Forensic analysis of TikTok application to seek digital artifacts on Android smartphone

Nghi Hoang Khoa[1,2]
[1]*Information Security Laboratory, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
khoanh@uit.edu.vn

Phan The Duy[1,2]
[1]*Information Security Laboratory, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
duypt@uit.edu.vn

Hien Do Hoang[1,2]
[1]*Information Security Laboratory, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
hiendh@uit.edu.vn

Do Thi Thu Hien[1,2]
[1]*Information Security Laboratory, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
hiendtt@uit.edu.vn

Van-Hau Pham[1,2]
[1]*Information Security Laboratory, University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam National University, Ho Chi Minh City, Vietnam*
haupv@uit.edu.vn

***Abstract*— Emerging with highlight features as a global phenomenon, TikTok - the international version of Douyin application in China market, is a social media video app for creating and sharing short lip-syncing. This social video platform has seen astounding growth by reaching 1.5 billion users in 2019 by capturing the cultural zeitgeist among teenage smartphone owners in unprecedented fashion. However, criminals have been paid attention to this platform where young users become the target of abusers or predators. In this paper, we introduce the forensic analysis of the artifacts left on Android phones by TikTok application. In more detail, we indicate a complete description of all the artifacts obtained in TikTok. Furthermore, by using the results discussed in the paper, an investigator can rebuild the list of followers, searching keywords, favorites, messages that have been exchanged by users. It is helpful to examine and determine which artifacts could be considered as evidence for criminal investigation on popular social networking application.**

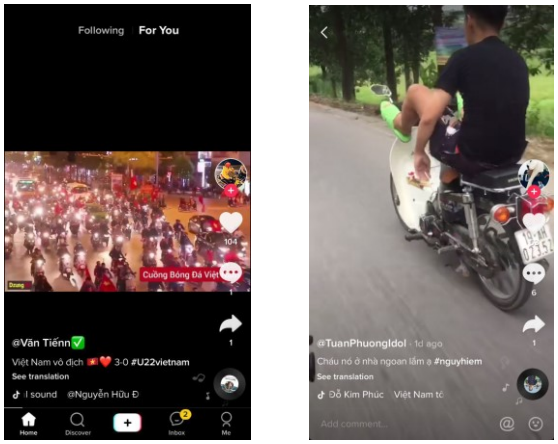
Keywords—TikTok forensics, mobile forensics, Android application, social networking

I. INTRODUCTION

Nowadays, social networking becomes a place for exchanging information between people around the world. Users can connect to social networks very easily on a smartphone which allows them to share their personal activities on a daily basis with their friends and acquaintances. This valuable information such as name, age, gender, location, audio, video files and messages plays a crucial role during digital forensics investigation.

Since its launch, TikTok application [1] is being promoted as a video-sharing social network with tremendous growth in popularity. It offers users the features of creating short-form videos ranging from not only lip-sync, challenges, dance videos to magic tricks and funny videos. TikTok users are able to gain a lot of followers and likes in a short period of time by taking and editing videos with special effects and filters right on their smartphones. Interaction related features in TikTok allow users to expose, share or exchange their reactions and thoughts in this social network (Fig. 1a). However, there is the fact that bad users can take advantage of this platform to broadcast and entice the youth to imitate dangerous actions via challenge videos (Fig. 1b) or chat messages.

Being an opening communication platform, users signed up to TikTok should note that their account information will be public. Moreover, TikTok provides the capability that your posted videos can reach everyone and anyone also can contact with you via messages. These features can lead to serious problems when the users are children or adolescents who have not fully understood the threat of social networking. TikTok enables users to create short singing and dancing videos, so do the children, which can expose a lot of critical information attracting many criminals. Besides that, messaging can become a useful mean for these criminals to approach and sexually abuse their targeted victims [2]. Hence, detecting suspicious behaviors on the TikTok application is in need, from both victim and criminal sides. The aim of this paper is to seek and obtain potential data and information left on Android devices by TikTok application that can be used later for digital forensics, to investigate and confirm whether any crime is taking place by taking advantage of this application.



(a) (b)

Fig. 1. GUI of TikTok application.
(a) Main features of TikTok. (b) Posts with bad influence.

This paper resolves the common questions that arise during examining the TikTok application to gather all the artifacts and data generated by TikTok application in Android smartphones. It includes: 1) how to acquire the user data of TikTok and how to reveal the content of them; 2) the friend and follower/following list of the user; 3) who the user communicated with and what they said with a timestamp. With this attempt, the researchers and investigators could find a better comprehensive understanding of artifacts in TikTok

application. It can be helpful in recovering user behaviors related to investigation cases.

The remainder of this paper is organized as follows. In Section II, we provide an overview of existing work whereas the methodology and tools for investigating are indicated in Section III. Finally, in Section IV we conclude the paper and outline the future work.

II. Related works

Digital forensic analysis on social networking and instant messaging, especially platforms having apps on Android devices, has gained more and more attention from many researchers and become a topic interested in many publications.

In [3], Awan et al. contributed a case study in forensics for famous social networking applications such as Facebook, Twitter and LinkedIn on devices running Android, iOS, Windows, BlackBerry operating systems. Upon extracting the image on the device where the application is installed, they analyze database files that the applications access and use.

In [4], focusing on WhatsApp social networking, the author has pointed out where contacts and messages are stored on the Android device to recreate list contact and chat content of this application. This article builds an Android Emulator called YouWave running Android version 4.4.4 as their analyzing environment. Then FTK imager tool is used to extract the disk image to obtain WhatApp's data for later analyzing step using SqliteMan.

A work by Songyang Wu et. al [5] is one of many articles focusing on forensic analysis for Wechat application. This publication not only is able to collect textual and non-textual content on devices having this application installed, but also has the capability of dealing with content stored in the database encrypted with SQLCipher, which makes the forensics with Wechat become more difficult than other applications. The author uses some analysis tools such as Apktool, dex2jar, JD-GUI tools to reverse and reconstruct the source code to learn how to encrypt data as well as the storing location of the encrypting key.

Meanwhile, the author of [6] has figured out that just like WeChat, most Telegram's data is stored and encrypted in a complicated form of binary string structures called Telegram Data Structures (TDS). To be able to decrypt the data, it is necessary to understand the behavior of the application in data encrypting and storing process through source code analysis, from where we can conduct the corresponding decrypting steps to obtain plaintext data.

While the above popular social networking applications have been discussed in many works, there is a lack of investigation on TikTok in the same manner. As mentioned, just the same as its counterparts, TikTok can make its users, especially the young ones, face multiple problems and threats due to the sharing and public environment. The findings in this paper can be helpful for crime investigators or parents who want to determine what happened with the victims in such a dangerous case relating to TikTok application.

III. TikTok forensics methodology and tools

In this study, we conducted several test cases to create real data on the application. The data were extracted after each test for analysis. There are four steps in digital forensics [7]: (1) identification (2) acquisition (3) examination (4) reporting. The article focuses on the first two steps which find information stored in the data directory of the TikTok application in the internal storage. However, to be able to access the data folder, root privilege is required.

A. Environment setup

To analyze TikTok application, in this paper, we set up an experiment environment having TikTok version 8.9.4 installed on a rooted Nox Player [8] with Android version 5.1.1. Many analyzing tools are used to find and extract data from the targeted devices, such as:

- ADB debugging [9] can connect to devices and provide a shell to interact with them by command line.
- SQLite Database Browser [10] supports reading database files.
- A code editor can read XML files.

B. TikTok forensics methodology

1) Data identification

The *data/data/com.ss.android.ugc.trill* is a crucial directory that contains the data file generated by the TikTok application. Once the application is installed, all interesting data for exploiting information on TikTok is located at *data/data/com.ss.android.ugc.trill/databases* and *data/data/com.ss.android.ugc.trill/shared_prefs*

- *data/data/com.ss.android.ugc.trill/databases* contains user data information such as messages, login sessions, etc.
- *data/data/com.ss.android.ugc.trill/shared_prefs* contains information related to the host device.

2) Data acquisition

In order to extract data from the device, we use ADB tool. ADB is a flexible command-line tool that supports communication with Android devices; It also plays an important role in digital investigation and information collection on Android devices. To pull data from Android devices via ADB debugging with the following command:

```
adb pull -p <path on Android> <path to save data>
```

However, the data collected from a device through ADB tool can have many differences based on its rooted or non-rooted status of the device. To access to related multilateral files, especially the TikTok application, most of the useful data for investigation is stored at *databases* and *shared_prefs* folders, which can only be read from the rooted phone.

a) General information

The *shared_prefs* directory contains a lot of potential information in form of XML files, as follows:

- The first opening time of the application
- The installation time of the application
- The last time that the application is updated
- The MAC address of the last connected SSID
- Recent searching history
- Language
- Region

Searching history is stored in *search.xml* file, as shown in Fig. 3.

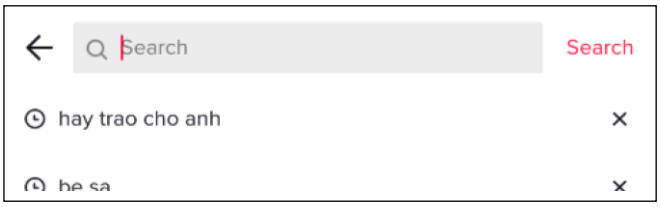


Fig. 2. A search action in TikTok application.

```
<map>
    <string name="place_holder">Search</string>
    <string
name="recent_history">[{&amp;quot;keyword&amp;quot;:&amp;quot;hay
trao                                                    cho
anh&amp;quot;,&amp;quot;int&amp;quot;:0},{&amp;quot;keyword&amp;quot;:&amp;quot;b
e sa&amp;quot;,&amp;quot;int&amp;quot;:0}]</string>
</map>
```

Fig. 3. Searching history in search.xml file.

The information of the account owner is located in *aweme_user.xml*. For example, Fig. 4 illustrates the information of an account can be obtained from this file such as name (*Nghi Hoàng Khoa*), nickname *(phapha_jian),* the URL of avatar (*http://{uri}/aweme/100x100/tiktok-obj/{uid}.webp),* the gender in form of a number (*1* for male), the timestamp of the latest logged in activity *(6540609301387952129),...* and other related information.

Fig. 4. Content of aweme_user.xml

Meanwhile, another XML file called *applog_stats.xml* provides some information such as device ID, the MAC address of last connected SSID, the timestamp of the last SSID connection…

Fig. 5. Device information in applog_stats.xml.

In fact, navigating to the *shared_prefs* directory, there are around 103 XML files which contain the clues of the user's behavior in TikTok application. Nevertheless, it is a time-consuming task to analyze all of them and retrieve suspicious evidence belonging to that user. In TABLE I. , we list XML files where the most common information can be found.

TABLE I. MOST COMMON INFORMATION IN XML FILES IN SHARE_PREFS

File name	Description
app_setting.xml	Application version
applog_stats.xml	• Fingerprint sending time • MAC address of device • MAC address of last accessed Wi-Fi SSID • Device id • Google Ads Key • Last check SSID time
aweme_friends.xml	Facebook access token
aweme_user.xml	• List of latest logged in UID • User account information (name, country, nickname, avatar URL, email, gender)
custom_channels.xml	Application install information (app channel, APK created time)
search.xml	Recent searching history

b) Information from database files

TikTok uses SQLite as its back-end database to store a lot of information and data for its operations. Its database can be found in form of files with *.db* extension at *data/data/com.ss.android.ugc.trill/database*. There are more than 20 files located in this directory, in which each file is a complete database containing a specific information. We use DB SQLite Browser to read these database files.

- **Messaging related database**

In the *database* directory, name of the messaging related database file consists of a 19-digit string and the postfix of *_im*, which is *{0-9}_im.db*. The most attractive information in this database is the ***msg*** table, which properties is listed in 0, records all sent and received messages by users.

TABLE II. MSG TABLE

Property names	Description
msg_uuid	Message identifier
msg_server_uuid	Message owner identifier
conversation_id	ID of conversation
created_time	Message creating time
sender	ID of sender
content	Message content
…	…

We can get some information of the message such as sending and receiving time, the sender and the receiver as well as the content of the message. Note that the message content is stored in the database in a JSON format, as shown in Fig. 8 with some additional information along with the main message content. Fig. 6 is an example of conversation in the TikTok application and what we obtain from the database file is shown in Fig. 7.

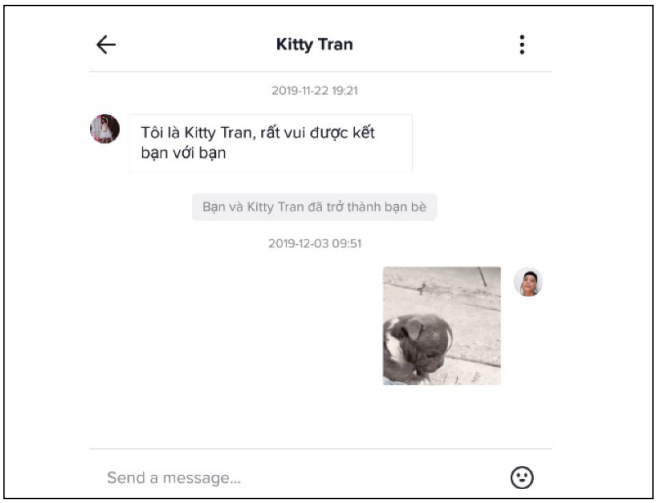


Fig. 6. An example inbox conversation.

	msg_uuid	conversation_id	created_time	sender	content
	Filter	Filter	Filter	Filter	Filter
1	a50451ca..47f...	0:1:647404015...	1574425294592	647404015268...	{'stickers':[{'id':65...
2	75a21fb6..8f4...	0:1:647404015...	1574425294561	647404015268...	{'text':'Tôi là Kitty T...
3	81da2985..f57...	0:1:647404015...	1575341516402	654060930138...	{'display_name':'cu...

Fig. 7. Information of conversation in database file.

```
{
    "text":"Tôi là Kitty Tran, rất vui được kết bạn với bạn",
    "aweType":701,
    "ext":
    {
        "tips":"Bạn và Kitty Tran đã trở thành bạn bè"
    }
}
```

Fig. 8. Structure of a message content in TikTok.

As we can see, not only message can be extracted, but database files also give us more information such as the sign of friend adding action of user in *tips* label.

With these message information and its timestamp, the whole conversation of users can be reconstructed via an analyzing process.

- **Contacts database**

Contact list of a user are stored in another database file called *db_im_xx*, with *.db* file format. The structure of contacts is described in the Fig. 9 and TABLE III.

	UID	SEC_UID	NICK_NAME	SIGNATURE	AVATAR_THUMB	UNIQUE_ID
	Filter	Filter	Filter	Filter	Filter	Filter
1	6748953635 11...	MS4wLjABAAAA...	李子柒	李家有女，人称...	{'height':0,'data...	cnliziqi
2	6727990991 02...	MS4wLjABAAAA...	jr__loera	喜歡就留下	{'height':0,'data...	user3z3g2ujjq7
3	6615106799 68...	MS4wLjABAAAA...	Hạ Vũ	❤Một bước đi sa...	{'height':0,'data...	havu2310
4	6610939366 88...	MS4wLjABAAAA...	Jay Nguyen		{'height':0,'data...	30660509177
5	6575096026 39...	MS4wLjABAAAA...	Đức Phong	ĐP..MN..29..10.....	{'height':0,'data...	cp_vn....29102.
6	6628490464 34...	MS4wLjABAAAA...	Gia Bao Nguy...	Bán guitar chất l...	{'height':0,'data...	giabaonguyennn
7	6741987625 18...	MS4wLjABAAAA...	photographe...	攝影師一名，記...	{'height':0,'data...	photographergu
8	70904421685	MS4wLjABAAAA...	摩登兄弟	刘宇宁	{'height':0,'data...	md45288
9	6648003223 79...	MS4wLjABAAAA...	30Shine	Đời thay đổi khi ...	{'height':0,'data...	30shinetv
10	6744269435 69...	MS4wLjABAAAA...	💋 Minna Na ...	Cảm ơn sự yêu t...	{'height':0,'data...	srvg917

Fig. 9. SIMPLE_USER table in contact database.

TABLE III. SIMPLE_USER TABLE

Property names	Description
uid	User identifier
nick_name	Username
avatar_thumb	Avatar
unique_id	User-defined format
…	…

Some information in this database can be used in later steps, such as to view the profile wall of a user, we can use the URL *https://www.tiktok.com/share/user/{uid}* with the UID obtained from above file. Besides that, with this link, videos in public mode of this user can also be viewed.

- **Video related databases**

As TikTok is a social networking with the aim of video-sharing, video is the most important object in this application, as well as its related databases. In *database* directory, there are multiple video related databases file. For example, the *video_header_http_t* table in *video.db* file contains records of viewed videos of user, including the URL of that video if it is still accessible (Fig. 10).

	_id	key	mime	contentLength	flag	extra
		Filter	Filter	Filter		Filter
1	1	A92814A4F888...	video/mp4	3255700	0	{'requestUrl':'http:\/\/v9.tiktokcdn.com\/ca
2	2	F180814E6F23...	video/mp4	3814189	0	{'requestUrl':'http:\/\/v9.tiktokcdn.com\/2
3	3	273358D46D7F...	video/mp4	1617565	0	{'requestUrl':'http:\/\/v9.tiktokcdn.com\/5
4	4	5DB006933FB4...	video/mp4	1285935	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
5	5	9676006ADDD...	video/mp4	913243	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
6	6	10497D24EC06...	video/mp4	1438301	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
7	7	410143235120...	video/mp4	1190241	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
8	8	02779B5FE802...	video/mp4	1515323	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
9	9	3C31B40EC0C2...	video/mp4	1908948	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/
10	10	D6BA2A973B81...	video/mp4	377885	0	{'requestUrl':'http:\/\/v16.tiktokcdn.com\/

Fig. 10. Viewed video information in TikTok application.

In another case, when users download any videos from other ones, the information of this action is also saved in a database file called *downloader.db* such as the video URL, the location of the video file on the device, hash code, size, etc. Note that, the URL of the video can be timeout due to the policy of TikTok application, so that the table also saves a permanent backup link (Fig. 11).

	_id	url	savePath	name	totalBytes	md5	backUpUrlsStr
	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	-2084787980	http://v16.tikto...	/storage/emulated/0...	676353184836...	3316886	4c74165a685b...	['http:\/\/v9.tik...
2	-1621157801	http://v16.tikto...	/storage/emulated/0...	676353184836...	3316886	4c74165a685b...	['http:\/\/v9.tik...
3	771323903	http://v16.tikto...	/storage/emulated/0...	676353184836...	3316886	4c74165a685b...	['http:\/\/v9.tik...

Fig. 11. Downloaded video information in TikTok application.

c) Media files

Along with textual data and information, TikTok also supports saving media files in the directory path of *data/data/com.ss.android.ugc.trill/files*. For example, edited and uploaded videos can be found here.

In TikTok, each created video will be saved into 3 different files with the name specification as *{datetime format YYYY-MM-DD}-{timestamp}-concat-{v or a}*, in which the one ends with *concat-v* is for video, *concat-a* for audio and a combining file with both video and audio will have the *mix-contact-a* postfix.

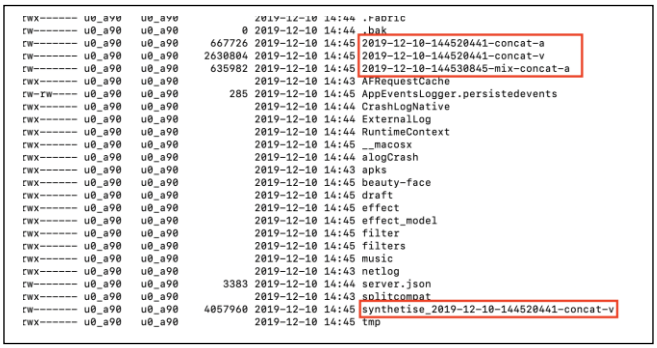


Fig. 12. Saved video files in TikTok application.

When a saved video is uploaded, TikTok then creates an additional file with the *synthetise_* prefix in the file name as *synthetise_{datetime format YYYY-MM-DD}-{timestamp}-concat-v.* All these 4 files can be extracted from the device and viewed as normal in video or audio player.

IV. Conclusion

The rapid development in the popularity of TikTok application makes it become a potential and useful data source for digital forensics. In this paper, we proposed an analysis process to find the location and obtain this valuable information on Android devices installing TikTok. The experiment results have proved the capability of using collected data to reconstruct the list of followers, searching keywords, favorites, messages that have been exchanged by users… in this application. It is important to note that the deploy environment of TikTok is not limited to Android but various other platforms (e.g. iOS), further work needs to be done due to the possible difference in storing format or location of platforms.

Acknowledgment

The authors would like to thank all colleagues, E8.1 members for their inspiration and help during this work. Thank all teachers who have helped us to develop the fundamental and essential academic competence.

This research is funded by University of Information Technology-Vietnam National University HoChiMinh City under grant number D1-2020-01.

References

[1] "TikTok," [Online]. Available: https://tiktok.com/en/. [Accessed December 2019].

[2] P. Knox, "APP TRAP What is TikTok, is the app dangerous for children and what should parents know?," 2019. [Online]. Available: https://www.thesun.co.uk/news/8805155/what-tiktok-app-dangerous-child-what-parents-know/. [Accessed December 2019].

[3] F. A. Awan, "Forensic examination of social networking applications on smartphones," in 2015 Conference on Information Assurance and Cyber Security (CIACS), 2015.

[4] C. Anglano, "Forensic analysis of WhatsApp Messenger on Android smartphones," Journal Digital Investigation: The International Journal of Digital Forensics & Incident Response, vol. 11, no. 3, pp. 201-213, 2014.

[5] S. Wu, Y. Zhang, X. Wang, X. Xiong and L. Du, "Forensic analysis of WeChat on Android smartphones," in The 16th Annual USA Digital Forensics Research Conference, USA, 2017.

[6] C. Anglano, M. Canonico and M. Guazzone, "Forensic analysis of Telegram Messenger on Android smartphones," Digital Investigation, vol. 23, pp. 31-49, 2017.

[7] P. T. Duy, H. D. Hoang, D. T. T. Hien, N. B. Khanh and V.-H. Pham, "SDNLog-Foren: Ensuring the Integrity and Tamper Resistance of Log Files for SDN Forensics Using Blockchain," in 2019 6th NAFOSTED Conference on Information and Computer Science (NICS 2019), Hanoi, 2019.

[8] "Free Android Emulator on PC and Mac - Download NoxPlayer," [Online]. Available: https://www.bignox.com/. [Accessed December 2019].

[9] "Android Debug Bridge (adb)," [Online]. Available: https://developer.android.com/studio/command-line/adb. [Accessed December 2019].

[10] "DB Browser for SQLite," [Online]. Available: https://sqlitebrowser.org/. [Accessed December 2019].