International Journal of Science and Technology (IJST)
ISSN: 3049-1118, Volume- 3, Issue- 2 (April - June 2026)

Novel Forensic Methodology for Extracting Geolocation Data from Browser Cache Artifacts

A. Jose Praveen[1], Sankara Narayanan S T[2]
[1]Department of Cyber Security, Dr. MGR Educational and Research Institute
Chennai, Tamil Nadu, India
[2]Department of Cyber Forensics and Information Security, Center of Excellence in Digital Forensics,
Chennai, Tamil Nadu, India

Abstract

Modern digital forensics make more use of geolocation evidence, which is helpful to understand where and when users have been active, as well as identify their possible association with a particular place or even an object. Conventional geolocation methods usually consider IP address tracking, GPS coordinates extracted from mobile devices and user-owned media, as well as location services on such devices. However, the current state of web technologies allows one to collect geolocation information not only from IP addresses but also by considering the process of automatic caching online content, including images, maps, and other relevant artifacts generated within the web browser environment. Thus, the purpose of this paper is to discuss a new forensic methodology aimed at collecting geolocation-related information using artifacts stored in the browser cache folder. The main aspects to be considered include cache image recovery, metadata extraction, map tile artifact analysis, as well as finding coordinates based on the information provided by web browser. An experiment was conducted to find out if the suggested approach is appropriate for geolocation investigation.

Keywords: Digital Forensics, Browser Cache Analysis, Geolocation Extraction, GPS Metadata, Cyber Forensics, Browser Artifacts, Cache Forensics, Map Tile Analysis.

Introduction

The increase in internet usage and the use of online maps, social media platforms, and multimedia services have resulted in significant amounts of geolocation-related data creation and storage. Modern digital devices such as smartphones regularly incorporate geographical location information in pictures and other forms of media using geo-tagging technologies. At the same time, browsers store website files in their cache to enhance browsing speed and efficiency.

Cached browser files may consist of cached images, map resources, thumbnails, icons, and other metadata related to websites. This information can be used by forensic experts to obtain evidence that is useful in determining user activity and geography-related preferences. One of

Published: 25 May 2026
DOI: https://doi.org/10.70558/IJST.2026.v3.i2.241265

the primary concerns during a digital forensics investigation is finding out the location of the suspect and linking their activities to a particular geographical region. IP-based location detection or the analysis of GPS-enabled media files remains one of the common practices. Nevertheless, both approaches face numerous limitations since suspects may mask IP addresses and there might be no media files available.

In this research, a new forensic approach for detecting and recovering geolocation data from browser cache data is introduced. The new forensic approach is based on several procedures that include extracting browser cache data, restoring images, analyzing metadata, examining map tiles, extracting coordinates, and visualization techniques.

Problem Statement

Contemporary web browsers inherently cache substantial amounts of internet-based materials, which include maps, pictures, and geographically relevant assets. While these artifacts hold significant value to forensics, there is no established and automatic approach to obtaining geolocation data from cache data. Common problems encountered during investigations include:

- Cache data encoding
- Damaged and fragmented cache artifacts
- Absence of any GPS data in the artifacts
- Inability to associate cache artifacts with specific locations
- Unstable nature of browser cache data

The current study focuses on developing a consistent forensic procedure that allows for the extraction, analysis, and visualization of geographically relevant artifacts from browser caches.

Objectives

The primary objectives of this research are:

1. To identify and analyse browser cache storage structures in modern web browsers.
2. To develop a forensic methodology for recovering geolocation-related artifacts from browser cache data.
3. To implement automated scripts for extracting GPS coordinates and metadata from recovered artifacts.
4. To visualize extracted geographic information using map-based forensic visualization.
5. To evaluate the reliability and limitations of browser cache geolocation analysis.

Literature Review

Researchers interested in digital forensics have examined browser artifacts and metadata analysis for more than ten years. Studies conducted earlier show that browser-derived information can yield useful insights into user actions and interactions.

Quick and Choo's 2018 study on geolocation in digital forensic investigations found that GPS coordinates extracted from digital artifacts can help investigators pinpoint user movement

behavior and link suspects to geographic locations. This research was mainly conducted using data collected from mobile devices and images.

Martini and Choo's 2020 research study explored forensic challenges faced with web technologies in cloud computing and browser forensics. Their findings revealed that browser optimization, compressed caches, and encoding have become more prevalent, complicating the analysis process. They pointed out the importance of forensic automation tools capable of analyzing vast amounts of browser data.

More recent studies in 2022 were done on browser-based geolocation investigations and forensic automation scripts. These studies found that the use of Python for forensic automation can increase efficiency in analysis, save time, and allow for repetition during artifact investigation. However, there is insufficient literature regarding the extraction of geolocation evidence in browser caches.

There is sufficient evidence in the literature that shows browser caching and metadata extraction techniques are common practices in digital forensics. Still, the combination of these practices, such as browser cache recovery, coordinate extraction, URL parsing, and geographic mapping, requires further exploration.

Methodology

Several scientific works in the field of digital forensics deal with browser artifacts and metadata extraction methods. It is established from earlier studies that user browsing history, cookies, and cache files can reveal some information about the actions performed by the computer owner. From the point of view of image metadata analysis, it has already been proven that information such as EXIF data contains GPS coordinates, timestamps, and other metadata related to device information. Image metadata has become an important source of information in many investigations and forensics.

There is a scientific basis stating that websites visited by users are saved temporarily within the cache folder by browsers. Modern browsers tend to save websites in encoded form and even compress the cache. In addition, Google Maps and similar location-based services rely on map tiles. Map tiles and dynamic rendering systems are used to visualize maps on the browser page, which makes it impossible to find any evidence of GPS in the cache. The main task is to develop an automated method of extracting geolocation information from cache artifacts.

To achieve this goal, the research will combine cache analysis, metadata extraction, URL analysis, and visualization into one methodology.

The above-mentioned forensic methodology was deployed using a systematic framework involving browser cache acquisition, artifact recovery, metadata extraction, coordinate parsing, CSV output creation, and geographic visualization. Initially, the browser cache artifacts were acquired from the Google Chrome cache folder. Cache data obtained from the Chrome cache folder was further analyzed to detect any image files, maps, or URL artifacts that have geographic information.

Artifacts recovered from cache data were filtered to determine whether there are any image files. Metadata of the recovered image files was further analyzed to discover EXIF data and potential GPS coordinates. Map tiles URL links and other location artifacts created by browsers were also analyzed. Python scripts were used to process extracted cache artifacts and automatically parse any coordinates. Geographic data generated in CSV format was later analyzed using an interactive map environment. This methodology allowed systematic processing of geolocation artifacts created by browsers.

1.Forensic Environmental Setup

Controlled conditions were created within the forensic setting to facilitate consistent research and proper storage of the collected evidence. The experiment was performed using a Windows operating system forensic workstation that had all the necessary forensic tools and scripting environment installed.

The software tools and technologies used in this research include:

- Google Chrome browser
- ChromeCacheView browser cache tool
- Python programming language
- Visual Studio Code
- Storage of forensic evidence in CSV files
- Folium geographic data visualization library

A project folder named "forensic automation" was created to contain the collected forensic artifacts, scripts, recovered images, URLs, and forensic evidence output. Dedicated folders were allocated for storing browser cache data, valid images, scripts, URLs, and forensic evidence output, respectively.

Such an organized environment helped to properly collect evidence during the experiment.

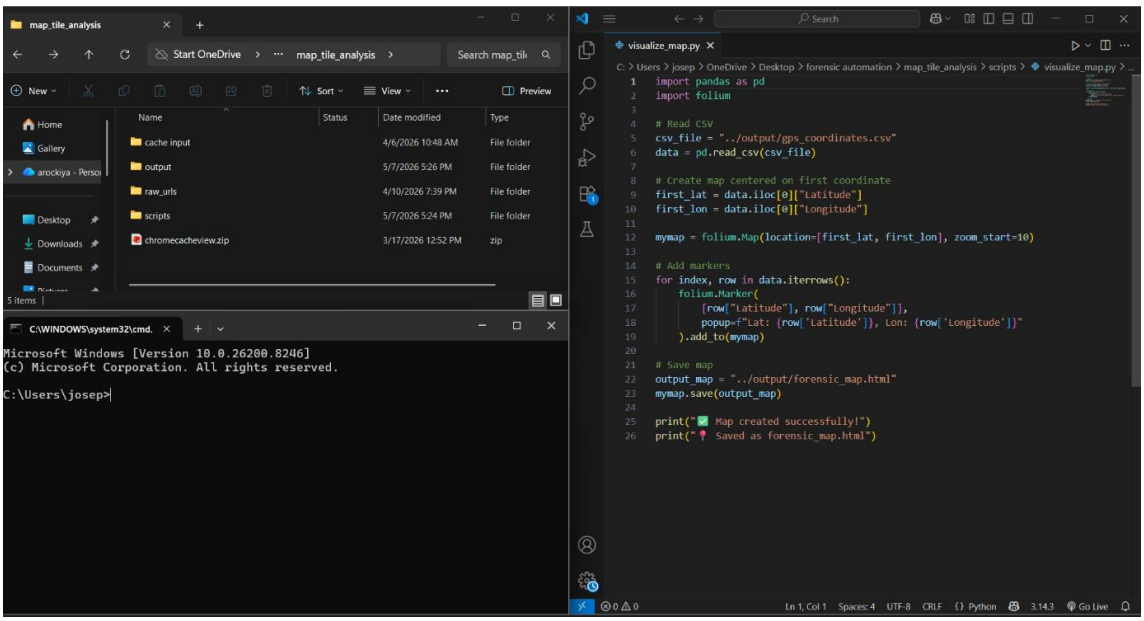

Figure 1: Controlled forensic analysis environment used for browser cache investigation.

2. Browser Cache Acquisition

The browser cache acquisition phase focused on collecting cached artifacts generated during web browsing activity. Google Chrome was selected as the primary browser due to its extensive usage and structured cache storage system.

The browser cache directory was accessed through the local application data path within the Windows operating system. Cache files generated by browser activity were copied into the forensic project environment for analysis.

To obtain artifacts, the following steps were taken:

- Launching Google Chrome and carrying out browser activities.
- Navigating map-related sites.
- Generating artifacts from browsing activity in the browser.
- Locating cache directory of Google Chrome.
- Copying cache artifacts into the forensic environment.

During the acquisition, the following items were collected:

- Image ache
- Thumbnails created by the browser.
- Temporary Internet Files.
- Cache tiles.
- Cache urls.
- Other web resources.

Browser cache provided the first artifact of evidence obtained for the forensic project.

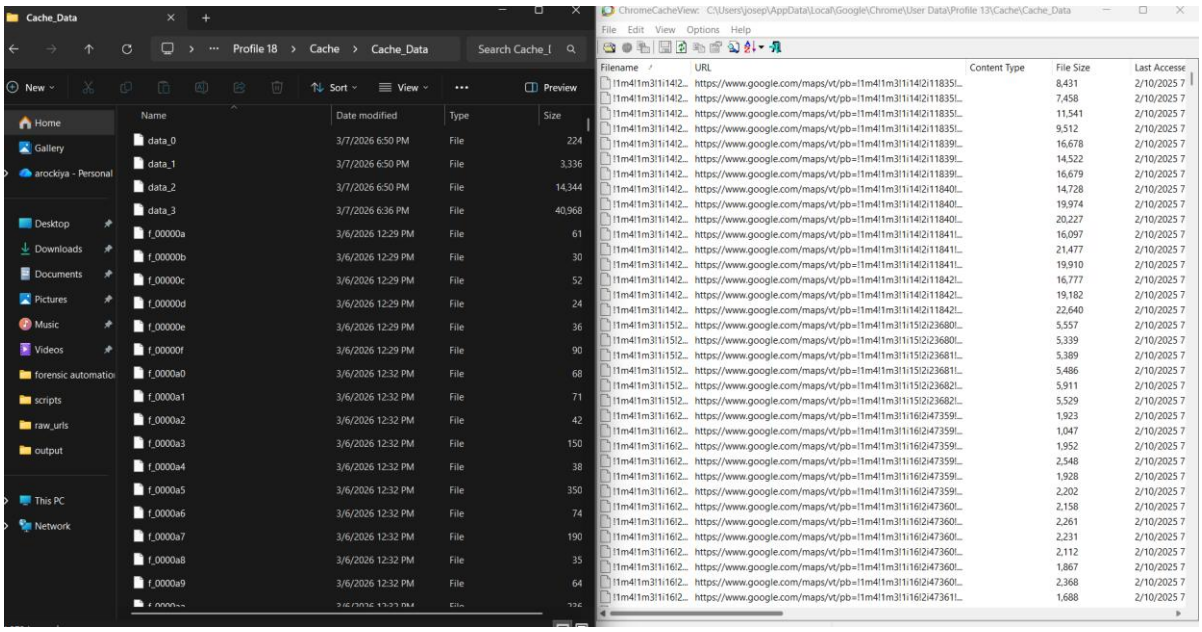

Figure 2: Google Chrome Cache Directory and Chromecacheview Cache Analysis

3. Image Recovery and Filtering

After cache recovery, analysis was conducted to determine which image files could be recovered from the data collected. The browser cache is known to hold both good and corrupt files as well as incomplete fragments of images and optimized thumbnails. Python script was written to facilitate the automated extraction of image files from browser caches. The script scans artifacts from browser caches and locates artifacts that have valid image signatures.

Image recovery entailed the following steps:

- Scanning browser cache artifacts.
- Locating image files.
- Recovering valid image artifacts.
- Removing corrupted artifacts.
- Eliminating thumbnails.
- Storing images in a designated folder.

Recovered images were kept in a folder called “valid_images”.

Analysis proved that many browser cache artifacts were corrupt because of browser optimization and temporary caching techniques. Several valid image files were recovered and kept for analysis.

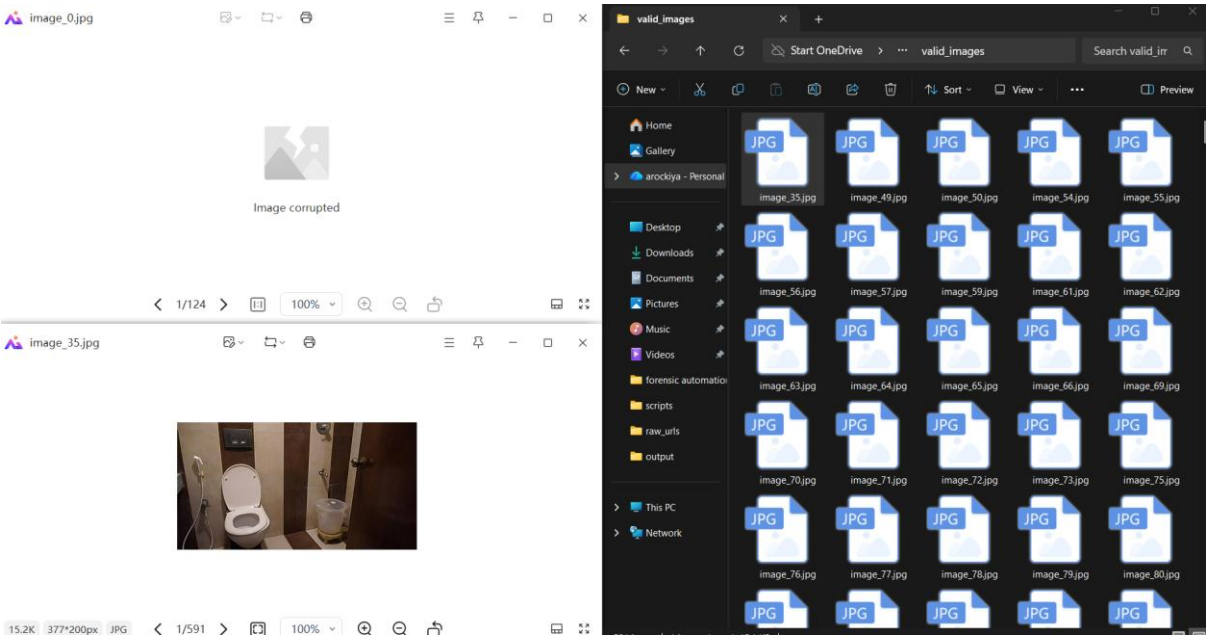

Figure 3: Corrupted and Valid Image Comparison

4. Metadata Extraction

The metadata extraction stage involved searching for embedded EXIF metadata as well as geolocation information present in recovered images. Metadata extraction Python scripts were written to perform automated extraction of EXIF information. Metadata fields in images were analyzed using the scripts.

The metadata extraction stage concentrated on extracting the following metadata items:

- Geographical coordinates
- Time stamps
- Device data
- Camera data
- Creation of images metadata

Results from experimental analysis showed that most images stored in browsers' caches lacked GPS metadata since contemporary webpages strip images of EXIF information when optimizing them online. Nevertheless, metadata extraction was an essential procedure in the forensic process because some images had valuable metadata information. Metadata analysis played a key role in distinguishing between browser-created thumbnails and original image artifacts.

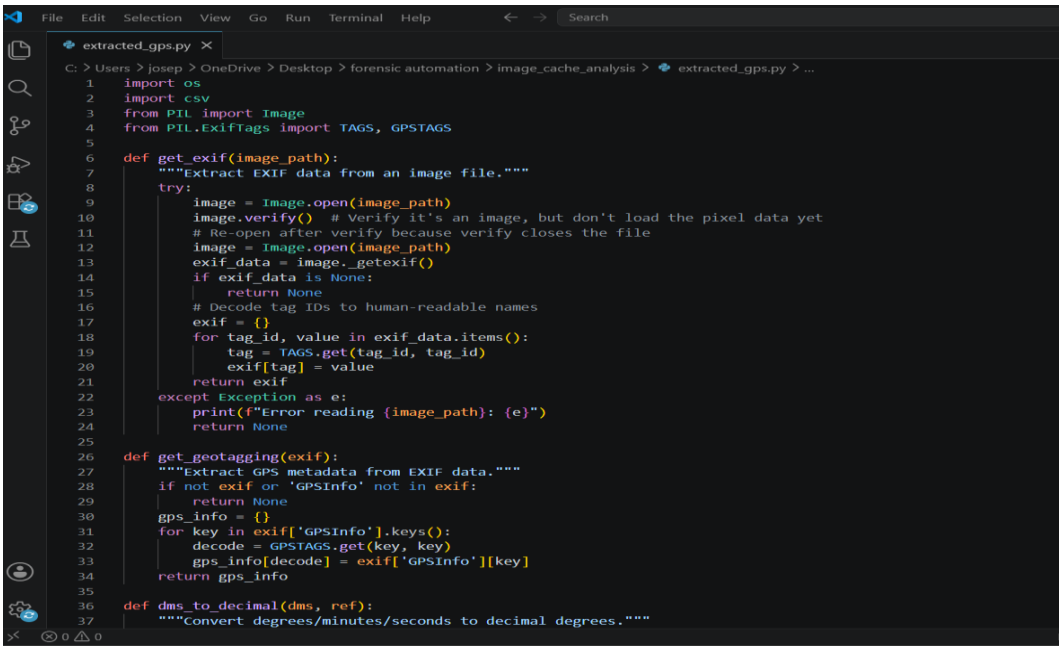

Figure 4: Python-based metadata extraction script used for EXIF and GPS analysis.

5.Map Tile Analysis

Analysis of map tiles was among the most crucial stages in the research methodology. In the experiment, visiting Google Maps created a large number of browser artifacts relating to maps.

ChromeCacheView was employed to detect browser cache artifacts related to geographical browsing activities, and the artifacts recovered were:

- Map tile URLs
- Resources for geographical browsing
- Cache of map icons
- URL structures based on geographical locations
- Geographical references embedded within the URLs

It was found that contemporary maps usually utilize tile structures for encoding geographic resources rather than GPS coordinates. While most of the map tile artifacts did not contain specific latitude and longitude coordinates, some URL structures contained coordinate data useful for forensics.

6.GPS Coordinate Extraction

A custom method of extracting geographic coordinates based on a Python framework was developed to extract geographic coordinates from map-related browser artifacts. Pattern matching based on regular expressions was used to process the extracted URLs and browser artifacts. Methods of URL decoding were also employed to decode browser artifacts.

Coordinate extraction included:

- Processing of browser-generated URLs obtained by extraction.
- URL decoding.
- Regular expressions.
- Detection of latitudes and longitudes.
- Exclusion of redundant coordinates.
- Forensic output generation.

The coordinates extracted during the experiment were automatically recorded using CSV. It was established experimentally that coordinate extraction is possible in cases where browser artifacts include geographic coordinates in URL structures. Nevertheless, significant limitations in terms of forensic investigations were found. Many map tile files included icons and encoded tiles but no coordinates themselves.

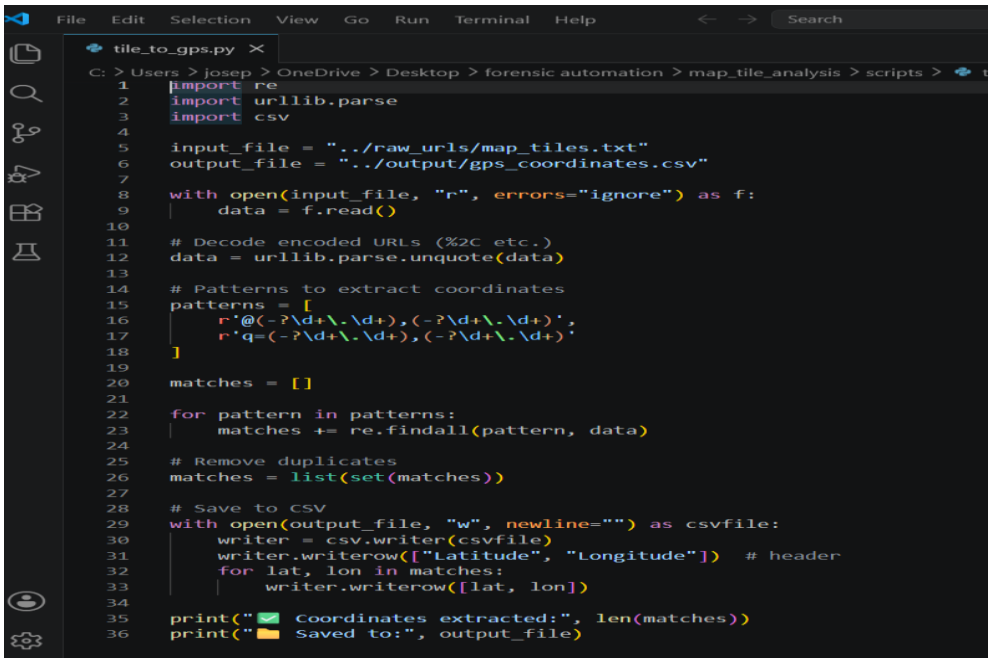

Figure 6: Command-line execution of automated GPS coordinate extraction script.

7.CSV Generation and Evidence Structuring

The obtained GPS coordinates were then transformed into CSV data for better documentation and analysis.

The CSV file consisted of:

- Latitude
- Longitude
- Coordinate structure storage

A CSV format for organizing evidence made it easier to transfer information and utilize it in visualizing software. The structured database made it possible for the investigators to maintain evidence in a readable format.

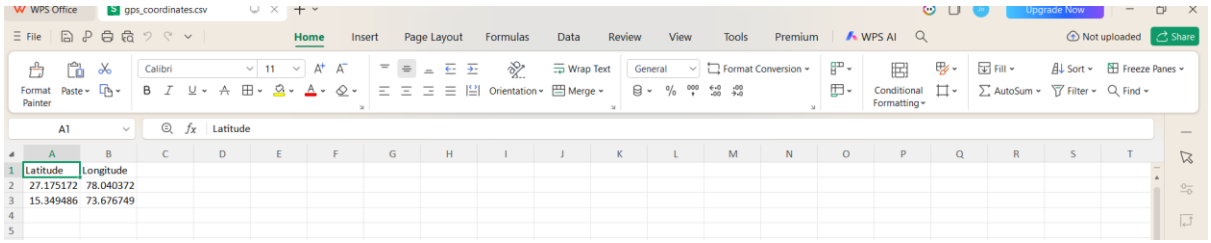

Figure 7: CSV evidence file for geographic analysis

8. Geographic Visualization

Finally, the methodology process was completed with the geographic visualization of the coordinates collected. Visualization using Python-based software packages allowed for plotting the geographic coordinates on an interactive map environment. This visualization helped investigators understand and analyze geographic relations to recover their locations easily.

The steps followed in the visualization process included:

- Loading the CSV coordinates.
- Setting up an interactive map environment.
- Placing geographic markers on the map.
- Producing forensic visualization output.
- Saving the visualization results as an HTML interactive map.

The map visualization formed the last forensic output produced by the methodology process.

Overall, the entire methodology revealed how useful geolocation evidence could be obtained from browser cache artifacts using forensic methods and automated extraction tools.

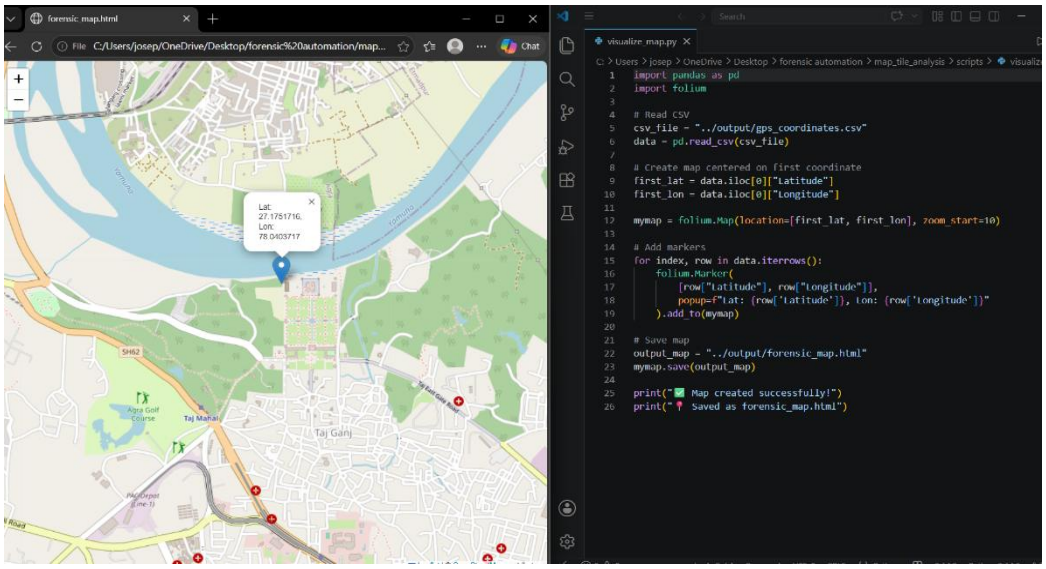

Figure 8: Interactive geographic visualization of extracted GPS coordinates.

Results and Analysis

LATITUDE	LONGITUDE
27.1751716	78.0403717
15.3494864	73.6767485

Table 1: GPS coordinates extracted from browser cache artifacts.

First, it was found out that browser cache artifacts can be used as an additional source for geolocation related information provided that the forensic process is carefully organized and conducted in special conditions. During research, browser cache artifacts that have been generated by Google Chrome browser have been acquired and thoroughly analyzed through special forensic tools and Python code.

The recovery process involved identification of some valid images. Nonetheless, quite a considerable part of recovered artifacts had turned out to be corrupted or partially damaged, whereas some other images were optimized by the caching process. It should be mentioned that analysis of metadata for those artifacts showed that modern platforms optimize images that do not possess any metadata anymore.

On the other hand, metadata extraction process helped to recover quite useful metadata elements including information related to software and images creation time. Analysis of browser artifacts resulted in identification of numerous valid map tiles which were analyzed later. The analysis was based on usage of ChromeCacheView and some automated Python scripts for URLs decoding and geocoding process.

Finally, the developed methodology allowed extracting valid pairs of GPS coordinates from browser artifacts. The extracted coordinates were automatically structured and stored in CSV

format for further forensic processing. The generated CSV dataset improved evidence organization and enabled efficient geographic analysis.

Discussion

The experiment findings show the forensic potential of artifacts produced by browsers as an important source of geolocation data. Nowadays, modern browsers work constantly with online services and create a huge amount of data stored in their cache. Unfortunately, most people do not know that they leave such kind of digital footprint on the Internet. One of the major findings of the current research is that, despite the absence of direct GPS data, artifacts produced by browsers may give investigators an indirect hint about user's interaction with some specific location using the cached map images and other data associated with geographic coordinates.

During the image reconstruction stage of the process, researchers found that the vast majority of image artifacts stored on the browser cache directory is either damaged or incomplete or does not contain necessary information to restore images successfully. During the experiment, more than 800 image artifacts have been found, but not all of them could be recovered successfully. Moreover, metadata extraction tests showed that the vast majority of browser-stored image artifacts no longer contain information about their GPS coordinates, because modern web platforms delete such metadata when optimizing images for faster loading. Nevertheless, in spite of these shortcomings, the research revealed that geographic coordinates may be discovered by means other than simple browser cache examination; for example, URL parsing and map-related cache analysis. During the test, map-related browser artifacts with coordinate-based URLs were discovered and analyzed using automated Python scripts.

It was determined that regular expression-based techniques may prove to be useful when searching for latitude and longitude information embedded within browser artifacts. Automation helped decrease the manual work involved, while also increasing repeatability of the process. Generated CSV files and maps clearly illustrated the potential usefulness of the technique. With the help of geographic visualization, investigators may identify locations of interest much more easily and establish geographic connections between them.

One of the most important findings of this research is the discovery of forensic limitations related to the analysis of browser cache data to establish geolocation information. The investigation proved that in contemporary browsers there are numerous map tile structures, compressed resources, and caching processes that prevent investigators from analyzing geolocation information from browser caches.

The results of the experiment also imply that when conducting forensic investigation, experts must correlate the use of browser cache analysis with other evidence sources. Browser data can serve as supplementary evidence but not a single piece of evidence confirming physical presence. One of the strengths of this study lies in its automation capability. Using the designed scripts, investigators can quickly filter, extract, parse, and visualize browser cache artifacts. Automation saves time and effort and enables processing vast amounts of data.

In conclusion, the suggested approach has proven successful in showing that browser cache artifacts have forensic value when conducting geolocation investigations. Despite certain

limitations, combining cache analysis, metadata extraction, URL parsing, and visualization offers a solid forensic technique for investigators.

Limitations

Limitations for this study included the following:

- Web browsers today typically strip or reduce metadata.
- Cache entries are prone to being modified or deleted.
- GPS coordinates cannot be reliably found within caches.
- Map tiles have encoded structure that makes coordinate recovery difficult.
- Findings are contingent on the user's browsing activities.

The above limitations point towards the need for more studies about browser artifact analysis.

Conclusion

This study has suggested and validated a new forensic approach in obtaining geolocation data using browser cache artifacts. The new forensic approach includes browser cache collection, image processing, metadata analysis, coordinate extraction, and map generation.

Findings revealed that geolocation can be obtained from browser caches under laboratory settings. This is shown through the capability of the automation scripts in extracting GPS coordinates from coordinate-based URLs.

However, it was also discovered that modern browser caching systems have several restrictions such as the use of coded map tiles and lack of GPS metadata in most artifacts.

The methodology presented in this paper provides valuable insights on the use of browser caches as sources of geolocation data around digital forensics.

Future work

Future enhancements may include:

- Support for multiple web browsers
- Automated KML/KMZ extraction
- Advanced tile decoding techniques
- Timeline-based geographic reconstruction
- Integration with forensic suites
- AI-assisted geolocation inference
- Real-time browser artifact monitoring

Reference

Adelstein, F. (2006). Live forensics: Diagnosing your system without killing it first. Communications of the ACM, 49(2), 63–66.

Al Mutawa, N., Baggili, I., & Marrington, A. (2012). Forensic analysis of social networking applications on mobile devices. Digital Investigation, 9, S24–S33.

Beebe, N. L., & Clark, J. G. (2005). A hierarchical, objectives-based framework for the digital investigations process. Digital Investigation, 2(2), 147–167.

Carrier, B. (2005). File System Forensic Analysis. Addison-Wesley Professional.

Casey, E. (2011). Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet. Academic Press.

Choo, K. K. R. (2011). Cloud computing: Challenges and future directions. Trends & Issues in Crime and Criminal Justice, 400, 1–6.

Garfinkel, S. (2010). Digital forensics research: The next 10 years. Digital Investigation, 7, S64–S73.

Google Developers. (2023). Google Maps Platform Documentation. Retrieved from Google Developers documentation.

Grispos, G., Storer, T., & Glisson, W. B. (2013). Calm before the storm: The challenges of cloud computing in digital forensics. International Journal of Digital Crime and Forensics, 5(2), 28–48.

Jones, K. J., Bejtlich, R., & Rose, C. W. (2006). Real Digital Forensics: Computer Security and Incident Response. Addison-Wesley.

Kent, K., Chevalier, S., Grance, T., & Dang, H. (2006). Guide to Integrating Forensic Techniques into Incident Response. National Institute of Standards and Technology (NIST).

Kohn, M., Eloff, M., & Olivier, M. (2013). Framework for a digital forensic investigation. ISSA Journal, 1(1), 1–7.

Martini, B., & Choo, K. K. R. (2012). An integrated conceptual digital forensic framework for cloud computing. Digital Investigation, 9(2), 71–80.

Mohay, G., Anderson, A., Collie, B., De Vel, O., & McKemmish, R. (2003). Computer and Intrusion Forensics. Artech House.

Microsoft Documentation. (2023). File System and Cache Storage Documentation. Microsoft Learn.

NirSoft. (2024). ChromeCacheView Utility Documentation. Available from NirSoft official documentation.

Palmer, G. (2001). A road map for digital forensic research. Digital Forensic Research Workshop (DFRWS).

Pollitt, M. (2007). An ad hoc review of digital forensic models. Digital Investigation, 4(1), 43–54.

Python Software Foundation. (2024). Python Documentation. Available at Python official documentation.

Quick, D., & Choo, K. K. R. (2014). Digital forensic intelligence: Data subsets and open source intelligence (DFINT+OSINT): A timely and cohesive mix. Future Generation Computer Systems, 29(2), 634–644.

Reith, M., Carr, C., & Gunsch, G. (2002). An examination of digital forensic models. International Journal of Digital Evidence, 1(3), 1–12.

Rogers, M. (2006). Computer forensics: Principles and practices. Digital Investigation, 3(1), 3–19.

Roussev, V. (2009). Building efficient digital forensic tools. Digital Investigation, 6, S146–S153.

Satvat, K., Saxena, N., & Stavrou, A. (2014). Automated digital forensic analysis of web browser artifacts. Proceedings of the IEEE Conference on Communications and Network Security, 68–76.

Zawoad, S., & Hasan, R. (2013). Digital forensics in the cloud. ACM Computing Surveys, 47(2), 1–36.