IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 39-44
www.iosrjournals.org

Forensic Analysis of Social Applications

Ambreen F.A.H1, C.N. Kayte[1]
1(Cyber Forensics, Govt. Institute Of Forensic Science, A'bad.(M.S.),India)

Abstract: *The increased use of social networking applications on smartphones makes these devices a goldmine for forensic investigators. Potential evidence can be held on these devices and recovered with the right tools and examination methods. This paper focuses on conducting forensic analysis on widely used social networking applications on smartphones such as Facebook, WhatsApp, etc. The forensic analysis consisted of determining whether activities conducted through these applications were stored on the device's internal memory. If so, the extent, significance, and location of the data that could be found and retrieved from the logical image of each device can be determined and a significant amount of valuable data that could be recovered and used by forensic investigators. This paper presents the forensic analysis of the artifacts left on Android devices by social applicationsand how it can be correlated together to infer various types of information that cannot be obtained by considering each one of them in isolation. By using the results discussed in this paper, an analyst will be able to reconstruct the list of contacts and the chronology of the messages that have been exchanged by users.*

Keywords: *Digital Forensics, Social Media Forensics, Mobile Forensics, Facebook Artifacts, WhatsApp Artifacts*

I. Introduction

The last several years have witnessed the rapid evolution of a new form of online communication known as ocial networking. By joining these services, users can interact and socialize, share information and ideas, post comments and updates, participate in activities and events, upload files and photos, and engage in real-time instant messaging and conversations. These applications attract millions of people from all over the world.

A study estimated that the number of users of online social networks worldwide was about 830 million at the end of 2009 (International Telecommunications Union, 2010). Despite being primarily used to communicate and socialize with friends, the diverse and anonymous nature of social networking makes them highly vulnerable to cybercrimes Phishers, fraudsters, child predators and others. In the recent years, social media apps have gained popularity in general public due to their ease of use. Faster access and informal way of signing up for the app are positive aspects which make them a preferred choice over traditional browser access. Few of the popular social media apps are Facebook, Twitter, Viber, WhatsApp etc. By using these apps, users voluntarily disclose large amounts of information ranging from their likes, dislikes to personal activities. In many cases, the privacy settings of these apps may allow this information to be easily retrieved by any individual, regardless of his relationship to the user. This has two major implications for a forensics investigator. Firstly, it may help him gather information about a suspect by acquiring public information on his social media profiles. Secondly, social media apps can be leveraged by malicious users to create fake and untraceable accounts through which they may perform illicit activities such as stalking, blackmailing, spamming or identity theft etc. This necessitates the investigators need for acquiring knowledge and skills required for collecting artifacts discharged by these apps.

There are a number of criminal cases where the case history is deeply rooted in the usage of social media. In an incident [8], a trainee nurse was killed by a person with whom she connected via social media. Investigation revealed that she was lured by the social media user to meet him at a certain place. Many other cases of murder, identity impersonation and blackmailing have also surfaced in the news[9]. Facebook is the most popular social networking application. Statistics of the first quarter of 2015 show that it has 1.44 billion active users followed by Twitter with 236 million users [6]. Viber is another application that supports calls and instant messaging services. Interestingly, many of the activities are logged on the hard disk of the device from which access is made. The artifacts may reveal details about private connections and the ongoing user activities. Features of these apps may include geolocation that can be used to identify the places from where the criminal accessed the service. Investigating Windows behavior has become imperative for forensic investigators due to increased usage of Windows OS on desktop, laptop and even on cellphones.

Focus of our research is to explore potential location of the remnants for Facebook and WhatsApp. The rest of the paper is organized as follows: we have mentioned related work in section II, followed by overview in section

III. Finally in section IV, we conclude our findings and discuss the direction for the future work.

II. Related Works

2.1. Mobile Device Forensics

Initial work in this field has focused on acquisition techniques and general forensic analyses of smart devices. In his paper, Burnette discussed the forensic examination of older versions of the BlackBerry and covered the hardware and software used for acquisition [1]. He also described several methods of examination, including the use of hex editors and emulators. Later research provided foundational concepts on forensic analysis of the new generations of smartphones (e.g., BlackBerry and iPhone). It outlined the technologies used, the handling procedures, and the common evidence storage locations for each device. The data that could be extracted from the internal memory of these devices included call logs, SMS, MMS, emails, webpage bookmarks, photos, videos, and calendar Notes.[3] Recent scientific research has focused on individual types of smartphones, investigating the methods that could be used to acquire and analyze the internal memory of the device and the data that could be extracted from each device. iPhone data could be acquired by either a physical or a logical method. The physical method requires jailbreaking the system, which causes a slight modification to the system‟s data [11]. However, the latest technique developed by Zdziarski acquires a physical-logical image of an iPhone without jailbreaking the phone [12]. It is considered the best forensic method for acquiring iPhone and has been evaluated by the National Institute of Standard and Technology (NIST.) [13] Similar to iPhones, Android-based smartphones can also be acquired using either a physical or a logical method. The physicaltechnique consists of obtaining a dd image of the phone‟s memory and requires root access to the device [14] Vidas et al. discuss an acquisition methodology based on overwriting the “Recovery” partition on the Android device‟s SD card with specialized forensic acquisition software [15].

2.2. Social Networking Forensic Artifacts

Scientific research has also included the investigation of artifacts left by social networking sites on computer systems and tools that assist in the extraction of these artifacts. Zellers has examined the unique data tags created in different MySpace source-code pages and used these tags to create focused artifact keyword searches [16] Other research discussed the process of recovering and reconstructing Facebook chat artifacts from a computer‟s hard disk [17]. Because many social networking applications are integrated into new smartphones, in cases involving social networks, forensic examiners may be able to find relevant evidence on a suspect‟s smartphone. A forensic examination of the iPhone 3GS (via a logical acquisition) showed that a database related to the Facebook application is stored on the phone‟s memory. The database stores data for each friend in the list, including their names, ID numbers, and phone number [18]. Two other directories related to the Twitter application were also found. These directories store information about Twitter account data, attachments sent with tweets, user names, and tweets with date and time values [20]. A forensic examination of an Android phone‟s logical image showed that basic Facebook friend information is stored in the contacts database (contacts.db) as the device “synchronizes contact‟s Facebook status updates with the phone book” [14]. It also showed that the device stores Twitter passwords and Twitter updates performed through the Twitter application in plain text [14].

Forensic research papers on BlackBerry phones and Windows smartphones, did not mention finding or recovering any data related to the use of social networking applications. Similar to computers, smartphones store data that can help determine how the device has been used or misused. Therefore, activities performed through social networks applications may be stored on smartphones. However, previous research has been limited to the recovery of very basic information related to the use of social networking applications.

It isclear that further experiments focusing on the recovery of artifacts related to the use of social networking applications are required to determine whether activities performed through these applications are stored and can be recovered from smartphones. The forensic analysis of IM applications on smartphones has been the subject of various works published in the literature. Compared with existing works, however, our contribution (a) has a wider scope, as it considers all the artifacts generated by WhatsApp Messenger (b) presents a more thorough and complete analysis of these artifacts, and (c) explains how these artifacts can be correlated to deduce various type of information having an evidentiary value, such as whether a message has been actually delivered to its destination after having been sent, if a user joined or left a group chat before or after a given time, and when a given user has been added to the list of contacts.

III. Analysis Of Social Applications

NIST defines any digital forensics case to consist of four main stages, namely identification, collection, organization and presentation. The identification phase refers to the identification of incident or the evidence. Evidentiary data is acquired and then carved which is followed by reducing the amount of data by discarding off any redundancies. In the organization phase, carved data is examined and correlated with the crime scene in order to reach solid conclusions. Finally, the presentation phase deals with bringing the data in a format that can be understood by the jury.

The three phases are entailed below:-

A. Identification

In digital forensic investigation, identification of evidence could mean a walkthrough of the crime scene and
identifying any hardware or software worthy of collection.

B. Collection

We refer to process of acquiring the disk image as the collection phase. We used FTK Imager for acquisition purpose (as well as for examination as discussed later on in this paper). The choice was made owing to the fact that FTK Imager is considered the fastest and most reliable imaging tool [4]. The disk size, as mentioned in section 3 was intentionally kept to 25 GB. A raw (dd) bit by bit (physical) image was acquired and saved onto another partition drive on the same system.

The identification of artifacts and their examination was easily carried out over the live system. However we chose to execute the entire examination process on storage medias image so as to reflect a real world scenario whereby integrity preservation is of considerable significance. To further preserve the integrity, we used digital hashes of the collected evidence.

Initially process monitoring was done in order to get to the artifact locations. Collection of only specific folders
identified by Process Monitor could be made, however we chose to image the entire disk space instead as our intention was to study the unallocated space for deleted artifacts as well.

C. Examination(Organization)

The image was explored for Facebook, whatsApp artifacts. The artifacts were examined and correlated in order to
analyse their usefulness in any real world case. Findings for each app are discussed below.

Facebook Artifacts:

Most of the Facebook artifacts were found from the same location as in windows 8.1 [2]. Number of SQLite
database files including Friends, Stories, Friend Requests, Messages, were at
\AppData\Local\Packages\Facebook.Facebook8xx8rvfyw5nnt\LocalState\*FacebookID\DB.
Most of these files were easily readable through the DB Browser for SQLite.
Most interestingly message body was displayed in plain text. Although only sender was displayed in the messaging database, but recipient‟s identification could easily be done by using threads.db in corroboration. Another interesting fact is that the location coordinates of the sender were also visible in messages.db. Stories.db showed the Facebook newsfeed visible on user‟s time line and also the permissions to the people in the list e.g. whether they were allowed to send friend requests or not[10].

WhatsApp Artifacts:

Each user is associated with its profile, a set of information that includes his/her WhatsApp name, status line, and avatar (a graphic file, typically a picture). The profile of each user is stored on a central system, from which it is downloaded by other WhatsApp users that include that user is in their contacts.
WhatsApp Messenger stores all the messages that have been sent or received into the chat database msgstore.db whose analysis makes it possible to reconstruct the chronology of exchanged messages, namely to determine when a message has been exchanged, the data it carried, the set of users involved in the conversation, and whether and when it has actually been received by its recipients.

The Structure of the chat database

The msgstore.db database contains the following three tables:

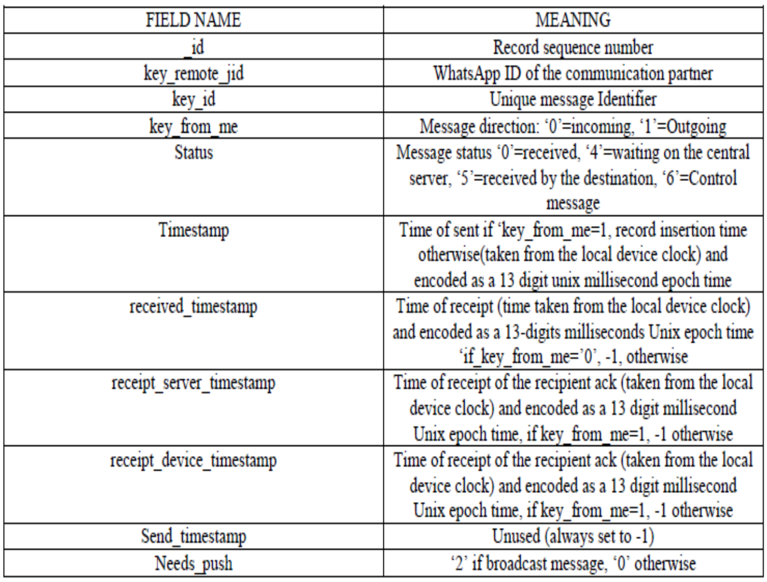
i. Messages, that contains a record for each message that has been sent or received by the user. To ease understanding, we classify the fields of these records in two distinct categories: those storing attributes of the message (listed in Table 1), and those storing the contents of the message and the corresponding metadata (listed in Table 2);

ii. Chat list, that contains a record for each conversation held by the user (a conversation consists into the set of messages exchanged with a particular contact), whose fields are described in Table3;

iii. Sqlite sequence, that stores housekeeping data used internally by WhatsApp Messenger, whose structure is not reported here since it does not have any evidentiary value.

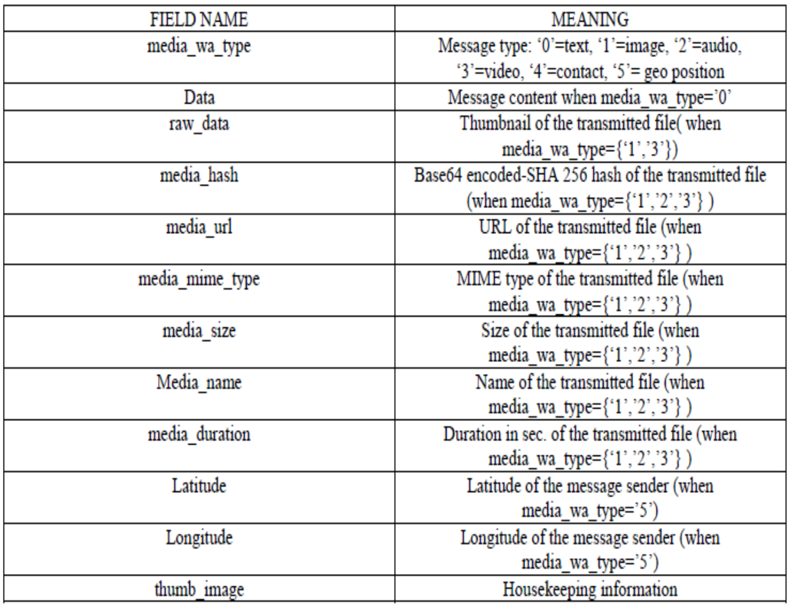
As earlier reported in [19], WhatsApp Messenger usually generates various backup copies of the msgstore.db database, These backups are full copies of the msgstore.db database, and are not kept synchronized with it. Therefore, they are particularly important from an investigative standpoint, since they may store messages that have been deleted from the main chat database. Backups are encrypted with the AES 192 algorithm, but they can be easily decrypted since, as discussed in [5], the same encryption key (namely, 346a23652a46392b4d73257c67317e352e3372482177652c) is used on all devices.

Table1: Structure of the messages table: fields storing message attributes. [7]



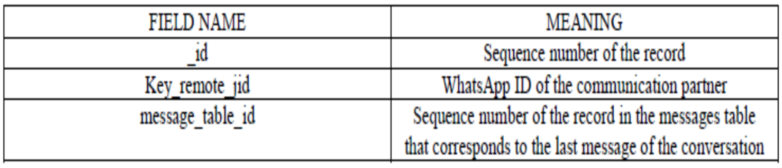
FIELD NAME	MEANING
_id	Record sequence number
key_remote_jid	WhatsApp ID of the communication partner
key_id	Unique message Identifier
key_from_me	Message direction: ‘0’=incoming, ‘1’=Outgoing
Status	Message status ‘0’=received, ‘4’=waiting on the central server, ‘5’=received by the destination, ‘6’=Control message
Timestamp	Time of sent if ‘key_from_me=1, record insertion time otherwise(taken from the local device clock) and encoded as a 13 digit unix millisecond epoch time
received_timestamp	Time of receipt (time taken from the local device clock) and encoded as a 13-digits milliseconds Unix epoch time ‘if_key_from_me=’0’, -1, otherwise
receipt_server_timestamp	Time of receipt of the recipient ack (taken from the local device clock) and encoded as a 13 digit millisecond Unix epoch time, if key_from_me=1, -1 otherwise
receipt_device_timestamp	Time of receipt of the recipient ack (taken from the local device clock) and encoded as a 13 digit millisecond Unix epoch time, if key_from_me=1, -1 otherwise
Send_timestamp	Unused (always set to -1)
Needs_push	‘2’ if broadcast message, ‘0’ otherwise

Table 2: Structure Of the messages - Fields storing information concerning message contents.[7]



FIELD NAME	MEANING
media_wa_type	Message type: ‘0’=text, ‘1’=image, ‘2’=audio, ‘3’=video, ‘4’=contact, ‘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	Base64 encoded-SHA 256 hash of the transmitted file (when media_wa_type={‘1’,’2’,’3’})
media_url	URL of the transmitted file (when media_wa_type={‘1’,’2’,’3’})
media_mime_type	MIME type of the transmitted file (when media_wa_type={‘1’,’2’,’3’})
media_size	Size of the transmitted file (when media_wa_type={‘1’,’2’,’3’})
Media_name	Name of the transmitted file (when media_wa_type={‘1’,’2’,’3’})
media_duration	Duration in 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

Table 3: Structure of the chat_list .[7]



FIELD NAME	MEANING
_id	Sequence number of the record
Key_remote_jid	WhatsApp ID of the communication partner
message_table_id	Sequence number of the record in the messages table that corresponds to the last message of the conversation

II. Discussion

Chronology of the messages exchanged can be re-constructed, the records stored in the messages table must be extracted and decoded. By examining the values in the fields, each message carries its own unique identifier in the key_id field: this value, set by the sender, is obtained by concatenating the timestamp corresponding to the last start time of WhatsApp Messenger (on the sender's device) with a progressive number (indicating the number of messages sent from that moment), and is used also by the recipient to denote that message. Therefore, by using this value, it is possible to correlate the records of the sender's and recipient's databases corresponding to the same message. In addition to plain text messages, WhatsApp allows its users to exchange messages containing data of various types, namely multimedia files (storing images, audio, and video), contact cards, and geo-location information. The type of data transmitted with a message is indicated (as reported in Table 2) by the media_wa_type field, while the information concerning message content is spread, for non-textual messages, over several fields (depending on the specific data type). As a matter of fact, while the content of textual messages (media wa type='0') is stored in the data field, for the other types of contents the situation is more involved, Multimedia files when the user sends a multimedia file, several activities take place automatically (i.e., without informing the involved users). First, WhatsApp Messenger copies the file. Then, it uploads the file to the WhatsAppserver, that sends back the URL of the corresponding location. Finally, the sender sends to the recipient a message containing this URL and, upon receiving this message, the recipient sends an acknowledgment back to the sender. When these steps have been completed, the sender stores into their *messages* table the record.

III. Conclusion

In this paper we have discussed the forensic analysis of the artifacts left by WhatsApp Messenger and Facebook on smartphones, and we have shown how these artifacts can provide many information of evidentiary value. In particular, we have shown how to interpret the data stored into the contacts and chat databases in order to reconstruct the list of contacts and the chronology of the messages that have been exchanged by users. More importantly, we have also shown the importance of correlating among them the artifacts generated by hats App Messenger in order to gather information that cannot be inferred by examining them in isolation. In future, we intend to investigate others apps as well, such as Twitter, Reddit and LinkedIn.

References

[1]. Burnette MW. Forensic examination of a RIM (BlackBerry) wireless device http://www.mandarino70.it/Documents/Blackberry%20Forensics.pdf.2002.

[2]. Parsons, A. ”Windows 10 Forensics: Conclusion” - Computer & Digital Forensics Blog, April 30. http://computerforensicsblog.champlain.edu/2015/04/30/windows- 10-forensics-conclusion 2015.

[3]. Punja SG., Mislan RP. Mobile device analysis. Small Scale Digital Device Forensics Journal June;2008 2(1).

[4]. Cellebrite LTD. Cellebrite Android Forensics. Available at http://www.cellebrite.com/mobileforensics/ capabilities/android forensics.2013

[5]. D. Cortjens., A. Spruyt., and W.F.C. Wieringa. WhatsApp Database Encryption Project Report. Available at https://www.os3.nl/ media/2011-2012/students/ssn project report.pdf. [6] Don Ho. Notepad++ Home, 2013. Available at http://notepad-plus-plus.org.

[6]. Cosimo Anglano. Forensic Analysis of WhatsApp Messenger on Android Smartphones. Digital Investigation Journal, Vol. 11, No. 3, pp. 201{213, September.2014}

[7]. Mohammad Iftekhar Husain., Ramalingam Sridhar. Forensic Analysis of Instant Messaging on Smart Phones. Springer Berlin Heidelberg, 2010.

[8]. S. Jeon., J. Bang., K. Byun., and S. Lee. A recovery method of deleted records for SQLite database. Personal and Ubiquotous Computing, 2012.

[9]. Asma Majeed., Haleemah Zia., Rabeea Imran and Shahzad Saleem. Forensic Analysis of three Social Media Apps in Windows 10, 12th International Conference on High-capacity Optical Networks and Enabling/Emerging Technologies (HONET) Kubasiak R., Morrissey S and Varsalone J. Macintosh OS X, iPod, and iPhone forensic analysis DVD toolkit. Burlington, MA: Syngress; 2009.

[10]. Zdziarski J. iPhone forensics: recovering evidence, personal data, and corporate assets. Sebastopol, CA: O‟Reilly; 2010.

[11]. NIST, S. 800-86. ”Guide to Integrating Forensic Techniques into Incident Response”, 2006, 800-86.

[12]. Lessard J., Kessler GC., Android forensics: simplifying cell phone examinations. Small Scale Digital Device Forensics Journal September 2010; 4(1).

[13]. Vidas T., Zhang C., Maloof M. Toward a general collection methodology for Android devices. In: Proceedings of the Eleventh Annual DFRWS Conference, vol. 8S; 2011. p. S14–23. New Orleans, USA, published in Digital Investigation.

[14]. Zellers F. MySpace.com forensic artifacts keyword searches. http://www.inlanddirect.com/CEIC-2008.pdf; 2008.

[15]. Al Mutawa N., Al Awadhi I., Baggili I and Marrington A. Forensic artifacts of Facebook‟s instant messaging service. International Conference for Internet Technology and Secured Transactions (ICITST); 2011. p. 771–6. Abu Dhabi, UAE.

[16]. Bader M., Baggili I+. iPhone 3GS forensics: logical analysis using apple itunes backup utility. Small Scale Digital Device Forensics Journal, September;2010 ,4(1).

[17]. N.S. Thakur. Forensic Analysis of WhatsApp on Android Smartphones. Master's thesis, University of New Orleans,. Paper 1706, 2013.

[18]. Morrissey S.. iOS forensic analysis for iPhone, iPad, and iPod touch. New York: Apress., 2010.