Whatsapp Messenger on the Android Platform: A Forensic Examination of a Physical Device

Abstract

The major objective of this research was to develop a methodology that might assist digital investigators to extract and analyze third party application data during a criminal investigation so that the following questions might be answered: Who has the user been communicating with and when? What was the content of the communication? What attachments were exchanged and where can they be found? A comprehensive description of the evidentiary artifacts created by a third-party application, WhatsApp, is provided by the authors. Results from this study shows that the contact database allows an investigator to discover which users contacted each other at the date and time of said contact. Furthermore, analyzing the content of the message database, especially the message table, allows a digital forensic investigator to create a chronological sequence of exchanged communications and determine if any attachments were exchanged and their location on the device.

Keywords

Android; Instant Messaging; Smartphones; Whatsapp Messenger; Whatsapp

Orr DA, Castro A (2018) Whatsapp Messenger on the Android Platform: A Forensic Examination of a Physical Device. Forensic Sci Methods Tech 2018(1): 04-19.

Correspondence should be addressed to
Douglas A Orr, USA
E-mail: douglas.orr@ung.edu

Copyright

Copyright © 2018 Douglas A Orr et al. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and work is properly cited.
being used by criminals to conduct illicit activities. It is for this reason that the forensic analysis of third-party applications has recently received considerable attention [3]. It is uncommon now for an investigator to lead a digital forensic investigation that does not include a cell phone. Smartphones have advanced significantly in recent years, making third party applications troublesome for examiner to discover data [4].

Valuable evidence would be missed if the digital investigator is not savvy in this regard where third-party applications are present. While third party applications store data in an application folder, how the data is stored will depend greatly on the digital device itself. Typically, Android devices use the default storage locations might often cause a digital forensic investigator to miss important artifacts [4].

As third-party applications turn out to be progressively more prevalent, the effect they have on digital forensic examination is substantial. To remain current and guarantee that cases are not compromised because of missed digital forensic evidence, digital forensic investigators must precaution to determine how and where third-party applications store data. In doing so, such skill would be profitable as the investigator might find key digital forensic evidence that others are overlooking [4].

The major focus of this research is to conduct a forensic analysis of the popular third-party application WhatsApp on an Android device [5]. The objectives are to identify, locate and recover forensics artifacts from WhatsApp on Android devices and to analyze and interpret the artifacts relating specifically to chronological chat data associated with contact information. Our methodology has greater explanatory value in that we outline the use of the Whatsapp Viewer to decrypt message databases. This methodology will assist investigators to extract and analyze third party applications data during a criminal investigation to answer the questions mentioned previously. While there are many third-party applications in use today, the methodology mentioned here will deal only with WhatsApp.

Literature Review

Ovens and Morison [6] conducted a study on the forensic analysis of the popular mobile application KIK messenger on iOS Devices [7]. The first study of its kind, said the authors, their main objective was to identify, analyze, and report artifacts left behind on iOS devices by KIK. They developed a methodology that consisted of installing the KIK application on two iOS devices and using both iOS devices to communicate with each other to simulate real user scenarios, such as one-to-one chats, image and video sharing and group messaging [6]. Once the authors created a data set for the study, the data was extracted from the iOS device using iTunes backup. The authors emphasized that the iTunes backup feature is not a forensic acquisition tool. However, they did point out that the feature has been used by different researchers to acquire iOS artifacts [6]. During the study the researcher found that communications using KIK messenger are stored on SQLite databases. To investigate the SQLite databases, the authors used the SQlite3 command utility and created a python script to export the results to a spreadsheet format for better viewing and analysis. They authors discovered that important artifacts can be recovered from KIK messenger such as messages and attachments shared by users. Moreover, they were able to document the location where key evidence can be located on iOS devices as it relates to KIK messenger. However, the authors were not able to locate any information that can provide the real identity of KIK messenger users. We speculate that this was not possible because KIK does not require that a phone number be associated with the KIK messenger account. While KIK requires that an email address be provided, however, a KIK user can offer a fictitious email account as KIK does not verify the email account provided the user [6].

Gudipaty and Jhala [8] developed a step by step methodology for the decryption of encrypted WhatsApp Databases on non-rooted Android Devices. The purpose of this step by step methodology was helpful to digital forensic investigators to parse WhatsApp encrypted databases in plaintext. For this experimental design, the authors used an unrooted Nexus 4 device (16 GB Model) Android version 5.0 (Lollipop) and running WhatsApp version 2.11.432 [8]. They connected the Android mobile device to the forensic workstation to extract the WhatsApp database file and used the WhatsApp Key/DB to obtain the msgstore.db file. This database contains undeleted conversations in decrypted format. Additionally, the WhatsApp Key/DB can be used to extract the key needed to decrypt the encrypted backup files [8]. According to Gudipaty and Jhala [8], the database that contains the WhatsApp messages is the msgstore.db and the wa.db stored users contacts information such as phone numbers, display name, timestamp, and other information provided by the user during the registration of WhatsApp. Using the authors methodology, a digital forensic investigator could decrypt the encrypted WhatsApp databases and analyze these databases for investigatory purposes.

Adebayo [9] developed a methodology to analyze artifacts of KIK Messenger on Android devices. The aim of this methodology was to identify and analyze potential evidence related to the popular KIK Messenger Application on Android devices. Their methodology consisted of three stages: 1) the preliminary device setup stage, 2) the acquisition stage, and
Two years later, Anglano et al., [3] conducted a study of the Chatsecure [14] instant messaging application on Android smartphones. The aim of their study was to determine exactly where the Chatsecure database stored artifacts on the device and how these artifacts can be recovered from the Chatsecure instant message application. During this study, the authors discovered that Chatsecure stores exchanged messages and files into two local databases both encrypted using AES-256 algorithm. They were able to develop a methodology allowing them to recover the passphrase used to encrypt the databases and, in turn, use the passphrase to decrypt the databases. Once the databases were decrypted, they were able to determine that Chatsecure stores many artifacts that can be of testimonial evidence to a digital forensic investigator. Exchanged messages using the Chatsecure Instant Messaging application are stored in the main database. Additionally, the main database also stores user information used to create the Chatsecure account. The main database consists of 21 tables. However, the authors concluded that only 11 out of the 21 tables are of evidential value to the digital forensic investigator. Additionally, the authors noted that they were not able to retrieve deleted messages from Chatsecure due to the deletion technique adopted by the SQLCipher [3].

Kausar and Alyahya [15] developed a methodology to extract pertinent databases used by the popular application Snapchat [16] using AXIOM Examine and Autopsy on Android Smartphones [17]. Snapchat is a popular online social network used by many users to communicate with friends and family. It allows users to send instant messages, share photos, videos, and stories. The main goal of this study was to see what types of artifacts can be extracted from Snapchat using two different tools, AXIOM Examine and Autopsy. To do this, they used a Samsung Galaxy Note GT-N7000, running Android 4.1.2 with Snapchat application (version 4.2.0). They launched a scenario that consisted of creating test account for Snapchat and added five friends to the test account. They exchanged messages, sent stories and photos, and deleted stories, images and photos. They used trial version of AXIOM Examine since AXIOM Examine is not free. Additionally, they used Autopsy, a free digital forensic tool. They concluded that both AXIOM Examine and Autopsy can be used to extract relevant digital artifacts from Snapchat. The study showed that AXIOM Examine is more user friendly than Autopsy because AXIOM Examine allows the examiner to extract digital forensics artifacts from the social media application rather than the entire android smartphone. The authors concluded, however that conducting manual analysis with the assistance of both tools produce better results meaning more artifacts were recovered [15].

Azfar et al., [18] developed a methodology known as the adversary model to collect information of evidentiary value to a digital forensic investigator from five popular Android social media apps - namely Twitter, POF dating, Snapchat, Fling, and Pinterest. The goal of this methodology was to develop a process that allows a digital forensic investigator, during the investigation of crimes or incidents, to identify potential digital forensic evidence that can be located on popular Android social media applications. In this study, the authors examined five different social media applications for Android mobile devices. They concluded that they were able to retrieve artifacts of evidential value from the five-test cases they set-up to for the study. Additionally, they were able to identify where these applications stored the evidentiary artifacts [18].

Aji et al., [19] developed a methodology to extract digital forensic evidence from Snapchat. The goal of the methodology was to identify potential digital evidence that can be located on Android and iOS mobile devices. Specifically, the authors developed a methodology to extract digital forensics artifacts using XML (Extensible Markup Language). They developed a simulated process using an Apple iPhone 6 (A1549), and a Samsung SM-N7505 Galaxy Note 3 Neo to generate messages between both devices. The authors concluded that they were...
able to retrieve valuable digital forensics artifacts from both mobile phones relating to Snapchat.

Barmpatsalou et al., [20] conducted a comprehensive review of seven years’ worth of mobile device forensic literature. The authors argued that Mobile Device Forensics is a field consisting of various methods that can be applied to a variety of computing devices, such as smartphones and satellite navigation systems. Additionally, the authors cited that over the years an extensive amount of research has been conducted relating to mobile devices, data acquisition and extraction methods. The goals of the authors were to conduct an exhaustive review of the field, by providing a detailed evaluation of the actions and methodologies used during the last seven years [20]. The authors concluded that the area of mobile device forensics is a rapidly growing field and is constantly changing. The authors identified different challenges in mobile device forensics and provided a comprehensive review that can be acquired by digital forensic practitioners interested in understanding the different areas of this rapidly growing field [20].

Scrivens and Lin [1] conducted a study relating to data extraction and analysis of Android mobile devices. The authors explained that smartphones are constantly identified during law enforcement investigations and these devices may require forensic analysis. Scrivens and Lin [1] argued that mobile device applications found on smartphones contain a significant amount of personal data. They developed a methodology aimed at answering the following questions: How much of this private/personal data do these applications store locally in smartphone storage that may be of significance to a forensic investigation? And how can this data be extracted? To answer these questions, they developed a methodology consisting of four steps. The first step was to understand the Android file-system structure. The second step was to understand the different data acquisition methods and choose the appropriate one. The third step was to conduct data extraction. And the final step was to analyze extracted data for fragments of interest [1].

To test their methodology, they used an Asus Zenfone 2 Laser with Android Version 6.0.1. installing Facebook Messenger version 86.0.0.17.70 and Google Hangouts version 11.0.130004787 on each. Using these applications, they exchanged various text messages, images, videos, voice-clips, location information, live voice/video calls, etc., in both groups and personal conversations [1]. Scrivens and Lin [1] concluded that a tremendous amount of data is stored locally by Messenger and Hangouts social media applications that may be relevant in a forensic investigation. The methodology developed by the authors was straightforward and can be applied to other Android smartphones [1].

Methodology

To date, the literature in this area lacks a further informed methodology that retrieves chronological chat data associated with contact information from an unrooted physical Android device using Whatsapp. The authors hoped to fill this void by employing the following approach. Our goal here is to analyze third-party applications data during a criminal investigation to answer the following questions: Who has the user been communicating with and when? What was the content of the communication? And what attachments were exchanged and where can they be found? To do this, we need to configure our workstation and tools, acquire an image of the target device, and analyze said image for artifacts.

Configuring the forensic workstation

Before conducting the examination of WhatsApp, a forensic workstation was configured with all the following hardware and software:

- Forensic tower operating with Windows 7 Professional, Service Pack 1.
- Samsung J3 Prime, Model SM-J327T smartphone
- Magnet ACQUIRE
- FTK Imager
- Autopsy
- SQLite DB Viewer,
- WhatsApp Viewer
- WhatsApp key /DB extractor

Device Setup: The primary stage included configuring the android smartphones. This setup consisted of conducting user activities on the test applications. The first step consisted of downloading and installing WhatsApp from the Google PlayStore [21]. During the second step, user accounts were created on the WhatsApp applications to associate themselves with different sets of activities in order to generate test data. The activities consisted of uploading profiles, creating group messages, exchanging pictures and video, text chatting, adding and deleting contacts, etc. Examining these types of activities is very important to the digital forensic examiner as they are the most common among users. This is especially important, as it allows an investigator to determine who the user contacted.

Acquisition

The second stage consisted of acquiring an image of the Android smartphones. Acquisition is the process of replicating digital evidence from a smartphone. The process of acquiring evidence in digital format from a smartphone and the associated
media file is known as imaging. For this experiment we are going to be using Magnet Acquire, a free imaging program for mobile devices by Magnet Forensics. For mobile devices, Magnet Acquire provides two types of images:

- A quick image is a comprehensive logical image that contains both user data and some native application data. Magnet ACQUIRE uses multiple acquisition methods to mine as much information as possible from the device as quickly as possible to expedite the investigation [11].
- A full image is a physical or file system logical image. During this type of acquisition, Magnet ACQUIRE copies the entire contents of a drive into a single file (either a .raw file or a .zip file) [11].

Another tool used during this experiment is the WhatsApp Key/DB extractor. WhatsApp databases are encrypted and cannot be viewed unless the databases are decrypted. To decrypt the databases, one would need the cipher key used by WhatsApp to encrypt/decrypt the databases. This cipher key is stored inside the RAM of the mobile device and it is not possible to access this part of the smartphone on an un-rooted device. WhatsApp Key/DB Extractor allows a digital forensic investigator to successfully extract the cipher key from the RAM of an unrooted android smartphone.

For this experiment, Magnet ACQUIRE was used to capture a full image of a Samsung J3 Prime, Model number SM-J327T, using Android Version 7.0. The Samsung J3 Prime, Model number SM-J327T was connected to the Forensic Workstation Operating under Windows 7, Service Pack 1. Magnet ACQUIRE software was launched and the software detected the device; a full image was selected.

A complete image of the Samsung SM-J327T smartphone was completed.
Next, we download the encryption key used by WhatsApp to encrypt the databases. The encryption key is stored inside the RAM of the mobile device. To extract the cipher key, the WhatsApp key/DB extractor was used.

![WhatsApp Key/DB Extractor](image)

**Figure 3:** WhatsApp Key/DB Extractor.

The WhatsApp key/DB extractor was able to successfully retrieve the encryption key from the RAM of the smartphone. This key was used during the examination process to decrypt the encrypted WhatsApp databases.

![WhatsApp cipher key](image)

**Figure 4:** WhatsApp cipher key.

**Whatsapp data analysis**

The third stage of this methodology is the analysis stage. Here, the WhatsApp application will be physically analyzed utilizing specific forensic tools to view and retrieve files and also to search for data related to the test application [9]. The goal of this stage is to analyze artifacts stored in the databases populated by the Whatsapp Messenger application. Such artifacts manifest themselves into sets of files having names and associations with content that includes contact information,
geolocation, plain text messaging and media.

WhatsApp is a cross-platform instant messenger service that has over one and a half billion users and continues to grow exponentially [22]. WhatsApp is more than a text messaging application. WhatsApp users are able to send each other images, videos, and audio media messages. WhatsApp auto syncs to the address book of the smartphone automatically showing all the contacts who are WhatsApp users [22]. From an investigative standpoint, WhatsApp’s popularity increases the likelihood that it will be used in criminal activities or for nefarious purposes. WhatsApp forensic analysis can yield valuable digital evidence to digital forensic investigators. For example, call and chat conversation data can assist a digital forensics investigator in answering the question of who has the user been communicating with and when? All of this is contained in the application’s file structure. Conceptually, the behavior of the user is stored in tables within various databases protected by encryption as outline in figure 5 below:

![Figure 5: Concept of Whatsapp Messenger file structure on Android Operating System.](image)

We conducted an analysis of the WhatsApp databases that were retrieved from the Samsung J3 Prime, Model SM-J327T using WhatsApp Version 2.18.177. In order to analyze these databases, three different tools were used. The first tool used was FTK imager from Access Data [23]. This tool was used to browse through the Samsung J3 Prime Model SM-J327T image in order to locate and extract the WhatsApp databases for analysis. The next tool used was the WhatsApp Viewer. WhatsApp viewer allows the digital forensics investigator to decrypt and read the contents of the databases, provided one has the cipher key. Using the FTK imager, the encrypted databases were located on the Samsung J3 Prime Model SM-J327T image in the following directory: /data/data/com.whatsapp/databases.

![Figure 6: Encrypted WhatsApp databases.](image)
The databases are encrypted and, in order to read the content of the databases, the databases must be decrypted using the cipher key that was extracted during the acquisition process. Using the cipher key extracted during the acquisition process, the WhatsApp Databases were decrypted using WhatsApp viewer.

![Decrypt WhatsApp Database](image1)

**Figure 7:** Decrypting WhatsApp databases.

The databases were decrypted and now ready for analysis.

![Example of decrypted WhatsApp chat database](image2)

**Figure 8:** Example of decrypted WhatsApp chat database.
WhatsApp messenger stores artifacts on an Android device into various databases and files. In isolation, these artifacts might seem unremarkable. However, associated with each other, they are of particular interest to the investigator. These databases and files are listed below in table 1 [10].

| Content                          | Directory                                      | File                                      |
|----------------------------------|-----------------------------------------------|-------------------------------------------|
| Contact database                 | /data/data/com.whatsapp/databases             | Wa.db (SQLite v.3)                        |
| Chat database                    | /data/data/com.whatsapp/databases             | Msgstore.db (SQLite v.3)                  |
| backups of the chat database     | /mnt/sdcard/Whatsapp/Databases                 | Msgstore.db.crypt                         |
|                                  |                                               | Msgstore-<date>.crypt                     |
| Avatars of contacts              | /data/data/com.whatsapp/files/Avatars         | UID.j, where UID is the identifier if the contact |
| Copies of contacts avatars       | /mnt/sdcard/WhatsApp/ProfilePictures          | UID.j, where UID is the identifier of the contact |
| Log files                        | /data/data/com.whatsapp/files/logs           | Whatsapp.log, whatapp <date>.log          |
| Received files                   | /mnt/sdcard/WhatsApp/Media                    | Various files                             |
| Sent files                       | /mnt/sdcard/WhatsApp/Media/sent               | Various files                             |
| User settings and preferences    | /data/data/com.whatsapp/files                 | Various files                             |

Table 1: WhatsApp Messenger artifacts [10].

Consequently, our methodology follows a certain “pipeline” as depicted in figure 9 below:

![Figure 9: Methodological Pipeline.](image-url)
**Analysis of contact information:** The analysis of contact database (`wa_db`) is of high importance when it comes to evidentiary value to a digital forensics investigator as it can answer the question of who the user has contacted. The contacts database contained 13 tables, namely:

- **wa_contacts**, which stores a record for each contact,
- **wa_group_admin_settings**, which stores the name of the group, and the contact information for each member of the group,
- **android_metadata ; SQlite_sequence ; system_contacts_version_table, wa_biz_profiles_wa_profiles; hours, wa_contact_capabilities; wa_profiles_websites, wa_contact_storage_usage; wa_group_admin_settings, wa_vnames,; wa_vnames_localized**, these tables have no evidentiary value to the digital forensics investigator.

The `wa_db` contains important information for the digital forensic investigator. For this experiment, communication was established with seven different individuals. Before communicating with these individuals their name and phone numbers were added to the phone book of the tested device the Samsung J3 Prime, Model SM-J327T. Using the SQLite database reader to open and analyze the contact database (`wa_db`), the contact names and phone numbers of each individual that we chatted with were discovered in the contact database. As shown by figures 10 and 11, the **field jid** stored the phone numbers of the individuals the user contacted.

WhatsApp requires that each WhatsApp user provide a valid phone number associated with their profile. **Field number** displays the phone number of the contacts stored in the phonebook of the smartphone. WhatsApp automatically synchronizes with the phonebook of the user and allows them to send messages to any contacts that have the same application installed on their device. **Field display name** displays the name of the contact and is added automatically by the WhatsApp during the synchronization of the user phonebook. The field **is whatsapp user** is used to distinguish between a WhatsApp user and someone who is not a WhatsApp user, the value of ‘1’ corresponds to a real user and a value ‘0’ corresponds to an unreal WhatsApp user. The field **given name** stores the given name of the users. The field **given family** name stores the family name of the users.

![Figure 10: Contact database fields.](https://doi.org/10.33513/FSMT/1801-05)
To better illustrate the contacts database fields and their meaning, please refer to table 2 below.

Data added to the contact database from the WhatsApp system

| Field name       | Meaning                                                                 |
|------------------|-------------------------------------------------------------------------|
| _id              | Sequence number of the record (set by SQLite)                           |
| _jid             | WhatsApp ID of the contact (a string structured as x@.whatsapp.net, where x is the phone number of the contact) |
| is_whatsapp_user | contains 1 if the contact corresponds to an actual WhatsApp user, 0 otherwise |
| unseen_msg_count | Number of messages sent by this contact that have been received, but still have to be read |
| photo_ts         | Unknown, always set to ‘0’                                              |
| thum_ts          | Unix epoch time (10 digits) indicating when the contact has set his/her current avatar pictures |
| Photo_id_timestamp | Unix million second epoch time (13 digits) indicating when the current avatar picture of the contact has been downloaded locally |
| wa_name          | WhatsApp name of the contact (as set in his/her profile)               |
| status           | Status line of the contact (as set in his/her profile)                 |
| sort_name        | Name of the contact user in sorting operations                         |

Data coming from the phonebook of the device

| Field name       | Meaning                                  |
|------------------|------------------------------------------|
| Number           | Phone number associated to the contact   |
| raw_contact_id   | Sequence number of the contact           |
| display_name     | Display name of the contact              |
| phone_type       | Type of phone                            |
| phone_label      | Label associated to the phone number     |
| given_name       | Given name of the user                   |
| family_name      | Famility name of the user                |

Table 2: Structure of the wa_contacts table [10].
Analysis of exchanged messages: All messages exchanged by WhatsApp users are stored in the message table of the msgstore.db. This database can contain valuable evidence for a digital forensics investigator. An examination of the message table allows a digital forensics investigator to recreate the sequence of events of exchanged communications. By examining this table, a digital forensics investigator can determine who the user has been communicating with and when. Specifically, the investigator can determine when a communication was initiated, the contents of the communication, and with whom the user was communicating. Additionally, the investigator can determine the dates and times the exchanged message was received by the parties involved [9].

Structure of the chat database: The structure of the msgstore.db for WhatsApp Version 2.18.177 contains 28 different tables. However, there are nine tables of evidentiary value to a digital forensic investigator. They are as follows:

- **Messages** - contains a record of each conversation that has been sent or received by the user.
- **chat_list** - contains a record for each conversation held by the user.
- **Frequents** - provides a message count by users.
- **Group_participants** - provide the WhatsApp ID of the contact
- **media_stream_sidecar** - provide a list of users and key message identifier.
- **message_thumbnail** - provide the WhatsApp ID of the communication partner.
- **message_fts** - contains a record of group conversation.
- **message_fts_content** - contain a record of group conversation (this table contains the same information as the message_fts table).
- **Receipts** - contains a receipt of all the messages sent and received. With a unique message identifier, and the WhatsApp ID of the communicating partner. As well as the date and time sent and received.

Reconstruction of the chat history: The reconstruction of the chat history is perhaps of the most value to a digital forensic investigator. Being able to reconstruct the time when a particular conversation took place, the content of the conversation, and the communicating parties can be extremely damaging to a guilty person or exculpatory for the innocent person. Anytime that a communication takes place between WhatsApp users, the communication is stored in the msgstore.db.

Specifically, the communication is stored in the message table that contains the information relating to the communication such as the body of the conversation and other metadata relating to the conversation [9]. Figure 13 demonstrates a conversation.

Figure 12: msgstore.db tables.
| Field name       | Meaning                                                   |
|------------------|-----------------------------------------------------------|
| _id              | Record sequence number                                   |
| key_remote_jid   | WhatsApp ID of the communication partner                 |
| key_id           | Unique message identified                                |
| key_from_me      | Message direction '0'=incoming, '1'=outgoing              |
| Status           | Message status: '0' = received, '4' = waiting on the central server, '5' = received by the destination, '6' = control message |
| Timestamp        | Time of send if key_from_me='1', record insertion time otherwise (taken from the local device clock, and encoded as a 13-digits millisecond Unix epoch time) |
| received_timestamp| Time of receipt of the central server ack (taken from the local device clock, and encoded as a 13-digits millisecond Unix epoch time) if key_from_me='1' otherwise |
| Receipt_device_timestamp | Time of receipt of the recipient ack (taken from the local device clock, and encoded as a 13-digits millisecond Unix epoch time) if key_from_me='1', '-1' otherwise |
| send_timestamp   | Unused (always set to '-1')                              |
| needs_puch       | '2' if broadcast message, '0' otherwise                  |
| Recipient_count  | Number of recipients (broadcast message)                 |
| remote_resource  | ID of the sender only for group chat messages            |

Table 3: Structure of the message table: field storing message attributes [10].

| Field name            | Meaning                                                                 |
|-----------------------|-------------------------------------------------------------------------|
| media_wa_type         | Message type: '0'=text, '1'=image, '3'=video, '4'=contact card, '5'=geo position) |
| Data                  | Message content when media_wa_type=0                                     |
| raw_data              | Thumbnail of the transmitted file when media_wa_type=['1','3']          |
| media_hash            | Based64-encoded SHA-256 hash of the transmitted file when media_wa_type=['1','2','3'] |
| media_url             | URL of the transmitted file (when media_type=['1','2','3'])             |
| media_mime_type       | MIME type of the transmitted file (when media_wa_type=['1','2','3'])   |
| media_name            | Name of transmitted file when (when media_wa_type=['1','2','3'])       |
| media_duration        | Duration is sec. of the transmitted file when media_wa_type=['1','2','3']) |
| Latitude              | Latitude of the message sender (when media_wa_type='5')                 |
| Longitude             | Longitude of the message sender (when media_wa_type='5')                |
| thumb_image           | Housekeeping information (no evidentiary value)                        |

Table 4: Structure of the message table: fields storing information concerning message contents [10].

On Saturday, June 16, 2018, at approximately 10:36:47 user 1585xxx6214-1528749173@g.us creates a group chat. On Saturday, June 16, 2018, at approximately 10:40 33 user 1585xxx0153@s.whatsapp.net responded “what’s up.” On Saturday, June 16, 2018, at approximately 10:41: 07 user 1585xxx2184@s.whatsapp.net responded “How are you doing bird?” On June 16, 2018, at approximately 10:41:44 user 1585xxx0077@s.whatsapp.net sent a message to the group “Hi Christopher.” On Saturday, June 16, at approximately 10:43:29 user 1585xxx0153@s.whatsapp.net responded “hi”, and the exchanged of messages continues between all of the involved parties. Accordingly, analyzing the data in this fashion builds an evident time line of events. More importantly, the digital forensics investigator could most likely determine the identity of the person involved in the communication as WhatsApp requires a valid and verifiable phone number to register the application. The phone number provided during the registration process becomes the WhatsApp user account number followed by the @s.whatsapp.net. For example, if one’s phone number is (585) 345-XXX and is used to create a WhatsApp account, the WhatsApp user account will be 585345XXX@s.whatsapp.
Extracting the contents of a message: WhatsApp users exchanged various types of messages, such as images, audio, video, contact cards and geo-location information. In order to determine if the message contains a multimedia file, an examination of the message table fields, media_url, media_wa_type, and the media name will reveal to the digital forensics investigator if the message contains any type of digital media. As shown by figure 14 below, the media_url field contains a URL, the media_wa_type field is equal to ‘1’, and the media name is IMG-20180616-WA0000.jpg, this indicates that an image (picture) was transmitted with this message. Furthermore, the media folder contains several subfolders with different media artifacts of evidentiary value to the digital forensics investigator (Figure 15). For example, the WhatsApp document stores any documents that were sent or received by the users. WhatsApp Images contains any and all photos sent or received by the user, WhatsApp Profile Photos, the profile picture of the user, and any profile pictures with whom the user has communicated. WhatsApp Video contains any video sent or received by the user.
Figure 15: Media folder.

Figure 16: WhatsApp images folder.

Figure 17: WhatsApp video folder.
Limitations of This Study

We are aware that our results extend only to physical Android platform and should not be generalized beyond the device and software used here. However, the authors suggest that the outcomes from our study using a physical device do not differ very much from studies focusing on the same platform but using an emulated device. Moreover, this study closely mirrors that of an actual digital investigation where commercial tools used by law enforcement officials would be used.

WhatsApp Forensics Conclusion

The forensic analysis of the artifacts located and identified on Android devices by the popular WhatsApp messenger is of evidentiary value for a digital forensic investigator. The analysis of the contacts database (wa_db) allows an investigator to determine who the user was in contact with and the date and time of the contact, a point of study the authors found underdeveloped in the current literature. Studying and analyzing the contact database allows an investigator to identify who the user communicated with and when. Additionally, analyzing the content of the message database (msgstore.db), specifically the message table, allows a digital forensic investigator to create the sequence of events and their respective exchanges of communication. The investigator can determine when a communication was initiated, the contents of the communication, and with whom the user was exchanging communications.

Moreover, the investigator can determine the dates and times the exchanged messages were sent and received by the parties involved. By examining the message table, the digital forensic investigator can determine if any attachments were exchanged and the location of these attachments on the device. Analyzing the digital artifacts found in the message table allows a digital forensic investigator to answer the questions of what was the content of the communication, what attachments were exchanged, and where can they be found. We agree with Anglano C [10] that future iterations of such studies should involve a comparison of artifacts found on the Android platform with newly discovered artifacts apparent on different platforms.

References

1. Scrivens N, Lin X (2017) Android Digital Forensics: Data, Extraction and Analysis. Oshawa, Ontario: University of Ontario Institute of Technology.
2. Walnycky D, Baggili I, Marrington A, Moore J, Breitinger F (2015) Network and device forensic analysis of Android Social-messaging applications. Digital Investigation 14: 77-84.
3. Anglano C, Canonico M, Guazzzone M (2016) Forensic analysis of the ChatSecure instant messaging application on android smartphones. Digital Investigation 19: 44-59.
4. Mahalik H, Quezada C (2014) Digging for Data, Finding Evidence in Third-Party Applications.
5. Whatsapp. (2018, July 17) Whatsapp Messenger.
6. Owens K, Morsion G (2016) Forensic Analysis of Kik messenger on iOS devices. Digital Investigations 17: 40-52.
7. KIK. (2018, July 17) KIK Instant Messenger.
8. Gudipaty L, Jhala K (2015) WhatsApp Forensics: Decryption of Encrypted WhatsApp Databases on Non Rooted Android Devices. J Inform Tech Softw Eng 5: 147.
9. Adebayo O, Sulaiman S, Osho O, Addulhamid S (2017) Forensic Analysis of KIK Messenger on Android Devices. 2nd International Engineering Conference.
10. Anglano C (2014) Forensic Analysis of WhatsApp Messenger on Android Smartphones. Digital Investigation 11: 201-213.
11. Magnet Forensics. (2018, July 17) Magnet ACQUIRE.
12. Cellebrite. (2018, July 17) Cellebrite: Digital Intelligence for a Safer World.
13. Oxygen Forensics. (2018, July 17) Oxygen Forensic Investigator.
14. Chatsecure. (2018, July 17) Chatsecure Messenger.
15. Alyahya T, Kausar F (2017) Snapchat Analysis to Discover Digital Forensic Artifacts on Android Smartphone. Procedia Computer Science 109: 1035-1040.
16. Snapchat. (2018, July 17) Snapchat.
17. Sleuth Kit. (2018, July 17) Autopsy.
18. Azfar A, Choo K-KR, Liu L (2016) An Android Communication App Forensic Taxonomy. Jour of Foren Sci 61: 1337-1350.
19. Aji M, Riadi I, Lutfhi A (2017) The Digital Forensic Analysis of Snapchat Application Using XML Records. Journal of Theoretical Applied Information Technology 95: 4992-5002.
20. Barmatsaluou K, Damopoulus D, Kambourakis G, Katos V (2013) A Critical review of 7 years of Mobile Device Forensics. Digital Investigation 10: 323-349.
21. Google. (2018, July 17) Google Play.
22. Neha ST (2013) Forensic Analysis of WhatsApp on Android Smartphones. University of New Orleans These and Dissertations.
23. Access Data. (2018, July 17) FTK Imager.