# Forensic Analysis towards the user behavior of Sina microblog

Long Chen[1,a], Yong-Qing Wang[2,b]

[1,2]Department of computer, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China.

[a]chenlong@cqupt.edu.cn, [b]wangyongqing123@163.com

**Keywords:** Microblog, user behaviour, iOS data acquisition.

**Abstract.** Microblog, a new network application in the era of Web 2.0, has become one of the major medium in China. Its main features are as followings: large number of users, frequent status of updating information, fast transmission speed of information. The writer took Sina Weibo iPhone App as an example to study the behavior of individual characteristics of microblog users by analyzing the data from samples generated by using microblog.

## Introduction

Due to the popularity and prevalence of smartphones, the number of third-party mobile applications increases rapidly. The number of mobile applications in Apple's official App Store has reached 1.49 million by January 2015 [1]. Many applications are making the feature-rich smartphones. There are many potential evidence for forensics workers. Foreign research in the field of third-party applications focuses on Facebook, Twitter and MySpace. The main study focuses on analyzing user's social networking activity and whether the data stored in the main memory and the mobile phone can be restored [2].Domestic research in this field focuses on Wechat and Sina Microblog. For Wechat, the main study focuses on analyzing the file directory structure [3] and getting the audio file[4]. For Sina Microblog, there are two methods of extracting the data of Sina Microblog: acquiring information based on Sina Micro-Blog Open Platform and acquiring information based on network data flow[5]. But with the development of mobile Internet, many Sina Microblog users begin to use mobile client other than PC client, and there is no relevant research on data extraction of Sina Microblog App.

The writer took Sina Microblog iPhone App as an example to extract some important data of Sina Microblog iPhone App, then analyzed the directory structure of microblog backup file and relevant important data. The method mentioned in this paper can help forensic investigator acquire some important data of Sina Microblog quickly and analyze the user behavior easily.

## Microblog User Behavior

Microblog, as a kind of new information communication platform, can satisfy our different requirements, such as information acquisition, information communication and information sharing etc[6].

On the Internet, there are three main typical behaviors of Microblog users: follow others, be followed by others, to tweet. The first one is a kind of behavior that the user acquires some information by following other users. The second one is a kind of behavior that the user affects other users through being followed by others. The third one is a kind of behavior that the user writes twitter and spreads information. The greater the number of Microblog being created and reposted, the larger the information being transferred by the user[7].

## Data Acquisition

There are three ways to acquire data from iOS devices: acquire data from backup file, acquire data by logical method and acquire data by physical method. This paper focuses on how to acquire data from iOS devices by backup file.

 

iPhone backups data by using iTunes according to some synchronous protocols about MAC OS, so we can acquire data from the backup directories stored in the computer. However, only the file data synchronized exactly by synchronous protocol can be acquire by this method. Different operating system has different storage location when iPhone backups data by using iTunes, the detail information is shown in table 1.

Table 1. backup file's storage location of using iTunes

| Operation system | location |
|---|---|
| Windows XP | C:\documents and setting\<user name>\ Application Data\Apple Computer\ MobleSync\Backup |
| Windows Vista/Windows 7 | C:\Users\<user name>\AppData\Roaming\Apple Computer \ MobleSync\Backup |
| Mac OS X | Users/<user name>/Library/Application Support/MobileSync/ Backup/ |

A large number of key information can be recovered by using the method mentioned above. Frequently-used data is usually stored in the SQLite database and some property list file, as synchronous protocol can support synchronous operation of the SQL database and some property list file.

## Forensic analysis of iPhone third-party application

The forensic analysis of the data generated by iPhone third-party application consists of three parts: analyzing file and directory structure, analyzing database/plist file, correlation analysis.

IOS device contains a large number of various types of data, including some data related with mobile phone and built-in applications, such as call log, contacts, short messages, photos and the cache files of Safari browser etc. In addition to this, iOS also contains the data generated by the third party applications which are from App Store. IOS device has two kinds of storage formats: one is property list file（plist） in binary form, it's used to store some setup information; another is SQLite database, it's used to store personal information[8].

**Analyzing file and directory structure.** Every iOS application has its own sandbox, the sandbox is a special file system directory which is separated from other file directories. It can prevent any application to exchange data with other applications.

The third-party apps of iPhone are usually stored in /private/var/mobile/Applications. Every third-party app has two directories: /Documents and /Library, the first directory contains some document information, the second directory contains preference settings and some cache files[9]. But different third party application has different storage location and format.

**Analyzing database/plist file.** SQLite database is one of the most common data type for storage, it's mainly found in the mobile application development. Many applications in the iOS use SQLite to store data. Many important data (such as Contacts, Short Messages, Call History etc) are stored in the form of SQLite database, these data are encoded in UTF-8.

Property List file is mainly used to store serialized objects. The filename extension is .plist, so it's usually called plist file. Plist file is usually to store user settings and extra information. Plist file is consist of three classes with hierarchical structure: Cocoa Foundation、Core Foundation and XML, all nodes are displayed in a list.

**Correlation analysis.** Although these files include many important information, such as the unique ID of visiting social network site, special data, where and when the event is taking place.

## Analyzing Sina Microblog

This paper will take Sina Microblog iPhone App as an example to discuss how to analyze Microblog users' behavior for forensic investigator. This work includes two steps: extract important backup file data related with Sina Microblog users' behavior, and analyze Sina Microblog directory structure, important database and plist file.

As the back files are all encrypted files, we can use some forensic tools to restore these encrypted files, two tools used in this paper are iBackupBot for iTunes.

Fig 1 shows the directory structure diagram of using iPhone Data Recovery to restore Sina Microblog, Sina Microblog has two directories: /Documents and /Library, the first directory is used to store document information , the second one is used to store preference settings and cache information.

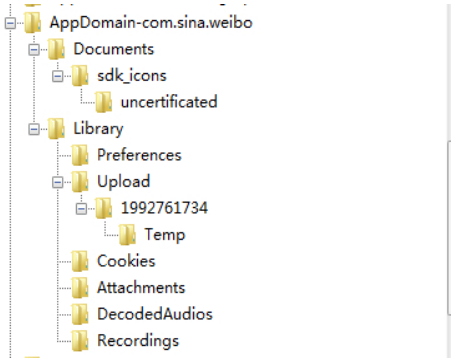


Fig 1. Directory Structure of Sina Microblog

Important information of Sina Microblog iPhone App is stored in a SQLite database called Documents/db_42500_1992761734.dat, the last ten digits(1992761734) is the unique id of the user. Then we can know that the filename of this database file in the backup files is 4ab36716f9ce19991ac7950591b2c06475e5d21e by computing the hash value(sha1) of ppDomain-com.sina.microblog-Documents/db_42500_1992761734.dat. Then we can find several tables in this database file, the detail information is shown in Fig 2.

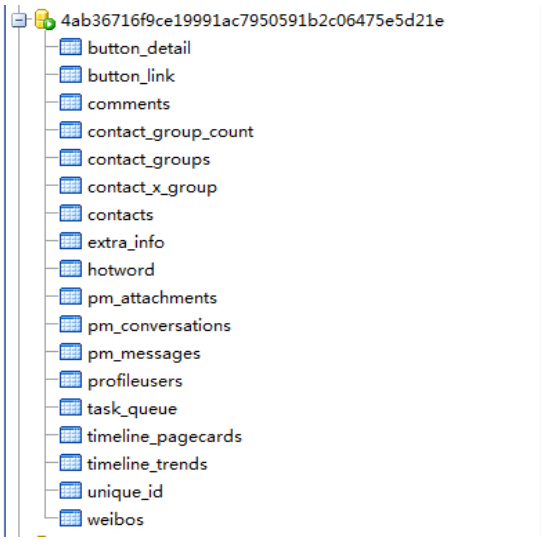


Fig 2. SQLite file

It's easy to analyze the relationship between the data content and corresponding Microblog information by viewing the structure of each table. Each table in the database (db_42500_1992761734.dat) has different functions, the detail information is shown as followings:

- contact_group_count: This table is used to record the amount of users in each group of folloing other users.
- contact_groups: This table is used to record some information about being followed by other users, including the GID and name of each group.
- contact_x_group: This table is used to record the users list of being followed by others, including userID and the groupID of each group.
- contacts: This table is used to record the users list of following others, including username and userID etc.

- pm_conversations: This table is used to record the list of user's Microblog private messages, including the newest record of private message with each user.
- pm_messages: This table is used to record the Microblog message list.
- microblogs: This table is used to record Microblog information by user's timeline, including the content of Microblog message, the author of Microblog message, userID, post time, the amount of forwarding, the name of Microblog client, geographical position information, the link of picture attached to Microblog message and so on.

Two kinds of typical user behaviors, "Follow" and "Be followed", form the basis of Microblog user's social network[10]. We can acquire the users list of following others from the contacts table. The contacts table can directly reflect the user's interests towards different kind of information. If we want to know the user's social network information, we should view the user's fans list from contact_x_group table.

User-posted Microblog messages are recorded in weibo table by timeline. The last 50 microblog records are stored in the mobile client, these records include the microblog messages posted or browsered by the user. The information recorded in the microblog table is very important for studying forensic investigation of Microblog user's behavior. he detail information of each field and the corresponding meaning is shown in Table 2

Table 2the information of weibo table

| Field | Stored information | Stored data type |
|---|---|---|
| nick | User nickname | NSString |
| uid | The unique ID of user | NSNumeber(intValue) |
| portrait | Image Information | NSString |
| concent | The body of posted Microblog | NSString |
| pic | Embedded picture in the Microblog | NSString |
| dateline | The date of posting Microblog | NSDate |
| rtrootuid | The unique ID of the posted Microblog | NSNumeber(intValue) |
| rtrootnick | The nickname of the posted Microblog | NSString |
| rtreason | The comment content of forwarded Microblog | NSString |
| source | The app of posting Microblog | NSString |
| longitude | Longitude | NSNumber(floatValue) |
| latitude | Latitude | NSNumber(floatValue) |
| url_structs | The link information embedded in the Microblog | NSDictionary |
| page_info | Page information(position,topic etc) | NSDictionary |
| topic_structs | Topic information(the link and title of the topic) | NSDictionary |
| pic_id_infos | The picture embedded in the posted Microblog | NSDictionary |
| extra_properties | Extra information(If the value of relation is 0,it indicates that this message is posted by the user;if the value of relation is 1,it indicates that this message is the public homepage's microblog message which is browsered by the user. ) | NSDictionary |

The forensic investigator can obtain many useful information by analyzing the important fields in weibo table, such as user-posted microblog messages, the public homepage's microblog messages which are browsered by the user, where the user posted the microblog message. In addition to this, the user's track during a period of time can be obtained by analyzing the information of longitude and latitude, then the forensic investigator can analyze Microblog users' behavior from the point of time and space relations.

Except for this, we can directly visit the user's Microblog homepage by entering the URL: http://microblog.com/0000000000 in the browser's address bar to validate whether the ID number in the URL belongs to the user. We can acquire all kinds of data of Microblog by calling API interface provided by Sina, including personal profile information, geographical position information, dynamic interaction information, user's fans information.

**Summary**

At present, the research of Microblog user's behavior and acquiring Microblog data are conducted separately, but they are inseparable for forensic workers. On this basis, this paper took Sina Microblog iPhone App as an example and proposed a new method: firstly extract data from Microblog app, then make analyze user behavior for the purpose of forensic analysis, this method can be applied to other Microblog app, too.

**Acknowledgements**

This work is supported by National Social Science Foundation Project of P. R. China(No. 14BFX156), Natural Science Foundation Project of CQ CSTC of P. R. China (No. cstc2011jjA40031).

**References**

[1] Information on http://www.pocketgamer.biz/metrics/app-store/

[2] Mutawa N A, Baggili I, Marrington A. Forensic analysis of social networking applications on mobile devices[J]. Digital Investigation, 2012, 9(15):S24–S33.

[3] Gao F, Zhang Y. Analysis of WeChat on iPhone[C]//2nd International Symposium on Computer, Communication, Control and Automation. Atlantis Press, 2013.

[4] Du Jiang, Wang Cong. iPhone third-party software forensics research[J]. Computer CD Software and Applications. 2013, (13):53-54.

[5] HUANG Yan-wei, LIU Jia-yong. Study on Sina microblog Data Acquisition Technology[J]. Information Security and Communications Privacy. 2013（06）：71-73.

[6] Zhao Ling, Zhang Jing. Multi-dimensional Analysis of Microblog User Behavior Research[J]. Information and Documentation Services.2013(05).

[7] Chen Peng, Shui Jinguang. Statistical Analysis of Microblog User Typical Behavior based on Individual Property[J].Knowledge Management Forum.2013(05).

[8] Chen C N, Tso R, Yang C H. Design and Implementation of Digital Forensic Software for iPhone[C]//Information Security (Asia JCIS), 2013 Eighth Asia Joint Conference on. IEEE, 2013: 90-95.

[9] Levinson A, Stackpole B, Johnson D. Third party application forensics on apple mobile devices[C]//System Sciences (HICSS), 2011 44th Hawaii International Conference on. IEEE, 2011: 1-9.

[10] XU Xiao-dong, XIAO Yin-tao, ZHU Shi-rui.Simulation Investigation of Rumor Propagation in Microblogging Community[J].Computer Engineering. 2011, 37(10): 272-274.