International Research Journal of Computer Science (IRJCS)
Issue 08, Volume 07 (August 2020)
ISSN: 2393-9842
https://www.irjcs.com/archives

A FORENSIC ANALYSIS OF WHATSAPP ON ANDROID SMART PHONE

Dr.Iyobor Egho-Promise*
Regional Technical Head, North/BA Regions
Glo Mobile Ghana Limited, Tamale, Region, Ghana
eghopromise@yahoo.com
Bamidele Ola
Technobeacon Consulting Ltd, London, UK
olawest@technobeacon.com
Aaron Arhin
Computer Science Department, Koforidua Technical University, Ghana
Arhin175@gmail.com
Richard Asuming
Computer Science Department, Koforidua Technical University, Ghana.
richardluvly@gmail.com

Publication History
Manuscript Reference No: IRJCS/RS/Vol.07/Issue08/AUCS10080
Received: 02, August 2020
Accepted: 12, August 2020
Published: 14, August 2020
DOI: https://doi.org/10.26562/irjcs.2020.v0708.001
Citation: Dr.Iyobor, Bamidele, Aaron,Richard (2020). A Forensic Analysis of Whatsapp on Android smartphone. IRJCS:: International Research Journal of Computer Science, Volume VII, 209-219.
https://doi.org/10.26562/irjcs.2020.v0708.001
Peer-review: Double-blind Peer-reviewed
Editor: Dr.A.Arul Lawrence Selvakumar, Chief Editor, IRJCS, AM Publications, India

Abstract: Android® forensics have progressed overtime providing vital opportunities as well as existing challenges. As an open source platform, Android provides freedom to developers to make contribution towards the rapid growth of the Android market. On the other hand, users of Android devices might not be familiar with the privacy and security implications of installing mobile application on their cell phones. Users might assume that a device that is password locked protects their personal information, but the devices might preserve the personal information on the devices in ways the end users might not be aware of. This research focuses on one of these applications 'WhatsApp®', a very popular social networking mobile application. This research outlines the ways forensic investigators can extract valuable information from WhatsApp and from similar mobile applications installed on the Android platform. The major focus is the extraction and analysis of the data of the application user from non-volatile external storage and the volatile memory (RAM) of an Android device.

Keywords: Analysis; Smart; Forensic; WhatsApp; Android;

I. INTRODUCTION

According to Anglano (2014) [1], WhatsApp Messenger is a registered, cross-platform instantaneous messaging application for the cell-phones. Users can send text messages along with the videos, images and audio media messages. The consumer software is accessible for Blackberry OS, Android, Series 40, Blackberry 10, iOS, Windows Phone, and Symbian (S60) and Series 40. WhatsApp Inc. was established in the year 2009, in Santa Clara, California, by Jan Koum and Brian Acton who are specialists of Yahoo!. WhatsApp showed a huge increase in dealing with messages, from handling two billion messages in April 2012 to deal with ten billion messages every day in August 2012. As per Financial Times, WhatsApp has carried out a similar thing to the SMS on the cellular phones that were done by Skype to the international calling on phone line. WhatsApp has also grown as a social networking app in the speed of light. On Android phones, the number of downloads for WhatsApp has surpassed one hundred million. It is among the top 30 application and among the top five free communication application on Google Play in just three years [2]. WhatsApp makes use of Wi-Fi or 3G of the user for messaging with family and friends. The size of the application is 10M and it also has an option for sending and receiving unlimited messages free of cost. In terms of its characteristics, WhatsApp can be utilised as an overall solution for simple and economical networking on mobile phones.

For accessing the WhatsApp services, the user has to provide the phone number that is utilised internally for creating a user account having user-id, for instance, ([poennumber]@s.whatsapp.net). WhatsApp can also auto-sync to the address book and shows all of the acquaintances making use of WhatsApp automatically. There has been no constraint on the length and the quantity of messages that can be interchanged along with this, there is no carrier IM fee applicable. Furthermore, the service user does not have to insert a sim-card for using WhatsApp, just an internet connection, supported cellphone, and storage space on mobile phone is required for downloading the WhatsApp application [3]. WhatsApp utilises a customised form of the open standard Extensible Messaging and Presence Protocol (XMPP) for exchanging information through the internet. Messages can be in terms of plain text, multimedia messages such as contact cards, audio, address book, video, icon and location. The excess of individual information which can be interchanged provides a motive for looking at WhatsApp by a forensic glass.

A. Scope

The major emphasis of this research is regarding WhatsApp Messenger enabled and how users of smartphone relate to the application.

B. Main Objectives

The main objective of this research is to carry out the forensic audit on android smartphones with its related WhatsApp Messenger data security breaches.

C. Sub Objectives

The sub-objectives of this research will sort for identifying the data security matters in WhatsApp on Android smartphones including

1. To understand Web Malware
2. To understand unencrypted backups
3. To understand encryption vulnerabilities

D. Problem Statement

With the advancement within the information communication technology, the security issue has gone far beyond the ordinary password usage. Following are some of the security issues regarding WhatsApp on an Android smartphone

- Facebook Intrusion and Data Sharing
- Web Malware
- Encryption Vulnerabilities
- Unencrypted Backup

II. LITERATURE REVIEW

The literature has a forensic examination of IM applications on cell phones as a matter of numerous works. In comparison to the previous researches, this research has an extensive scope since it takes into account, all of the artifacts produced by the WhatsApp Messenger that is the log files, contacts database, preference files and the avatar pictures. This research also provides a more comprehensive and detailed analysis of these artifacts. Moreover, it also provides an explanation regarding how these artifacts can be associated for comprehending different types of material having evidentiary cost, for instance, whether a message has been really conveyed to the end point after being directed, in case a service user left or joined a group chat prior or after a particular time period and when a specific service user has been included within the contact list. Husain & Sridhar (2010) [4] has emphasised on the forensic investigation of three IM applications including Yahoo!, AIM, Google Talk and Messenger on the iOS platform.

Their research work is different from this research in terms of both the IM applications and the smartphone platform that has been taken into consideration. The study of Kumar & Sharma (2016) [5] emphasises on the evaluation of different IM applications along with WhatsApp Messenger on different smartphone platforms, involving Android with the intent of determining the algorithms for encryption utilised by them. However, their research has not dealt with identifying, analysing and associating all of the artifacts produced by the WhatsApp Messenger. Tso et al. (2012) [6] has focused on the evaluation of iTunes backups for the iOS smartphones having the objective to determine the artifacts left by numerous social network application involving WhatsApp Messenger. Their research work is different in a sense that it is focused towards iOS and iTunes and the chat database of WhatsApp is taken into consideration meanwhile just this artifact involved in the iTune backup. Moreover, the information that is stored within the chat database is evaluated just in parts. The research work of Thakur (2013) [7] and Mahajan et al. (2013) [8] is related to this research as they have emphasised on the forensic analysis of the WhatsApp Messenger on Android. On the other hand, theses researches have focused majorly on the forensic acquirement of the artifact left by the WhatsApp Messenger and deals with their evaluation just in parts. Their research is limited to the chat databases and partial analysis. Similar considerations are applicable to the WhatsApp Xtract tool by Sangiacomo & Weidner (2012) [9] which excerpts certain data kept within the chat database and particularly within the contracts database, though, without giving any kind of description regarding the ways these databases are explained.

A. WhatsApp Database – hardware acquisition

Mahajan, Aditya, M.S. Dahiya and h.P. Sanghvi (2013) [8] made use of a physical analyser for analysing the instantaneous messenger applications including Viber and WhatsApp on the Android gadgets.

Within the context WhatsApp, timestamps, chat messages artifacts and names of documentations received and sent were identified. On the contrary, the location of those files where they were stored was not identified. After the File System Extraction, on the manual investigation of WhatsApp application, the database records including wa.bd and mgstore.db were identified along with the minutiae of the chat conferences. Mostly, for the existing messages, the databases extraction and well comprehensive analysis were performed. For the WhatsApp application data remnants, the analysis of RAM was not taken into consideration. Along with this, the recovery of the deleted information was stated as a forthcoming implication. This research has taken the extraction of deleted messages from RAM into consideration and has also been successful in doing so. Furthermore, the database extraction was carried out by utilising the UFED physical analyser, however, an unencrypted version can be attained on rooting the phone.

B. WhatsApp Database - Software Acquisition and Analysis

Picasso (2012) [10] also made a contribution towards WhatsApp forensic by writing a tool for decrypting and organising the SQLite databases files within the organised HTML format. The tool functions for the decrypted and encrypted database files. The WhatsApp Database Encryption Project [11] has prepared a well-known susceptibility within the Android application of the AES cypher: the 192-bit key can be noticed carrying out both active and static investigation on the software package and the result is:

346a23652a46392b4d73257c67317e352e3372482177652c.

The python script utilises this key for decrypting and encrypting the db file and provides the results within the HTML page. The study also infers that the equal encryption key has been utilised for all kind of WhatsApp connexions on Android. This research has made use of python tool for decrypting as well as for interpreting the encrypted database which was performed in a successful way. Figure 1 shows the output of the research. This tool has enabled to read the database files alternately by the 'SQLite browser', however, the data representation and timestamps are not straightforward. Moreover, one more benefit of this instrument is that the exchange of data related to media is exhibited on the HTML page. An individual does not have to explore the media folder distinctly. This tool can also valuable within the comparison of the data that this research analyses. All these features of this tool epitomise it within a useful tool; however, after messages are removed from the database, the instrument cannot recover or represent them. Only the static information can be presented which is existent within the record. The databases that are on the external storage particularly the SD card is updated occasionally resulting in the representation of old data. The aim of this research was to acquire the deleted messages; therefore, the decision was made on the volatile memory acquisition and analysis. Moreover, for acquiring the updated user information, live analysis on the device should be carried out and volatile memory must be acquired for the advance analysis.

III. RESEARCH METHODOLOGY

The fundamental objective of this research was testing the WhatsApp application from the scientific perspective on the Android phone. The simple method towards the acquirement of Android memory was taken into account within this research and the steps were taken for reducing any sort of human impressions on the retrieved data. Anglano (2014) [1] WhatsApp 2.9 was initially installed on more Android phone through Google play store. The applications are put in storage within the phone's internal memory. The application routinely synchronises with the contacts in phone showing the individuals who are already utilising WhatsApp. When the phone in which WhatsApp is installed is opened, the 'com.whatsapp' the procedure obtains the signal for initiating the 'External Media Manage' and 'Message Service' service that runs within background till the phone is on. The messages that have been exchanged have been stored within 'msgstore.db' and 'wa.db' which are the SQLite's databases. In order to access the data faster, the databases are loaded within RAM. Generally, all of the content might not be persevered or might be overwritten because of swapping within RAM; however, this might not be accurate for the Android.

On the basis of the lifecycle process of Android, the application performs for as long as it is possible. Mahajan et al. (2013) [8] Android carry out garbage collection on the basis of app by app and is established on the process precedence. If a higher precedence process requires additional memory resources and the RAM is occupied, then the data might persevere within memory for the long time period. This characteristic has proved to be beneficial within this research for the extraction of WhatsApp matters from the memory. The WhatsApp app is an immediate messenger service, therefore, the users get informed regarding messages by the push-mechanism once the messages are attained, and hence WhatsApp upholds the high precedence within memory, generally the visible procedure. This provides convenience to the users for constantly receiving the messages within the background without the requirement of downloading them from a particular web server similar to the email service. The main issue after acquiring the file msgstore.db.crypt is its decryption. The tool made by Francesco Picasso for organising and decrypting the SQLite database files within the organised HTML form was helpful. This tool worked for the decrypted and encrypted files. The WhatsApp Database Encryption Project has recognised a liability within the Android employment of the AES Cipher: the 192-bit key can be perceived carrying out the active or static examination on the software package [12] (Nations, 2013).

The python script utilises the similar key for decrypting the encoded db file and offers the results within the well-organised HTML page. The research indicates that for all of the installation of WhatsApp on Android, the same encryption key is utilised. This research has utilised the Python tool for decrypting and reading the encrypted the database which was successfully performed with the latest variety of WhatsApp 2.11.186. The database files can be read alternately by the 'SQLite browser', however, the data representation and timestamps are not direct. Within WhatsApp Xtract tool, all of the media subjects which are interchanged are exhibited on the HTML page itself and there is no need to search separately within the media folder. This tool can also be beneficial in the comparison of the data that has been analysed in this research.

A. Finding the Information

All of the chats of WhatsApp are stored on the SQLite database. The path of the database file varies from one platform to another platform.

Android

(/sdcard/WhatsApp/Databases/mgstore.db.crypt)

iOS

(Application/net.whatsapp.WhatsApp/Documents/ChatStorage.sqlite)

How to use:

Step 1: Download WhatsApp Xtract package on the computer and excerpt it.

Step 2: Download and install Python programming language environment on the computer.

Step 3: Open the file where the WhatsApp Xtract archive is downloaded. Discover the file named as !installpyCrypto.bat, right-click on it and click run as administrator. This bat file will implement the pypm install pycrypto Python command. This common set up the pycrypto library automatically on the computer that will be utilised for decrypting the backup information of WhatsApp.

Step 4: Within the similar folder, run either whatsapp_xtract_iphone.bat, whatsapp_xtract_android_crypted.bat or whatsapp_xtract_android.bat based on the backup folder that has been utilised. In order to execute any of these files, just right click on it and click run as an administrator as done in the above step. In order to specify the WhatsApp backup file manually, whatsapp_xtract_console.bat can also be run.

```
/*\For Android DB:*/
Python whatsapp_xtract.py -i msgstore.db -w wa.db
/* If wa.db is unavailable */
python whatsapp_xtract.py -i msgstore.db
/*For crypted DB*/
python whatsapp_xtract.py -i msgstore.db.crypt
/*For iPhone DB*/
python whatsapp_xtract.py -i ChatStorage.sqlite
```

B Results

After the completion of command or execution of the bat file, all of the WhatsApp backup data will be decoded and will be shown within the default browser of the computer.

IV. DATA ANALYSIS

WhatsApp offers different communication forms to the users that are broadcast messages, the user to user communications as well as group chats. During communication, the users might exchange the plain text messages and multimedia files comprising of video, images and audio, geolocation information and contract cards.

Table 1: WhatsApp Messenger Artifacts

Row#	Content	Directory	File
1	chat database	/data/data/ com.whatsapp/databases	msgstore.db (SQLite v.3)
2	Contacts database	/data/data/ com.whatsapp/databases	wa.db (SQLite v.3)
3	copies of contacts avatars	/mnt/sdcard/WhatsApp/ProfilePictures	UID.j, where UID is the identifier of the contact
4	avatars of contacts	/data/data/com.whatsapp/files/Avatars	UID.j, where UID is the identifier of the contact
5	backups of the chat database	/mnt/sdcard/Whatsapp/Databases	msgstore.db.crypt msgstore-<date>.crypt
6	received files	/mnt/sdcard/Whatsapp/Media	various files
7	log files	/data/data/com.whatsapp/files/Logs	whatsapp.log, whatsapp-<date>.log
8	user settings andpreferences	/data/data/ comm.whatsapp/files	various files
9	sent files	/mnt/sdcard/Whatsapp/Media/Sent	various files

Every user is related to its profile, a set of data that involves his/her name on WhatsApp, status line and avatar that is a visual file generally an image. Every users' profile is kept on the central system from where it is downloaded by other users of WhatsApp which involve users within their acquaintances list. The central systems also offer other facilities, such as authentication, registration of user, and message transmission. As stated in [8], the artifacts produced by WhatsApp Messenger on the Android gadgets are put in storage within the set of records, whose site, content and name are listed in Table 1 above. The artefacts of the Whatsapp Messenger are given in the above table. In the next section, Analysis and Comparison of these artefacts will be provided to gain the insight of different types of data: Initially the discussion will start contact information (Sec. 4.1), then the analysis will move on to returned messages (Sec. 4.2), and lastly we will evaluate the settings and user precedence (Sec. 4.3).

A. Analysis of Contact Information

The contact information's evidential value is prominent and contentious, as it enables an analyst to acquire the knowledge about with whom the user was interconnecting or in correspondence with. Initially, this unit will give an account of the data which is gathered and preserved in the contact database. Furthermore, we will address how the data could be evaluated and analyzed to demonstrate:

- The record of the contacts
- When the user was registered in the database
- When (only if) a given contact has been blocked
- How to deal with the deleted contacts

1) Retrieving Contact Information

The contact database (wa.db) encompasses three listings, to be more specific, first is the wa contacts (which gathers and stores the record of every contact), second is the Android Metadata, and lastly Sqlite Sequence (both, the android metadata and SQLite sequence perform the function of storing housekeeping data which have no evidential value. The configuration and arrangement of the record of wa contacts are demonstrated in the table provided below (Table 2). In that table, we will identify the areas which contains the data extracted from the Whatsapp system and which possess some evidential value, from the storing data which have been retrieved from the user's phonebook (which is reserved by the user, not the Whatsapp, is not related to the research study). As it can be identified from the given table below, that every record caches the Whatsapp ID (field jid) of the user, a series systematized as 'x@s.whatsapp.net', in which 'x' is denoted as the contact's cell phone number (on account of the intelligibility, consequentially the user is signified via phone numbers rather than the entire WhatsApp IDs). Moreover, every best report reserves the figuration name (field wa name), and the status series (data status) of the respective user. The data is used to distinguish the actual WhatsApp users from the invalid or fake ones. WhatsApp messenger incorporates every phone number in the database record which is saved in the contact list of the user, despite being the numbers is not authorized in the WhatsApp structure. The display picture in the WhatsApp plays an integral part to determine the actual identity of the user.

Table 2: Structure of the WA Contacts Table

Data deriving from the WhatsApp System	Data Deriving from the device's phonebook
Field Name	**Meaning**
ID	The arrangement of the record (arranged by SQLite)
JID	Contact's Whatsapp ID (a series arranged as 'x@s.whatsapp.net', where 'x' is denoted as the contact's cell phone number
IS Whatsapp User	Integrates '1' if the contact is associated with an actual WhatsApp user, or else '0'
Count of unseen messages	Amount of texts transmitted by another contact which were received but not yet read
Photo TS	Unfamiliar, every time set to '0'
Thumb TS	Unix epoch time (10 digits) demonstrating after the user has set their recent display picture
Photo ID timestamp	Unix millisecond epoch time (13 digits) demonstrating when the present-day display photo of the user has been locally downloaded
WA name	WhatsApp name of the user which is set in the WhatsApp profile of the user
Status	Current status track of the user (as set in their profile)
Sort Name	Name of the user utilized in assorting operations
Number	Cell Phone number linked with the user
Raw Contact ID	The arrangement number of the contact
Display name	Display name of the user

Data deriving from the WhatsApp System	Data Deriving from the device's phonebook
Field Name	**Meaning**
Phone type	Sort of the phone i.e. android, ios
Phone label	Brand connected with the phone number
Provided name	Provided name of the handler
Family name	The family name of the handler

Structure of the table could have particularly consorted with that individual. The display image of the user x@s.whatsapp.net is mobilized as, as a JPEG folder labelled x@s.whatsapp.net.j, in the records specified in Table 1, row 4 and 5. The timestamps are reserved in the thumbs TS and image ID. The field of thumbs TS demonstrate when the current display picture is uploaded or set by the user, and photo ID indicated when the existing image of the user has been loaded locally.

2) Determining When a Contact Has Been Added

In a few analyses, it is essential to identify at what time a given operator has been registered in the contact database 2. This data is hardly arranged in the table of WA users, nevertheless it could be derived or detected through the evaluation of the log files which have been advanced by WhatsApp Messenger (these are arranged in the records scheduled in Table 1, row 6). When user is registered in the database of the wa.db, WhatsApp Messenger transcripts numerous actions which are accompanied with the period of their transaction and to the Whatsapp ID associated with the user. Examples of these actions or affairs, correlated to the incorporation of user 39331xxxxxxx, are recorded in Figure 1, it can be observed from the Figure 1 those subsequent occurrences are transcribed every while a new consumer is registered:

(a) The finding that the operator is hardly intervening up till now in the contact database (line 4)
(b) The analyses to the central structure to obtain numerous data related to the contact (line 7,10, and 14).
(c) The download completion of the correlated display picture (line 17).

When these events occur, it demonstrates that the consumer was registered to the contact database (which is September. 25, 2013 at 14:14:24, in the specimen provided).

3) Dealing with Blocked Contacts

WhatsApp Messenger provides the feature to its user to block anybody from their phonebook, which prevents every connection or message with the blocked user till they are unblocked. In an analysis, it is significant to identify if the user was blocked or not at a specified period, to authenticate or eliminate the manipulation or acceptance of a message delivered at that time. The records associated to the blocked consumer is neither reserved in the contact database nor anywhere else on the device's memory (it can be speculated that the record of blocked contact is located on the fundamental structure of WhatsApp, as when then blocking is happening, messages are exchanges by WhatsApp Messenger with it). However, blocked users could be determined, within some conditions, through assessing and analyzing the logbook files. Once a user is blocked, an occurrence of recording the WhatsApp ID of the blocked user and the exact period of the action's proceeding is certainly reported within the record file Inopportunely, once the blocked user is unblocked, the occurrence which is transcribed is hardly reported as the WhatsApp ID of the linked user, and it is aggregated (i.e. it might be regarded as the set of users being unblocked at once) Therefore, it is every time practicable to identify regardless if and when a giver contact X was blocked, but if the user is still not unblocked at the provided period, it could only be perceived either;

(a) No actions of blocking are reported in the record file once the occurrence of blocking, or
(b) The blocking occurrence is existing, but consumer X was only blocked at that particular period.

It refers that whether a number of consumers are blocked at once, and one or more than one events of unblocking is transcribed, it is never practicable to identify in particular the users who is still blocked and who was unblocked. It is significant to emphasize that the interferences mentioned above could only be executed with the availability of record files, recording, blocking and unblocking actions (i.e. is the WhatsApp messenger have not deleted the previous records to free up space for the newer ones). Concluding to that, it could be observed that if there is no data present at all on the side of the user who has been blocked, so it is imperative within the analysis to declare whether the user was blocked or not by anybody from their phonebook.

4) Managing deleted contact

Usually, contact is deleted by a user for the purpose of concealing an interaction of the past. Hence, when a contact is deleted, the conforming data is removed from the WhatsApp contact table. In some instances, the retrieval of the deleted information is possible by using certain appropriate techniques and tools. This recovery of the record is made certain when the above table has not been vacuumed by the SQ lite engine [13]. The possibility of retrieving the deleted contact information is indicated by our experiments that are carried out employing Oxygen Forensic SQLite Viewer. Nevertheless, generally, at the time of analysis, it might be the case that the recovery of deleted contact is impossible because the deleted information have been cleaned and vacuumed. In the following circumstances, the determination of removed data might be possible.

And this can be accomplished by first regenerating the contacts list that have been in connections earlier (the list can be reconstructed by the analysis of log files as described in section 4.1.2), followed by the comparison of this list with the wa contacts table' contents: the contacts that exist in the list and are absent from the record are the ones that have been eradicated. However, the recovery of deleted contacts is attained if the log file in which the addition of a concerned contact is reported is still accessible at the time of analysis [4].

B) Inspection of exchanged messages

All of the text messages that have been conveyed or attained in the chat database msgstore.db are stored by Whatsapp Messenger (revealed in the directory that is listed in Table 1, row 2). The investigation of the mentioned database enables to restore exchanged messages chronology, specifically to determine the time of message exchange, the group of users that are engaged, the information that it possessed, and also when and whether the recipient has actually received that message. However, every single step will be described individually in the following analysis: we begin with the elucidation of the chat database structure i.e. sec. 4.2.1 followed by the description of how to

(1) Regenerate the chat history in section 4.2.2.
(2) Discover and extricate the content of the message in section 4.2.3.
(3) Determine the message status in section 4.2.4.
(4) Identify user's set among which the message has been reciprocated in section 4.2.5.
(5) Manage deleted messages in section 4.2.6.

1) The structure of the chat database

Following the three tables are associated with msgstore.db database.

- Messages, which possess data for each and every message that have either been conveyed or attained by the user. The fields of these records have been classified into two categories in order to make it clearer. Two of these categories include: Firstly, those preserving the features of messages- Listed in table 3, secondly, those preserving the matters of messages and metadatset that is consistent.
- SQLite sequence, that preserves housekeeping data which is utilized by the Whatsapp Messenger internally. It lacks the value pertaining to evidence because its structure is left unreported. As outlined in [13], distinguishing backup copies of the msgstore.db database is generated are usually produced by Whatsapp Messenger. These backup copies are reserved in the directory listed in
- Chat list, the information of the conversation that is held by the user is contained in the chat list (a conversation is contained in the set of messages that are traded with a specific contact). Table 5 is a description of its fields.

Table 3: Structure of the Messages Table: Fields Storing Message Attributes

Field name	Meaning
Id	Record sequence number
Key id	unique message identifier
Key remote jid	Whatsapp id of the contact (a string constructed as 'x@s.whatsapp.net', where phone number of the contact is denoted by 'x')
timestamp	time of sending if the key from me='1', record insertion time otherwise (taken from the local device clock, and encoded as a 13-digits millisecond Unix epoch time)
Status	message status: '0'=received, '4'=waiting on the the central server, '5'=received by the destination, '6'=control message
received timestamp	time of receipt (taken from the local device clock, and encoded as a 13-digits millisecond Unix epoch time) if the key from me='0', '-1' otherwise
key from me	message direction: '0'=incoming, '1'=outgoing
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 the key from me='1', '-1' otherwise
receipt server 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 the key from me='1', '-1' otherwise
send timestamp	unused (always set to '-1')
remote resource	The ID of the sender (only for group chat messages)
recipient count	number of recipients (broadcast message)
needs push	'2' if the broadcast message, '0' otherwise

2) Extricating the contents of a message

Along with the exchange of plain text messages, messages that contain different types of data namely: contact cards, multimedia files (videos, audios, and storing images) and record of geo-location are permitted to be exchanged between the users by Whatsapp Messenger.

Media wa type field is the indicator of the kind of data that is transferred with the message. Whereas, data about the content of the message is proliferated over various fields, for messages that are non-textual (relying on the specific type of data). In actual fact, whilst the data field preserves the content of the textual messages, for various other contents and data types, as elucidated below the situation is more complicated.

Multimedia Files- As soon as the multimedia file is sent by the user, simultaneously various activities are initiated without even the users being informed about it.

Step1- The file is copied into the folder by Whatsapp Messenger. Table 1, row 8 is its representation

Step2- The multimedia file is uploaded to the Whatsapp server. The function of the Whatsapp server is to return the URL of the corresponding location.

Step3- The URL confined in the message is sent to the recipient by the sender. Finally, when the message containing the URL is received by the recipient, its acknowledgement from the recipient's side is sent back to the sender.Upon the completion of the above-mentioned steps, the record is preserved into the sender's messages table. The fields which are relevant to the message content that is recorded. As it is clear from Figure 1a below that media mime type file indicates the file's type. Media name field is the location where its name is saved. Its size in bytes by media size (40267 in the example), and its thumbnail in the raw data field (as a blob, i.e. a binary large object) is stored. The URL location on the server that is central where the file storage is temporary is kept in the media URL field. The server is responsible to name the file and the last part corresponds to that prescribed name. Lastly, the media hash field stores the base64-encoded SHA-256 hash of the file that is transmitted. The Exchange of Multimedia File: sender side on the recipient side, after when the message is received, Whatsapp Messenger showcase the file's thumbnail that is transmitted. Only on the request of the recipient, the downloading of the actual file can take place afterwards. Immediately after the file is received by the recipient, stores in their table of messages a record like the one displayed in Figure. 1a.

Figure 1a: Uniformity of fields

Exchange of Multimedia File: recipient side to those kept by the sender (in precise, wa media type, media hash media size, raw data, and media mime type,). With the exception of the name assigned to a file, the URL media content is distinguishing. The media name field is vacant, unlike the sender. Therefore, Whatsapp Messenger assigns a local name to that file and this name is unknown. But the identification of the file is accomplished by the comparison of all the files that have been received to the SHA-256 hash preserved in the corresponding record. Finally, it is notified that the comparison between these two files yields the correlation of the file that is received by the recipient and the file that is sent by the sender. (that are saved, as explained above, in the media URL and media hash fields of the corresponding records).

Contact Cards- Some messages contain contact cards. These contact cards are usually extricated from sender's phonebook and correspond to records of messages that preserves the data which is conveyed in VCARD format into the data field. It also corresponds to the name assigned to that contact present in the media name field by the sender. Figure 1b shows an example VCARD.

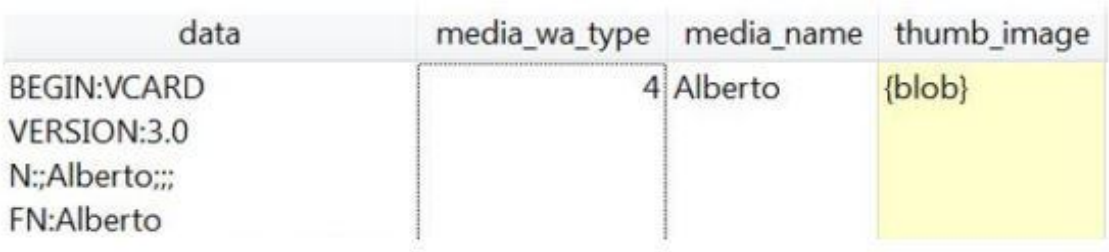

Figure 1b: Media file contents in VCARD

Geolocation coordinates- The geographic location coordinates are mostly acquired by the Android Location Services that run on the devices. The exchange of topographical coordinates between the users is enabled by Whatsapp Messenger [14]. Senders as well as on the recipient's side, geographic coordinates containing messages corresponds to {to messages data that encompasses the longitude and the latitude values into the longitude fields, latitude and a JPEG thumbnail of the Google Map showing the above coordinates in the field of raw data. Figure 2 is a case of this kind of record.

longitude	latitude	raw_data
8,6186183	44,923975	{blob}

(a)

(b)

Figure 2: Record of geologica coordinates

3) Determining the state of the message

The pattern of message exchange in WhatsApp is such that it is first transmitted to the central server [15]. The central server then conveys it to the recipient if they are available or store it otherwise until it can be delivered. Hence, there is no direct exchange of messages among the users.

5) Determining the Partners of a Message

WhatsApp, along with user to use communication, provide two kinds of collective communications to the users, involving:

- Broadcast (i.e. one to many) communication, by which a user who is the source user conveys the same message to a set of other users i.e. the destination users who are not familiar with each other and whose probable answers are transmitted to just the source user;
- Group Chats, offering numerous communication services, by which every message conveyed by any of the user who is a part of a chat is attained by all of the users who are members of that chat. Although the WhatsApp ID of the communication companion within a user to user communication is simply recovered from the key remote jid field, for identifying the set of users included within a broadcast or group chat message, numerous sections have to be associated which are discussed as follows:

Broadcast Messages – When a user conveys a broadcast message, a separate folder is developed within his or her message table for every recipient along with one for himself or herself as illustrated within Figure 3(a) which depicts the folders that are produced by a broadcast message transmitted to the users 3920xxxxxxx, 39335xxxxxxx and 39333xxxxxxx. As depicted in Figure 3(a), all of the records corresponding to the same broadcast message have the similar message identifier (stored within the key id field), so that they can be identified without any difficulty. Every record stores within the key remote jid field the WhatsApp ID of the recipient (the sender makes use of the keyword broadcast for denoting himself as a recipient).

	key_id	key_remote_jid	remote_resource	recipient_count	needs_push
1	1382694005-1	39320 @s.whatsapp.net	39320 @s.w...	3	2
2	1382694005-1	39335 @s.whatsapp.net	39320 @s.w...	3	2
3	1382694005-1	39333 @s.whatsapp.net	39320 @s.w...	3	2
4	1382694005-1	broadcast	39320 @s.w...	3	2

(a)

key_id	key_remote_jid	remote_resource	recipient_count	needs_push
%~1382694005-1	39320 @s.whatsapp.net		(null)	0

(b)

Figure 3 Records produced for a broadcast message transmitted to three recipients on (a) the sender, (b) one of the recipients. Only the sections which make contribution to the identification of the associates are shown.

Whereas the recipient count fields and remote resource stores the WhatsApp ID of the set of destinations and their respective number (field requires to push in its place it always store the value '2'). The state of every destination is rather diverse (Figure 3(b)) as each one of them store within his or her messages table, just a single folder which is created when it gets the broadcast message. This folder can be differentiated from those consistent to the non-broadcast messages by examining the value that is stored within the key id field which entails within the concatenation of the %~ characters with the message identifier fixed by the sender.

Group Chat Communication – When a message is conveyed in a group chat, a folder is created within the messages table of all of the group members (along with the sender). Every record stores within key remote jid field, the identifier of the group (the group id), a string formatted as fcreator's phone numberg-fcreation timeg@g.us (Where the creation time is encoded as a Unix epoch time). For illustration, take into account a group chat comprising of three members that is 3933xxxxxxx, 3936xxxxxxx and 3932xxxxxxx (in the subsequent represented as A, B and C, correspondingly for brevity) where every user, sequentially, sends a message to the group with the textual information 'Message from X' (Where 'X' is the user's name). Let us concentrate on the files stored within the messages table of user A at the end of this interchange which is illustrated within Figure 4 (the situation for the other users is identical)

	key_remote_jid	remote_resource	key_from_me	status	timestamp	data
1	3933 -1363078943@g.us	(null)	1	4	1363079028764	Message from A
2	3933 -1363078943@g.us	3936 @s.whatsapp.net	0	0	1363079064000	Message from B
3	3933 -1363078943@g.us	3932 @s.whatsapp.net	0	0	1363079078000	Message from C

Figure 4: Records corresponding to three messages exchanged within a group

As it is evident from the above image, all of these files store the similar group id 3933xxxxxxx-1363078943@g.us within the key remote-id field. From this value, the initiator of the group (user A) and the date and hour of the formation of the group (March 12, 2013, at 09:02:23) can be identified. Moreover, the WhatsApp ID of the message originator is stowed within the remote resource field. Although the time of the receiving message is kept within the timestamp field. It must also be noted that A also stores the records consistent to the message that he or she has conveyed to the group (record no. 1 in Figure 4). The files similar to this can be recognised effortlessly by just examining the contents of their position and remote resource fields which stores the value '4' and 'null' correspondingly. It should also be noted that the set of receivers, i.e. of the set of members of groups at the time of conveying is not kept at any place in the record. On the other hand, it can be identified incidentally by investigating the files conforming to the control messages which are exchanged routinely by different group members each time a user leaves or joins the group. These messages which are also kept within the messages table continuously comprises of value '6' within the status field and code within the media size field the particular operation consistent to the message (in particular, the values 'l', '4' and '5' indicate creation of the group, joining and leaving respectively).

V. CONCLUSION AND RECOMMENDATION

WhatsApp has turn into a well-known application for social networking on which the individuals might be interchanging their personal information and business-related data. This research has depicted that an individual can get whole access to all of the material in WhatsApp as well as in other alike social networking applications, for instance, "Viber". Majority of chat applications survey the same pattern to store messages within the database and periodically bringing up-to-date database. The method taken in this research provided a general plan for all the same applications which run on the android gadgets. This research was able to attain its aim effectively. One must be aware of the fact that a password-locked cellphone is not a black box and one can excerpt valued application user information from the file as well as volatile memory [7]. The results of this research can be valuable for Live Forensic Analysis on Android Smartphones. The databases are just updated once each day, therefore, the information obtained might not be up-to-date at the investigation time, while live acquisition and evaluation of the volatile memory can provide current information. While performing forensic investigation, the existence of the most current messages for the purpose of investigation can play a significant role [6]. Along with the recent messages, an individual can also look at the deleted messages. Therefore, recovering the artefacts after the factory reset of the phone or recovering the deleted data can be considered as the future characteristic [16]. Within the future, additional work can be carried out on the explanation of the RAM data within the human-readable form. The tool presented in this research can be customised for displaying the user-specific information on the basis of the requirement of an individual. As of now, the technique highlights three significant aspects of user data namely user's phone numbers, exchanged messages as well as database enquiries providing the basic database framework for WhatsApp.

REFERENCES

1. Anglano, C., 2014. Forensic analysis of WhatsApp Messenger on Android smartphones.. Digital Investigation, , 11(3), pp. 201-213..
2. Developers, 2020. Application Fundamentals. [Online] Available at: https://developer.android.com/guide/components/fundamentals.html [Accessed 31 7 2020].
3. Developers, 2020. Processes and Treads. [Online] Available at: https://developer.android.com/guide/components/processes-andthreads.html [Accessed 31 7 2020].
4. Husain, M. & Sridhar., R., 2010. iForensics: Forensic Analysis of Instant Messaging on Smart Phones. . In: In Sanjay Goel, editor, Digital Forensics and Cyber Crime, volume 31 of Lecture Notes of the Institute for Computer Sciences, Social InformInformatics and Telecommunications Engineering.. s.l.:Springer Berlin Heidelberg.
5. Kumar, N. & Sharma, S., 2016. Survey Analysis on the usage and Impact of Whatsapp Messenger.. Global Journal of Enterprise Information System,, 8(3), pp. 52-57..
6. Tso, Y.-C., Wang, S.-J., Huang, C.-T. & Wang., W.-J., 2012. iPhone Social Networking for Evidence Investigations Using iTunes Forensics. In Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication, ICUIMC '12, New York, NY, USA.
7. Thakur, N., 2013. Forensic Analysis of WhatsApp on Android Smartphones.. Master's thesis, University of New Orleans, , Volume 1706..
8. Mahajan, A., Dahiya, M. & Sanghvi., H., 2013. Forensic Analysis of Instant Messenger Applications on Android Devices. International Journal of Computer Applications, 68(8).
9. Sangiacomo, F. & Weidner., M., 2012. WhatsApp Xtract (v. 2.1),. [Online] Available at: https://code.google.com/p/hotoloti/downloads/list. [Accessed 31 7 2020].
10. Picasso, F., 2012. Zena Forensics "WhatsAppXtract 2012". [Online] Available at: http://code.google.com/p/hotoloti/downloads/list http://blog.digital-forensics.it/2012/05/whatsapp-forensics.html
11. Cortjens, D., Spruyt, A. & Wieringa., W. F. C., n.d. "WhatsApp Database Encryption Project, s.l.: s.n.
12. United Nations, 2013. The United Nations Office on Drugs and Crime. Comprehensive. [Online] Available at: http://www.unodc.org/documents/organizedcrime/UNODCCCPCJEG.42013/CYBERCRIME_STUDY/ [Accessed 31 7 2020].
13. Krynski, L., Goldfarb, G. & Maglio, I., 2018. Technology-mediated communication with patients: WhatsApp Messenger, e-mail, patient portals. A challenge for pediatricians in the digital era.. Arch Argent Pediatr,, 116(4), pp. 554-559..
14. Barghuthi, N. A. & Said., H., 2013. Social Networks IM Forensics: Encryption Analysis.. Journal of Communications, 8(11).
15. Cortjens, D., Spruyt, A. & Wieringa., W. F. C., n.d. "WhatsApp Database Encryption Project, s.l.: s.n.
16. Dorwal, P. et al., 2016. Role of WhatsApp messenger in the laboratory management system: a boon to communication.. Journal of medical systems,, 40(1), p. 14.
17. Sangiacomo, F. & Weidner., M., 2012. WhatsApp Xtract (v. 2.1). [Online]. Available at: https://code.google.com/p/hotoloti/downloads/list. [Accessed 27 7 2020].

