 좁힐 수 있는 것이다. 또한 의심 행위로 판단된 시간을 필터링 한 결과 '무음 카메라'와 같이 추가적인 의심 정황 기록이 남아 있다면, 용의자의 계획 범행 정황 증거로 사용될 수 있다.

이와 같이 분석 시점을 파악한 후 필터링을 수행함으로써 용의자의 특정 범죄 행위를 빠르게 파악할 수 있다. 또한 중요 이벤트를 날짜와 시간대로 분석하여 용의자가 특정 시간대에 어떤 행위를 많이 하는지 등 생활 패턴 파악도 가능하다.

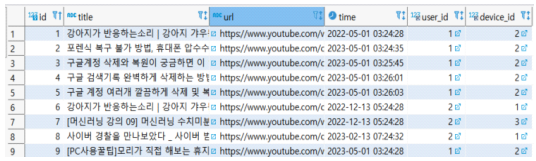

Fig. 11. YouTube View Log Data

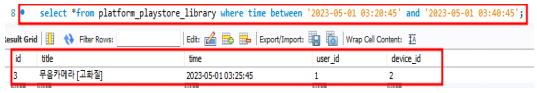

Fig. 12. Google Play Store Data

5.6 Visualization Analysis

구글 아카이빙 데이터는 검색 기록, 결제 정보, 위치 정보 등 다양한 데이터를 보관하기 때문에 분석 결과를 시각화하여 제시함으로써 다음과 같은 의의를 얻을 수 있다. 우선 대량의 데이터를 다루는 과정에서 사건의 상관관계 등을 시각적으로 파악할 수 있다. 시각화를 통해 복잡한 정보를 단순화함으로써 수사관은 데이터를 쉽게 이해하고, 중요 정보를 빠르게 판단할 수 있다.

두 번째로 데이터의 새로운 연관 관계를 도출할 수 있다. 사건의 재구성은 수사에서 중요한 과정이며 사건의 순서와 발생 위치 등을 시각적으로 나타내는 과정에서 데이터에 숨어있던 패턴을 찾아낼 수 있다. 따라서 수사적 관점에서 데이터 시각화는 유용하며, 이를 활용해 효율적인 수사를 진행할 수 있다. 데이터 시각화에는 다양한 유형이 있으며, 사건의 목적에 맞는 차트를 활용해 도형의 크기, 위치, 색상을 이용해 데이터의 분포를 파악할 수 있다.

본 연구에서는 구축한 DB와 연동이 가능한 웹 기반 시각화를 위해 Grafana 도구를 활용한다. Grafana는 데이터를 실시간으로 시각화하고 모니터링하기 위한 오픈소스 데이터 시각화 도구로 데이터를 시각적으로 표현해 대시보드 형태로 제공해 준다. 이를 통해 사용자는 데이터를 쉽게 이해하고 분석할 수 있다. 다양한 데이터 소스와의 통합을 지원하며, 대표적으로 Prometheus, Elasticsearch, MySQL 등이 있다. 이에 구성된 DB를 연결해 쿼리를 통한 시각화를 할 수 있어 유용하다.

Fig. 13.은 위치 데이터에서 추출된 위/경도를 시각화하여 나타낸 것이다. 이를 통해 사용자가 특정 시각에 어디에 있었는지, 어느 장소에 자주 방문했는지 한눈에 파악할 수 있다. Fig. 14.는 사용자의 유튜브 검색 키워드를 시각화한 것으로 어떤 키워드를 얼마나 검색했는지 확인할 수 있다. 이렇게 시각화를 수행함으로써 직관적으로 데이터를 분석하고, 중요한

Table 5. Time Data

Contents	File	Time Event Name	Type	Detail
Chrome	Browser History	time_usec	Epoch Time	Search Time
	Search Engines	last_modified	Epoch Time	Search engine access time
	Device Information	last_updated_timestamp	Epoch Time	Last accessed time
	Omnibox	visits	Epoch Time	Connection time via omnibox
Drive	〔File name〕	last_modified_by_any_user	ISO 8601	Last modified time by someone else
		last_modified_by_me	ISO 8601	Last modified time by user
		content_last_modified	ISO 8601	File Modification Time
		created	ISO 8601	File creation Time
YouTube	Search Log	Time	ISO 8601	Search Time
	View Log	Time	ISO 8601	View Time
Google Play Store	Installs/ Library	firstInstallationTime	ISO 8601	Application Installation Time
		lastUpdateTime	ISO 8601	Application Last update Time
	Devices	DeviceRegistrationTime	ISO 8601	Device registration Time
		lastTimeDeviceActive	ISO 8601	Device Last Activity Time
Google chat	messages	created_date	ISO 8601	1:1 / 1:N Chatting Time
Google Pay	Transaction history on google	Time	ISO 8601	Payment Time

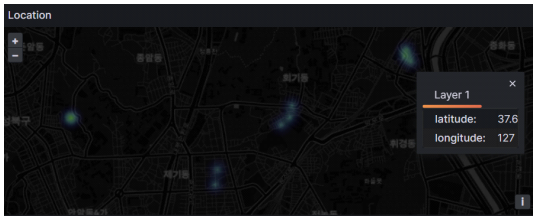

Fig. 13. Visualization - Location

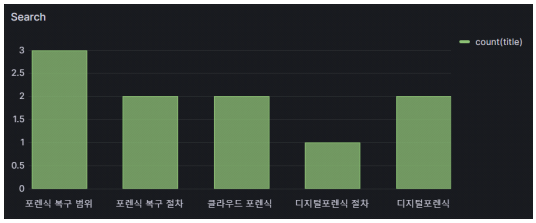

Fig. 14. Youtube Search Data Visualization

정보를 빠르게 판단할 수 있게 된다.

VI. 가상 시나리오 기반 사용자 행위 분석

제안한 분석 방법의 실용성 검증을 위해 가상의 범죄 행위 시나리오를 구성하고, 이를 기반으로 수사 시뮬레이션을 수행하고자 한다. Table 6.은 아카이빙 데이터 기반의 수사 측면 활용 과정을 보이기 위해 기업 내 사원이 구글 계정으로 로그인한 후 행한 가상의 시나리오를 작성한 것이다. 계정 접속 후 드라이브 접근, 기밀 유출을 위한 인터넷 및 유튜브 검색, 구글 플레이스토어 등 애플리케이션 및 서비스를 이용한 상황을 가정한다.

Table 6. Scenario

Scenario
(Google Account) Login Google
(Google Drive) Drive Access
(Chrome and YouTube) Search
(Google Play Store) App Download
(Google Photo) Take a picture

6.1 SQL Query/Visualization Analysis

6.1.1 분석 기준시간 기반 멀티 행위 분석

디지털 포렌식 수사 시 발생한 사건의 행위 시간 분석 및 타임라인 구성은 중요한 작업 중 하나이다[16]. 기준 시간 파악은 각 이벤트에서 용의자 행위를 분석하는 데 유용하다. 특정 이벤트 발생 시점 전·후로 어떤 일이 발생했는지 파악 가능하며, 정밀 분석 대상을 신속하게 선별할 수 있다.

이를 위해 우선 특정 테이블에서 의심 행위가 이루어진 시간을 파악한 후 다른 이벤트 테이블의 결합을 통해 검색을 수행한다. 이후 분산된 이벤트 테이블을 통합하여 기준 시점 전·후에 다른 이벤트에서는 어떤 행위가 이루어졌는지 확인한다. Fig. 15.는 인터넷 검색 및 드라이브 접근 로그를 확인할 수 있는 usageinfo_chrome과 애플리케이션 다운로드 정보를 확인할 수 있는 Google Play Store를 조인해 분석한 결과이다. 분석 기준 시간 전·후로 용의자가 인터넷 검색으로 남긴 기록과 해당 시간 범위에서 다운로드한 애플리케이션 정보를 함께 확인할 수 있다. 이를 통해 이벤트 단위로만 확인이 가능했던 구글 클라우드 데이터를 추가적으로 조합해 사용자가 동일 시간대에 어떤 행동을 했는지에 대한 멀티 행위 분석이 가능해진다. Fig. 16.은 Grafana 시각화 도구의 status history 템플릿을 활용해 크롬과 구글 플레이스토어, 구글 포토에 남겨진 사용자의 멀티 행위를 시각화한 결과이다. 분석 시간 기준으로 용의자가 유출 관련 검색과 무음 카메라 앱 설치, 해당 시점에 찍은 사진에 대한 행위 정보를 통합해 한눈에 확인할 수 있다.

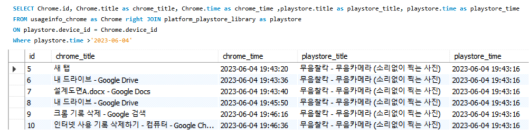

Fig. 15. Determining the analysis Time

Fig. 16. Multi-Behavior Visualization - Google Chrome/Play Store/Photo

6.1.2 사용자 / 기기별 분석

구성된 데이터베이스에는 사용자 정보가 담긴 user 테이블과 사용자가 로그인한 기기 정보가 담긴

user_device 테이블이 있다. 이를 각 이벤트 테이블과 연결해 분석하면 특정 행위가 누구의 계정과 어떤 기기로 접속된 것인지 확인할 수 있다. 구글은 계정만 있다면 다중 기기 및 멀티 프로필 접속이 가능하기 때문에 수사 시 해당 정보를 통해 사용자와 접속 기기가 구분될 수 있어야 한다.

만일 6.1.1에서의 인터넷 검색 기록 행위가 누구로부터, 어떤 기기로 인해 일어났는지에 대한 판단을 하기 위해서는 크롬 이벤트와 user, user_device와의 조합을 통해 가능하다. Fig. 17.은 크롬 검색 기록에 대한 사용자의 기기 및 계정 정보를 분석한 결과이며, Fig. 18.은 이를 시각화로 나타낸 것이다. 이를 통해 특정 이벤트 행위에 대한 주체의 정보를 확인할 수 있다.

```
SELECT Chrome.id, Chrome.title as chrome_search, user.account, user.name, user_device.modelName
FROM usageinfo_chrome as Chrome inner join user
on Chrome.user_id = user.id
inner join user_device
ON user.id = user_device.user_id;
```

	id	chrome_search	account	name	modelName
▶	1	NAVER	[redacted]22@gmail.com	김구글	SM-G986N
	2	[redacted] Educational System	[redacted]22@gmail.com	김구글	SM-G986N
	3	[redacted] - (2)	[redacted]22@gmail.com	김구글	SM-G986N
	4	[redacted]ducational System	[redacted]22@gmail.com	김구글	SM-G986N
	5	새 탭	[redacted]an0@gmail.com	김보안	SM-F936N
	6	내 드라이브 - Google Drive	[redacted]an0@gmail.com	김보안	SM-F936N
	7	설계도면A.docx - Google Docs	[redacted]an0@gmail.com	김보안	SM-F936N
	8	내 드라이브 - Google Drive	[redacted]an0@gmail.com	김보안	SM-F936N
	9	크롬 기록 삭제 - Google 검색	[redacted]an0@gmail.com	김보안	SM-F936N

Fig. 17. User/Device Analysis

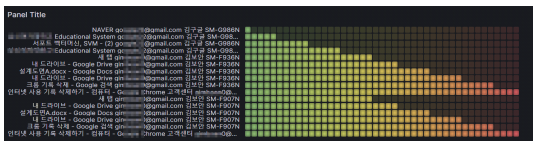

Fig. 18. User/Device Visualization

6.1.3 사진 촬영 시 위치 분석

만일 용의자가 범행 시간에 다른 장소에 있었다는 주장을 한다면, 위치 정보를 통해 주장을 확인하거나 반박할 수 있다. 시나리오에서는 구글 포토의 위/경도 데이터를 활용해 위치 정보를 확인할 수 있으며, 사진을 촬영한 시점에 용의자가 어느 위치에 있었는지 파악할 수 있다. Fig. 19.는 사진 촬영 시 용의자의 위치를 분석한 결과를 나타낸 것이다. 두 데이터 모두 촬영된 이미지의 위/경도를 확인할 수 있지만, 쿼리를 통한 분석의 경우(Fig. 19. 윗부분) 용의자의 위치를 단번에 파악하기 어렵다. 이에 시각화를 통한 정보 분석을 수행함으로써 데이터의 직관적 판단이 가능하도록 할 수 있다.

Fig. 19.의 아랫부분은 Grafana의 Geomap 템플릿으로 지도상에 위치를 표시하고 시각적으로 확인한 결과이다. 용의자가 사진을 찍었던 모든 위치 정보가 초록 원으로 표시되어 있으며, 해당 지점에서 찍은 사진 정보를 한눈에 확인할 수 있다. 이를 통해 지리 데이터에 대한 시각화의 필요성 가치를 확인할 수 있다[17].

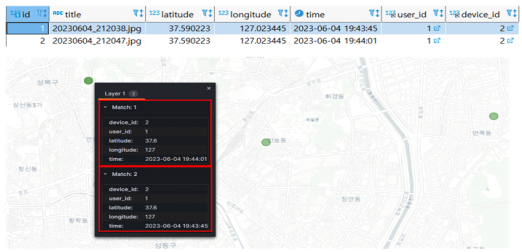

Fig. 19. Location Visualization - Google Photo

VII. 결 론

클라우드 서비스의 수요가 증가함에 따라 클라우드 데이터 분석은 디지털 포렌식 수사에서 중요한 역할이 되었다. 특히 구글 클라우드에는 사용자가 업로드 및 동기화한 파일뿐만 아니라 유튜브, 플레이스토어, 구글 챗 등 사용자의 구글 서비스에 대한 행위, 위치, 시간 정보까지 저장되기 때문에 수사에 유용하게 활용될 수 있다. 구글이 제공하는 서비스의 종류와 이를 활용하는 사용자는 증가하고 있기 때문에 새롭게 생성된 데이터와 기존의 방대한 데이터를 사건 유형에 따라 효율적으로 분석할 수 있는 연구가 지속적으로 필요하다. 그러나 현재 아카이빙 데이터에 대한 처리 및 분석 과정 연구는 부족한 상황이다.

본 연구에서는 이러한 중요성을 고려해 구글 아카이빙 서비스 데이터 중 수사적 관점에서 필요한 데이터를 처리 과정 중심으로 분석하고, 데이터의 시간 정보를 기준으로 이벤트를 연계한 시각화를 제안하였다. 이를 통해 구글 이벤트 기반의 사건 재구성과 시간 기준 검색을 통해 사건의 선·후 관계 및 멀티 행위를 파악하는 데 도움이 될 수 있음을 확인하였다. 또한 선행 연구에서는 분석되지 않은 구글 챗, 구글 페이를 추가로 분석하였으며, 기본 계정에 주어지는 아카이빙 데이터 중 수사적 가치가 있다고 판단된 데이터의 시간 정보를 정리하였다. 이러한 아카이빙 데이터의 처리 과정 및 분석 연구를 통해 얻을 수 있는 의의는 다음과 같다. 우선 높은 점유율 차지하는 구

글 서비스의 증거수집과 멀티 아티팩트 연계가 가능한 클라우드 포렌식의 기반을 마련할 수 있다. 구글 서비스 활동 로그는 연동된 계정마다 자동 동기화되기 때문에 같은 시간에 사용자 행위 데이터가 복합적으로 저장될 수 있다. 따라서 분석된 데이터를 기반으로 사용자 행위 증거를 수집하면 용의자가 같은 시간에 접근했던 서비스의 멀티 행위 분석이 가능하며, 수사를 효과적으로 이끌어 나갈 수 있다. 다음으로 각 이벤트 별로 분석된 시간 정보와 계정 아이디를 기반으로 멀티 프로필 분석이 가능하다. 구글은 한 명의 사용자가 여러 계정으로 접속할 수 있는 멀티 프로필 기능을 제공한다. 구성된 DB는 단일 계정뿐만 아니라 다른 계정에 대한 데이터 수집 및 저장도 가능케 해 멀티 프로필에 대한 이벤트 연계 분석도 가능하다. 이를 통해 클라우드 포렌식의 한계점이었던 아티팩트의 연계성 문제를 해결할 수 있다.

그러나 아카이빙 데이터 수집을 위해서는 클라우드 스토리지 접속을 위한 사용자 계정 정보가 필요하다. 현재 국내에서는 피압수자의 동의 없이 계정 정보를 수집하는 것에 법적 한계가 있다. 이에 데이터의 활용성을 높이기 위해 제도적 개선이 요구되며 아카이빙 데이터의 가치를 보이는 연구를 통해 개선 필요성을 강조하고자 한다. 향후 연구에서는 제안한 프레임워크를 자동화하여 행위 분석을 진행하고자 하며, 다양한 카테고리의 범주화된 시나리오를 기반으로 실질적인 행위 탐지가 가능하도록 구체화할 것이다.

References

[1] Seunghee Seo, Jueun Kim and Changhoon Lee "Trends in Cloud Storage Analysis from a Digital Forensics Perspective", Journal of The Korea Institute of information Security & Cryptology, 32(23), pp. 29-36, Apr. 2022.

[2] Northern District of Ohio. "Northeast Ohio Priest Pleads Guilty to Charges of Sex Trafficking of a Minor, Sexual Exploitation of a Child and Possession of Child Pornography", https://www.justice.gov/usao-ndoh/pr/northeast-ohio-priest-pleads-guilty-charges-sex-trafficking-minor-sexual-exploitation, last modified Jun 16.2021, accessed Jun 09.2023.

[3] Sunglim Kang, Jungheum Park and Sangjin Lee, "The Trace Analysis of SaaS from a Client's Perspective", Korea Information Processing Society, 19(1), pp. 1-8, Feb. 2012.

[4] Boyu Zhang, Anis Zaman, Rupam Acharyya, Ehsan Hoque, Vincent Silenzio and Henry Kautz, "Detecting Individuals with Depressive Disorder from Personal Google Search and YouTube History Logs", arXiv preprint arXiv:2010.15670, Oct. 2020.

[5] Anis Zaman, Boyu Zhang, Vincent Silenzio, Ehsan Hoque and Henry Kautz "Individual-level Anxiety Detection and Prediction from Longitudinal YouTube and Google Search Engagement Logs", arXiv preprint arXiv:2007.00613, Jul. 2020.

[6] Dohyun Kim, Junki Kim and Sangjin Lee, "An Analysis of Google Cloud Data from a Digital Forensic Perspective", Journal of the Korea Institute of Information and Communication Engineering, 24(12), pp. 1662-1669, Dec. 2020.

[7] Hyun-Woo Kim and Sangjin Lee, "Visualization of User Behavior With Google User", Journal of Digital Forensics, 13(4), pp. 287-301, Dec. 2019.

[8] Dongho Kim and Sangjin Lee, "A Study on the Usage of Investigation of Google Cloud Data (Smartphone user-oriented)", Journal of Digital Forensics, 12(3), pp. 107-118, Dec. 2018.

[9] Statcounter, "browser-market-share" , https://gs.statcounter.com/browser-market-share/all/mauritius, last modified Sep 26. 2023, accessed Sep.

26 2023.
〔10〕 “September 2022, 41.83 million ‘YouTube’ users, 81% of users in Korea”, Platum, last modified Oct 12. 2022, accessed Aug 03. 2023, https://news.zum.com/articles/78755836.
〔11〕 Techopedia. “Google Data Liberation Front”, https://www.techopedia.com/definition/26784/google-data-liberation-front, last modified Oct 24. 2012, accessed Jun 09. 2023.
〔12〕 MarketSplash, “Over 200 Influential Google Statistics in 2023”, https://marketsplash.com/ko/interesting-google-statistics/, last modified Jun 2023, accessed Aug 03. 2023.
〔13〕 Platum, “Worldwide app downloads reach 35 billion in the first quarter of 2023 AI-based productivity apps grow noticeably”, https://platum.kr/archives/208845, last modified Jun 15. 2023, accessed Aug 03. 2023.
〔14〕 Google, “Google Chat Manual”, https://docs.google.com/document/d/1V9HBCcTxj8Ml8myMWfkdwJ3QIjKNNnwdORJZ7eZs1sU/edit, last modified Jun 13. 2020, accessed Aug 03. 2023.
〔15〕 Víctor Eduardo Vásquez-Ortiz, Marco Milla, Joaquín Verástegui and Oscar Romero, “FPGA-based GPS controlled timing system with nanosecond accuracy and leap second support”, IEEE Fourth Ecuador Technical Chapters Meeting(ETCM), pp. 1-6, Nov. 2019.
〔16〕 Keun-Gi Lee, Seong-Jin Hwang, Chang-Hoon Lee and SangJin Lee, “Study on advanced analysis method based on timeline chart for Digital Forensic Investigation”, Journal of Advanced Navigation Technology, 18(1), pp. 50-55. Apr. 2014.
〔17〕 Kyeong Park, Eunmi Chang, Jun-Beom Park and Sungja Choi, “Analysis of Service Demand on Geological and Mine-related Thematic Maps and Available Contents in Geological Maps and Mine Maps”, Journal of the Korean Cartographic Association, 15(1), pp. 13-24, Apr. 2014.

〈저자소개〉

김 예 은 (Yeeun Kim) 학생회원
2022년 8월: 성신여자대학교 융합보안공학과 학사
2022년 9월~현재: 성신여자대학교 미래융합기술공학과 석사과정
〈관심분야〉 정보보호, 클라우드 컴퓨팅, 디지털 포렌식

홍 사 라 (Sara Hong) 학생회원
2021년 2월: 성신여자대학교 융합보안학과 학사
2023년 8월: 성신여자대학교 미래융합기술공학과 석사과정
〈관심분야〉 클라우드 컴퓨팅, 서버리스 컴퓨팅, 컨테이너 보안

김 성 민 (Seongmin Kim) 종신회원
2012년 2월: 한국과학기술원 전기 및 전자공학과 졸업
2014년 2월: 한국과학기술원 전기 및 전자공학과 석사
2019년 2월: 한국과학기술원 정보보호대학원 박사
2019년 9월~2020년 8월: 삼성전자 삼성리서치 Staff Engineer
2020년 9월~현재: 성신여자대학교 융합보안공학과 조교수
〈관심분야〉 신뢰 실행 환경, 클라우드 컴퓨팅, 시스템 보안