Forensic Analysis of Data Leakage Indicators in Mobile Gemini User Content Interactions*

Yewon Kim[1], Minjung Yoo[1], Seunghyun Park[1], and Seongmin Kim[1,*]

Sungshin Women's University, Seoul, Korea
{20240932, 20211079, 20211058, *sm.kim*}@sungshin.ac.kr

Abstract

Generative AI poses risks of confidential data exposure, as user inputs and files are processed on external servers. In particular, Google's Gemini integrates with Google Workspace to manage files and content; however, forensic insight into these interactions remains limited. This study proposes a forensic methodology that analyzes artifacts generated during Gemini's file and content interactions to trace potential pathways of confidential data leakage.

Keywords – Mobile Forensics · Google Gemini · Digital Trace Analysis · Google Workspace

1 Introduction

Generative AI processes user inputs on external servers, posing risks of unauthorized exposure when sensitive information becomes part of model training data. Artifacts generated during file uploads or image creation may persist on user devices or associated cloud services, increasing the likelihood of confidential data leakage. For example, Samsung Electronics reported an incident in which internal source code was inadvertently uploaded to ChatGPT, raising concerns about the external transmission of proprietary data in 2023 [1].

To address these concerns, this study analyzes ADB Logcat and Google Takeout data generated during file and content interactions in the Gemini mobile app. It identifies residual artifacts left during user data transmission and content generation, demonstrating how these traces can support data leakage detection and behavioral reconstruction from a digital forensic perspective.

2 Identifying Artifacts in Gemini Mobile

In this study, a Samsung Galaxy Z Flip3 device was used to execute Gemini's attachment summarization, Google Drive document processing, and image generation functions. During the experiments, system logs were collected via ADB Logcat and Gemini activity records were extracted in JSON format using Google Takeout.

Information Observable via Logcat. The ADB Logcat analysis reveals how the Gemini app interacts with the device file system, network, and linked cloud services during file and content operations. Representative examples from the collected logs are summarized in Table 1. Log 1 illustrates how Gemini accessed and uploaded a PDF file from the device's local storage. The `MediaProvider` tag appears when Android interacts with the file system, and the path "`/storage/emulated/0/Download/namefile.pdf`" confirms direct file access and transmission from

[*] Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-11, December 16-18, 2025, Sapporo, Japan. © The copyright of this paper remains with the author(s).

the local device. Also, Log 2 shows the transmission of user queries and attachments to Gemini. During this process, Gemini aggregated attachment information by source and type; however, the number of files was masked, preventing identification of specific attachments.

Log 3 captures Gemini's integration with Google Workspace for Drive-based document processing. Immediately after login, the Cello engine—Google Workspace's synchronization module—was activated, and Gemini issued duplicate responses to the same query. Note that no direct API calls to Google Drive were observed, suggesting that the Android system pre-established the session during authentication, which Gemini later leveraged for Workspace access. Finally, Log 4 records repeated `OpenGLRenderer` entries following an image-generation request, indicating GPU rendering during content creation or preview. Although no explicit image-generation log appears, this pattern strongly implies that Gemini performs image creation internally.

Table 1: Gemini related log entries extracted from ADB Logcat

No	**log detail**
1	MediaProvider: Open with lower FS for /storage/emulated/0/Download/namefile.pdf. Uid: 10302
2	arjl : #sendRobinQuery arjl : (REDACTED) Logging streamz with featuremode: %s, fileCount: %s, localImageCount: %s, driveImageCount: %s, ...
3	Cello : [32053:NonCelloThread] content_cache.cc:425:InitializeOnIOThread Cache index at: /data/user/0/com.google.android.apps.docs/app_cello/...
4	OpenGLRenderer: — Failed to create image decoder with message 'unimplemented'

Information Observable via Google Takeout. The Gemini activity data extracted through Google Takeout contained multiple artifact types, including JSON, PDF, and JPG files. These files are organized under the "`/Takeout/My Activity/Gemini`" directory. Among them, MyActivity.json serves as the primary record, storing Gemini usage history in a structured format [2]. It includes user queries, precise transmission timestamps, and Gemini's responses, as well as metadata for both attached and generated files [3], enabling the identification of original content exchanged during conversations.

3 Conclusion

This study identified digital artifacts generated during the attachment and creation of user data within the Gemini application, providing forensic evidence for detecting data leakage and reconstructing user behavior. The findings are expected to support the forensic reconstruction of generative AI usage patterns and contribute to the identification of potential confidential information exposure.

References

[1] Muhammad Arsal, Bilal Saleem, Sommia Jalil, Muhammad Ali, Maila Zahra, Ayaz Ur Rehman, and Zia Muhammad. Emerging cybersecurity and privacy threats of chatgpt, gemini, and copilot: Current trends, challenges, and future directions. 2024.

[2] Kyungsuk Cho, Yunji Park, Jiyun Kim, Byeongjun Kim, and Doowon Jeong. Conversational ai forensics: A case study on chatgpt, gemini, copilot, and claude. *Forensic Science International: Digital Investigation*, 52:301855, 2025.

[3] Sonali Tyagi, Yufeng Gong, and Umit Karabiyik. Forensic analysis and privacy implications of llm mobile apps: A case study of chatgpt, copilot, and gemini. *Forensic Science International: Digital Investigation*, 54:301974, 2025.