Shadi Zakarneh, 2021, 9:4
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Forensic Investigation of WhatsApp on Android Smartphone’s

M.Sc. Scholar Shadi Zakarneh
Department of Information Technology,
Palestine Technical University,
Kadoorei, Tulkarem, Palestine,
zakarnehshadi@gmail.com

Abstract- After the rapid and exponential development in communication technology and the internet, also accelerate development in Smartphone and their data connectivity like 3G and 4G. Social networking and instant messaging (IM) companies developed their mobile applications. Other IM mobile applications were developed such as WhatsApp (WA), Viber, and IMO. WA is considered the most popular IM application. WA enables users to exchange messages with different types of contents such as text, audio, video, and documents. Also, WA enables audio and video calls for its users over the internet. The popularity and the widespread of WA helped in different cybercrime cases commitment. While the messages and exchanged files and call logs are stored in Smartphone memory, WA usage leaves different types of artifacts that can be extracted and analyzed to determine the digital evidence. Besides, the android platform is the most platforms used in smart phones. Therefore, forensic investigation tools and methods are required for the investigation process. One of these methods is the National Institute for Standards and Technology (NIST) method. NIST method divides the investigation process into four stages; Evidence collection, examination, analysis, and reporting. Here, the NIST method was used in an investigation process over an android Smartphone, where a digital crime was assumed to be committed through WA. First, during the collection stage, the Smartphone was identified, preserved, documented, secured, and disconnected from all types of networks. The Examination stage was completed by phone imaging process and the hash value (MD5) was calculated. The analysis stage was completed by accessing chats details and logs, call logs, and multimedia to determine the evidence. Finally, the investigation process and evidence were reported. As a conclusion, WA forensic artifacts could be analyzed and discovered successfully using the NIST method. The deleted chats could be restored only when the WA database and its backups were existing.

Keywords- WhatsApp; Forensic; Android; Smartphone; Analysis; Investigation.

I. INTRODUCTION

As a result of the rapid and massive development of communications and the Internet, many virtual communication applications have appeared, such as social networks and instant messaging (IM) applications such as Facebook, Twitter, Snapchat, Skype, and WhatsApp (WA).

These tools have provided easy, fast, and free means of communication, in addition to neutralizing geographical borders in communication between people(Han and goleman, Daniel; boyatzis, Richard; Mckee, 2019).

Due to the rapid development in Smartphone’s and their applications resulting from the development of technology and communication, and due to their ease of mobility, cost, and energy consumption, Smartphone's have replaced computers in many uses (Umar, Riadi, and Zamroni, 2018).

Also, IM applications have replaced some basic services in mobile phones, such as SMS services (Umar, Riadi, and Zamroni, 2018).

Among the many IM applications on Smartphone's, WA has become the most popular IM application (Ubaidillah et al., 2019).

According to a statistic made by Statista in October 2020 for the most popular and widespread IM application based on the number of active users per month (Clement, 2020), WA is the most used by 2 billion active users, followed by Facebook Messenger 1.3 billion, WeChat 1.206 billion, QQ 648 million, Snapchat 433 million, and the Telegram has 400 million active users, as shown in figure 1.

With the widespread use of WA, and the variety of content that can be exchanged through it, such as text messages, pictures, audio files, video, and voice and video calls, it has become impossible to prevent the misuse of WA for criminal purposes (Umar, Riadi and Zamroni, 2018).

Many artifacts result from WA messages and calls. These artifacts are analyzed by forensic investigation to achieve the evidence information (Ubaidillah et al., 2019).

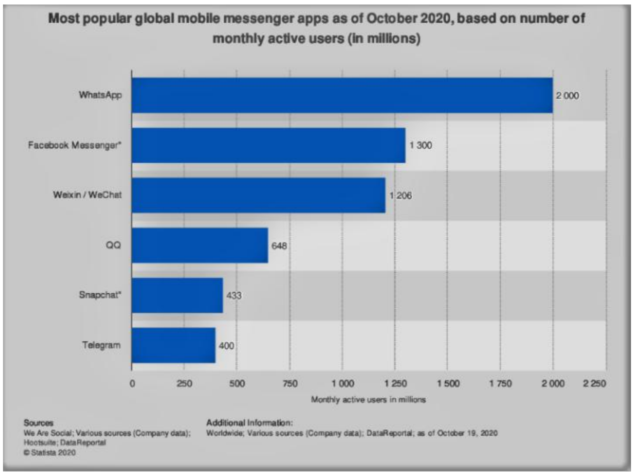


Fig 1. Mobile instant messaging applications statistic (Clement, 2020).

This research aims to explore the artifacts in WA on the android platform and the use of tools related to the digital investigation to access digital evidence in WA by extracting messages and calls, analyzing them, and linking them in a chronological sequence to reach the digital evidence of the case. The forensic method used in this research is the National Institute of Standards and Technology (NIST) digital forensic method which includes the collection, examination, analysis, and reporting (Umar, Riadi, and Zamroni, 2018).

II. LITERATURE REVIEW

WA is independent on Smartphone platforms, as it works on different platforms such as IOS, Android, Windows Phone, and Symbian. In addition to that WA is a free charge application for most platforms, and it is charge-free in sending text, voice, images, and video messages (AlHidaifi, 2018).

Android is an operating system for mobile devices developed by Android Inc., then Google bought it, and it was bought finally by the Open Handset Alliance.

Android is based on the Linux kernel. Android Platform can be used on different hardware with different mobile phone manufacturers. Because android integrates seamlessly and robustly with Google products and strongly supports cloud computing, making it the best operating system for mobile devices (AlHidaifi, 2018).

1. Android Architecture:

Android OS is an open system architecture using a hierarchical structure. As shown in figure 2 android structure divided into five layers as follows (AlHidaifi, 2018):

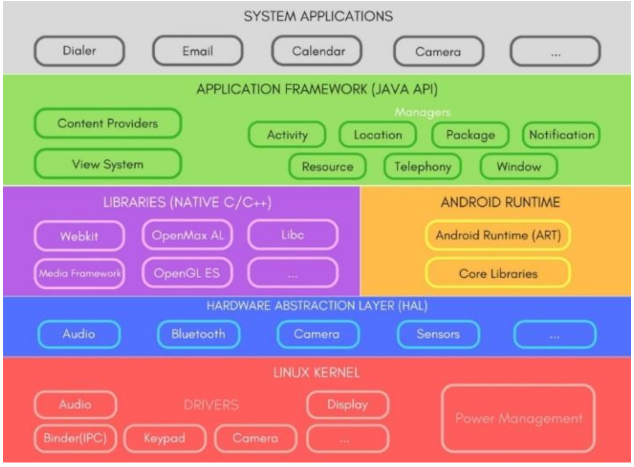


Fig 2. Android Architecture (Study tonight, 2020).

1.1 Linux kernel Layer: Linux 2.6 kernel is the android OS base layer, as this layer is responsible for the basic functions of the system such as memory management, process management, device management, power management. Linux kernel provides better interaction with the peripheral devices in the Smartphone (Ekanayake, 2018).

1.2 Hardware Abstraction Layer (HAL): This layer is above the Linux kernel layer. This layer provides an interface between system services and device drivers for those services. It makes Android neutral concerning low-level drivers (Khan and Shahzad, 2016).

1.3 Native Libraries Layer: The Android system includes a set of main basic libraries that are exposed to developers through the Android application framework; these libraries are SQLite, FreeType, Webkit, OpenGL ES, and Media Framework. These libraries are written in C ++, and they enable the device to deal with different types of data (Khan and Shahzad, 2016).

1.4 Application Framework: The basic functions of the device managed by the application layer. This layer provides user applications with application programming interfaces (APIs) that are used by applications for several purposes including getting notifications, accessing the telephony system, and sharing data. The application framework consists of an activity manager that works to manage application activity, content providers responsible for managing data sharing between applications, a location manager that works to manage locations through GPS, a telephony manager responsible for managing voice calls in applications, and managing the various resources used In applications by the resource manager (Ekanayake, 2018).

1.5 Applications LAYER: This layer is the top layer in android architecture; it includes the basic applications such as the contacts manager application, the SMS application, the dialer application, and the web browser application. This layer also includes applications that are developed by third parties. Since third-party developers have access to this layer, they can re-develop some basic features and applications such as the user interface and other applications to replace the basic applications of the system; this is a strong feature of the open-source Android system. Applications are written by developers in java. The developed applications are interpreted by the Dalvik virtual machine, which is replaced by Android Runtime (ART) (Ekanayake, 2018).

2. Android File Systems:

Android OS uses a file system to organize the data on storage; the efficiency of the file system depends on the speed of storing, reading, and retrieving data. The file allocation system 32 (FAT32), yet another flash file system 2 (YAFFS2), and extended file system (EXT) file systems are used in the Android platform. These systems are used to operate the device, boot, store, and retrieve data. Also, These systems are used to organize data and files on SD memory. On flash memory, the YAFFS2 file system is used (Khan and Shahzad, 2016).

3. WhatsApp:

WA is a popular IM application in Smartphone’s. WA allows users to exchange text messages, images, video files, audio files, and many other types of files such as document files, pdf files, and others. WA allows users to create user groups to send messages and files to all group members. WA also allows the user to control personal profile information such as name, profile photo, and information about the user (Udenze and Oshionebo, 2020). WA messaging is done in end-to-end encryption, and therefore no man-in-the-middle can read messages between two WA users (Shidek, Cahyani and Wardana, 2020).

WA stores its data in the mobile phone's internal memory. WA automatically connects to the phone contacts database and detects the contacts that use the WA application and adds them to its database. WA application also includes a procedure named "com.whatsapp", which is a procedure for operating the external media management service and the messaging service that runs in the background, as this procedure works when turning on the Smartphone (Alhassan et al., 2017).

Old versions of WA used the SQ-Lite database "msgstore.db" to save messages that were exchanged between users, this database was unencrypted. Because unencrypting the database, led to easy access to the details of the messages stored in it, to bypass this problem and to achieve better protection for users' privacy, an encryption mechanism was developed for the WA database on the Android platform using Advanced Encryption

Algorithm (AES) with an encryption key 192-bit length.

As a result, the database name has changed to msgstore.db.crypt, msgstore.db.crypt5, megastore. db.crypt7, and msgstore.db.crypt8. In recent versions of WA, the AES algorithm was used with a 256-bit key and the database became msgstore.db.crypt12 (Alhassan et al., 2017).

Research studies mostly deal with forensic methodologies on various mobile applications, such as forensic analysis of contact lists, SMS messages, and social media. Some researchers have compared several analysis tools by applying them to the analysis of processes for obtaining WA messages and files. Other researchers determined different artifacts on android platforms generated by WA such as contacts database, messages database, and the database encryption key. (Shidek, Cahyani, and Wardana, 2020).

Other studies focused on the decryption of WA encrypted database. In this study, the researcher used five different tools to achieve his goals; these tools are WA key/db extractor, WA viewer, WA extract, SQLite spy, and android backup extractor. Some of these tools are written in python code, so a python compiler is also needed in their experiment (Jhala KY, 2015).

III. RESEARCH METHODOLOGY

Two basic steps to complete this study. The first step is to collect data and information from literature studies, as a literature study is a method of collecting data through reading books, thesis, journals, and other related resources. The second step is to design a scenario to simulate the case to implement the digital forensic stages and then analyze the evidence extracted from WA. And finally, get to write a digital forensic analysis report.

Through the assumed scenario, the mobile forensic steps that are followed in this study are clarified. In this study, a case of cyber stalking by the suspect against the victim via WA, using text messages, pictures, and voice calls is used.

The scenario depicts a suspect stalking the victim through text messages, photos, and WA voice calls. The investigator will follow the forensic processes in dealing with the evidence to secure it and preserve it from any changes by taking an image of the evidence.

The data evidence is then extracted from the image of the evidence to explore the conversations, voice calls, and photos that took place between the suspect and the victim via WA. In this scenario, the Smartphone is in an active state and can be switched on as well as not rooted.

In the simulation process, the method of the NIST on digital forensic was used, as shown in Figure 3. All steps were taken to obtain valid and admissible evidence to be presented in a witness report that can be submitted to the court to decide the case.

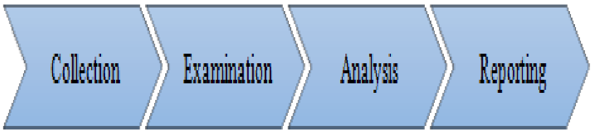


Fig 3. NIST Digital Forensic Method.

In the collection stage, the physical evidence is identified and collected, which is the Smartphone used in the crime, also the evidence information is recorded and documented. This stage includes the process of securing the evidence to maintain the integrity of the data and prevent any changes to it.

The data processing and verification of the evidence are then carried out backup and imaging data on the Smartphone using forensic tools. In this study, the Final Mobile Forensic tool was used to implement the data imaging process. This process is carried out in the examination stage.

In the Data Acquisition process from a non-rooted Smartphone, a logical ADB backup was used, where a backup of the data was taken from the internal memory and the Android system, the backup data image was saved in the storage media to protect it and perform analysis process later.

When this process is completed, the Acquisition process information is saved and documented in a report. This information includes investigator information, acquiring time and information, and image information which includes a file name, file size, and the MD5 hash value for the file, see appendix 1.

The investigator used the image taken in the previous stage to perform the analysis process to obtain more evidence related to the crime. Evidence is collected and verified by exploring and reading the conversations stored in the WA database.

Before the beginning of the analysis stage, the investigator checking the integrity of the evidence image by recalculating the image hash value (MD5)and comparing it with a hash value calculated in the acquisition process.

After the analysis stage, the results and data obtained from all stages of the investigation are reported. The report contains information about the investigator and an introduction to the crime.

The report includes documentation of all the evidence that was reached, including the documentation of the chronology of the events that were collected. The conduct of all operations must be within the law and the approved procedures for the report to be admissible to the court.

IV. RESULTS AND DISCUSSION

The results of the study that was conducted are successful in getting evidence from the Android Smartphone. The devices and programs shown in Table 1 were used in this study.

Table 1. Devices, Tools and specifications.

No.	Name	Specification
1.	Laptop	Toshiba Satellite, Intel Core I5, RAM: 16 GB. Windows 10
2.	USB Cable	USB Cable Type C
3.	Forensic Tool	Final Mobile Forensic Version 4
4.	Smartphone	Xiaomi Redmi Note 7

The Smartphone in Table 1 is used as evidence in the study. The forensic tool “Final Mobile Forensic” is used to carry out the evidence imaging process and get the MD5 value to ensure the evidence integrity before the analysis process.

1. Collection:

This stage is the stage of identifying, gathering, and documenting evidence and data. The evidence used in this study is a Smartphone that was photographed as shown in Figure 4. The evidence data has been documented as shown in Table 2.

Fig 4. Smartphone Evidence.

Table 2. Smartphone Evidence Specification.

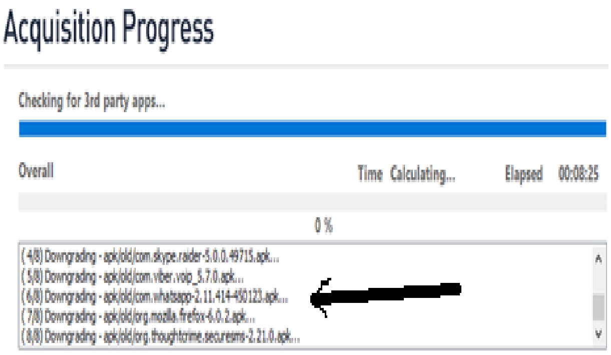
Name	Xiaomi Redmi Note 7
Model Number	M1901F7G
IMEI Number	866294042697419
OS Version	MIUI Version: 12.0.1 Android Version: 10 QKQ1.190910.002

Also at this stage, the evidence and dataare securedto preserve its integrity and protection from any change or destruction, as well as to ensure that this evidence or the data contained in it are not misused.

2. Examination:

At this stage, data acquisition on smartphones is carried out. Data acquisition is done to obtain evidence on the Smartphone which is not rooted using the Final Mobile Forensic tool. The acquisition method used is the Logical ADB backup method as shown in appendix2.

To verify that the acquisition results are identical to the original file that is on the Smartphone, the MD5 value is calculated and shown in the acquisition report as shown in appendix 1.



Acquisition Progress

Checking for 3rd party apps...

Overall Time Calculating... Elapsed 00:08:25

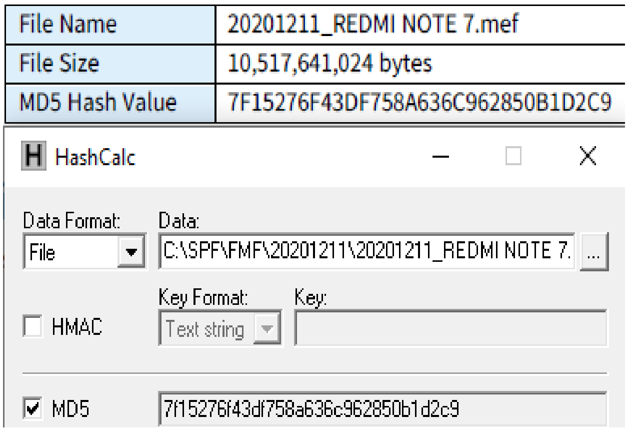
0 %

```
( 4/8) Downgrading - apk/old/com.skype.raider-5.0.0.49715.apk...
( 5/8) Downgrading - apk/old/com.viber.voip_5.7.0.apk...
( 6/8) Downgrading - apk/old/com.whatsapp-2.11.414-450123.apk...
( 7/8) Downgrading - apk/old/org.mozilla.firefox-6.0.2.apk...
( 8/8) Downgrading - apk/old/org.thoughtcrime.securesms-2.21.0.apk...
```

Fig 5. Downgrade WA to an older version.

While the Smartphone evidence contains a WA version 2.20.206.24 that used crypt12 database which used AES encryption algorithm with 256-bit key length, the acquisition tool downgraded the WA to version 2.11.414 to extract and decrypt the database automatically as shown in figure 5. After finishing the acquisition stage the forensic tool restored the original version of the WA application.

The image hash value recalculated using the hash calculator tool, the result was compared with the hash value (MD5) generated from the acquisition stage to ensure the integrity of the evidence before the analysis stage started.



File Name	20201211_REDMI NOTE 7.mef
File Size	10,517,641,024 bytes
MD5 Hash Value	7F15276F43DF758A636C962850B1D2C9

☑ MD5 7f15276f43df758a636c962850b1d2c9

Fig 6. Comparing the MD5 Value.

3. Analysis:

The analysis process succeeded as the evidence was found by reviewing WA chats and calls, as well as reviewing audio files, video files and photos exchanged through the WA application.

Using the final mobile forensic tool, the messages lists on WA were reviewed, the details of the existing messages were read as shown in appendix 3, in addition to viewing deleted messages, reading their texts, and viewing their attachments, including images and files as shown in appendix 4. Multimedia files exchanged via WA have also been reviewed as shown in appendix 5.

4. Reporting:

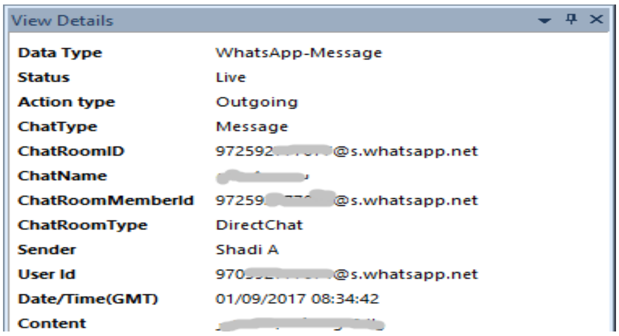
The reporting stage is the documenting and reporting of the analysis results relating to the case that had been investigated. The findings from the forensic process using the Final Mobile Forensic tool shown in table 3.

The details in table 3 represent the found data that is the same as the data in the Smartphone that was used as evidence.

Table 3. Evidence Findings.

Information	Findings
Mobile Phone Number	00972592777***
User Name	Shadi
WhatsApp Version	2.20.206.24
Contacts	784
Messages	59814
Deleted messages	530
Calls	423
WhatsApp photos	4046
WhatsApp audio files	4
WhatsApp video files	175
WhatsApp documents	18

The results obtained are from the WA database and its backups on the Smartphone, where text messages were found and read, and their characteristics such as sending time, contacts, photos, audio, and video clips, and documents as shown in figure 7.



Data Type	WhatsApp-Message
Status	Live
Action type	Outgoing
ChatType	Message
ChatRoomID	972592[redacted]@s.whatsapp.net
ChatName	[redacted]
ChatRoomMemberId	97259[redacted]@s.whatsapp.net
ChatRoomType	DirectChat
Sender	Shadi A
User Id	970[redacted]@s.whatsapp.net
Date/Time(GMT)	01/09/2017 08:34:42
Content	[redacted]

Fig 7. WA Message Details.

V. RESEARCH LIMITATION AND CHALLENGES

The investigation process that is performed in the study used a working and unlocked Smartphone. Other tools and mechanisms must be developed if the suspected Smartphone was locked or the screen not working, or the Smartphone dead.

While the investigation process is carried out on an existing WA database, a challenge will appear when the database's files were deleted.

VI. CONCLUSION

WA is the most popular application for IM, where people can exchange text messages, audio files, video files, and documents. This research has focused on forensic investigation of WA on the android platform. NIST digital forensic method is used in the research. The Final Mobile Forensic tool is used in the examination and analysis stages. While the WA database is encrypted, the used tool decrypted the databases by downgraded the WA version and then extracted the databases.

In the analysis stage the messages, call logs, photos, audio files, videos, and documents are viewed and analyzed to determine the evidence related to the assumed crime.

REFERENCES

[1] Alhassan, J. K. et al. (2017) 'Forensic Acquisition of Data from a Crypt 12 Encrypted Database of Whatsapp', 2nd International Engineering Conference, (October).

[2] AlHidaifi, S. (2018) 'Mobile Forensics: Android Platforms and WhatsApp Extraction Tools', International Journal of Computer Applications, 179(47), pp. 25–29. doi: 10.5120/ijca2018917264.

[3] Clement, J. (2020) most popular global mobile messaging apps 2020, Statista. Available at: https://www.statista.com/statistics/307143/growth-of-whatsapp-usage-worldwide/ (Accessed: 17 December 2020).

[4] Ekanayake, N. (2018) 'Android Operating System', (July), pp. 1–11. doi: 10.13140/RG.2.2.20829.72169.

[5] Han, E. S. and goleman, daniel; boyatzis, Richard; Mckee, A. (2019) 'Is Whatsapp The Future Of Workplace Communication?: Investigating The Use of Whatsapp In Decision- Making Episodes', Journal of Chemical Information and Modeling, 53(9), pp. 1689–1699.

[6] Jhala KY, G. L. (2015) 'WhatsApp Forensics: Decryption of Encrypted WhatsApp Databases on Non Rooted Android Devices', Journal of Information Technology & Software Engineering, 05(02), pp. 2–5. doi: 10.4172/2165-7866.1000147.

[7] Khan, J. and Shahzad, S. (2016) 'Android Architecture and Related Security Risks', Asian Journal of Technology & Management Research, 05(December 2015), pp. 2249–892.

[8] Shidek, H., Cahyani, N. and Wardana, A. A. (2020) 'WhatsApp Chat Visualizer: A Visualization of WhatsApp Messenger's Artifact Using the Timeline Method', International Journal on Information and Communication Technology (IJoICT), 6(1), p.1. doi: 10.21108/ijoict.2020.61.489.

[9] Studytonight (2020) Android Architecture - Software Stack of Android, Study tonight Technologies Pvt. Ltd. Available at: https://www.studytonight.com/android/android-architecture# (Accessed: 27 December 2020).

[10] Ubaidillah et al. (2019) 'Analysis whatsapp forensic and visualization in android Smartphone with support vector machine (SVM) Method', Journal of Physics: Conference Series, 1196(1). doi: 10.1088/1742-6596/1196/1/012064.

[11] Udenze, S. and Oshionebo, B. (2020) 'Investigating "WhatsApp" for Collaborative Learning among Undergraduates', Etkileşim, 3(5), pp. 24–50. doi: 10.32739/etkilesim.2020.5.92.

[12] Umar, R., Riadi, I. and Zamroni, G. M. (2018) 'Mobile forensic tools evaluation for digital crime investigation', International Journal on Advanced Science, Engineering and Information Technology, 8(3), pp. 949–955. doi: 10.18517/ijaseit.8.3.3591.