Hindawi
Security and Communication Networks
Volume 2021, Article ID 4973404, 9 pages
https://doi.org/10.1155/2021/4973404

Research Article

Data Privacy in Wearable IoT Devices: Anonymization and Deanonymization

Semi Park, Riha Kim, Hyunsik Yoon, and Kyungho Lee

School of Cybersecurity, Korea University, Seoul 02841, Republic of Korea

Correspondence should be addressed to Kyungho Lee; kevinlee@korea.ac.kr

Received 3 September 2021; Revised 27 October 2021; Accepted 11 November 2021; Published 27 November 2021

Academic Editor: Ilsun You

With the development of IoT devices, wearable devices are being used to record various types of information. Wearable IoT devices are attached to the user and can collect and transmit user data at all times along with a smartphone. In particular, sensitive information such as location information has an essential value in terms of privacy, and therefore some IoT devices implement data protection by introducing methods such as masking. However, masking can only protect privacy to a certain extent in logs having large numbers of recorded data. However, the effectiveness may decrease if we are linked with other information collected from within the device. Herein, a scenario-based case study on deanonymizing anonymized location information based on logs stored in wearable devices is described. As a result, we combined contextual and direct evidence from the collected information. It was possible to obtain the result in which the user could effectively identify the actual location. Through this study, not only can a deanonymized user location be identified but we can also confirm that cross-validation is possible even when dealing with modified GPS coordinates.

1. Introduction

With the development of IoT devices, the intimacy between device users and devices has increased. Previously, IoT devices were strongly perceived as being complex and difficult to use. Typical examples are wearable IoT devices, intelligent buildings, and smart cities. However, there are pros and cons to the development of such devices. Although IoT technology has made people's lives easier and more prosperous, there may be privacy issues with the information collected through such IoT devices [1]. In particular, sensors possessed by IoT devices will become more refined and diversified, and more diverse information will be exchanged through high-performance networks. Under this situation, in the case of a wearable device attached to a person 24 h a day, there is a much greater possibility of infringing on privacy. In line with this idea, wearable device manufacturers try to collect only the minimum amount of data necessary for a service by using anonymization techniques to prevent possible invasion of privacy of the data they collect [2]. In this paper, a method for deanonymizing anonymized location information is described using a scenario studied based on an IoT device investigation. Through four scenarios, the location information, which is sensitive information among the personal information of users, was collected or reprocessed for the development of meaningful data.

The structure of this thesis is as follows: Section 1 describes the direction of this research. In Section 2, we describe related studies on data analysis and data privacy processing methods for wearable IoT devices. Section 3 describes the experimental methodology used. In Section 4, case studies conducted based on the methodology described in Section 3 are detailed. Section 5 presents the significance and limitations of the experiments and follow-up studies and provides some concluding remarks.

2. Related Works

IoT devices are closely related to available devices and their need in people's lives; therefore, their analysis is becoming more critical. In particular, in the case of a wearable device, a user carries the device along with a smartphone at all times, which collects various logs. Based on this perspective, many studies have investigated the internal information of wearable devices.

In particular, Kasukurti and Patil [3] presented a methodology for investigating wearable devices along with a case study. In this study, data such as geolocation information, the user's physical and health information, activity logs, account details of social media interaction, calendar details, media files, key generation mechanisms, and KeyGen logs can be gathered. We can assert the practicality of a wearable device based on data.

In [4], an investigation method was conducted on three devices: Fitbit, Garmin, and HETP watches, all of which are popular low-cost devices. The methods for collecting forensic data through these devices are sound, and the identification of files of interest and location data on such devices has been studied.

In [5], the authors conducted a case study on a TomTom Spark 3 device. After installing the application on a companion Android-based device, the authors confirmed that data such as the device information, past user activities, and audio files could be acquired through an analysis.

Yoon and Karabiyik [6], Almogbil et al. [7], and Almogbil et al. [8] conducted a study using Fitbit, which is a wearable device. Fitbit supports functions such as the GPS itself. In addition, there is a possibility of manipulated or altered GPS-tracked activities, and thus there is a need for an open-source tool for dealing with sensitive information such as social media message notifications, credit card information, and health-related data. Fitbit saved this kind of data as plaintext. However, some other devices have stored sensitive information with masking or encryption to protect user privacy.

In [9], the authors researched sensitive information stored in wearable devices. They used a Samsung Gear S3, which is also a device used in a study by Becirovic and Mrdovic [10]. The authors confirmed that some data, including SMS messages and information, are stored on companion devices, determining the device owner's contextual environment. Baggili et al. [11] and Hassenfeldt et al. [12] also established a research methodology for wearable devices.

To understand the subtle nature of an insider threat, this paper reviews previous studies in this area. It examines the different types of insider threats based on insider characteristics and activities. Moreover, it explores sensors that can detect insider threats in an automated manner and the public datasets available for research. Finally, it examines the detection approaches used in related studies from various perspectives. In particular, IoT devices are the main threat to insider detection [13].

Because mobile devices are used in various areas, the risk of cyberattacks targeting them is becoming critical. In this paper, a threat intelligence evaluation is proposed for mobile malware from the viewpoint of situational awareness through the extraction of features that can detect Android malware using machine learning [14].

Location spoofing is a problem in mobile devices. In the present experiment, the authors used the Bluetooth discovery functionality to collect information about nearby devices and learn about the surrounding environment, which can be used to verify the genuineness of the GPS data. From this perspective, we are trying to discover masked location GPS data [15].

Because the problem of SCA in IoT devices has increased, the trends of SCAs used in such devices are introduced in this paper, and the urgency of developing countermeasures to single-trace attacks that only use side-channel information is suggested [16].

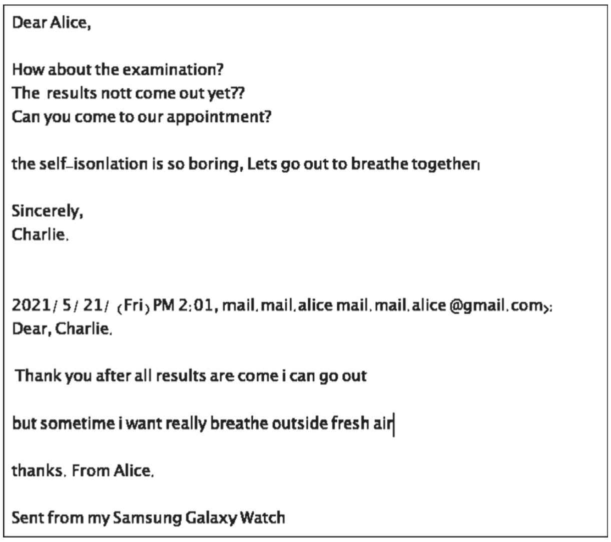
In addition, we consider deanonymization for a broader approach. This paper proposes an inference attack that can re-identify anonymous mobility data. The attack is based on a mobility model called the mobility Markov chain (MMC). Gambs et al. designed several distances between MMCs to evaluate their impact on anonymization. Experiments on real datasets demonstrate the efficiency of the attack. The results showed that anonymizing mobility data is a difficult task [17].

These studies collected various types of data, and the research methodologies described wearable devices. In some cases, devices such as Fitbit store sensitive information using plain text with a high risk of privacy violation, whereas devices such as Samsung Gear S3 partially mask the information using GPS. This study attempted to re-identify attempts to process such sensitive information, such as through masking.

3. Methodology

3.1. Scenario. We used a specific scenario for the case study because the primary target data in the study are location information, which establishes a situation in which location information can be sensitive. We got inspired from a famous cyber competition in South Korea. With the outbreak and spread of COVID-19 in 2021, the concept of a "self-quarantine" began to emerge. In South Korea, a law was enacted by which a person suspected of being infected with COVID-19 cannot leave the house for two weeks and must have their symptoms observed [18]. Under this scenario, Alice wants to go to the movie theatre to eat something using only a wearable device while leaving her smartphone at home. Law enforcement authorities can investigate such cases and track the self-quarantine location. Figure 1 shows some content related to the introductory part of the scenario. Based on this situation, an attempt was made to re-identify the masked location information in the wearable device.

3.2. Experimental Environment. The device used for the experiment was a Samsung Gear S3, and information about its companion device is shown in Figure 2. As previously confirmed in the related works section, we confirmed that some devices such as the Fitbit store sensitive location information such as GPS in plain text. The related information can be checked, as shown in Figure 3. Moreover, we executed a python-based anonymizer to generate masked sensitive anonymized data.



Dear Alice,

How about the examination?
The results nott come out yet??
Can you come to our appointment?

the self_isonlation is so boring, Lets go out to breathe together!

Sincerely,
Charlie.

2021/ 5/ 21/ (Fri) PM 2:01, mail.mail.alice mail.mail.alice @gmail.com>:
Dear, Charlie.

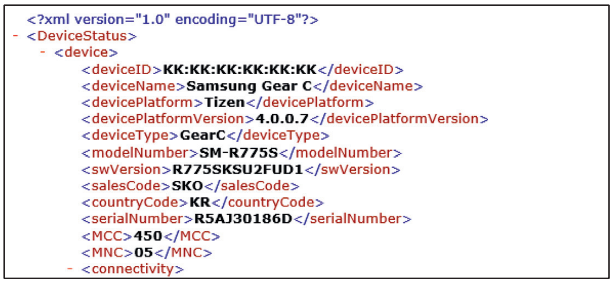
Thank you after all results are come i can go out

but sometime i want really breathe outside fresh air

thanks. From Alice.

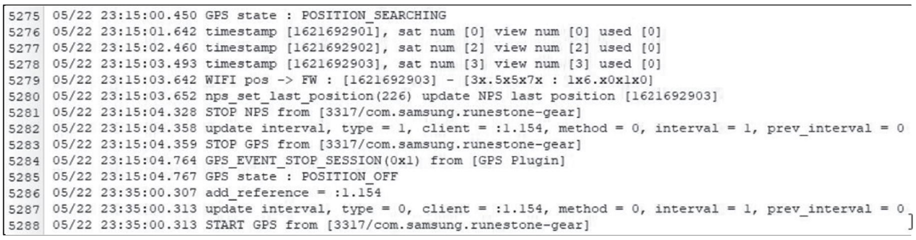
Sent from my Samsung Galaxy Watch

Figure 1: Scenario setup with e-mail.



```
  <?xml version="1.0" encoding="UTF-8"?>
- <DeviceStatus>
   - <device>
        <deviceID>KK:KK:KK:KK:KK:KK</deviceID>
        <deviceName>Samsung Gear C</deviceName>
        <devicePlatform>Tizen</devicePlatform>
        <devicePlatformVersion>4.0.0.7</devicePlatformVersion>
        <deviceType>GearC</deviceType>
        <modelNumber>SM-R775S</modelNumber>
        <swVersion>R775SKSU2FUD1</swVersion>
        <salesCode>SKO</salesCode>
        <countryCode>KR</countryCode>
        <serialNumber>R5AJ30186D</serialNumber>
        <MCC>450</MCC>
        <MNC>05</MNC>
      - <connectivity>
```

Figure 2: Information about wearable and companion devices.



```
05/22 23:15:00.450 GPS state : POSITION_SEARCHING
05/22 23:15:01.642 timestamp [1621692901], sat num [0] view num [0] used [0]
05/22 23:15:02.460 timestamp [1621692902], sat num [2] view num [2] used [0]
05/22 23:15:03.493 timestamp [1621692903], sat num [3] view num [3] used [0]
05/22 23:15:03.642 WIFI pos -> FW : [1621692903] - [3x.5x5x7x : 1x6.x0x1x0]
05/22 23:15:03.652 nps_set_last_position(226) update NPS last position [1621692903]
05/22 23:15:04.328 STOP NPS from [3317/com.samsung.runestone-gear]
05/22 23:15:04.358 update interval, type = 1, client = :1.154, method = 0, interval = 1, prev_interval = 0
05/22 23:15:04.359 STOP GPS from [3317/com.samsung.runestone-gear]
05/22 23:15:04.764 GPS_EVENT_STOP_SESSION(0x1) from [GPS Plugin]
05/22 23:15:04.767 GPS state : POSITION_OFF
05/22 23:35:00.307 add_reference = :1.154
05/22 23:35:00.313 update interval, type = 0, client = :1.154, method = 0, interval = 1, prev_interval = 0
05/22 23:35:00.313 START GPS from [3317/com.samsung.runestone-gear]
```

Figure 3: Example of sensitive anonymized data.

3.3. IoT Device Investigation. The device investigation used in this study utilized the tools applied in digital auditing. The Samsung Gear S3 used in the scenario applies to Tizen OS based on Android. The most commonly used file format for examining Android devices is the SQLite format [19]. In addition, various third-party applications running on Tizen OS use WebView independently [9]. To this end, we used the WEFA tool to analyze web-related artifacts and cache data. In addition, we conducted an investigation and deanonymization of wearable IoT devices and used basic Linux commands such as cat, grep, and awk for a log analysis [20].

In addition, before proceeding with the case study, classification was carried out using the concept of generating and storing evidence based on the Criminal Procedure Act of the Republic of Korea [21]. Generic evidence is data automatically generated by a system or application and is evidence with relatively little human intervention. Typical examples of its creation include web cache logs, system logs, and event-related logs. Archival evidence is data created to express a person's thoughts or feelings, representative examples of e-mails, SMSs, posts, or photographs. In this study, the generated evidence was directly related to the user's location information. This is called direct evidence. The archived evidence is indirectly related to the user's location information; therefore, only contextual evidence can be known.

The process of the IoT device investigation is shown in Figure 4. We considered masking-based anonymization techniques and encryption, which make it harder to identify specific users.

3.4. IoT Data Acquisition Investigation. The Android Debug Bridge (ADB) was used to acquire data from the Galaxy Gear S3. Because Tizen OS has components of Android, data can be collected using the pull function of the ADB [22]. The collection was generated only under the/data folder related to the user's behavior and the integrity of the collected data, which we can ensure using ad1 of FTK Imager [23].

4. Case Studies

4.1. Contextual Evidence (Personal E-Mail). Based on this scenario, we assumed that Alice escaped self-quarantine to meet Charlie; it is shown in Figure 5. We checked this type of data through an e-mail record. We determined the e-mail log using an SQLite database labeled/com.samsung.wemail/data/dbspace/.wemail.db. Alice received that e-mail and asked whether she had left her self-quarantine to go to a movie theatre for something to eat, which remains the limit of the contextual evidence. However, she later stated that she could serve as background knowledge that she could use direct evidence.

4.2. Contextual Evidence (Map Application and Web Browser). As shown in Figure 6, we confirmed that the device owner executed the GoogleMaps application by checking the given wearable device. An artifact labeled/usr/home/owner/applications/dbspace/.context-app-history.db contains the history of the applications executed in TizenOS in SQLite format.

In addition to the previously checked wemail, we can check the execution traces related to the Gearbrowser and Google Map. Google Map-related information can be found in/usr/home/owner/apps_rw/Eb12CGjuFc and/usr/apps/Eb12CGjuFc. After checking the directory structure before running the analysis, we can estimate that it operates based on chromium by checking the folder called chromium-efl, as shown in Figure 7 [24].

Next, as a result of a string search for files in the folder using the command, data estimated as GPS coordinates could be obtained in Figure 8.

In addition, by examining the cache data using ChromeCacheView, as shown in Figure 9, it is possible to check the search history of some GPS coordinates (37.535180, 126.903378) and the search near subway station.

To analyze the Gearbrowser log in the same way,/usr/apps/com.fin10.tizen.gearbrowser and/usr/home/owner/apps_rw/com.fin10.tizen.gearbrowser were targeted; it is shown in Figure 10.

As a result of checking the search history and string, traces related to CGV, a Korean movie brand, were found in Mok-dong. We cannot provide direct evidence due to contextual evidence's limitation, but it can act as background knowledge for use as future direct evidence.

4.3. Direct Evidence (Anonymized GPS Coordinates). The GPS data recorded in the device, that is, the system log related to the coordinates, leave a record of the location of the actual device, independent of user actions, such as the map, mail application, and web applications, which we looked at previously. We found GPS-related coordinates on TizenOS Gear S3 in/usr/data/location/dump_gps.log. Only those corresponding to the coordinate data were extracted separately using regular expressions and grep. The extraction results are as shown in Figure 11.

In Gear S3, we have seen that every even number of digits of the GPS coordinates is masked. Through the analysis of 4.1 and 4.2, we can infer that the relevance to Seoul, Korea, is high. In Seoul, Korea, the range is from longitude 37.715133 to 37.413294 and latitude 126.734086 to 127.269311. We observed that the first masked numbers of latitude and longitude were 7 and 2, respectively.

Because the GPS log is data from the GPS sensor, there may be errors in the GPS values depending on the surrounding environment and situation [25]. To prevent risk owing to such errors, we attempted deanonymization using data up to the second masking, which is far from the error range. As a result of checking the statistics based on the masked number, we observed almost no change in latitude and longitude except on 5/22, the day in which the individual escaped self-isolation, as shown in Figure 12.

In particular, checking the GPS masked based on the e-mail timeframe (after 2 pm) conducted in 4.1 on May 22, the individual was estimated to have escaped self-isolation, which we confirmed through the highest number occurring for the situation about Alice. We confirmed this through contextual evidence. The related images are shown in Figure 13.

It is not easy to specify the original GPS log owing to the masking of the original system log. However, we succeeded in deanonymizing the location visited by the actual user by correlating the information on the country where the incident occurred.

4.4. Direct Evidence (WI-FI SSID Names). We can check the logs related to the Wi-Fi connection in/usr/data/snlp/snlp_dump on TizenOS. In particular, a list of searched APs related to the Wi-Fi search is shown. In many stores and subway stations, the name of the Wi-Fi connection is the name of the store or subway station. In general, the Wi-Fi range is approximately 50–100 m. We can estimate that the user's device exists within 1 of 100 locations, the location of which can be determined using the store name or subway station that appears on the SSID [26]. We can infer this from the current location through/usr/data/snlp/snlp_dump, which records the SSID search result, as shown in Figure 14. In addition, awk and grep were used to extract related logs from/usr/data/snlp/snlp_dump. Subsequently, cases that could identify the actual location using heuristics were confirmed.

The first case is the location identified when contextual evidence is used with other evidence. Under this scenario, we confirmed through the Wi-Fi search, log-based on circumstantial evidence that we determined through the

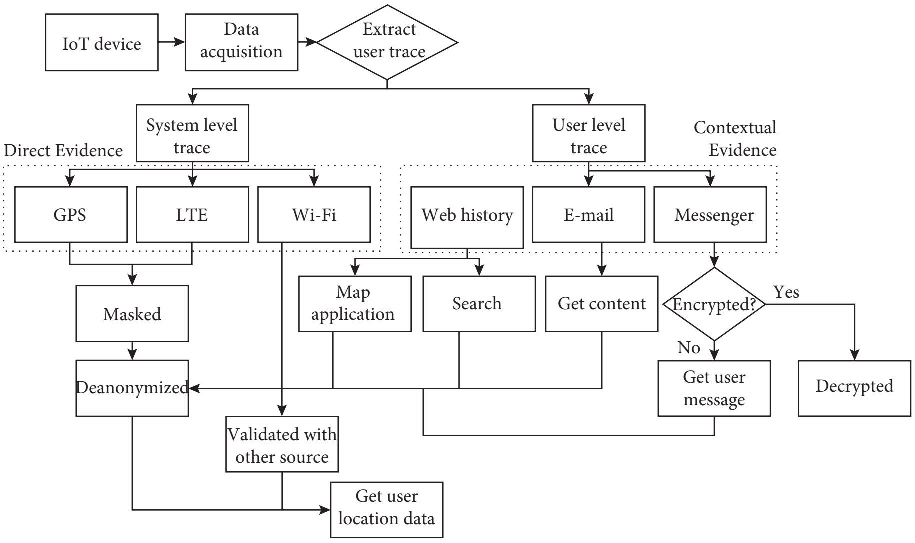
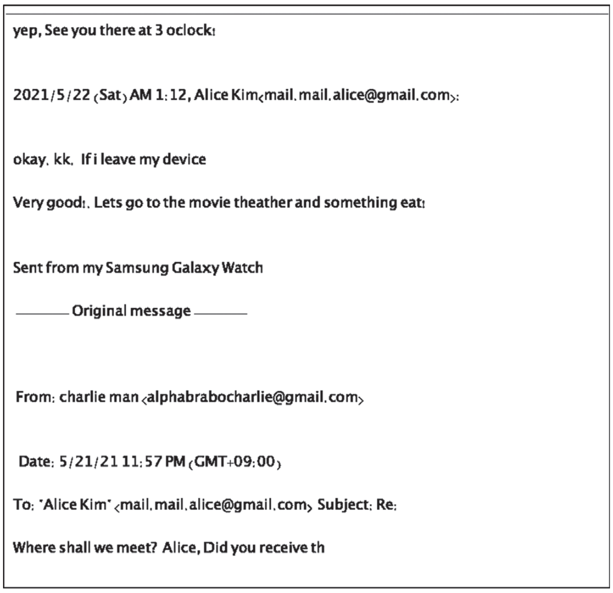


Figure 4: Research process of IoT device investigation.



yep, See you there at 3 oclock!

2021/5/22 (Sat) AM 1:12, Alice Kim<mail.mail.alice@gmail.com>:

okay. kk. If i leave my device

Very good!. Lets go to the movie theather and something eat!

Sent from my Samsung Galaxy Watch

——— Original message ———

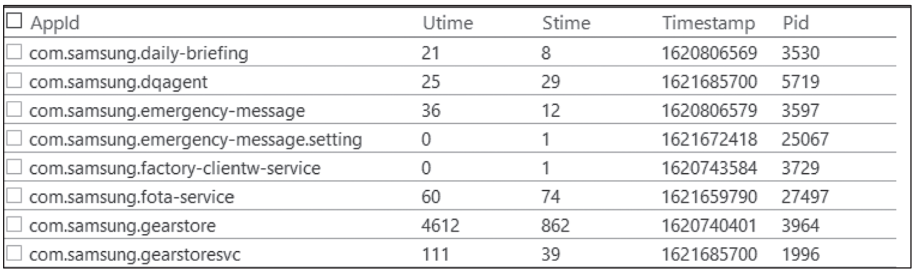
From: charlie man <alphabrabocharlie@gmail.com>

Date: 5/21/21 11:57 PM (GMT+09:00)

To: "Alice Kim" <mail.mail.alice@gmail.com> Subject: Re:

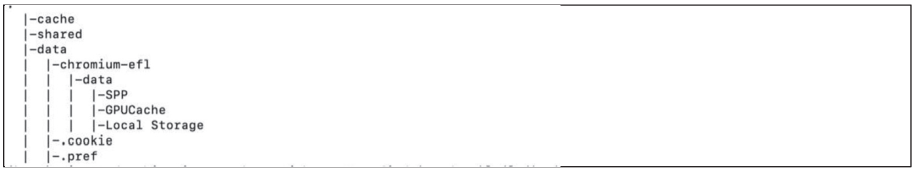
Where shall we meet? Alice, Did you receive th

Figure 5: Information regarding Alice’s appointment.



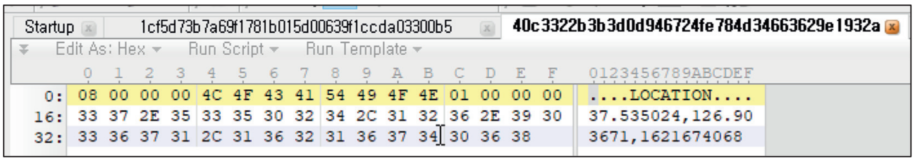
☐ AppId	Utime	Stime	Timestamp	Pid
☐ com.samsung.daily-briefing	21	8	1620806569	3530
☐ com.samsung.dqagent	25	29	1621685700	5719
☐ com.samsung.emergency-message	36	12	1620806579	3597
☐ com.samsung.emergency-message.setting	0	1	1621672418	25067
☐ com.samsung.factory-clientw-service	0	1	1620743584	3729
☐ com.samsung.fota-service	60	74	1621659790	27497
☐ com.samsung.gearstore	4612	862	1620740401	3964
☐ com.samsung.gearstoresvc	111	39	1621685700	1996

Figure 6: Application execution log.



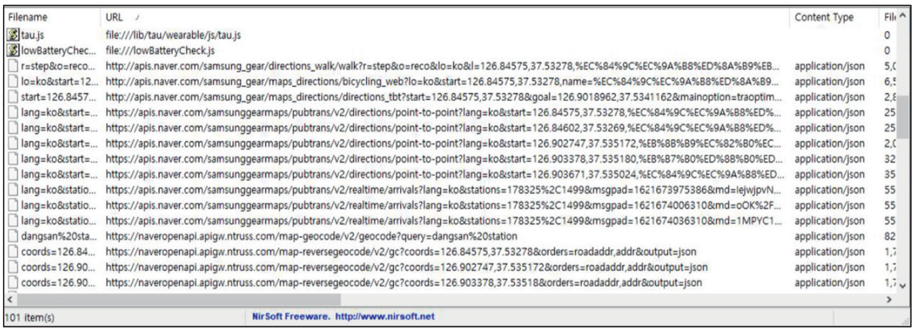
```
|-cache
|-shared
|-data
|  |-chromium-efl
|  |  |-data
|  |  |  |-SPP
|  |  |  |-GPUCache
|  |  |  |-Local Storage
|  |-.cookie
|  |-.pref
```

Figure 7: Structure of folder data (Google Maps).


Startup
1cf5d73b7a69f1781b015d00639f1ccda03300b5
40c3322b3b3d0d946724fe784d34663629e1932a
Edit As: Hex
Run Script
Run Template

```
     0  1  2  3  4  5  6  7  8  9  A  B  C  D  E  F   0123456789ABCDEF
 0: 08 00 00 00 4C 4F 43 41 54 49 4F 4E 01 00 00 00   ....LOCATION....
16: 33 37 2E 35 33 35 30 32 34 2C 31 32 36 2E 39 30   37.535024,126.90
32: 33 36 37 31 2C 31 36 32 31 36 37 34 30 36 38      3671,1621674068
```

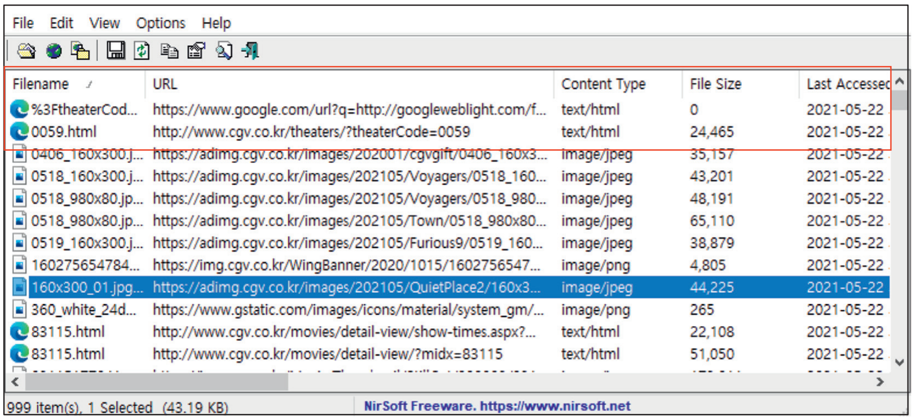
Figure 8: Results of string search in Google Maps folder.



Filename	URL	Content Type	Fil
tau.js	file:///lib/tau/wearable/js/tau.js		0
lowBatteryChec...	file:///lowBatteryCheck.js		0
r=step&o=reco...	http://apis.naver.com/samsung_gear/directions_walk/walk?r=step&o=reco&lo=ko&l=126.84575,37.53278,%EC%84%9C%EC%9A%B8%ED%8A%B9%EB...	application/json	5,(
lo=ko&start=12...	http://apis.naver.com/samsung_gear/maps_directions/bicycling_web?lo=ko&start=126.84575,37.53278,name=%EC%84%9C%EC%9A%B8%ED%8A%B9...	application/json	6,5
start=126.8457...	http://apis.naver.com/samsung_gear/maps_directions/directions_tbt?start=126.84575,37.53278&goal=126.9018962,37.5341162&mainoption=traoptim...	application/json	2,8
lang=ko&start=...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/directions/point-to-point?lang=ko&start=126.84575,37.53278,%EC%84%9C%EC%9A%B8%ED%...	application/json	25
lang=ko&start=...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/directions/point-to-point?lang=ko&start=126.84602,37.53269,%EC%84%9C%EC%9A%B8%ED%...	application/json	25
lang=ko&start=...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/directions/point-to-point?lang=ko&start=126.902747,37.535172,%EB%8B%B9%EC%82%B0%EC...	application/json	2,(
lang=ko&start=...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/directions/point-to-point?lang=ko&start=126.903378,37.535180,%EB%B7%B0%ED%8B%B0%ED...	application/json	32
lang=ko&start=...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/directions/point-to-point?lang=ko&start=126.903671,37.535024,%EC%84%9C%EC%9A%B8%ED...	application/json	35
lang=ko&statio...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/realtime/arrivals?lang=ko&stations=178325%2C1499&msgpad=1621673975386&md=lejwjpvN...	application/json	55
lang=ko&statio...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/realtime/arrivals?lang=ko&stations=178325%2C1499&msgpad=1621674006310&md=oOK%2F...	application/json	55
lang=ko&statio...	https://apis.naver.com/samsunggearmaps/pubtrans/v2/realtime/arrivals?lang=ko&stations=178325%2C1499&msgpad=1621674036310&md=1MPYC1...	application/json	55
dangsan%20sta...	https://naveropenapi.apigw.ntruss.com/map-geocode/v2/geocode?query=dangsan%20station	application/json	82
coords=126.84...	https://naveropenapi.apigw.ntruss.com/map-reversegeocode/v2/gc?coords=126.84575,37.53278&orders=roadaddr,addr&output=json	application/json	1,7
coords=126.90...	https://naveropenapi.apigw.ntruss.com/map-reversegeocode/v2/gc?coords=126.902747,37.535172&orders=roadaddr,addr&output=json	application/json	1,7
coords=126.90...	https://naveropenapi.apigw.ntruss.com/map-reversegeocode/v2/gc?coords=126.903378,37.53518&orders=roadaddr,addr&output=json	application/json	1,7

101 item(s) NirSoft Freeware. http://www.nirsoft.net

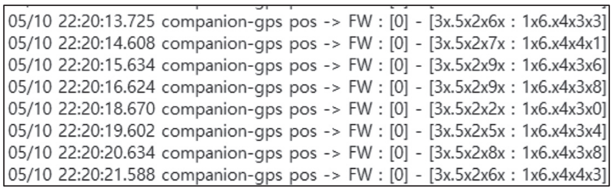
Figure 9: Web cache result.



Filename	URL	Content Type	File Size	Last Accessed
%3FtheaterCod...	https://www.google.com/url?q=http://googleweblight.com/f...	text/html	0	2021-05-22
0059.html	http://www.cgv.co.kr/theaters/?theaterCode=0059	text/html	24,465	2021-05-22
0406_160x300.j...	https://adimg.cgv.co.kr/images/202001/cgvgift/0406_160x3...	image/jpeg	35,157	2021-05-22
0518_160x300.j...	https://adimg.cgv.co.kr/images/202105/Voyagers/0518_160...	image/jpeg	43,201	2021-05-22
0518_980x80.jp...	https://adimg.cgv.co.kr/images/202105/Voyagers/0518_980...	image/jpeg	48,191	2021-05-22
0518_980x80.jp...	https://adimg.cgv.co.kr/images/202105/Town/0518_980x80...	image/jpeg	65,110	2021-05-22
0519_160x300.j...	https://adimg.cgv.co.kr/images/202105/Furious9/0519_160...	image/jpeg	38,879	2021-05-22
160275654784...	https://img.cgv.co.kr/WingBanner/2020/1015/1602756547...	image/png	4,805	2021-05-22
160x300_01.jpg...	https://adimg.cgv.co.kr/images/202105/QuietPlace2/160x3...	image/jpeg	44,225	2021-05-22
360_white_24d...	https://www.gstatic.com/images/icons/material/system_gm/...	image/png	265	2021-05-22
83115.html	http://www.cgv.co.kr/movies/detail-view/show-times.aspx?...	text/html	22,108	2021-05-22
83115.html	http://www.cgv.co.kr/movies/detail-view/?midx=83115	text/html	51,050	2021-05-22

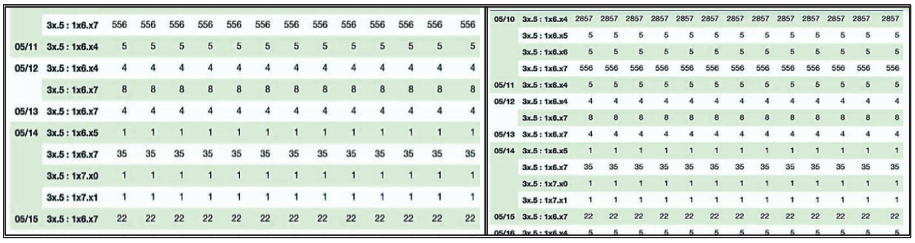
999 item(s), 1 Selected (43.19 KB) NirSoft Freeware. https://www.nirsoft.net

Figure 10: Related information about Gearbrowser.



```
05/10 22:20:13.725 companion-gps pos -> FW : [0] - [3x.5x2x6x : 1x6.x4x3x3]
05/10 22:20:14.608 companion-gps pos -> FW : [0] - [3x.5x2x7x : 1x6.x4x4x1]
05/10 22:20:15.634 companion-gps pos -> FW : [0] - [3x.5x2x9x : 1x6.x4x3x6]
05/10 22:20:16.624 companion-gps pos -> FW : [0] - [3x.5x2x9x : 1x6.x4x3x8]
05/10 22:20:18.670 companion-gps pos -> FW : [0] - [3x.5x2x2x : 1x6.x4x3x0]
05/10 22:20:19.602 companion-gps pos -> FW : [0] - [3x.5x2x5x : 1x6.x4x3x4]
05/10 22:20:20.634 companion-gps pos -> FW : [0] - [3x.5x2x8x : 1x6.x4x3x8]
05/10 22:20:21.588 companion-gps pos -> FW : [0] - [3x.5x2x6x : 1x6.x4x4x3]
```

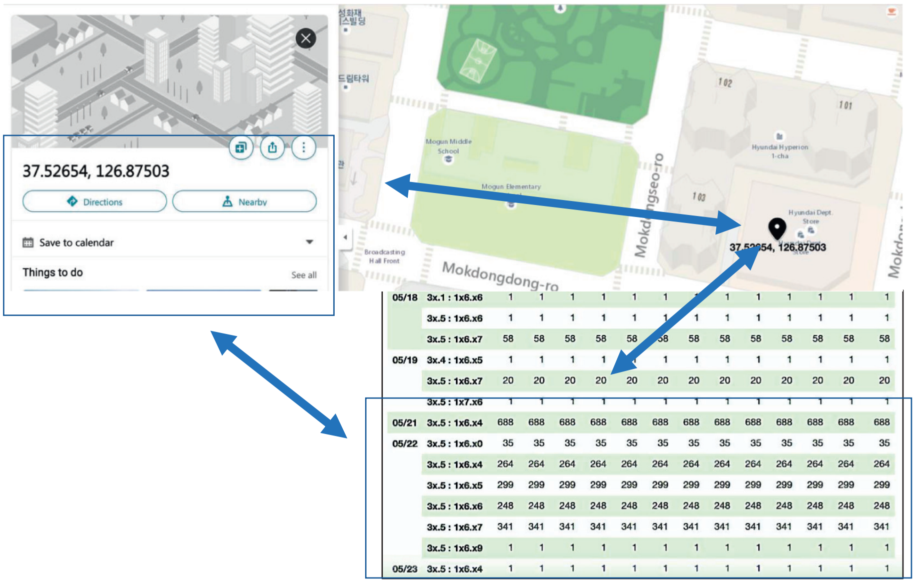
Figure 11: Extracted masked location data.



	3x.5 : 1x6.x7	556	556	556	556	556	556	556	556	556	556	556	556	556
05/11	3x.5 : 1x6.x4	5	5	5	5	5	5	5	5	5	5	5	5	5
05/12	3x.5 : 1x6.x4	4	4	4	4	4	4	4	4	4	4	4	4	4
	3x.5 : 1x6.x7	8	8	8	8	8	8	8	8	8	8	8	8	8
05/13	3x.5 : 1x6.x7	4	4	4	4	4	4	4	4	4	4	4	4	4
05/14	3x.5 : 1x6.x5	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x6.x7	35	35	35	35	35	35	35	35	35	35	35	35	35
	3x.5 : 1x7.x0	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x7.x1	1	1	1	1	1	1	1	1	1	1	1	1	1
05/15	3x.5 : 1x6.x7	22	22	22	22	22	22	22	22	22	22	22	22	22

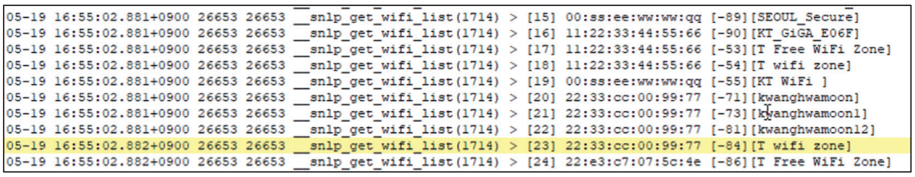
05/10	3x.5 : 1x6.x4	2857	2857	2857	2857	2857	2857	2857	2857	2857	2857	2857	2857	2857
	3x.5 : 1x6.x5	5	5	5	5	5	5	5	5	5	5	5	5	5
	3x.5 : 1x6.x6	5	5	5	5	5	5	5	5	5	5	5	5	5
	3x.5 : 1x6.x7	556	556	556	556	556	556	556	556	556	556	556	556	556
05/11	3x.5 : 1x6.x4	5	5	5	5	5	5	5	5	5	5	5	5	5
05/12	3x.5 : 1x6.x4	4	4	4	4	4	4	4	4	4	4	4	4	4
	3x.5 : 1x6.x7	8	8	8	8	8	8	8	8	8	8	8	8	8
05/13	3x.5 : 1x6.x7	4	4	4	4	4	4	4	4	4	4	4	4	4
05/14	3x.5 : 1x6.x5	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x6.x7	35	35	35	35	35	35	35	35	35	35	35	35	35
	3x.5 : 1x7.x0	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x7.x1	1	1	1	1	1	1	1	1	1	1	1	1	1
05/15	3x.5 : 1x6.x7	22	22	22	22	22	22	22	22	22	22	22	22	22
05/16	3x.5 : 1x6.x4	5	5	5	5	5	5	5	5	5	5	5	5	5

Figure 12: Statistics about masked location log (day by day).



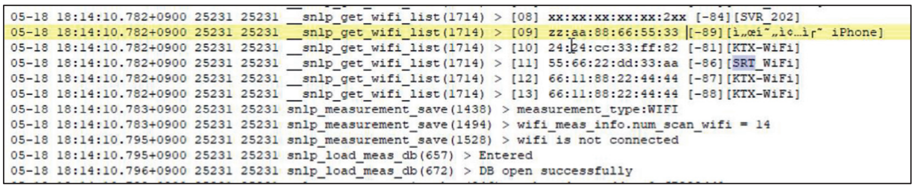
05/18	3x.1 : 1x6.x6	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x6.x6	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x6.x7	58	58	58	58	58	58	58	58	58	58	58	58	58
05/19	3x.4 : 1x6.x5	1	1	1	1	1	1	1	1	1	1	1	1	1
	3x.5 : 1x6.x7	20	20	20	20	20	20	20	20	20	20	20	20	20
	3x.5 : 1x7.x6	1	1	1	1	1	1	1	1	1	1	1	1	1
05/21	3x.5 : 1x6.x4	688	688	688	688	688	688	688	688	688	688	688	688	688
05/22	3x.5 : 1x6.x0	35	35	35	35	35	35	35	35	35	35	35	35	35
	3x.5 : 1x6.x4	264	264	264	264	264	264	264	264	264	264	264	264	264
	3x.5 : 1x6.x5	299	299	299	299	299	299	299	299	299	299	299	299	299
	3x.5 : 1x6.x6	248	248	248	248	248	248	248	248	248	248	248	248	248
	3x.5 : 1x6.x7	341	341	341	341	341	341	341	341	341	341	341	341	341
	3x.5 : 1x6.x9	1	1	1	1	1	1	1	1	1	1	1	1	1
05/23	3x.5 : 1x6.x4	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 13: Deanonymization result on GPS coordinates.



```
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [15] 00:ss:ee:ww:ww:qq [-89][SEOUL_Secure]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [16] 11:22:33:44:55:66 [-90][KT_GiGA_E06F]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [17] 11:22:33:44:55:66 [-53][T Free WiFi Zone]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [18] 11:22:33:44:55:66 [-54][T wifi zone]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [19] 00:ss:ee:ww:ww:qq [-55][KT WiFi ]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [20] 22:33:cc:00:99:77 [-71][kwanghwamoon]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [21] 22:33:cc:00:99:77 [-73][kwanghwamoon1]
05-19 16:55:02.881+0900 26653 26653 __snlp_get_wifi_list(1714) > [22] 22:33:cc:00:99:77 [-81][kwanghwamoon12]
05-19 16:55:02.882+0900 26653 26653 __snlp_get_wifi_list(1714) > [23] 22:33:cc:00:99:77 [-84][T wifi zone]
05-19 16:55:02.882+0900 26653 26653 __snlp_get_wifi_list(1714) > [24] 22:e3:c7:07:5c:4e [-86][T Free WiFi Zone]
```

Figure 14: Deanonymization result with SSID name and GPS coordinates.



```
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [08] xx:xx:xx:xx:xx:2xx [-84][SVR_202]
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [09] zz:aa:88:66:55:33 [-89][ì„œí˜„ì¢…ìƒ˜ iPhone]
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [10] 24:24:cc:33:ff:82 [-81][KTX-WiFi]
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [11] 55:66:22:dd:33:aa [-86][SRT_WiFi]
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [12] 66:11:88:22:44:44 [-87][KTX-WiFi]
05-18 18:14:10.782+0900 25231 25231 __snlp_get_wifi_list(1714) > [13] 66:11:88:22:44:44 [-88][KTX-WiFi]
05-18 18:14:10.783+0900 25231 25231 snlp_measurement_save(1438) > measurement_type:WIFI
05-18 18:14:10.783+0900 25231 25231 snlp_measurement_save(1494) > wifi_meas_info.num_scan_wifi = 14
05-18 18:14:10.795+0900 25231 25231 snlp_measurement_save(1528) > wifi is not connected
05-18 18:14:10.795+0900 25231 25231 snlp_load_meas_db(657) > Entered
05-18 18:14:10.796+0900 25231 25231 snlp_load_meas_db(672) > DB open successfully
```

Figure 15: Location detection based on SSID names.

.pref of Google Maps, that the device exists in the corresponding location, which indicated transportation infrastructure (KTX), as shown in Figure 15.

At approximately 17 : 34, we inferred that the device was passed near Kwanghwamoon Station. However, the error range was too broad with a masked GPS log, and we could not conduct cross-validation. In the deanonymization method using the SSID name, we found that the location specification was not as precise as in the anonymized GPS data. However, we found a case in which location identification was sufficiently possible when combined with contextual evidence.

5. Conclusions

This paper presented a deanonymization method for location information in an anonymized wearable IoT environment. As expected from the existing related studies on wearable devices, we checked various records related to user behavior and a system log. We conducted a case study using these user behavior data by building a scenario that could occur with an actual IoT device in the current COVID-19 environment, which is a hot topic.

In the case study, we divided the data collected from a Samsung Gear S3 into contextual evidence based on user behavior and direct evidence related to the system logs. We found contextual evidence, e-mail data, and application data. Through an email, users can exchange information such as a promise to go to a specific place. Such evidence is consistent with the nature of contextual evidence. There is a limitation in that which is necessary to determine whether the user went to the location separately. We stated that only an appointment had been made and that it had not occurred. As traces related to the application, searching for information about movie theaters through the map application and the Internet was found. These traces are also data related to searching for a route, and there is a limitation in which they cannot be used as direct evidence that the actual user went to the location.

However, this does not mean that contextual evidence is meaningless. We used contextual evidence and direct evidence as a later case study. First, through anonymization, GPS coordinates were set as direct evidence, and masking was applied to the log itself, making it impossible to specify the user's location and path. However, based on the information obtained through contextual evidence in advance, the location can be specified using information that the user is in Seoul and will see a movie. In addition, contextual evidence can help with direct evidence related to the Wi-Fi, which is called snlp dump. Logic reinforcement was possible through Wi-Fi SSID-based distance estimation for contextual information obtained through the map application.

This case study examined the meaning of contextual and direct evidence collected through wearable devices. From the perspective of direct evidence, if privacy such as masking is not processed, it can be used as the most powerful type of intelligence. However, in the case of masking, as in the current scenario, contextual evidence is indispensable. In addition, when using such a GPS manipulation application, if we conduct cross-verification with contextual evidence, we can also characterize whether the attacker applied actual GPS data.

However, the argument in this study has certain limitations. First, we found that it was impossible to generalize the results because only one type of wearable device was masked. However, regardless of the masking process, there are ways in which contextual and direct evidence can be used sufficiently as a general approach. Moreover, to apply these methodologies, there must be a guarantee that no log loss or manipulation exists during the collection phase [27].

In future research, we plan to anonymize sensitive information in anonymized IoT devices. In particular, in this paper, geolocation is limited as sensitive information; however, we intend to research more diverse devices and sensitive information in future research.

Data Availability

Because the data contained private information about an individual, it is difficult to share such data through open access. If someone wants the data for further research, please contact the authors through e-mail (semi0502@korea.ac.kr).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by a grant of the Korean Heath Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (HI19C0866), Ministry of Education of the Republic of the Korea, and National Research Foundation of Korea (NRF-2018S1A5B6075594).

References

[1] F. Servida and E. Casey, "IoT forensic challenges and opportunities for digital traces," *Digital Investigation*, vol. 28, pp. S22–S29, 2019.

[2] T. Zia, P. Liu, and W. Han, "Application-specific digital forensics investigative model in Internet of Things (IoT)," in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, pp. 1–7, Reggio Calabria, Italy, September 2017.

[3] D. H. Kasukurti and S. Patil, "Wearable device forensic: probable case studies and proposed methodology," in *Proceedings of the International Symposium on Security in Computing and Communication*, pp. 290–300, Bangalore, India, September 2018.

[4] Á. MacDermott, S. Lea, F. Iqbal, I. Idowu, and B. Shah, "Forensic analysis of wearable devices: Fitbit, Garmin and HETP watches," in *Proceedings of the 2019 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, pp. 1–6, Canary Island, Spain, June 2019.

[5] L. Dawson and A. Akinbi, "Challenges and opportunities for wearable IoT forensics: TomTom Spark 3 as a case study," *Forensic Science International: Report*, vol. 3, 2021.

[6] Y. H. Yoon and U. Karabiyik, "Forensic analysis of Fitbit versa 2 data on android," *Electronics*, vol. 9, no. 9, pp. 1431–1448, 2020.

[7] A. Almogbil, A. Alghofaili, C. Deane, and T. Leschke, "Digital forensic analysis of Fitbit wearable technology: an investigator's guide," in *Proceedings of the 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp. 44–49, New York, NY, USA, August 2020.

[8] A. Almogbil, A. Alghofaili, C. Deane, and T. Leschke, "The accuracy of GPS-enabled Fitbit activities as evidence: a digital forensics study," in *Proceedings of the 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge*

Computing and Scalable Cloud (EdgeCom), pp. 186–189, New York, NY, USA, August 2020.

[9] N. R. Odom, J. M. Lindmar, J. Hirt, and J. Brunty, "Forensic inspection of sensitive user data and artifacts from smartwatch wearable devices," *Journal of Forensic Sciences*, vol. 64, no. 6, pp. 1673–1686, 2019.

[10] S. Becirovic and S. Mrdovic, "Manual IoT forensics of a samsung gear S3 frontier smartwatch," in *Proceedings of the 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–5, Split, Croatia, September 2019.

[11] I. Baggili, J. Oduro, K. Anthony, F. Breitinger, and G. McGee, "Watch what you wear: preliminary forensic analysis of smart watches," in *Proceedings of the 2015 10th International Conference on Availability, Reliability and Security*, pp. 303–311, Toulouse, France, August 2015.

[12] C. Hassenfeldt, S. Baig, I. Baggili, and X. Zhang, "Map My Murder: a digital forensic study of mobile health and fitness applications," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, pp. 1–12, New York, NY, USA, August 2019.

[13] A. R. Kim, J. Oh, J. Ryu, J. Lee, K. Kwon, and K. Lee, "SoK: a systematic review of insider threat detection," *Journal of Wireles Mobile Networks Ubiquitous Computer Dependable Applications*, vol. 10, no. 4, pp. 46–67, 2019.

[14] M. K. Park, J. W. Seo, J. Han, H. Oh, and K. Lee, "Situational awareness framework for threat intelligence measurement of Android malware," *Journal of Wireless Mobile Networks Ubiquitous Computer Dependable Application*, vol. 9, no. 3, pp. 25–38, 2018.

[15] S. K. Wong and S. M. Yiu, "Identification of device motion status via Bluetooth discovery," *Journal of Internet Service and Information Secuity*, vol. 10, no. 4, pp. 59–69, 2020.

[16] B. Sim and D. Han, "A study on the side-channel analysis trends for application to IoT devices," *JJournal of Internet Service and Information Secuity*, vol. 10, pp. 2–21, 2020.

[17] S. Gambs, M.-O. Killijian, and M. Núñez del Prado Cortez, "De-anonymization attack on geolocated data," *Journal of Computer and System Sciences*, vol. 80, no. 8, pp. 1597–1614, 2014.

[18] D. Lee and J. Lee, "Testing on the move: South Korea's rapid response to the COVID-19 pandemic," *Transportation Research Interdisciplinary Perspectives*, vol. 5, pp. 100111–100120, 2020.

[19] A. Hoog, "Android software development kit and android Debug Bridge," in *Android Forensics: Investigation, Analysis and mobile Security for Google Android*, pp. 65–104, Elsevier, Waltham, MA, USA, 2011.

[20] X. Lin, "Keyword forensics," *Introductory Computer Forensics*, Springer, New York, NY, USA, 2018.

[21] S. G. Kim and J. S. Park, "Legal limits of search and seizure for digital forensic in Korea," *Journal of Computer Virology and Hacking Techniques*, vol. 10, no. 2, pp. 157–163, 2014.

[22] P. Feng, Q. Li, P. Zhang, and Z. Chen, "Logical acquisition method based on data migration for Android mobile devices," *Digital Investigation*, vol. 26, pp. 55–62, 2018.

[23] F. Carbone, "Working with FTK imager," in *Computer Forensics with FTK*Packt Publishing, Birmingham, UK, 2014.

[24] F. Immanuel, B. Martini, and K. K. R. Choo, "Android cache taxonomy and forensic process," in *Proceedings of the 2015 IEEE Trustcom/BigDataSE/ISPA*, pp. 1094–1101, Helsinki, Finland, August 2015.

[25] H. Zhang, J. Zheng, and H. Zhou, "GPS positioning error analysis and outlier elimination method in forestry," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 41, no. 5, pp. 143–147, 2010.

[26] S. Al-Kuwari and S. D. Wolthusen, "A survey of forensic localization and tracking mechanisms in short-range and cellular networks," in *Proceedings of the International Conference on Digital Forensics and Cyber Crime*, pp. 19–32, Albany, NY, USA, October 2009.

[27] B. Boeck, D. Huemer, and A. M. Tjoa, "Towards more trustable log files for digital forensics by means of trusted computing," in *Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications*, pp. 1020–1027, Perth, Western Australia, April 2010.