Identification of WhatsApp Digital Evidence on Android Smartphones using The Android Backup Application Package Kit (APK) Downgrade Method

Sulisdyantoro Deny*, Marza Ihsan Marzuki
Department of Electrical Engineering, Faculty of Engineering, Universitas Mercu Buana, Indonesia

Abstract
The use of WhatsApp for actions that lead to unlawful acts is a serious matter that needs to be proven in court. Android and the WhatsApp messaging application continue to update their features and security to provide maximum service and protection to its users, such as the WhatsApp database encryption using crypt14. With crypt14 encryption on the WhatsApp database, investigations of WhatsApp digital evidence against Electronic Evidence (BBE) require an acquisition and extraction method to identify artefacts relevant to digital evidence needs. The National Institute Standard Technology (NIST) reference methodology, from the collection, examination, and analysis to reporting stages, has become a widely used framework for digital forensics against BBE. The Android Backup Application Package Kit (APK) Downgrade method can decrypt the WhatsApp database crypt14 to become a solution that can be used in the framework of mobile forensics to answer the needs of investigations into certain criminal cases, including data that users have deleted. With the Cellebrite tools, the Android Backup Application Package Kit (APK) Downgrade method can identify approximately 651% more artefacts than the Android Backup and logical acquisition methods using the FinalData and MobilEdit tools.

Keywords:
Android; APK downgrade; Cellebrite; Smartphone; WhatsApp;

Article History:
Received: July 27, 2022
Revised: February 23, 2023
Accepted: February 28, 2023
Published: March 25, 2023

Corresponding Author:
Sulisdyantoro Deny
Department of Electrical Engineering, Faculty of Engineering, Universitas Mercu Buana, Indonesia
Email: deny.sulisdyantoro@gmail.com

INTRODUCTION
Currently, the role of the internet is increasingly important in the global world's social, economic and political life. Indonesia also experienced this increase. The survey results of the Indonesian Internet Service Providers Association (APJII) 2019-2020 Q2 [1], the penetration of the number of internet users in Indonesia is 196.71 out of a total population of 266.91 million Indonesians or around 73.7%. And in the survey, in 2019, smartphone devices dominated internet use, with a percentage of 95.4%. Instant Messenger is an application often used to communicate, replacing the role of Short Message Services (SMS) [2]. WhatsApp is one of the most popular instant messenger applications and can be used on mobile devices and computers [3]. The number of smartphone users is increasing, especially on the Android platform [4].

WhatsApp has many features, such as telephone, group chat, video calling, file sending, and voice messaging [5]. WhatsApp places the messaging app far ahead of all other messaging apps like Facebook Messenger, WeChat, Viber, Apple Business Chat, or Telegram regarding user count. WhatsApp has many features, such as telephone, group chat,
messaging, video calling, file sending, and voice messaging. WhatsApp has become a reference in digital forensics in Indonesia [6]. Figure 1 shows that WhatsApp is in the top three as a widely used platform.

Android smartphone forensics has evolved, offering significant opportunities and exciting challenges. Some of these crimes include using the sophistication of the Android smartphone to commit crimes such as fraud, gambling, pornography, corruption, drug networks, to murder cases. In several recent crime cases, such as in the trial of Nainggolan's investigation [7] and in a follow-up hearing on the false news of Sarumpaet. [8], using WhatsApp conversations as evidence in court. This shows that from a forensic investigation perspective, the WhatsApp application can store evidence data that can be used in court as evidence. Therefore, using the Forensic approach, it is very important to have a methodology and framework to parse WhatsApp application data on both active and deleted Android devices. This study will describe the mechanism for opening database files and backup files from the WhatsApp application on an Android smartphone. This research is expected to contribute to a framework that can be applied to perform smartphone forensics by finding and analyzing artefacts in the WhatsApp application on Android smartphones using the Android Backup APK Downgrade acquisition method. With the forensic approach used, it is not only possible to restore existing conversations and data but also data and conversations that have been deleted.

MATERIAL AND METHODS

Related Work

Several previous studies to identify WhatsApp digital evidence have been carried out previously on Android-based smartphones. Zhang [9] explained that there is no encrypted database on a rooted phone while the database is encrypted on a non-rooted phone. Rahadhian [10] using the Integrated Digital Forensic Investigation Framework (IDFIF) version 2 (two) and Oxygen Forensic Suite, MobilEdit Express, and Andriller tools, can identify WhatsApp Desktop content that contains cybercrime. Unlike the case with Umar [3] comparing the Belkasoft Evidence tools, WhatsApp Key DB/Extractor uses the NIST framework to make acquisitions with physical and logical methods. As a result, Belkasoft Evidence has the highest index number, and WhatsApp Key/DB Extractor is superior in terms of cost. At the same time, Oxygen Forensic is superior in obtaining WhatsApp artefacts through logical and physical acquisitions.

Yuliani [11], using Oxygen forensics, and Andriller obtained evidence of smartphone artefacts such as chat sessions, avatars, and contacts in the WhatsApp application.

![Figure 1. WhatsApp User Statistics](image-url)
Then, Widiandana [12] acquired with Oxygen Forensic, analyzing the acquisition results using Text Mining, Cosine Similarity and NIST framework methods. It was found that there are similarities between words in digital evidence and words that are negative will help identify cyber bullying actions.

Several studies using the FTK Imager tool by Marfianto [13] have obtained artefacts in the form of WhatsApp text message chat conversation sessions and can also get other media files encrypted by crypt12. Riadi [14], with the live forensic method on WhatsApp desktop, obtained digital evidence in the form of texts of WhatsApp conversations that occurred between the suspect and the victim, which can be used as digital evidence related to the online shop fraud case that occurred. Then Kumang [15] by using ProDiscover Basic Tools, AccessData FTK Imager, WhatsApp Viewer, and DB Browser for SQLite and WhatsApp encryption on crypt8, getting evidence artefacts in the form of chat sessions, avatars, contact numbers on the WhatsApp application, voice notes, profile photo, the identity of the owner of the WhatsApp account and also can get other media files and most importantly encrypted backup database files. Meanwhile, Saputra [16], with FTK Imager and live forensic methods, used the extracted data from the acquisition, explored the characteristics of WhatsApp Messenger users according to the chat content and labelled using the crowdsourcing method. Finally, Wirara [17] used XRY and Encase Mobile Forensic to obtain WhatsApp digital evidence even though the device was not rooted/jailbroken first.

Anggraini [18] uses recovery tools Dr Fone for Android to recover deleted WhatsApp messenger data, but the media URL cannot be opened because it is encrypted by WhatsApp, while WhatsApp messenger database analysis uses DB Browser for SQLite. On the other hand, a study by Akbar [19] explained that it could generate the required data, such as log timestamps, photos sent, call logs, and messages sent and received. Tools used Using Internet Protocol with Wireshark and Live Memory. In contrast to Mirza [20] who proposed the possibility of anti-forensic techniques for the WhatsApp application through two hypothetical case scenarios, namely by disabling the delete feature for everyone when the sender is blocked from the contact list and conducting an application demo that can recover deleted messages even after the status is blocked.

Meanwhile, for research using Cellebrite UFED tools, Hussein [21] conducted experiments on Android smartphones with internal storage, no-rooting and Cellebrite UFED tools. This study explains how to extract the Crypt key from the WhatsApp application to decrypt the database and extract artefacts on the android system without rooting the device. With these conditions and the development of WhatsApp application updates to improve services, features and security, it is inversely proportional to mobile forensics tools which have limitations with the large variety of Android phones with the development of each of the modifications of the Android version according to the characteristics of each brand. So in this study, it is proposed to use the Android APK downgrade backup method by extracting WhatsApp application data using Android backup (android backup) which contains the Android file system. In the acquisition process, the version of the WhatsApp application used (*.apk file) is temporarily downgraded to the previous version so that data can be extracted. Then the version used will be restored at the end of the extraction process. Thus, using the Android APK Downgrade backup method through the extraction process of the database encrypted by WhatsApp can be made possible to run after the WhatsApp version is downgraded. The acquisition results will be analyzed using the Cellebrite Physical Analyzer tool.
Material
The specifications of the software used for digital evidence analysis, in this case, are as follows: Android Smartphone Samsung Galaxy Note 9 (OS. Android 10), Data cable, Windows 10 and platforms for acquisition and analysis, Cellebrite UFED version 7.50.0.137 for acquisition, Cellebrite Physical Analyzer version 7.52.0.36 for analysis, and WhatsApp application version 2.22.3.75. Table 1 lists the devices used.

Methods
The National Institute of Standards and Technology (NIST) is a non-regulatory body of the Technology Administration section of the United States Department of Commerce, which issues the publication of NIST SP 800-86, which divides the process of digital forensics into four stages. The four stages are as follows [27, 28, 29, 30]:

Figure 2 shows the stages of the digital forensic process, starting from collection, examination, analysis and reporting. The collection is the first stage in the digital forensics process to identify potential data sources before acquiring the data. At this stage, physical and digital evidence will be safeguarded, and its integrity will be ensured at every stage of the investigation process by following established guidelines and procedures. Identification by preserving BBE is the most important procedure in digital forensic investigations [9]. Forensic investigators must document every event and activity without or with little change to the evidence in digital forensic rules. Every change must be well documented so that the validity and integrity of the evidence are maintained and can be accounted for in court. The examination is the stage of obtaining evidence in accordance with standard procedures, which is preceded by examining the data, which includes the assessment and extracting of information from previously collected data by using a combination of automatic and manual methods to assess and extract interesting data while maintaining data integrity. The analysis is the process of analyzing data that can identify people, places, items, and related events so that conclusions can be drawn by interpreting the digital evidence that has been identified. Unreadable digital evidence is secured or used directly for presentation in court. Therefore, forensic investigators need to use forensic tools to analyze the collected data [9]. This stage aims to obtain useful information to answer the questions that are the driving force in conducting the collection and examination, namely by analyzing the data obtained in the previous stage to identify the source of the crime and motive using justifiable methods/techniques. Legally, supported by adequate/qualified evidence, it can be accepted and ultimately prove the person responsible for the crime and/or deny the alleged crime.

Table 1. Simulation device

<table>
<thead>
<tr>
<th>NO</th>
<th>Simulation device</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Samsung Galaxy Note 9 SM-N960F</td>
<td>Android 10</td>
</tr>
<tr>
<td>2</td>
<td>WhatsApp</td>
<td>Instant messaging application version 2.22.3.75</td>
</tr>
<tr>
<td>3</td>
<td>Workstation</td>
<td>Microsoft Windows 10 Pro 64 Bit, Intel i9-7900X, 64 GB RAM</td>
</tr>
<tr>
<td>4</td>
<td>USB Cable</td>
<td>connect a smartphone to a workstation with a USB type C cable</td>
</tr>
</tbody>
</table>

Figure 2. Number and Figure Caption
Finally, Reporting is preparing and presenting the information from the analysis phase. This final phase reports the results of the analysis, describes the actions used, explains how tools and procedures were selected, determines what other actions need to be taken (e.g., forensic examination of additional data sources, securing identified vulnerabilities, improves existing security controls), and provides recommendations for improvement of policies, guidelines, procedures, tools, and other aspects of the forensic process [27, 28, 29, 30].

RESULTS AND DISCUSSION

This research was conducted using an Android smartphone installed with the WhatsApp messaging application version 2.22.3.75. In the scenario that investigators find evidence of a Samsung Galaxy Note 9 smartphone with an unrooted version of Android 10 used in this research.

Simulation

This research simulation was carried out at the Indonesian Attorney's Digital Forensic Laboratory. In the simulation, it takes an Android smartphone device, Samsung Galaxy Note 9, installed the WhatsApp messaging application version 2.22.3.75 then, devices communicate and send chats, pictures and video calls. The next stage is to carry out the acquisition and extraction of the two smartphone devices using a PC or laptop with Windows OS. After imaging, the analysis process is carried out, and then a report is made on the evidence.

Stages of collection

At the collection stage, the initial stage in the mobile forensics’ method, what is being done is to search, collect and document evidence [31, 32, 33]. For testing the research that became the sample of the evidence analyzed was in the form of two smartphones which were screened as evidence in a crime case. Both smartphones are not rooted with the condition that the password security feature is active and the screen security is active. At this stage, documentation of matters relating to the smartphone is carried out. The specifications for evidence are listed in Table 2.

In addition to collecting and documenting, preparation and planning are also carried out on how the smartphone will be analyzed and what tools and tools are needed to support the process. In the design of the framework flow for analyzing smartphones to get digital artefacts related to the WhatsApp application in the condition of both smartphones that are not rooted with active screen security features. The workflow complies with forensic rules by taking steps that minimally alter the evidence.

For testing, an Android Smartphone is used in an unrooted state where some of its features have been disabled to prevent the user from damaging the operating system. The rooted state can remove these limitations so that full access to the system is allowed. For the condition of a rooted Android phone, the user will have more control over the settings, features and performance so that the process of accessing system files for forensic analysis will be easier. However, for forensic procedures on unrooted Android phones, it is recommended to avoid rooting permanently because it is very risky to change the evidence and can cause data to be overwritten.

<table>
<thead>
<tr>
<th>MERK</th>
<th>SERIES</th>
<th>MODEL</th>
<th>ANDROID OS</th>
<th>VERSION WHATSAPP</th>
<th>CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP1</td>
<td>Samsung Galaxy Note 9</td>
<td>SM-N960F</td>
<td>10</td>
<td>2.22.3.75</td>
<td>unrooted</td>
</tr>
</tbody>
</table>

S. Deny & M. I. Marzuki, Identification of WhatsApp Digital Evidence on Android Smartphones … 11
Examination Stages
At this stage, the acquisition and extraction process for internal and external memory from smartphones. Early Detection of Smartphones with Cellebrite UFED, as shown in Figure 3, using a console and a USB data cable. In carrying out the data collection process, the method used is an android backup and Android APK Downgrade backup. The tools used are MobilEdit, FinalData, and Cellebrite UFED. This process allows physical extraction by bypassing the decrypting bootloader (BTL) to make acquisitions with the android backup method using the Android Debug Bridge (ADB) as a command line tool in Cellebrite UFED that allows communicating with devices.

In Figure 4, it is explained that the acquisition and extraction process on Samsung smartphones is divided based on the chipset used, namely Qualcomm, Exynos and Generic chipsets. The Samsung Galaxy Note 9 smartphone uses the Exynos chipset. The adb command facilitates various device actions, such as installing and debugging applications. To use adb with devices connected via USB, USB debugging is enabled in the device's system settings, in the Developer options section. On Android 4.2 and up, the Developer options screen is hidden by default. To make it visible, go to Settings > About phone, then tap Version number seven times. Return to the previous screen to find the developer options at the bottom. Setting parameters on the smartphone before starting the acquisition and extraction process is done by ensuring that the display screen is always ON, the password, pin, and biometric are inactive, developer options are ON and USB setting is in data transfer condition.

Acquisition with an android backup method
To perform internal memory imaging on the two smartphones, researchers used the Cellebrite UFED tool, which will update the Android bootloader to retrieve data on the Android system partition and internal memory without having to root, and enable USB Debugging. The ongoing acquisition process is shown in Figure 5, which shows that the WhatsApp_backup.ab file is being acquired.

![Device Information](image1)

![Acquisition and extraction process on android Samsung exynos](image2)

![File list acquisition result of Android backup method](image3)
The acquisition results in Figure 5 show that three files were acquired using the Android backup method. A file with the type UFED Dump with a size of 9,505,848 KB will be extracted using the Cellebrite Physical Analyzer for later analysis of its artefacts.

**Acquisition with Android APK Downgrade backup method**

The acquisition process is carried out by first lowering the WhatsApp APK version to the standard version where the encrypted database file from WhatsApp can be decrypted and then the backup process is carried out. The APK version of WhatsApp installed on the smartphone will be saved in anticipation of failure during acquisition. Figure 6 shows that the WhatsApp apk is temporarily downgraded and the WhatsApp version installed (2.22.3.75) on the smartphone is stored on the Cellebrite device as a backup apk which can be downloaded for needs if there are problems during the acquisition process.

The acquisition results in Figure 6 show that two files were acquired using the Android APK Downgrade backup method. A file with the type UFED Dump with a size of 9,578,479 KB which will be extracted using the Cellebrite Physical Analyzer for later analysis of the artefacts contained in it. The dump file is larger than the file acquired by the android backup method.

**Extract WhatsApp Data from Image Data**

The steps to extract data from the image are to take data from the external and internal memory backup images for the target smartphone. The data is stored in a folder labelled according to the backup date. The data sought is a dump file resulting from the acquisition process, as shown in Figure 5 dan Figure 6, which is then extracted. The extract results were analyzed to export the WhatsApp folder and the com.WhatsApp folder. To decrypt the encrypted database using the Cellebrite Physical Analyzer application, some artefacts can be seen in Hex View. The whole process is carried out by a hashing mechanism to maintain the integrity of the digital data. Figures 7 show the hash value using the SHA 256 algorithm for files acquired using the android backup method and the Android Backup APK downgrade method.

The results of the acquisition and extraction with the Android backup method cannot see the contents of WhatsApp communications on smartphones due to the ability of the tools that can only describe the WhatsApp database up to a certain version that the mobile forensics device can decrypt [34, 35, 36, 37].

Figure 8 described that this research was carried out using the acquisition method by downgrading the application version where WhatsApp installed on the smartphone (version 2.22.3.75).

<table>
<thead>
<tr>
<th>Name</th>
<th>Date modified</th>
<th>Type</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung GSM_SM-N960F Galaxy Note 9.udd</td>
<td>2/2/2022 23:48</td>
<td>UFED Dump</td>
<td>1 KB</td>
</tr>
<tr>
<td>Samsung GSM_SM-N960F Galaxy Note 9.zip</td>
<td>2/2/2022 23:47</td>
<td>WinRAR ZIP archive</td>
<td>9,578,479 KB</td>
</tr>
</tbody>
</table>

Figure 6. File list acquisition result of Android Backup APK Downgrade method

Figure 7. The acquisition hash value of android backup method and APK Downgrade
Figures 9 explained that the extraction file category is data content that can be extracted for further analysis according to the needs of digital evidence investigations. There is a clear difference in the extraction results of the two methods, where chat content is only found in the Android Backup APK Downgrade method.

Analysis Stages
This stage aims to uncover and analyze the results of the acquisition stage to obtain data related to the WhatsApp application. In this study, the Cellebrite Physical Analyzer tool is used to analyze the imaging results that have been carried out previously.

Call Log
Figure 10 shows that by analyzing the artefacts in the call log category, you will see the caller, recipient, timestamp, duration, status and source of file information extracted from the WhatsApp database msgstore.db.

Figure 8. WhatsApp APK downgrading process

Figure 9. Acquisition results with an android backup method and APK downgrade

Figure 10. Data Analysis on Call Log Category
Chats

Figure 11 shows that by analyzing the artefacts in the chats category, participants, timestamp, last activity, and source of file information extracted from the WhatsApp msgstore.db database will be seen.

Contacts

Figure 12 shows that by analyzing the artefacts in the contacts category, names, phone numbers, e-mail, participants, notes, sources, accounts and sources of file information extracted from the WhatsApp database msgstore.db are shown.

Device Locations

Figure 13 shows that by analyzing the artefacts in the device locations category, timestamps, address GPS coordinates, and file information sources are extracted from the WhatsApp database msgstore.db.

Web History

In Figure 14, it is shown that with the analysis of artefacts in the web history category, when was the last time the url was accessed, and the source of the file information extracted from the WhatsApp database msgstore.db.

Figure 11. Data Analysis on Chats Category

Figure 12. Data Analysis in Contacts Category
Audio

Figure 15 shows that the artefact analysis in the audio category will show file names, storage locations, file sizes, modified timestamps and file information sources extracted from WhatsApp data.

Databases

Figure 16 shows that the analysis of artefacts in the databases category will show file names, storage locations, file sizes, timestamps and file information sources extracted from WhatsApp data. It is also seen that the encrypted WhatsApp database can be decrypted by the Cellebrite tool so that artifacts can be displayed for further analysis and search according to digital evidence needs.

Figure 13. Data Analysis in Device Locations Category

Figure 14. Data Analysis in the Web History category

Figure 15. Data Analysis in Audio category
Documents

Figure 17 shows that the analysis of artefacts in the documents category will show file names, storage locations, file sizes, timestamps and file information sources extracted from internal and external memory on smartphones.

Images

In Figure 18, it is shown that the analysis of artefacts in the images category will show file names, storage locations, file sizes, timestamps, image senders and file information sources extracted from the WhatsApp database and the smartphone’s internal/external memory.

Videos

Figure 19 shows that the analysis of artefacts in the videos category will show file names, storage locations, file sizes, timestamps, image senders and file information sources extracted from the WhatsApp database and smartphone internal/external memory.

From the results of the acquisition and extraction above, this study provides an overview of how to carry out acquisition, extraction and forensic analysis of artefacts in the Android-based WhatsApp application to obtain forensic evidence information related to data that shows details of file names, file locations, file sizes, GPS locations, and data related to other investigation needs.

Figure 16. Data Analysis in the Databases category

Figure 17. Data Analysis in the Documents category

Figure 18. Data Analysis in the Images category
It can be seen that the acquisition and extraction using the Android backup method cannot describe the encrypted WhatsApp database compared to the Android APK downgrade backup method proposed in this study. From the artefacts that have been analyzed, there are deleted files found in WhatsApp communication. For a cross, it means that the data is in a deleted condition where deleted files can be recovered or not. In the artefact analysis of the acquired file, there is a hash value with the MD5 algorithm, which guarantees the integrity of the data that is not the result of a modification, meaning that there is no change during the confiscation until the completion of the digital forensic process on the smartphone. Figure 20 shows that chat conversations are in the form of text that can be recovered, while for file attachments such as pdf documents, the video contained in the communication can be seen as thumbnails and file names.

Table 3 compares the acquisition and extraction results with the Cellebrite tools using the Android Backup and Android Backup APK Downgrade methods. In this study, acquisition and extraction were also carried out using MobilEdit and FinalData tools. Table 3 shows that the Android backup APK downgrade method can extract all categories of artifacts, especially the chat category except the calendar category and has the highest number of artefacts that can be extracted.

Table 4 shows the percentage of the number of artefacts compared to other methods. The table shows that the Android Backup APK Downgrade method can retrieve 651% more artefacts than the Android Backup Method, 851% more than the Logical-MobilEdit method and 854% more than the Android Backup Final Data method.

![Figure 19. Data Analysis on Videos category](image-url)

![Figure 20. Deleted data](image-url)
Deleted artefact

Recovery of deleted files/data in the WhatsApp application is an important thing that needs to be pursued in every mobile forensic action. The acquisition of WhatsApp digital evidence using the Android Backup APK Downgrade method can recover deleted data in WhatsApp text conversation communications, while for file attachments, only the thumbnail appears, as shown in Figure 21 and Figure 22.
CONCLUSION

The procedural approach used to obtain WhatsApp artefact data on an Android smartphone can be in different ways depending on several things, such as the type of smartphone manufacturer, smartphone security features, the transfer protocol used, and the Android version. There is a hash value with the SHA256 algorithm on the results of the acquisition and extraction using the android backup method and the Android backup APK downgrade method, and there is a hash value with the MD5 algorithm on the artefacts of the acquisition and extraction files. The hash value guarantees the integrity of the data that is not the result of a modification, meaning that there is no change during the confiscation until the digital forensic process on electronic evidence is completed. With the steps of the forensic analysis procedure carried out in this study, we succeeded in obtaining evidence artefacts in the form of chat sessions, avatars, contact numbers on the WhatsApp application, voice notes, profile photos, the identity of the WhatsApp account owner and also being able to get other media files and most importantly database files.

A common challenge for forensic investigators is the ever-evolving WhatsApp encryption standard to protect backups from unauthorized access. Therefore, it is very important for forensic investigators to always update technological developments related to WhatsApp backup databases to extract chat sessions that may exist on the suspect's device. Another challenge is that WhatsApp has added end-to-end encryption facility for all messages. Therefore, research is needed to conduct forensics on conversational sessions that utilize end-to-end encryption. The WhatsApp database is encrypted using crypt14 with WhatsApp
technology which is continuously updated by the developer, so further research and testing need to be carried out regarding forensic procedures for the latest encryption features of the WhatsApp messaging application in the future. Given that the WhatsApp application is a cross-platform application, it is also important to conduct a forensic analysis of WhatsApp artefacts on other platforms.

REFERENCES


