Master Thesis

In order to obtain the

Professional Master 2
In
MDFS option: Forensic Science

Presented and defended by:

Louaye Lamaa

On Monday July 22, 2019

Forensic Analysis of WhatsApp SQLite Databases on the Unrooted Android Phones

Supervisor:
Dr. Hasan Kazan

Reviewers:
Colonel Albert Khoury
Major Hany Kallassy

Lebanese University - Faculty of Sciences

Acknowledgement

This thesis would not have been achievable without the support and assistance from many people. I would like to thank all of them for being part of this journey and making this thesis possible.

I would like first to express my deepest gratitude to my supervisor Dr. Hasan Kazan for his guidance, patience, and support of my study.

I would like to thank the members of the committee for offering their time to read my thesis and for being able to attend my presentation. I honestly appreciate it.

Finally, I would like to recognize the biggest support that came from my family. I am deeply thankful for their love, sacrifices, and encouragement.

Abstract

WhatsApp is the most popular instant messaging mobile application all over the world. Originally designed for simple and fast communication, its privacy features, such as end-to-end encryption, eased private and unobserved communication for criminals aiming to commit illegal acts.

In this paper, we present the forensic analysis of the artifacts left by the encrypted WhatsApp SQLite databases on the unrooted Android devices.

In order to provide a complete interpretation of the artifacts, we perform a set of controlled experiments to generate them. Once generated, we identify their storage location and databases structure on the device. Since the data is stored in an encrypted SQLite database, we first discuss its decryption. Then, we show how to analyze the artifacts and how they can be correlated to cover all the possible evidence.

The results show how to reconstruct the list of contacts, the history of exchanged textual and non-textual messages as well as the details of their contents. Furthermore, this paper shows how to determine the properties of both the broadcast and the group communications in which the user has been involved as well as of the feature called status. Finally, we show how to reconstruct the logs of the voice and video calls.

These results show that the reconstruction of the WhatsApp data from the SQLite databases is possible and this data persists on the phone even after the uninstallation of WhatsApp.

Keywords: Mobile forensics, WhatsApp messenger, Instant messaging, Android, Unrooted devices, Data recovery, SQLite databases.

Contents

Acknowledgement **i**

Abstract **ii**

1 Introduction **1**

2 Related Works **4**

3 SQLite Forensics **5**

3.1 What is SQLite 5

3.2 SQLite file format 5

3.2.1 Pages 5

3.2.2 Database header 7

4 The Analysis Methodology **8**

4.1 Requirements 9

4.2 Workflow 9

4.3 Sets of experiments 9

4.3.1 Experiments concerning contacts 9

4.3.2 Experiments concerning the private chat communication between user and contacts 10

4.3.3 Experiments concerning the messages state 11

4.3.4 Experiments concerning the broadcast and group messages 11

4.3.5 Experiments concerning voice and video calls 12

5 Forensic Analysis of WhatsApp Messenger **13**

5.1 Analysis of WhatsApp functionalities 13

5.2 Location and types of WhatsApp artifacts 14

5.3 SQLite databases decryption 15

5.4 Analysis of the contacts database "wa.db" 19

5.4.1 The structure of the contacts database "wa.db" 20

5.4.2 Reconstruction of the contacts list 20

5.4.3 Blocked contacts 22

5.5 Analysis of the chat database "msgstore.db" 24

5.5.1 The structure of the chat database "msgstore" 24

5.5.2 Determination of the chat history 25

5.5.3 Analysis of the messages content 26

Multimedia files 26

Contact cards 29

Geolocation coordinates 29

Attachments 30

5.5.4 The determination of the messages state 31

5.6 Multiple message destinations 32

5.6.1 Broadcast messages 32
5.6.2 Group chat 33
5.7 Voice and call logs 35
5.8 WhatsApp status analysis 37
5.9 Deleted data 38

6 WhatsApp Security 40

7 WhatsApp in the Court 41
7.1 Admissibility 41
7.2 Authenticity 41
7.3 Relevance 42
7.4 Digital forensics report 43

8 Conclusion 46

List of Figures

1.1 Instant Messaging applications surpass the Social Network apps. 1
1.2 WhatsApp monthly active users compared to others' IM apps. 2

3.1 SQLite page layout. 6
3.2 SQLite is set of data pages of fixed size. 6
3.3 Set of free list pages that contain deleted data. 7
3.4 The SQLite database header. 7

4.1 Workflow of the analysis methodology. 10

5.1 WhatsApp messenger artifacts. 14
5.2 The encrypted databases backups of msgstore.db stored in the internal memory. 15
5.3 The encryption process. 16
5.4 The PowerShell script used to decrypt the databases. 16
5.5 WhatsApp databases and key extraction. 17
5.6 DBs and the key are extracted from the phone. 17
5.7 WhatsApp viewer tool. 17
5.8 The database is successfully decrypted. 18
5.9 The decrypted database "msgstore" is opened in DB Browser for SQLite. . . 18
5.10 Damaged database: Disk image is malformed. 19
5.11 Recovered table "messages" from the damaged database "msgstore". 19
5.12 "wa_contacts" table - the individual contacts records. 21
5.13 "wa_contacts" table - the group contacts records. 21
5.14 "wa_block_list" table - the blocked contacts records. 22
5.15 Extracting the blocked contact information using a SQL query. 22
5.16 Reconstruction of the chat history. 25
5.17 The three types of a multimedia message. 26
5.18 The image message content: the sender and the recipient records. 27
5.19 LEFT JOIN Venn diagram. 27
5.20 LEFT JOIN operation between messages and message_thumbnails tables. . . 28
5.21 "message_thumbnails" table - the thumbnail of an image message extracted using LEFT JOIN query. 28
5.22 The contact card message. 29
5.23 The geolocation message. 29
5.24 "message_thumbnails" table - the thumbnail of a location message. 30
5.25 The attachments messages records. 30
5.26 "message_thumbnails" - the thumbnail of the document message extracted using LEFT JOIN. 31
5.27 The possible states of a message. 32
5.28 The broadcast messages records. 33
5.29 The broadcast message - recipient side. 33
5.30 The group chat records. 34
5.31 The group records created when a member leaves the group. 35

5.32 Timeline of the chronology of the group composition. 35
5.33 "call_log" table - voice and video calls records. 36
5.34 "jid" table - the phone number of the caller 36
5.35 The WhatsApp status records. 37
5.36 "message_thumbnails" table - identification of the status picture using SQL query. 38
5.37 Deleted messages: the *NULL* value. 38
5.38 The structure of the B-tree pages. 38
5.39 The output of the script used to extract deleted data. 39