Retrieving Cloud Conversation Records of Xiaomi Smart Speaker

Lu Zhao[1,2,*], Qian Zhang[2], Bo Wang[2], Yanrong Kang[2], Guo Lili[2], Xing Guidong[2], Xianbo Jiang[2], Yanhui Du[1,*]

1. *School of Information Network Security, People's Public Security University of China*
2. *Institute of Forensic Science, Ministry of Public Security, Beijing, China*

Abstract: **Smart home devices, such as Xiaomi smart speakers, generate valuable cloud-based conversation records that can serve as critical evidence in forensic investigations. However, accessing these records directly from the manufacturer's servers is often hindered by encryption and restricted access. To address this challenge, we developed a method for retrieving cloud conversation records from Xiaomi smart speakers by simulating the communication protocol. This approach uses app backup data to reconstruct a protocol model that mimics the interaction between the app and the server, enabling the extraction of cloud-stored conversations without requiring the app's login credentials or internet connectivity of the source phone. Forensic examiners can leverage this technique to obtain relevant evidence while ensuring the integrity of data on the source device.**

Keywords: **smart home devices, smartphone management app, simulation protocol model, cloud data**

1. Introduction:

Current methods for acquiring smart home cloud data require user credentials or SIM card access, limiting their use when such information is unavailable. We propose a new approach using app backup data to simulate app-server communications, enabling cloud data extraction without credentials or SIM cards. The method's workflow is shown in Fig.1.

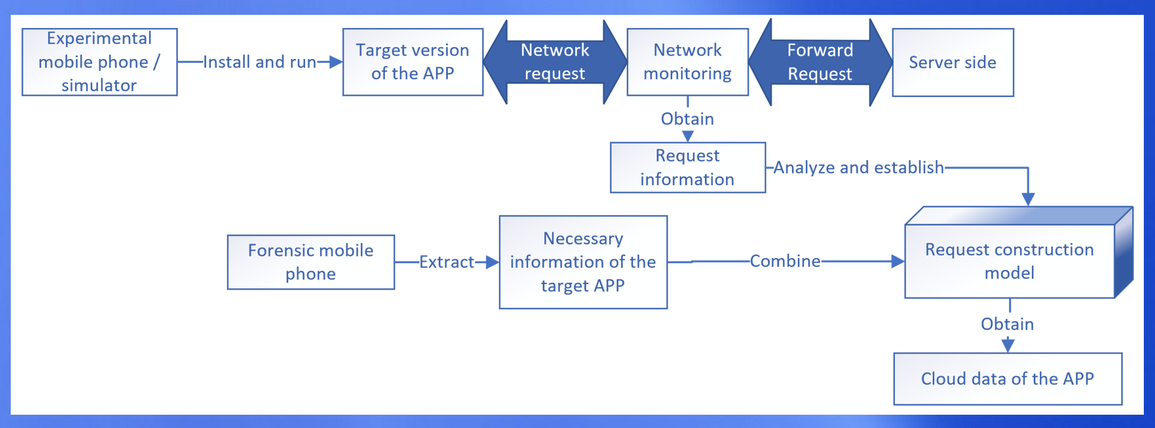

Fig.1. Diagram of Method Implementation

2. Materials and Tools:

By analyzing network packets with *Fiddler* and replicating app operations on an emulator, the method extracts data while preserving device integrity. We demonstrate this using Xiaomi smart speaker conversation records via protocol simulation. Details of materials and tools are in Table 1.

Table.1. Information of materials and tools

Type	Name	Version
Smart Speaker	Xiaomi Sound	-
Management APP	Xiaoai Speaker	v2.4 21
Emulator	MEMu Play	V8.0.2
Packet Capture Tool	Fiddler	v5.0.20204

3. Methodology

3.1 Request Monitoring- Identify Key Elements

Install the Xiaoai Speaker app on the emulator, log in to the account, launch Fiddler, retrieve conversation records through the app, and analyze the corresponding data requests in Fiddler . The intercepted request content is shown in Fig.2. The cloud server URL, *userID, serviceToken, deviceId, sn,* and *deviceSNProfile* are all essential elements for constructing the simulation protocol.

```
GET
https://userprofile.mina.mi.com/device_profile/conversation?timestamp=1655713129099&limit=20&requestId=Quhgo1EGp8u60qX2M4qA HTTP/1.1
User-Agent: MICO/AndroidApp/@SHIP.TO.2A2FE0D7@/2.4.10
Host: userprofile.mina.mi.com
Connection: Keep-Alive
Accept-Encoding: gzip
Cookie: userId=910036666;
serviceToken=7H56ukNFkPwgXs6C6AYnzLLCpG8+4YPmbINufSTBV7UW9gzjCZAU+zrjkOSo9vhsw+kCyYGhcdWgNo52uhCfyDOrR+JhC0Sz55RvsnICq+zvvvaAtlqc6QqpioooohDbSKhLxtLVvQm5DApwtAvTBgvokKQ8tfuXxS2w6HPIYDdSa+Yv38BHNQK8ZL7gPeXfuYLie7ToPgA8n6FJrBPLCDQ==;
userId=910036666; deviceId=e15b61fc-577d-45d4-af70-dfa914a129f5; deviceId=e15b61fc-577d-45d4-af70-dfa914a129f5;
sn=28605/A0R800651;
serviceToken=7H56ukNFkPwgXs6C6AYnzLLCpG8+4YPmbINufSTBV7UW9gzjCZAU+zrjkOSo9vhsw+kCyYGhcdWgNo52uhCfyDOrR+JhC0Sz55RvsnICq+zvvvaAtlqc6QqpioooohDbSKhLxtLVvQm5DApwtAvTBgvokKQ8tfuXxS2w6HPIYDdSa+Yv38BHNQK8ZL7gPeXfuYLie7ToPgA8n6FJrBPLCDQ==;
deviceSNProfile=eyJzaWduIjoiM2Q4ODMwY2E4MDgwZDJlNjQzN2RjZWQwYTViOGM4NTE3NzhlYjY0ZmIyMzg0YTBjYWRkODBhMGM0M2QyY2JkMCIsInNuIjoiMjg2MDUvQTBSODAwNjUxIn0=; sn=28605/A0R800651;
deviceSNProfile=eyJzaWduIjoiM2Q4ODMwY2E4MDgwZDJlNjQzN2RjZWQwYTViOGM4NTE3NzhlYjY0ZmIyMzg0YTBjYWRkODBhMGM0M2QyY2JkMCIsInNuIjoiMjg2MDUvQTBSODAwNjUxIn0=
```

Fig.2. GET request command intercepted by fiddler

3.2 Search and Communication– Find Out All Key Elements

(1) Local Search

We matched request data with '*com.xiaomi.mico*', the *'Xiaoai Speaker*' app's expanded backup. For instance, searching for "*userId*" (e.g., 910036666) located the miliaosdk database, where *XIAOMI_ID* and *SERVICE_TOKEN* in the *USER_ACCOUNT* table corresponded to *userId* and *serviceToken*, as shown in Fig.3.

UUID	XIAOMI_ID	SERVICE_TOKEN
543548806	910036666	_S001_C4pKvn+p+nTC0U1I5BbZpeA6JsK4aAIiRdh5eHMoq2aoX1+8lWt2zKgdFtT0r+RT

Fig.3. Data in the USER_ACCOUNT table

(2) Get request for more information

After obtaining stable user-related elements locally, a simulated GET request was constructed using the *userId* and *serviceToken* to retrieve device-related information from the cloud server. The data retrieved included details of Xiaomi Speaker device bound to the app, enabling the extraction of *deviceId*, *sn*, and *deviceSNProfile* elements for the device, as shown in Fig. 4.


```
{"code":0,"message":"Success","data":[{"deviceID":"9197a6eb-3547-4771-b0f0-712a8fc303d8","serialNumber":"34471/A1T238416","name":"Xiaomi Sound","alias":"XiaomiSound","current":false,"presence":"offline","address":"221.223.199.92","miotDID":"501578339","hardware":"L16A","romVersion":"1.76.11","capabilities":{"multiroom_music":1,"content_blacklist":1,"school_timetable":1,"user_nick_name":1,"night_mode":1,"player_pause_timer":1,"dialog_h5":1,"child_mode_2":1,"stereo_mode_v2":1,"voip_signal":1,"ai_instruction":1,"classified_alarm":1,"mesh":1,"ai_protocol_3_0":1,"voice_print_multidevice":1,"child_mode":1,"night_mode_detail":1,"baby_schedule":1,"tone_setting":1,"earthquake":1,"alarm_repeat_option_v2":1,"nearfield_relay":1,"xiaomi_voip":1,"nearby_wakeup_cloud":1,"continuous_dialogue":1,"bluetooth_option_v2":1,"answer_tone":1,"stereo_mode":1,"mico_current":1,"yunduantts":1,"voip_used_time":1},"remoteCtrlType":"","deviceSNProfile":"eyJhdWRpb19yZWxheV9pZG1fdmVyc2lvbiI6IjI1IiwiYXVkaW9fcmVsYXlfbWlsaW5rX3ZlcnNpb24iOiIxMjA0
```


Fig.4. The device data returned by the request

4. Acquiring Conversations

After obtaining all the essential information, we can use these elements to construct a GET request script to send to the Xiao Ai Speaker server. By doing so, we can retrieve the conversation records. This script is similar to Fig.2. The script targets the server userprofile.mima.mi.com, and part of the conversation returned by the server is stored is shown in Fig. 5. We can see that there are conversations, timestamps, and so on.


```
{"timestamp":1638797033103,"recordGroup":"{\"xiaoai\":{\"instructions\":\"[{\\\"header\\\":{\\\"namespace\\\":\\\"Nlp\\\",\\\"name\\\":\\\"StartAnswer\\\",\\\"id\\\":\\\"7f0db8ab34ad4902b22ab43c651c7d72\\\",\\\"dialog_id\\\":\\\"a3c867ea1e324dd2ac8320038c858671\\\"},\\\"payload\\\":{}},{\\\"header\\\":{\\\"namespace\\\":\\\"Execution\\\",\\\"name\\\":\\\"InstructionControl\\\",\\\"id\\\":\\\"ccd953ffc3494a9fad1fbc7db95c7f94\\\",\\\"dialog_id\\\":\\\"a3c867ea1e324dd2ac8320038c858671\\\"},\\\"payload\\\":{\\\"behavior\\\":\\\"INSERT_FRONT\\\"}},{\\\"header\\\":{\\\"namespace\\\":\\\"Template\\\",\\\"name\\\":\\\"Toast\\\",\\\"id\\\":\\\"c4baab6207204290b3502f2fde00655b\\\",\\\"dialog_id\\\":\\\"a3c867ea1e324dd2ac8320038c858671\\\"},\\\"payload\\\":{\\\"text\\\":\\\"现在是晚上9点23分\\\",\\\"query\\\":\\\"几点了\\\",\\\"display\\\":{\\\"f
```


Translation: It's 9:23 p.m

Translation: What's the time

Fig.5. The conversation data returned by the request