Kazuki Nomoto*, Mitsuaki Akiyama, Masashi Eto, Atsuo Inomata, and Tatsuya Mori

On the Feasibility of Linking Attack to Google/Apple Exposure Notification Framework

Abstract: Digital contact-tracing (DCT) applications have been installed on more than 188 M smartphones worldwide as an effective mechanism for monitoring contact with COVID-19 infected individuals. DCT is promising not only for COVID-19, but also for preparing for a possible future large-scale pandemic. The DCT framework is unique in that it combines Bluetooth Low Energy (BLE) communications with cryptography techniques to track exposure on a large scale while protecting user privacy. The objective of this study is to assess the risk of the *linking attack* to the DCT frameworks; i.e., linking individuals to the identifiers contained in BLE broadcast frames that are supposed to be anonymized. Specifically, we target Google/Apple's Exposure Notification (GAEN), which is the representative implementation of DCT. Our extensive experiments demonstrate that passively collected rolling proximity identifiers (RPIs) contained in the BLE frames can be linked to face photos which could lead to the exposure of privacy information with high accuracy, including infection status. We also demonstrate that an attacker with a few number of devices can correctly link RPIs and the images of the target person with a success rate of 86% at a rate of 5,000 users per hour. Based on these results, we propose countermeasures to reduce the inherent privacy risk of the GAEN framework.

Keywords: Privacy risk, Exposure Notification, BLE, COVID-19

DOI 10.56553/popets-2022-0103
Received 2022-02-28; revised 2022-06-15; accepted 2022-06-16.

***Corresponding Author: Kazuki Nomoto:** Waseda University, E-mail: nomotokazuki@nsl.cs.waseda.ac.jp
Mitsuaki Akiyama: NTT, E-mail: akiyama@ieee.org
Masashi Eto: Ministry of Internal Affairs and Communications (MIC), E-mail: m2.etou@soumu.go.jp
Atsuo Inomata: Osaka University, E-mail: inomata@mail.osaka-u.ac.jp
Tatsuya Mori: Waseda University/NICT/RIKEN AIP, E-mail: mori@nsl.cs.waseda.ac.jp

1 Introduction

As a promising technology aimed at understanding and preventing the spread of infectious diseases such as COVID-19, smartphone-based digital contact-tracing (DCT) frameworks have been developed and are being deployed worldwide. DCT frameworks take advantage of the fact that most of the world's population carry smartphones with Bluetooth Low Energy (BLE) communication capabilities to provide a means of quantitative assessment of the spread of infectious diseases and the individual risk of infection. Swaminathan et al. conjectured that the DCT approach will be used in future pandemics of infectious diseases [1]. Several studies, including a study by Chowdhury et al. [2], examined ways to improve DCT technology with the goal of preparing for future pandemics.

Although DCT frameworks offer the benefit of providing quick insight into the spread of infection, they should not invade people's privacy. In the worst-case scenario, privacy leakage may lead to unlawful discrimination or abuse. For instance, during the COVID-19 pandemic, discrimination and prejudice against infected individuals and healthcare workers became a serious social problem [3–5]. Aduhammad et al. discussed the acceptability and ethical issues associated with the use of DCT [6]. According to their study, DCT technology raises concerns for people, including those related to privacy, voluntariness, and the beneficence of data usage. Kaspar et al. found that users are negatively motivated to use DCT because of their perception of the severity of data misuse and their vulnerability to data misuse [7].

Given these observations, it is critical to implement built-in user-privacy protection mechanisms in the DCT. The Google/Apple's Exposure Notification (GAEN) is representative implementation of DCT. It is a distributed system in which contact tracings are performed collectively from user smartphone apps; cryptographic techniques are applied to protect personal privacy. In principle, it is impossible for GAEN developers and operators to obtain data on an individual's contact and location history. As of 2022, GAEN is a standard feature in the world's leading mobile operating systems: Android and iOS. Furthermore, it is expected to be ef-

fectively utilized in the event of a new pandemic in the future.

However, the GAEN framework has inherent privacy risks that can lead to the unauthorized disclosure or theft of personal information. In this study, we focus on the risk of a *face photo linking attack* (hereafter referred to as a linking attack), which links a person's face photo to his or her infection status.The principle of the attack is described below. A smartphone installed with GAEN advertises a unique identifier to neighboring smartphones using BLE. As described later, this identifier can be derived from the temporary key that a person who tests positive for COVID-19 registers on a server operated by a health authority. Using commodity hardware setup, consisting of BLE receivers and cameras, an attacker collects the identifier sent from the GAEN user's smartphone and the person's face photo at the same time, and links them as a pair. The attacker then derives the identifier from the temporary key of the positive person disclosed by the server of the health authority and matches it with the collected pair. Consequently, the attacker can obtain a face photo of the person with a positive COVID-19 status from the collected pairs. Because many people use social media today, there is a risk that a face photo search engine can be used to identify the positive person's personal information and social media account from the positive person's face photo [8–10].

The threats of linking attacks targeting DCT frameworks have been discussed in previous studies [11–14]. While these studies suggest that linking attacks are possible in principle, they do *not* provide any experimental evaluation nor quantitative assessment of the feasibility or scalability of the attacks. Thus, there is no fundamental data to conduct a proper risk assessment of the threat posed by an attack. In [14], Boutet et al. qualitatively discussed the risks posed by a privacy breach due to a linking attack in GAEN, noting that if the attack is feasible, it could hinder GDPR compliance.

Based on this background, this study addresses the following three research questions:

RQ1: *Is a linking attack targeting the GAEN framework feasible?* (Section 4)

RQ2: *Is the attack scalable?* (Section 5)

RQ3: *What are the effective approaches to mitigating the attack?* (Section 6)

To address these research questions, we conducted extensive field experiments and large-scale simulations targeting the GAEN framework.

The contributions of this work are as follows. First, through extensive field experiments using smartphone devices, we demonstrate that an attacker can establish a linking attack with high accuracy by using a simple attack device that includes a directional antenna. Experiments using smartphones with various BLE transmission power profiles demonstrate that linking attacks are successful even for smartphones with relatively weak transmission power. We also show that attack is successful even when the distance between the target and the attacking device is far (up to 7 m) and when the smartphone is placed in a pocket or a bag.

Second, we demonstrate the scalability of the linking attack through realistic simulation experiments. Our simulation model accurately incorporates the propagation characteristics of the radio waves emitted by directional antennas. The simulation model adopts a 3-dimensional (3D) model of a walking person to evaluate the impact of pedestrian flow, number of devices used by the attacker, and behavior of the pedestrian on the success rate of the linking attack. To model the movement of people in the simulation, we leverage open data that contain records of the trajectories of pedestrians in a city. The experimental results verified that an attacker can achieve a high linking attack success rate of approximately 86% against a high pedestrian flow of 5,000 people per hour by deploying only a few attack devices. Based on the findings of our experiments, we discuss realistic attack scenarios for linking attacks in Section 7.1.

Finally, we propose mechanisms that aim to mitigate the threats posed by the linking attacks. The effectiveness of these countermeasure techniques is clarified by simulation evaluation and field experiments.

2 GAEN Framework

The GAEN framework is a decentralized scheme in which smartphone devices use BLE signals to exchange anonymous identifiers that are updated periodically [15]. GAEN was developed based on the decentralized privacy-preserving proximity tracing (DP-3T) framework [16]. As of February 2022, contact-tracing applications using the GAEN framework were in operation in 38 countries worldwide [17–20].

A brief overview of the scheme for generating anonymous identifiers exchanged in the GAEN framework is shown in Figure 1. First, each client device generates a Temporary Exposure Key (TEK), which is randomly generated every 24 h. The TEK is stored inside the device for 14 days and is not exposed to the public unless

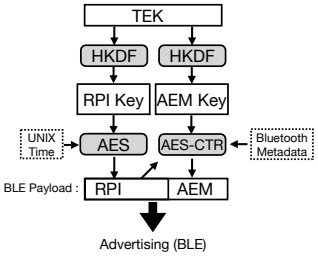


Fig. 1. Derivation of rolling proximity identifiers [21].

the owner of the device registers as a positive person on the health authority's key management server. From the TEK and timestamp, the client derives an Rolling Proximity Identifier (RPI) every 10–20 min.

Additionally, the client derives associated encrypted metadata (AEM) based on TEK, RPI, and Bluetooth metadata and advertises a BLE frame with the RPI and AEM on the payload to its neighbors [21]. The client advertises BLE frames with a period of 200–270 ms [15]. While the client advertises its own RPI and AEM, it also receives the BLE frames advertised by other neighboring client devices. The client receives BLE frames at a period of 4 s every 5 min and stores the payload inside the device [22].

A user who tests positive for COVID-19 uploads a diagnosis key to the health authority's key management server using the GAEN-powered app. The diagnosis keys contain the TEKs for the past 14 days and timestamps corresponding to the times when the TEK was valid. The health authority's key management server disseminates the collected diagnostic keys to GAEN-enabled clients in the country. The client that receives the diagnosis key of a positive person achieves exposure notification by matching the diagnosis key with the RPIs received in the past 14 days.

In this study, we examine the Contact-confirming Application (COCOA), which was developed and operated by the Ministry of Health, Labor, and Welfare (MHLW) in Japan, as a DCT application implemented using the GAEN framework [17]. COCOA incorporates a function that adjusts the TEK to be uploaded based on the date that the user reported infection-related symptoms of COVID-19 [23]. Specifically, TEKs generated from 2-days prior to the date of a positive test for COVID-19 to the date of positive registration on the MHLW server using COCOA were uploaded as the diagnosis key, where the maximum backward period was set to 14 days.

3 Linking Attack

In this section, we present an overview of the linking attack that targets the GAEN framework. We first describe our threat model. We then outline the detailed attack procedure.

3.1 Threat Model

The goal of an attacker performing a linking attack is to associate the anonymous identifiers (i.e., RPIs) advertised by the smartphones running the GAEN framework using BLE with the face photos of the corresponding smartphone owners and to identify the RPIs of users who have declared themselves positive for COVID-19 using the GAEN framework. In other words, an attacker aims to automatically collect images of COVID-19 positive individuals. Given the face photos of COVID-19 positive people, the attacker can search for their face photos using facial image search engines to identify them through various publicly available web resources such as social media [8, 9].

The primary equipment used by the attacker is an off-the-shelf BLE receiver and a camera. The attacker sets up these devices on a street full of pedestrians and continuously collects data. As we see later, the attacker can effectively increase coverage by using multiple devices. An attacker uses a BLE receiver to monitor the frames generated by the GAEN-activated smartphones in their vicinity and extracts the RPIs from the frames. Simultaneously, the attacker takes photos of the pedestrians using a camera. To extract the images of a person from the captured images, an object recognition algorithm is applied.

Supposing that an attacker collects a large number of BLE signals containing RPIs emitted from the smartphones of target users and their images (photos), the collected data would include a timestamp recorded when the data were measured. The attacker downloads the diagnosis keys of the COVID-19 positive person, which can be collected from the health authority's key management server, and derives the TEK, date information, and the corresponding RPIs. By analyzing these data, the attacker identifies the RPIs matching the diagnosis keys of the COVID-19 positive person and successfully extracts their images.

Note that this attack does not emit any radio signals, i.e., the attack is completely passive and cannot be detected by a third party that monitors BLE signals.

3.2 Attack Procedure

The procedure of the linking attack consists of the following two steps: **Step 1**: Linking the RPI with the target's image and **Step 2**: Linking the COVID-19 positive person with the corresponding RPI. By achieving these two linking steps, the attacker can automatically collect images that are eventually linked to COVID-19-positive persons. The two linking steps are outlined below.

Step 1: Linking RPIs and Images of Person

Both the BLE signal and the camera image collected by the attacker are continuous data, with one target arriving after another. The technical challenge faced by the attacker is detecting the exact time when both data were acquired in order to link the RPI contained in the received BLE signal to the image. In other words, by knowing the timing at which the target approaches the BLE receiver and camera, it is possible to accurately extract the RPI contained in the signal and the image of the target person.

The key idea to pinpointing the time when the target approaches the attacking device is to use the information of radio-wave strength. Theoretically, the signal strength decays inversely proportionally to the square of the distance. Therefore, the timing at which the strength of the radio wave carrying the BLE frame containing the RPI reaches its maximum should coincide with the moment when the owner of the smartphone that emitted the RPI is closest to the attacking device. As shown later, such a measurement can be realized with high accuracy by using commercially available off-the-shelf parabolic antennas with high directivity. The time of maximum signal strength is identified by applying a peak detection algorithm to the time-series data of the signal strength.

The timing at which the signal strength of the BLE frame containing the RPI is maximized is determined as follows. The attacker collects the BLE frames and constructs a time series of the signal strength data for each RPI, i: $\mathbb{S}_i = \{S_i(t_1), S_i(t_2), \ldots\}$, where $t_1, t_2, \ldots$ is the time when the BLE signal is measured, and $S_i(t_m)$ is the signal strength of the BLE frame containing the RPI, i, at time t_m. According to the GAEN specification, the RPI advertised by BLE frames is updated every 10–20 min. We assume that the target passes in front of the attack setup once during the period when the RPI takes a certain value. For each RPI, i, the time at which the signal strength reaches its maximum value can be calculated as follows:

$$t_{\max}(i) = \underset{t \in \{t_1, t_2, \ldots\}}{\arg\max}\ \mathbb{S}_i. \tag{1}$$

The target image, corresponding to the RPI, i, is expected to be contained in the image taken at this time, $t_{\max}(i)$. The validity of this assumption is examined in Section 4.2.

Following the assumption made above, the attacker extracts the image taken at the identified time, $t_{\max}(i)$, and applies the object detection algorithm to crop-out the image of the target person. To this end, the attacker can employ an object detection algorithm such as You Only Look Once (YOLO) [24], which is a general-purpose open-source object detection algorithm. Notably, sometimes images of multiple people are acquired at the same time, such as when two or more people walk in tandem. We evaluate such cases in Section 5.

Step 2: Linking RPIs and Positive Person

In the GAEN framework, the diagnosis keys of COVID-19 positive person are distributed by key-management servers operated by the health authorities in each country [25]. The GAEN-based exposure notification app and key management server are deployed in different ways in each country. In some cases, the diagnosis keys can be downloaded by accessing a specific universal resource locator (URL), whereas in other cases (e.g., Northern Ireland and Scotland), a refreshToken/authToken is required to download the diagnosis keys; the token can be extracted from a running app [26]. A survey conducted by the Testing Apps for Contact Tracing (TACT) project [27] shows that as of February 2022, TEKs were accessible on the Internet in at least 21 countries, including Canada, England, Germany, Italy, and Spain, and in four states in the United States. The researchers published their script for collecting TEKs on github [28]. As we mentioned in Section 2, we adopted COCOA, which is implemented based on the GAEN framework and is operated in Japan by MHLW. In addition to the regions listed in [27], we confirmed that the TEKs published by the MHLW are also available for download in other regions. These observations imply that it is easy for an attacker to obtain diagnosis keys corresponding to COVID-19 positive people.

The diagnosis key is encoded by Protocol Buffers (Protobuf), and by decoding the diagnosis key using a tool published by Google, metadata such as the BASE64-encoded TEK and a timestamp corresponding to the date and time the TEK was valid can be obtained [29]. The following is the procedure for deriving the RPI from the diagnosis keys.

Let Tk be a TEK extracted from a diagnosis key, and let Rk be the RPI key corresponding to Tk. Using the Hashed Message Authentication Code (HMAC)-based extract-and-expand key

derivation function (HKDF), Rk is computed as $Rk = \texttt{HKDF}(Tk, \texttt{NULL}, str, 16)$, where the arguments are `HKDF(Key,Salt,Info,OutputLength)` as defined in RFC5869 [30], and str is a string "EN-RPIK" encoded by UTF-8. Next, we derive the RPIs. I_t, from Rk, where t is the Unix epoch time at the moment a roll occurs. By using Rk as a key and applying the 128-bit Advanced Encryption Standard (AES) encryption algorithm to the data D_t, we obtain the rolling proximity identifier as $I_t = \texttt{AES}(Rk, D_t)$, where D_t is the data obtained by adding the rolling time epoch, t. The attacker can identify the RPIs of the COVID-19 positive person and the corresponding images by comparing the RPIs derived according to the procedure shown above with the collected RPI.

We collected the TEKs of positive persons from the MHLW's server, as described above, and confirmed that it is possible to derive the RPI generated at a specific time from the obtained TEKs. By matching the derived RPI of a positive person with the RPI collected in Step 1, the attacker can obtain the face photo of the person linked to the RPI.

4 Feasibility of the Linking Attack

In this section, we address **RQ1**: "*Is a linking attack targeting the GAEN framework feasible?*." Specifically, we focus on step 1 of the attack described in Section 3.2: linking RPIs with the images of target users. We clarify the feasibility of the linking attack through field experiments conducted under various conditions. We note that as mentioned above, it is obvious that an attacker can link the RPIs derived from the TEKs of the positives and the RPIs recorded in Step 1; thus, we can safely skip the verification of Step 2.

In the following, we clarify the feasibility of a linking attack through extensive field experiments. First, we demonstrate the feasibility of the attack through field experiments (Sec 4.2). Next, we verify the robustness of the attack by changing conditions such as the distance between the attacker and target, the type of smartphone, and the manner in which the smartphone is held (Sec 4.3.)

4.1 Experimental Setup

We first describe the equipment used in the experiments. The equipment used by the attacker is a personal computer, a Bluetooth receiver, an antenna, and two universal serial-bus (USB) cameras (main and sub). Table 1 lists the equipment used in the experiment. Ubertooth One is a USB dongle-type Bluetooth receiver that can stably receive and record BLE frames. As an antenna connected to Ubertooth One, we used a directional parabolic antenna compatible with the 2.4-GHz band. For reference, we show the evaluation results when using an omnidirectional antenna in Section 7.2. As a USB camera, we used a device with a 120° viewing angle and a resolution of $1,920 \times 1,080$ pixels. All equipment is inexpensive and can be purchased from online shopping sites, such as Amazon.

Table 1. Attacker's setup.

Equipment	Model
Computer	Apple Macbook Pro 2021 (macOS 12.0)
BLE Receiver	Ubertooth One (firmware: 2020-12-R1) [31]
Antenna	ANT-GRID-24dBi [32]
USB Camera	BUFFALO BSW505MBK

Table 2. Target's smartphones. BLE transmission power (dBm) and antenna gain (dBi) were obtained from Refs [33–37].

Vendor	OS	Model	Power	Gain
Apple	iOS 14.6	iPhone 8	20	-0.44
Apple	iOS 14.6	iPhone XR	16	-4.9
LG	Android 10	G8X ThinQ	4.65	-5.03
Huawei	Android 9	P20 Lite	5.36	n/a
Motorola	Android 7.1.1	Nexus 6	6.57	-3.00

We used the five devices listed in Table 2 as the smartphones used by the target users. All had COCOA 1.2.4, which is a DCT application based on the GAEN framework operated in Japan [17]. As we will see, these devices differ in the strength of the BLE signal transmission. There is also a slight difference in the implementation of the GAEN between the iPhone and Android.

The experimental setup is shown in Figure 2. The left side of the figure shows a schematic diagram of the setup, and the right side shows a photograph of the actual setup. Here, we assume that the target person crosses the area captured by the two cameras. A parabolic antenna was pointed at that area to collect the BLE frames emitted from the target's smartphone. The reason for using two cameras is as follows. A camera was placed at the same position as the antenna and was used as the main camera. The camera captures the target in the middle of the captured image at the moment when the BLE signal strength is at its maximum. The goal is to link the RPI and the image of a person with high accuracy. However, because the main camera can only capture the side view of a target person, it may

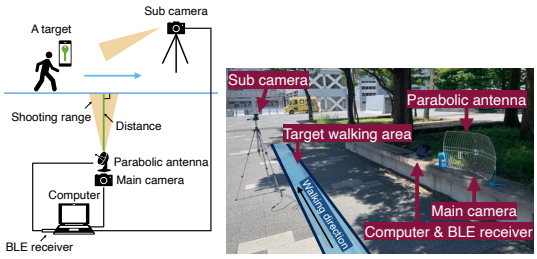


Fig. 2. Overview of the attacker's setup.

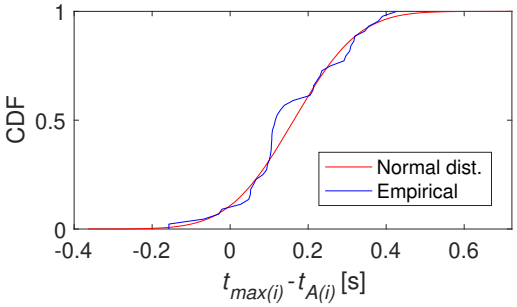


Fig. 3. Empirical distribution of $t_{\max}(i) - t_A(i)$.

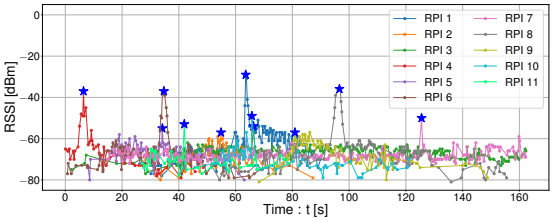


Fig. 4. Time series of BLE signal strengths for observed RPIs. The star symbols correspond to $t_{\max}(i)$ at which the signal for each RPI, i, is maximum.

be difficult to identify the individual. Therefore, we also collected images of the target taken from the front (see Fig. 6 for example shots). The sub-camera was used to achieve the shot. The images taken by the sub-camera are likely to include more than one person in the image because pedestrians walking in front of and behind the target are also captured simultaneously. In such a case, the frontal image of the person captured by the sub-camera can be identified based on the information of the person identified by the main camera.

In our experiments, the two cameras and parabolic antenna were visible to pedestrians; however, in an actual attack, these devices will be hidden from view. Both the camera and the parabolic antenna are connected to a monitoring computer, which continuously records the BLE signals and image data. The attacker leverages these data to establish a linking attack.

4.2 Feasibility of the Attack

We empirically verify that the linking attack (step 1) shown in Section 3 is feasible. The key to the success of the linking attack is the simultaneous acquisition of the RPI data contained in the BLE frames generated by the target person's device and the images taken by the person. The objective of our approach is to identify when the target is closest to the setup using a parabolic antenna/camera, as shown in Figure 2. Doing so, we assume that the target arrives in front of the camera at the moment the signal strength of the BLE frame containing the RPI generated by the target's device is highest and that the target is captured in the image taken at that moment. In this experiment, we used iPhone XR as the target smartphone.

In reality, the time when the signal strength reaches its maximum ($t_{\max}(i)$ defined in Eq 1) and the time when the target arrives in front of the camera (let the time be $t_A(i)$) may differ, owing to factors such of radio-wave reflection and movement of the handheld smartphone. Therefore, we conducted an experiment to measure the time difference between $t_{\max}(i)$ and $t_A(i)$. In this experiment, the target placed the smartphone in a pocket and passed in front of the attacking device. The distance between the target and the antenna and the camera was set to 2 m. The number of trials was set to 45. There was one obvious outlier; therefore, we excluded it. The empirical distribution of the time difference is shown in Figure 3. The mean value of the time difference was 0.162 s, and the standard deviation was 0.130. The results demonstrate that the timing when the signal strength of the target is at its maximum and the time when the target passes in front of the camera are very close, indicating that the timing-based linking attack is feasible. We can also see that the distribution can be approximated by a normal curve. The parameters of the normal distribution in the figure were fitted using the maximum likelihood estimation.

Figure 4 shows the signal strength of the BLE frame containing all received RPIs and the time at which each signal strength takes its maximum value with star symbols. In the figure, we can observe 11 unique RPIs and the corresponding times at which the BLE signal strength for each RPI reaches its maximum value. For privacy reasons, the RPIs were stored as hash values during data collection.

Figure 5 plots the time series corresponding to `RPI 1` shown in Fig. 4. BLE frames containing `RPI 1` were generated by a smartphone owned by one of the

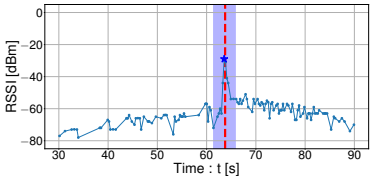


Fig. 5. Time series of BLE signal strength for the RPI corresponding to a target person (Fig. 6). The distance between the antenna and the target was set to 2 m.

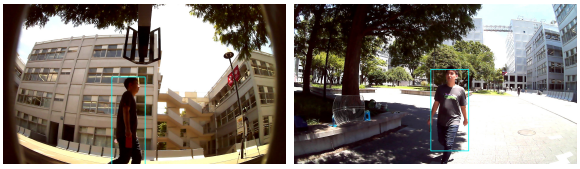
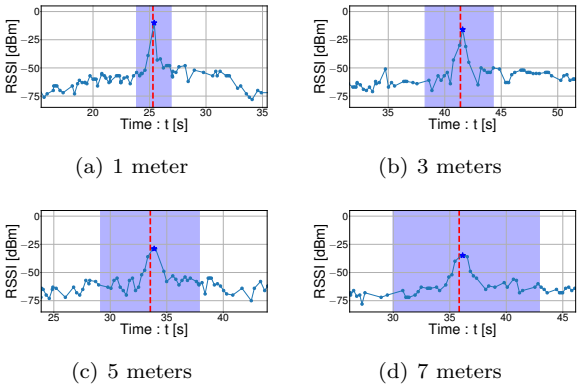


Fig. 6. Captured images corresponding to $t_{\max}(i) = 63.56$ s. YOLO v3 was applied to detect a person in the images, as shown by the rectangular frames colored in light green.



(a) 1 meter (b) 3 meters

(c) 5 meters (d) 7 meters

Fig. 7. Time variation of BLE signal strength for various distances.

authors who conducted the experiment. We can see that the BLE signal strength has a maximum value with a sharp peak at $t_{\max}(i) = 63.56$ s. In the figure, the light purple area indicates the time from when the target entered the camera's field of view to when it exited the frame, and the red dotted line indicates the time when the target was directly in front of the camera. It is clear that the target was fully included in the image taken at time $t_{\max}(i)$, as indicated by the stars.

Figure 6 shows the images corresponding to the BLE frame captured at $t_{\max}(i) = 63.56$ s, which is shown with the star symbol in Fig. 5. Although the two cameras continuously captured shots, the timings deviated slightly from $t_{\max}(i)$; hence, we adopted the images captured at the closest timings; the corresponding image frames were captured at 63.56 s (main camera, left) and 63.53 s (sub-camera, right), respectively. We can see that the target was correctly captured in both cases and that YOLOv3 successfully detected the image of the person. Furthermore, we confirmed that it is possible to extract facial parts from detected pedestrian images by applying YOLOv5 [38] and a trained face recognition model [39]. These experimental results clearly demonstrate that a linking attack that links the targets' RPIs to their face photos is feasible.

4.3 Robustness of the Attack

To examine the robustness of the attack, we experimentally clarified the effects of the distance between the target and attack device, the type of smartphone (i.e., BLE TX power), and the manner in which the target holds the smartphone for a successful attack.

4.3.1 Impact of the Distance

We examined the impact of the distance between the target smartphone and the parabolic antenna that measures the BLE signal on the attack. The maximum distance was set to 7 m to match the typical sidewalk width [40]. In the following experiment, the target walks in front of the attack device while holding the smartphone in their hand on the antenna side. We adopted iPhone XR as the target's smartphone.

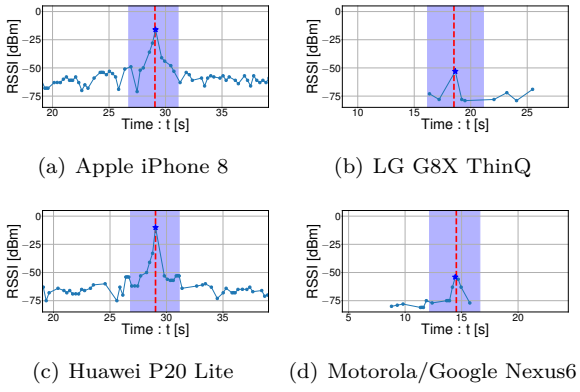
Figure 7 shows the time evolution of the radio signal strength of the BLE frames containing the RPI of the target user when the distances are set to 1, 3, 5, and 7 m. As the distance increases, the time region indicated by the light purple area (i.e., the time when the target is included in the image) becomes longer. This observation corresponds to the fact that the range of the target in the picture increases with the distance. On the other hand, regardless of distance, the time indicated by the red dotted line (i.e., the time when the target came directly in front of the camera) and the time indicated by the star symbol (i.e., the time when the signal strength reached its maximum) were extremely close. Thus, the image taken at the time when the BLE signal strength is at its maximum has the target in the middle of the image, even if the image is a wide angle and contains multiple people, the target contained in the image can be identified with high accuracy. These results demonstrate that linking attacks are feasible in the range of 1 to 7 m, which covers the width of a typical sidewalk.

4.3.2 Impact of Smartphone Models

We examined the impact of different smartphone models in terms of the transmission power strength of the BLE signal and the gain of the antenna on the success of an attack. As shown in Table 2, different smartphone models have different BLE transmission powers and antenna gains. In fact, the GAEN framework calibrates the values of the transmission power to account for the differences between devices. The devices used in our experiments cover a variety of transmission power and antenna gain profiles.

In our experiments, a pedestrian passed in front of the attacking device while holding a smartphone in a hand. The distance at which the pedestrian passed in front of the attacking device was fixed at 2 m. [1]

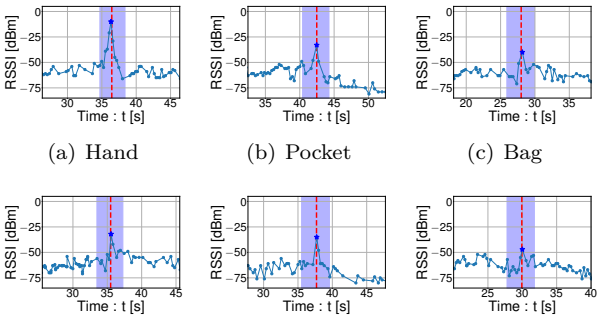
The results are shown in Fig. 8. We can see that in the case of all four devices, the attack was successful; the time when the BLE signal strength was at its maximum and the time when the target arrived directly in front of the main camera nearly coincided. In the figure, we can also see that only a part of the signal was observed for G8X ThinQ and the Nexus 6. This result reflects the fact that as the transmission strength of the two devices was weak, and the gain of the transmitting antenna was not high enough; hence, no signal was observed until they were close to the target. The devices with weak transmission strength show a clear signal peak when they pass in front of the antenna. These results clearly demonstrate that our attack is successful even for devices with weak transmission power.



(a) Apple iPhone 8 (b) LG G8X ThinQ

(c) Huawei P20 Lite (d) Motorola/Google Nexus6

Fig. 8. Time variations of BLE signal strength for four smartphone devices, each having different BLE transmission power. The distance was fixed at 2 m.



(a) Hand (b) Pocket (c) Bag

(d) Hand (e) Pocket (f) Bag

Fig. 9. Time variation of BLE signal strength for six types of smartphone carrying patterns (Top : Attacker side, Bottom : Opposite side).

4.3.3 Impact of How the Smartphone is Held

Finally, we examined the effect of the way the smartphone was held on the success of the attack. We focused on cases wherein the target held the smartphone in a hand, in a pocket, or in a bag. Then, we categorized the cases according to whether the smartphone was on the attacking device side or the other side. For six cases, as in the previous experiments, we measured the temporal variation of the signal strength of the BLE signal containing the RPI data. Here, the distance when a pedestrian passed in front of the attacking device was fixed at 2 m, and the iPhone XR was used as the device held by the target.

The results are shown in Fig. 9. As expected, the signal strength was highest when the device was held in the hand, followed by when it was in the pocket, and lowest when it was in the bag. Additionally, the signal strength was weaker when the smartphone was on the opposite side of the attack device. Despite these differences, we can see that the linking attack worked in all cases. In particular, when the device was placed in a bag, the variation in the signal strength decreased, but the parabolic antenna accurately detected the time when the signal strength was at its maximum.

5 Scalability of the Attack

In this section, we perform simulation experiments to address **RQ2**: "*Is the attack scalable?*" The field ex-

1 We omit the experiments with other conditions because of space limitations. The findings obtained from experiments conducted under other conditions were similar.

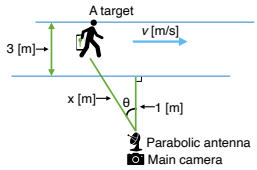


Fig. 10. Simulation model.

Fig. 11. A screenshot of the 3D model.

periment presented in Section 4 employed a scenario in which a single pedestrian passes in front of the attacking device[2]. In reality, it is expected that a large number of pedestrians will continuously pass in front of an attack device. Implementing such a large-scale situation in a field experiment is expensive in terms of preparing devices, recruiting experiment participants, and obtaining permits to use the streets. In addition, it is not easy to change parameters such as the number of pedestrians and street width. Therefore, we employed simulations that realistically model radio propagation and pedestrian movement.

In our experiments, we first evaluated the attack success rate when the number of pedestrians increases (Sec 5.2) and then showed that the success rate can be further improved by increasing the number of attack devices (Sec 5.3). Then, we evaluated the success rate of the attack, particularly on congested streets, using a simulation model and real-world pedestrian data (Sec 5.4). Based on the results obtained in this section, we discuss realistic attack scenarios in Section 7.1.

5.1 Simulation Model

This section provides a description of the simulation model. First, we describe the overview of the entire simulation. We then describe a model that governs the dynamics of pedestrians to be simulated. We also describe the propagation model of BLE radio waves generated by the smartphones carried by pedestrians. Finally, we present the criteria for determining a successful attack, and define the attack success rate.

Overview of the Simulation Model: The components of the simulation model are shown in Figure 10. As shown in the figure, the width of the street was set to 3 m[3] An attacker installs an attack device 1 m in front of the edge of the street. Each target person is assigned a sequential ID as a pseudo RPI. The target person moves in a linear motion with a constant velocity of v [m/s] for both directions, i.e., from the left/right to the right/left in the figure. The velocity, v, is assumed to follow a normal distribution, as shown later. For each direction, we modeled the arrival process of the target using the Poisson process. The trajectory of the target is represented using 3D modeling, which is built with Panda3D [41], which is an open-source engine for real-time 3D games, visualizations, simulations, and experiments. Figure 11 presents an example shot in which five persons were captured in an image. Thus, employing 3D modeling enabled us to reproduce the images captured by the camera.

Let θ be the angle between the target and the parabolic antenna, as shown in Fig. 10. For simplicity, we used a two-dimensional projection onto the ground to calculate the angle between the attacker and target. This assumption is reasonable considering that the variation in the height at which each pedestrian is holding the smartphone is sufficiently small compared with the distance between the target and attacker. The directivity property of the parabolic antenna used in our experiment [42] can be leveraged to derive the relationship between the angle and the absolute gain of the receiving antenna, as shown in Fig. 12. The concrete method of calculating the received signal strength is described later. Based on the time variation of the radio signal strength calculated for each target, we identified the time when the strength reached its maximum value. We then analyzed the image taken at the detected time, $t_{\max}(i)$. We noted that the image can be captured by 3D modeling. We then linked the image of the person having the smallest θ in the image with the RPI corresponding to the received radio wave. The success rate of the attack as calculated by comparing the results of the link with the ground truth.

The common parameter settings throughout the simulation are summarized in Table 3.

Modeling the Dynamics of Pedestrians: Pedestrians were assumed to arrive randomly according to the Poisson process, where the inter-arrival time of the pedestrians was exponentially distributed. In this study, the parameter of the exponential distribution, λ, was set as $\lambda = N/3,600$, where N is the total number of per-

2 For reference, we demonstrate that the attack is also valid when two consecutive pedestrians pass in front of the attacking device within a short period of time in Appendix A.

3 Increasing the width of the street did not significantly affect the results (see Appendix C).

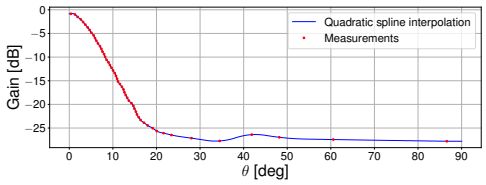


Fig. 12. Gain map of the parabolic antenna (ANT-GRID-24dBi) [42]

Table 3. Common parameters used in the simulation.

	Description	Value
G_t	Absolute gain of the target's BLE transmission antenna	-4.9 [dBi]
G_r	Maximum absolute gain of the receiving a parabolic antenna	23 [dBi]
P_t	Target's BLE transmission power	16 [dBm]
λ	Radio wavelength of BLE (2.4 GHz)	0.125 [m]
τ	Time period of BLE frame generation	0.270 [s]
$x_{\min}$	Minimum distance between target and attack device	1 [m]

sons arriving in 1 h; the unit of the arrival rate, λ, is the number of arrivals per second. N is a parameter that will be explored later. The walking speed of individual pedestrians was given by the normal distribution with a mean of 1.30 m/s and standard deviation of 0.22, based on the measurement results reported in Ref. [43].

We assumed that the position of pedestrians on the street, the lateral position relative to the direction of movement, follows a normal distribution centered in the middle of the street. Let the width of the road be $w = 4\sigma$, where σ is the standard deviation, i.e., 95% of the pedestrians will be distributed within the width of the street w. If the random variable representing the position exceeds 2σ, we assumed that they walk along the street edge. In the simulation, we assumed that pedestrians walk in a straight line without swinging to the left or right once their position on the street is determined.

Modeling the Strength of Received BLE Signals: The BLE radio-signal strength received by the attacker's parabolic antenna is computed using Friis transmission equation [44]. Let the absolute gain of the BLE transmit antenna be G_t [dBi], the absolute gain of the parabolic antenna be G_r [dBi], the BLE transmit power be P_t [dBm], the distance between the target and the antenna be d [m], and the radio wavelength be λ [m]. Note that the numerical values set for G_t, G_r, P_t, and λ are listed in Table 3. The received power, P_r [dBm], at the parabolic antenna is a function of d and θ and is

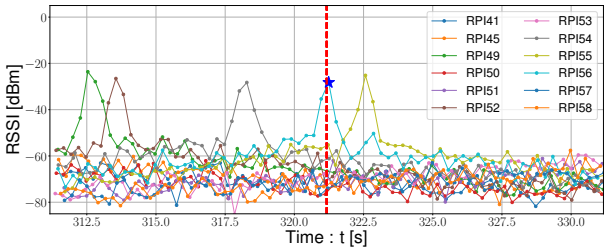


Fig. 13. Simulated BLE signal strengths.

calculated as follows:

$$P_r(d,\theta) = 10\log_{10}\left(\frac{\lambda}{4\pi d}\right)^2 + G_t + G_r + P_t, \quad (2)$$

where $G_r(\theta)$ is given by the gain map of the parabolic antenna as shown in Fig. 12. Note that the units of gain and radio strength are expressed in decibels, not watts; hence, in logarithmic calculations, product/quotient is expressed as addition/subtraction.

Although the above equation provides the theoretical received signal strength, in real spaces, there are many disturbing factors, such as urban noise, pedestrian movement, and transmission losses. Therefore, this study introduced additive white Gaussian noise, loss corrections, and the attenuation factor to represent such factors, incorporating them into the simulation using the eq.: $P_r^*(d,\theta) = P_r(d,\theta) + \epsilon + \eta + \kappa$, where we empirically set the additive Gaussian noise as $\epsilon \sim \mathcal{N}(0, 3^2)$ and the loss correction term as $\eta = -12$ (See Appendix B for the details). Following the empirical results reported in Ref. [45], we set the attenuation factor to be $\kappa = -15$ [dBm] when another pedestrian overlaps between the antenna and the target smartphone.

Figure 13 presents an example of the strengths of the simulated BLE signals received by the parabolic antenna.

Definition of Attack Success: The following is the definition of attack success. As mentioned earlier, an attacker measures the BLE signal strength for each RPI and links the RPI with the face photo of the person most centrally located in the image taken when the signal strength reaches its maximum value. After the completion of the simulation, we define the attack as a perfect success if there is a one-to-one correspondence between the linked results and the ground truth. If, after the simulation, multiple RPIs are linked to the target and one of them matches the ground truth, the attack is also considered successful but is evaluated separately from a perfect success. If no correct RPI is associated with the target, the attack is defined as a failure. Based on the above definitions, we define the (perfect) attack suc-

cess rate as the percentage of pedestrians who passed in front of the attacking device and had a (perfect) successful attack.

5.2 Number of Pedestrians (Targets)

We first study the impact of the arrival rate of pedestrians on the success of the linking attack. As the arrival rate increases, the distance between pedestrians decreases. For example, when $N = 5,000$, the arrival rate becomes $\lambda = 5,000/3,600$ [arrivals/s], and the distance between pedestrians follows an exponential distribution with a mean of 0.94 m. When the distance between pedestrians is small, it becomes difficult to uniquely identify the exact time when the signal strength is at its maximum because of the overlap of BLE signals. Additionally, there will be multiple people in the captured image Thus, the attack success rate will become low.

We performed a simulation experiment to evaluate how the success rate of linking attacks changes when the number of pedestrians per hour, N, is increased. We varied N from 200 to 10,000. The results are shown in Table 4, where the simulation was performed three times under the same conditions with different random-number seeds, and the average of the results is shown. Let the variable m be the number of RPI candidates linked to the target person. $m = 1$ indicates perfect attack success, i.e., a target person's face photo is linked to a specific RPI. $m = 2$ indicates that the linking attack has succeeded in narrowing down the number of RPIs associated with a target person's face photo to two candidates. Note that the sum of each line does not equal 100%, because cases with incorrect links are not counted. From the table, we can see that when $N = 1,000$, the linking attack is perfectly successful for 78% of the pedestrians. We can also see that when $N = 5,000$, the percentage of perfect attack success is 38% and the percentage of narrowing down the number of candidate RPIs to two or less is 57%.

On the other hand, as mentioned, when $N = 5,000$, the average distance between pedestrians is roughly 1 m, implying that the street is congested as if there were a protest demonstration going on. It is noteworthy that even in such cases, the linking attack can be successful (1.9 K targets for the perfect attack success). As we see in the next subsection, we can also increase the attack success probability by increasing the number of attack devices. We also note that the number of successful attacks can be increased by extending the time required to conduct the linking attack. In our simulation experiments, the simulation time was set to $T = 1,200$ s, which could be further extended to an actual attack.

Table 4. N (# pedestrians/h) vs. attack success rate (%). The values in parentheses are standard deviations.

N	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m \geq 5$
200	97.9 (2.9)	1.0 (1.5)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
400	92.6 (0.5)	3.6 (0.4)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
800	83.3 (2.2)	7.0 (1.3)	0.3 (0.2)	0.0 (0.0)	0.0 (0.0)
1K	78.4 (2.5)	8.8 (1.4)	0.7 (0.5)	0.0 (0.0)	0.0 (0.0)
2K	66.7 (1.5)	13.3 (0.8)	0.8 (0.4)	0.0 (0.0)	0.0 (0.0)
3K	54.3 (1.2)	16.6 (0.6)	1.6 (0.2)	0.0 (0.0)	0.0 (0.0)
5K	38.2 (1.9)	18.5 (0.4)	2.8 (0.3)	0.4 (0.1)	0.0 (0.0)
10K	21.4 (0.5)	15.5 (0.1)	5.0 (0.2)	0.8 (0.3)	0.2 (0.0)

5.3 Number of Attack Devices

In Section 5.2, we evaluated the attack success rate when an attacker uses a single device. The results indicated that the attack success rate decreased as the number of pedestrians increased. The reason for the decrease in the attack success rate is that, as the number of pedestrians increases, more pedestrians walk closer to each other, resulting in multiple pedestrians passing in front of the attack device at almost the same time. We expect that an attacker can increase the attack success rate by installing an attack device at multiple locations. In other words, by increasing the number of attacking devices, the attacker increases the number of attack trials, resulting in multiple linking results. The attack success rate is expected to increase, based on the majority vote of multiple linking results.

Given these observations, we evaluated the extent to which the success of the linking attack increased when we increased the number of attack devices. The attack devices were placed at equal intervals of 100 m along the straight line that the pedestrians followed. We studied the attack success rate when the number of attack devices, L, increased. The simulation time is set to $T = 1,200$ s. The results are shown in Table 5, where the number of pedestrians per hour was set to $N = 5,000$.

As the number of attack devices increased, the attack success rate increased. When there was only one attack device, the perfect attack success rate was 41%, whereas by setting the number of attack devices to $L \geq 4$, the perfect attack success rate can be increased to 86%. Thus, by increasing the number of attack devices, the number of successfully extracted RPI-image pairs increases. If an attacker keeps collecting data on this street for 12 h during a day, over a week, they can ob-

Table 5. L (# attack devices) vs. attack success rate (%). The values in parentheses are standard deviations.

L	$m=1$	$m=2$	$m=3$	$m=4$	$m\geq 5$
1	38.2 (1.9)	18.5 (0.4)	2.8 (0.3)	0.4 (0.1)	0.0 (0.0)
2	59.7 (0.8)	15.2 (0.3)	5.6 (0.6)	1.2 (0.1)	0.2 (0.0)
3	75.2 (2.2)	9.1 (0.4)	4.3 (0.5)	1.6 (0.0)	0.6 (0.2)
4	85.7 (0.3)	4.7 (0.5)	3.0 (0.4)	1.3 (0.1)	0.7 (0.1)
5	90.1 (0.8)	3.0 (0.3)	1.9 (0.4)	1.1 (0.2)	0.5 (0.2)
6	93.0 (0.8)	2.1 (0.5)	1.5 (0.2)	0.9 (0.2)	0.5 (0.3)

tain the RPI-image pairs for 360 K corresponding people. Of course, the same person could be included in the data because of the regular pattern of human activities. Still, we can expect that a few sets of attack devices can collect data in the order of 10^5. Based on these findings, we discuss realistic attack scenarios in Section 7.1

Additional Results: For reference, the results of changing the width of a street are shown in Appendix C.

5.4 Attack Evaluation on a Wide and Congested Street

In this section, we evaluate the attack success rate, particularly on a wide, congested street. An attacker can carry out a linking attack targeting a large number of people in a short period by collecting data on a crowded street where a large number of pedestrians are moving. In addition to the theoretical pedestrian arrival model, we evaluate the attack success rate based on realistic pedestrian arrival data measured on a wide crowded street. We extend the theoretical model by introducing the batch arrival Markov process because some people arrive in groups on a wide street.

Setup (Real Pedestrian Arrival Data): As real pedestrian arrival data, we adopted the open data published by Japan's Ministry of Land, Infrastructure, Transport and Tourism [46]. These data consist of the coordinates of people measured every second on a 4.8 meter wide street located underneath Tokyo Station. Each pedestrian is assigned a unique ID. Figure 14 illustrates the number of pedestrians measured at 0:00–23:59 on February 12, 2021. In this study, we adopted pedestrian data measured at 12:00–13:00, which was the most congested time, to simulate human traffic flow on a congested street. In the simulation, the average walking speed was calculated from the recorded coordinates of each person, and the person was assumed to move at that speed. In addition, the pedestrian's lateral position

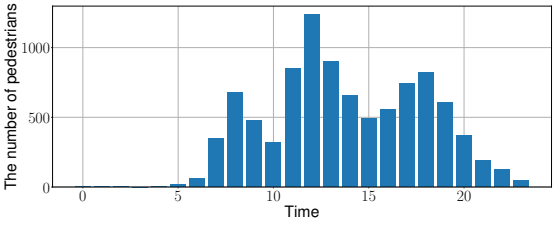


Fig. 14. Number of observed pedestrians in a day [46]

Table 6. Attack success rate (%) for the real-world pedestrian simulation vs theoretical pedestrian simulation. The values in parentheses are standard deviations.

Model Type	$m=1$	$m=2$	$m=3$	$m=4$	$m\geq 5$
Real data	66.8 (0.5)	11.6 (0.3)	1.3 (0.3)	0.2 (0.1)	0.1 (0.0)
Theoretical	53.9 (1.5)	14.7 (0.5)	3.0 (0.6)	0.8 (0.2)	0.2 (0.1)

relative to the direction of movement on the street was maintained in its initial state.

Setup (Theoretical Model): We describe the group arrival model adopted in the theoretical model. Generally, most pedestrians may walk alone, but on a wide street, some may walk in groups of two, three, or more. In this study, following the statistics reported in Ref. [43], we assumed that 66% of the population walks alone, 25% of the population walks in groups of two, 6% of the population walks in groups of three, and 3% of the population walks in groups of four or more. The simulation was performed assuming that each group arrived randomly according to the Poisson process. In a group, people are assumed to walk at equal intervals of 0.5 m in the depth direction.

Simulation Results: Using the setups shown above, we performed simulation experiments. The simulation based on the theoretical model was performed using the number of pedestrians observed in the real data ($N = 1,236$) and the measurement time ($T = 3,600$ s). The results are shown in Table 6. We can see that the probabilities of a completely successful attack are 67% (real data) and 54% (theoretical model), and the probabilities of successfully narrowing down to two candidates are 78% (real data) and 69% (theoretical model). These results show that the evaluation using the theoretical model provides more stringent conditions than the evaluation using real data. The simulation results using real data reveal that the probability of a completely successful attack is approximately 70%, which supports the practicality of the attack on a crowded street. A discussion of the results obtained from the simulations presented above is provided in Section 7.1.

6 Countermeasures

In this section, we address **RQ3**: "*What are the effective approaches to mitigating the attack?*" We discuss potential countermeasures against linking attacks to GAEN framework and derive several recommendations for stakeholders, such as developers of the GAEN framework and operating-system vendors.

6.1 Interacted-RPI

We present a new procedure to match the diagnosis key with the RPIs, which uses multiple consecutive RPIs. In the proposed procedure, the attacker must observe the target's RPIs continuously for at least longer than the RPI-update period, which consequently makes the linking attack more difficult. This countermeasure makes it impossible for an attacker to narrow down the Top-K positive candidates without observing two distinct I-RPIs in consecutive time windows. In the current GAEN implementation, RPI is uniquely calculated from TEK and a current timestamp and is advertised on a BLE frame. On the other hand, our proposed procedure uses the interacted-RPIs (I-RPIs) as a new identifier calculated by multiple consecutive RPIs. The current I-RPI is calculated using the exclusive-OR (XOR) of the current RPI and the previous I-RPI. To implement this procedure, the key management server does not require any changes. It can be implemented by simply updating the GAEN framework.

In the current GAEN specification, the value of the RPI varies every 10–20 min. Under this update frequency, for an attacker to observe two consecutive I-RPIs, the attacker needs to keep observing the I-RPIs sent by the target for at least 15/2=7.5 min on average, and at most 20 min. Therefore, the attacker must place the devices at an average distance of approximately 586 m and a maximum distance of 1562 m on the walking path of the possible target, as the average walking speed of pedestrians is 78.09 m/min [43].

Generally, pedestrians have a wide variety of walking paths, which makes this attack more difficult.

When I-RPI is introduced, the changes in the procedure on the receiver and sender sides are as follows: The sender device sends I-RPIs instead of RPIs and stores the previous I-RPI, which is required to calculate the current I-RPI. The receiver calculates the XOR values of all combinations from the RPIs (calculated from the diagnosis key) and I-RPIs (received from senders' devices) and compares whether the XOR-ed values match with I-RPIs when identifying close contacts once a day.

Performance Evaluation: In the following section, we evaluate the impact of introducing the I-RPI mechanism on the data processing time through empirical experiments. We implemented the matching process adopted by the current GAEN and the matching process after the introduction of I-RPI on two Android smartphones and compared the processing times. We used LG G8X ThinQ as an example of a high-end device and Huawei P20 Lite as an example of a low-end device. For the implementation, we used the source code of GAEN [47].

Based on the statistics, the number of TEKs to be matched was set to $N = 1,192,452$, and the number of identifiers to be received in one time window (15 min) was set to $c = 118$. Details of the parameter derivation are provided in Appendix D. In each condition, the time required for the matching process was measured ten times, and the average value was calculated. As a result of the experiments using high-end/low-end smartphones, the average matching time in the current GAEN was 14.9/50.2 s, and the average matching time after the introduction of I-RPI was 16.8/59.5 s, respectively. With the introduction of I-RPI, the increase in the computation time required for the matching process was approximately 10–20%, indicating that the overhead was within an acceptable range.

6.2 Intermittent Signal Transmissions

We propose a method that thwarts an attack by increasing the signal transmission interval. In GAEN, RPIs are advertised by transmitting BLE frames at short intervals of 200–270 ms [15], which enables an attacker to obtain a signal strength graph with a high resolution to accurately detect the peak. By increasing the signal transmission interval, the resolution of the signal strength graph can be reduced, making peak detection infeasible. However, if we increase the interval between beacon transmissions, there is a risk that the RPI will not be received because of frame errors. To solve this problem, we adopted the approach of intermittent burst transmission; i.e., for every time period (T [s]), we transmit m consecutive frames with an interval of τ s. For example, the parameters are defined as $T = 30$ s, $\tau = 100$ ms, and $m = 30$.

We ran a simulation to verify the effectiveness of the proposed countermeasure – increasing the signal transmission interval. For this simulation, we set $N = 800$ pedestrians/h. An attacker uses one attack device. Fig-

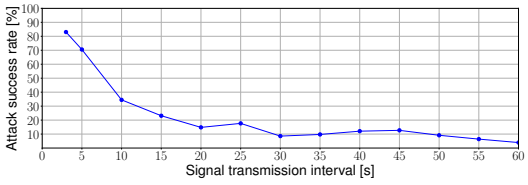


Fig. 15. Signal transmission interval–attack success rate

ure 15 presents the simulation results. We can see that by increasing the transmission interval, the attack success rate decreases. For example, if we set the BLE frame transmission interval to 30 s, the attack success rate reduces from 83% to 9%. However, caution must be exercised because an excessively long transmission interval will result in a no contact confirmation.

To implement this countermeasure, we need to tune the setting of the BLE receiver. In the GAEN framework, a client scans the BLE frames for four seconds every five minutes. When the intermittent transmission scheme is adopted, the client may miss BLE frames if the scanning window is only four seconds in the five minutes interval. To overcome this problem, the client must scan the frames at least T seconds every five minutes to ensure that the advertised BLE frames are received. In this way, the most fundamental functionality of the GAEN framework is maintained.

We note that currently, the setting of signal transmission intervals of BLE is defined by the smartphone operating system and it cannot be changed arbitrarily by users; iOS does not provide a way to change the setting, and Android provides a way to change it, but only with limited values (e.g., 100, 250, and 1,000 ms). With the above constraints, we implemented a method to adjust the signal transmission period and confirmed the effectiveness of the countermeasures. The details are provided in Appendix E.

7 Discussion

In this section, we discuss realistic attack scenarios and ethical considerations.

7.1 Realistic Attack Scenarios

In the following, we discuss realistic scenarios for linking attacks targeting GAEN users based on the findings obtained through the experiments presented in Section 5.

Scenario 1 (Random mass attack): As shown in Section 5.3, an attacker can obtain 360,000 pairs of identifiers and face photos by running a few attack devices for seven days. Based on an estimate of the population of Tokyo and the number of people with COVID-19 as of January and February 2022 [48, 49], this result means that an attacker can collect face photos of approximately 8,600 people with COVID-19. Thus, it is possible for an attacker to conduct a linking attack on a large number of random targets. The motivation for an attacker to conduct a linking attack against random targets is to satisfy their own curiosity or harass someone else. For example, an attacker might use the information obtained from a linking attack to commit so-called "doxing" [50], the act of publishing a private information of a victim on the Internet. In this scenario, we expect that an attacker can increase the attack success rate by performing an attack in the vicinity of a PCR testing facility or hospital.

Scenario 2 (Phishing attack): An attacker can leverage information obtained from a linking attack to conduct a phishing attack. For example, suppose an attacker targets a specific company and obtains information on employees who tested positive for COVID-19. As mentioned earlier, an attacker can use a face photo search engine to identify the personal information and social account of an employee with a positive COVID-19 test result using photos of their faces [8, 9]. Because employees who test positive are likely to be absent, an attacker can conduct social engineering or phishing attacks by impersonating those employees. If a specific individual is found to have tested positive, a phishing attack can be conducted by sending an email or social networking message to that individual, impersonating a health authority. Today, people are forced to make significant changes in their behavior based on the results of COVID-19 tests. Attackers can stealthily collect information regarding people's health statuses through a linking attack.

7.2 Impact of Antenna Design

Linking performance can be improved via careful selection and orientation of the antenna. In this section, we first show that linking performance decreases when an omnidirectional antenna is selected. Next, we show that the performance is retained even when the directional antenna is placed at an angle to the pedestrians.

Impact of Antenna Characteristics: In this study, we adopted a directional antenna to increase the at-

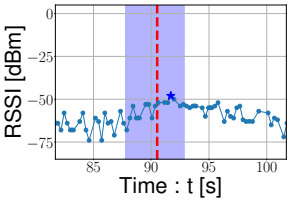


Fig. 16. Time–Signal strength graph.

Fig. 17. Captured images corresponding to $t_{\max}(i) = 91.70$ s

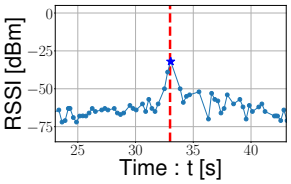


Fig. 18. Time–Signal strength graph.

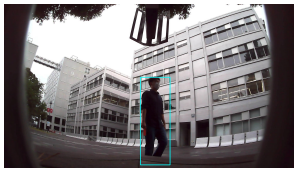


Fig. 19. Captured images corresponding to $t_{\max}(i) = 33.05$ s

tack success rate. In the following section, we evaluate how the characteristics of the antenna affect the results. In other words, we study cases in which the attack is conducted using a dipole antenna, which is an omni-directional antenna. In this experiment, a pedestrian walked with an iPhone 8 in his/her hand. The other conditions were the same as those in the experiments in Section 4.2. Figure 16 presents the measured time–signal strength graph. The BLE signal strength reached its maximum value at $t_{\max}(i) = 91.70$ s. Compared with that of the directional antenna (Fig. 5), the time–signal strength graph of the dipole antenna did not exhibit a sharp peak. The photo image captured at the time of maximum signal intensity is shown in Figure 17. A person can be seen in the image but is far off center. This result implies that, if there are multiple pedestrians in the image, it may not be possible to link them correctly. The pedestrian took 90.51 s to reach the camera, and the difference between the time of maximum signal strength and the time when the target passed in front of the camera was 1.18 s. These results suggest that a directional antenna is preferable to the receiving antenna used by the attacker.

Impact of Antenna Direction: To verify that the linking is achieved when the antenna direction is changed, we performed field and simulation experiments. First, a field experiment was conducted. The directional antennas were installed at an angle of 45° from the walking direction of the pedestrians. The other experimental conditions were the same as those described in Section 4.2. Figure 18 shows the time evolution of the radio signal strength of BLE frames containing the RPI of the target. The signal strength had a maximum value at $t_{\max}(i) = 33.05$ s. Figure 19 shows images corresponding to the BLE frame captured at $t_{\max}(i) = 33.05$ s. We can see that the target is correctly captured, and the linking is successful. Appendix F evaluates the impact of antenna direction through simulation. Simulation results show that the attack success rate remains at a high value even when the antenna is at an angle of 45°. These experimental results show that the attack is feasible even when a directional antenna is placed diagonally to the walking direction of the pedestrians.

We note that Appendix G evaluate the attack success rate in the condition where there are devices with different BLE transmission powers, by means of a simulation study. The results show that this condition did not cause a decrease in attack success rate. In theory, as the attack leverages the changes in the signal strength transmitted by a target device, differences in the absolute magnitude of the signal strength transmitted by devices should not affect the attack success rate. Our experimental results are validated to be consistent with the theory.

7.3 Ethical Considerations

We carefully designed our experiments to protect user privacy. Authors and colleagues with their consent participated in our experiments and performed the experiments based on the prepared behavior scenarios. The experiment to capture the target's photo was conducted only for the authors. The BLE data collected did not include directly personally identifiable information, but we kept the data strictly confidential because it could include the BLE data of third-party users. Please note that, according to the institutional review board's (IRB's) preliminary review procedure, our study was exempt from further IRB review because it did not use personally identifiable information of third-party users.

Through our experiment, we demonstrated that privacy threats can become apparent under certain conditions. We reported our findings and actionable countermeasures to Google. We received a response from Google that "From the perspective of the exposure notifications program, we do not consider this attack in scope. This was considered during the design phase." As Google mentioned, the attack was not caused by a specific application's vulnerability, but by the design of the DCT framework based on BLE. We believe that publishing this paper will provide insights and benefit

broad stakeholders. Our findings and practices can be used to review the existing designs of frameworks and protocols, which should contribute to reducing potential privacy risks.

8 Related Work

Health Information Privacy: The Health Insurance Portability and Accountability Act (HIPAA) [51] is a federal law in the United States that requires protection of sensitive patient health information from being disclosed without the patient's consent or knowledge. In HIPAA, protected health information (PHI) is individually identifiable information relating to the health status of an individual, which includes information our experiments successfully extracted, such as medical test results (i.e., COVID-19 positive) and face photographic images. HIPPA provides a privacy rule, called the Safe Harbor method, for de-identification to limit the possible uses and disclosures of PHI. However, Benitez and Malin illustrated such protection rules leave different organizations vulnerable to re-identification [52]. To guarantee confidentiality, scalability, and flexibility of health information management at a third party (e.g., cloud services), privacy-preserving and patient-centric models based on cryptography for the storage and exchange of health information have been studied [53, 54].

Re-identification Attacks: Re-identification or de-anonymization is a class of privacy attacks that identify users from anonymized user data. Health information datasets have often been targeted by many studies in this research area, as well as those mentioned [52]. Other types of user data have been targeted. Narayanan and Shmatikov demonstrated that an attacker can identify the Netflix records of known users by using the Internet Movie Database as the source of background knowledge [55]. Wondracek et al. proposed a new de-anonymization attack that exploits group membership information on social networking services [56]. The attack can be accomplished by stealing a browser history of a victim for certain URLs that reveal group memberships on a social network and combining this information with previously collected group membership data. Su et al. proposed a de-anonymization attack that links web browsing histories to pages on social networking services [57]. Our attack focuses on a new area of de-anonymization that bridges the gap between digital data and physical data, that is, combining pseudonymized identifiers of COVID-19 positives with their photographic images.

Privacy Attacks on GAEN Framework and DCT Apps: The ways to track specific users (i.e., COVID-19 positives) and expose their behavioral history on the GAEN framework have been extensively studied since the framework was released [12, 58–60]. The basic idea of these studies is to deploy multiple BLE receivers to collect the identity of pedestrians and match them against COVID-19 positives, so that the behavioral history (e.g., location and movement) of COVID-19 positive can be tracked. However, these studies neither discussed nor examined how to identify specific COVID-19 positives, for example, taking photographic information. Most relevant to our study, the linking attack to combine COVID-19 positives with their images has been discussed as a proof of concept [11, 13, 61]. Our study is the first to establish a scientific and reproducible methods for evaluating linking attacks and their countermeasures.

9 Conclusion

The primary objective of this study was to evaluate the feasibility and scalability of linking attacks that target GAEN, which is the most representative DCT framework. To this end, we conducted a field experiment using real equipment and a realistic simulation experiment that incorporates radio characteristics and a 3D model of a human being. The results demonstrated that the linking attack is highly feasible and that an attacker can succeed with a probability of 86% against a high pedestrian traffic flow of 5,000 people per hour by installing a few attack devices; i.e., the attack is scalable. In addition, as a mechanism to suppress the linking attack, we proposed a new contact tracing mechanism using multiple RPIs and a method to appropriately adjust the advertisement period of BLE frames in GAEN, and clarified their effectiveness by simulation.

The DCT framework represented by GAEN is expected to be a promising solution for preventing new infectious diseases in the future. Also, the widespread use of DCT frameworks should not compromise user privacy. We hope that the findings of this study will contribute to the design of a DCT framework with improved privacy protection.

Acknowledgment

A part of this work was supported by JSPS Grant-in-Aid for Scientific Research, Grant Number 19H04111. A part of this work was also supported by the Security Innovator Training Program, SecHack365 [62], which is operated by the National Institute of Information and Communications Technology (NICT).

References

[1] WHO. WHO and ECDC launch indicator framework to evaluate public health effectiveness of digital proximity tracing solutions. https://www.euro.who.int/en/health-topics/Health-systems/digital-health/news/news/2021/6/who-and-ecdc-launch-indicator-framework-to-evaluate-public-health-effectiveness-of-digital-proximity-tracing-solutions, Jun 2021.

[2] Chowdhury et al. COVID-19 contact tracing: Challenges and future directions. *IEEE Access*, 8:225703–225729, 2020.

[3] WHO UNICEF and IFRC. Social stigma associated with the coronavirus disease (COVID-19). https://www.unicef.org/documents/social-stigma-associated-coronavirus-disease-covid-19, Mar 2020.

[4] Chii-Chii Chew, Xin-Jie Lim, Chee-Tao Chang, Philip Rajan, Nordin Nasir, and Wah-Yun Low. Experiences of social stigma among patients tested positive for COVID-19 and their family members: a qualitative study. *BMC Public Health*, 21(1):1623, 2021.

[5] Tiziana Ramaci, Massimiliano Barattucci, Caterina Ledda, and Venerando Rapisarda. Social stigma during COVID-19 and its impact on HCWs outcomes. *Sustainability*, 12(9), 2020.

[6] Sawsan Abuhammad, Omar Khabour, and Karem Alzoubi. COVID-19 contact-tracing technology: Acceptability and ethical issues of use. *Patient preference and adherence*, 14:1639–1647, 09 2020.

[7] Kai Kaspar. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: Quantitative survey study. *J Med Internet Res*, 22(8):e21613, Aug 2020.

[8] PimEyes. PimEyes: Face recognition search engine and reverse image search. https://pimeyes.com/en, jan 2022.

[9] Inc Clearview AI. Clearview AI | the world's largest facial network. https://www.clearview.ai/, jan 2022.

[10] The New York Times. A face search engine anyone can use is alarmingly accurate. https://www.nytimes.com/2022/05/26/technology/pimeyes-facial-recognition-search.html, may 2022.

[11] Tijmen Schep. Corona detective & Corona milder. https://www.tijmenschep.com/corona-detective-corona-milder/, Dec 2020.

[12] Vincenzo Iovino. Immuni Detector & Paparazzi attack+camera. https://sites.google.com/site/vincenzoiovinoit/immuni, Jul 2020.

[13] Serge Vaudenay. Analysis of DP3T. Cryptology ePrint Archive, Report 2020/399, 2020. https://eprint.iacr.org/2020/399.

[14] Antoine Boutet, Nataliia Bielova, Claude Castelluccia, Mathieu Cunche, Cédric Lauradoux, Daniel Le Métayer, and Vincent Roca. Proximity Tracing Approaches - Comparative Impact Analysis. Research report, INRIA Grenoble - Rhone-Alpes, April 2020.

[15] Apple and Google. Exposure Notification Bluetooth specification. https://blog.google/documents/70/Exposure_Notification_-_Bluetooth_Specification_v1.2.2.pdf, Apr 2020.

[16] Carmela Troncoso et al. Decentralized privacy-preserving proximity tracing, 2020.

[17] Labour Ministry of Health and Welfare. COVID-19 contact-confirming application (COCOA). https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/cocoa_00007.html, Jul 2021.

[18] Italian Ministry of Health. Immuni - sito ufficiale. https://www.immuni.italia.it/, Jul 2021.

[19] NHS. NHS COVID-19 - NHS. https://www.nhs.uk/apps-library/nhs-covid-19/, Sep 2020.

[20] Corona-Warn-App open-source project. Open-source project Corona-Warn-App. https://www.coronawarn.app/en/, Jul 2021.

[21] Apple and Google. Exposure Notification cryptography specification. https://blog.google/documents/69/Exposure_Notification_-_Cryptography_Specification_v1.2.1.pdf, Apr 2020.

[22] Google. Exposure Notifications BLE attenuations. https://developers.google.com/android/exposure-notifications/ble-attenuation-overview, Nov 2020.

[23] Labour Ministry of Health and Welfare. [COVID-19 Contact-Confirming Application ver.1.2.1] Notice of change in the scope of Exposure Notification (in Japanese). https://www.mhlw.go.jp/content/10906000/000705649.pdf, Dec 2020.

[24] Joseph Redmon et al. You only look once: Unified, real-time object detection, 2016.

[25] Guillaume Kessibi et al. Analysis of Diagnosis Key distribution mechanism in contact tracing applications based on Google-Apple Exposure Notification (GAEN) framework. working paper or preprint, July 2020.

[26] Stephen Farrell. October 2020 survey of GAEN app key uploads. https://down.dsg.cs.tcd.ie/tact/survey10.pdf, Oct 2020.

[27] Stephen Farrell and Doug Leith. Testing apps for COVID-19 tracing (TACT) - TEK survey. https://down.dsg.cs.tcd.ie/tact/tek-counts/, Jul 2021.

[28] Testing Apps for COVID-19 Tracing (TACT) project. Collecting TEKs. https://github.com/sftcd/tek_transparency/.

[29] Google. google / exposure-notifications-server. https://github.com/google/exposure-notifications-server/blob/main/tools/export-analyzer/main.go, Feb 2021.

[30] Internet Engineering Task Force (IETF). HMAC-based extract-and-expand key derivation function (HKDF). https://datatracker.ietf.org/doc/html/rfc5869, May 2010.

[31] Michael Ossmann. greatscottgadgets / ubertooth. https://github.com/greatscottgadgets/ubertooth/, Jul 2021.

[32] Premiertek. Outdoor 2.4GHz 24dBi directional high-gain n-type female aluminum die cast grid parabolic antenna part#: ANT-GRID-24dBi. http://www.premiertek.net/

products/networking/ANT-GRID-24dBi.html.

[33] UL Verification Services Inc. MPE report. https://bit.ly/3leUBq4, Sep 2017.

[34] UL Verification Services Inc. MPE report. https://bit.ly/3igifAQ, Sep 2018.

[35] Dt&C Co Ltd. Bluetooth LE test report. https://fccid.io/ZNFKA1935/Test-Report/Bluetooth-LE-Test-Report-4450952.pdf, Sep 2019.

[36] SGS-CSTC Standards Technical Services Co Ltd Shenzhen Branch. RF test report of LTE B26. https://bit.ly/3ff5VyS, May 2018.

[37] Sporton International INC. FCC RF Test Report. https://fccid.io/IHDT56QD2/Test-Report/Test-Report-BT-LE-2421227.pdf, Oct 2014.

[38] Ultralytics. ultralytics / yolov5. https://github.com/ultralytics/yolov5, feb 2022.

[39] Milad Soltany. miladsoltany / face-detection. https://github.com/miladsoltany/Face-Detection, may 2021.

[40] Meli Harvey. sidewalkwidths-nyc. https://github.com/meliharvey/sidewalkwidths-nyc, Mar 2021.

[41] Carnegie Mellon University. Panda3D | open source framework for 3D rendering & games. https://www.panda3d.org/, Jul 2021.

[42] TP-LINK. TP-LINK 2.4GHz 24dBi grid parabolic antenna TL-ANT2424B data sheet.

[43] M Amanda and NYC Director. New York City pedestrian level of service study phase I, 2006.

[44] H.T. Friis. A note on a simple transmission formula. *Proceedings of the IRE*, 34(5):254–256, 1946.

[45] B. J. Jang, S. H. Wi, J. G. Yook, M. Q. Lee, and K. J. Lee. Wireless bio-radar sensor for heartbeat and respiration detection. *Progress In Electromagnetics Research C*, 5:149–168, 2008.

[46] Transport Ministry of Land, Infrastructure and Tourism. Open data on human flow in Otemachi, Marunouchi, and Yurakucho area (in Japanese). https://www.geospatial.jp/ckan/dataset/human-flow-marunouchi, may 2021.

[47] Google. exposure-notifications-internals. https://github.com/google/exposure-notifications-internals, mar 2021.

[48] Bureau of Social Welfare and Tokyo Metropolitan Government Public Health. Details of the announcement of the new COVID-19 positive patient in Tokyo (in Japanese). https://catalog.data.metro.tokyo.lg.jp/dataset/t000010d0000000087, jan 2022.

[49] Bureau of General Affairs (Tokyo Metropolitan Government) Statistics Division. Summary of population of Tokyo (estimated). https://www.toukei.metro.tokyo.lg.jp/jsuikei/2022/js221f0000.pdf, jan 2022.

[50] David M Douglas. Doxing: a conceptual analysis. *Ethics and Information Technology*, 18(3):199–210, 2016.

[51] U.S. Department of Health & Human Services. Health information privacy. https://www.hhs.gov/hipaa/index.html.

[52] Kathleen Benitez and Bradley Malin. Evaluating re-identification risks with respect to the HIPAA privacy rule. *Journal of the American Medical Informatics Association*, 17(2):169–177, 03 2010.

[53] Shivaramakrishnan Narayan et al. Privacy preserving EHR system using attribute-based infrastructure. In *Proceedings of the 2010 ACM Workshop on Cloud Computing Security Workshop*, CCSW '10, page 47–52, New York, NY, USA, 2010. Association for Computing Machinery.

[54] Ming Li et al. Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE Transactions on Parallel and Distributed Systems*, 24(1):131–143, 2013.

[55] Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, 2008.

[56] Gilbert Wondracek et al. A practical attack to de-anonymize social network users. In *2010 IEEE Symposium on Security and Privacy*, pages 223–238, 2010.

[57] Jessica Su et al. De-anonymizing web browsing data with social networks. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, page 1261–1269, Republic and Canton of Geneva, CHE, 2017. International World Wide Web Conferences Steering Committee.

[58] Jianwei Huang et al. On the privacy and integrity risks of contact-tracing applications, 2020.

[59] Lars Baumgärtner et al. Mind the gap: Security privacy risks of contact tracing apps. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 458–467, 2020.

[60] O. Seiskari. BLE contact tracing sniffer PoC. https://github.com/oseiskar/corona-sniffer, Mar 2020.

[61] Yaron Gvili. Security analysis of the COVID-19 contact tracing specifications by Apple Inc. and Google Inc. Cryptology ePrint Archive, Report 2020/428, 2020. https://ia.cr/2020/428.

[62] NICT. Young security innovator training program SecHack365 (in Japanese). https://sechack365.nict.go.jp/.

[63] Johannes K Becker, David Li, and David Starobinski. Tracking anonymized Bluetooth devices. *Proceedings on Privacy Enhancing Technologies*, 2019(3):50–65, 2019.

[64] Aveek K. Das, Parth H. Pathak, Chen-Nee Chuah, and Prasant Mohapatra. Uncovering privacy leakage in BLE network traffic of wearable fitness trackers. In *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications*, HotMobile '16, page 99–104, New York, NY, USA, 2016. Association for Computing Machinery.

[65] Google. Exposure Notification: Risks and mitigations FAQ. https://github.com/google/exposure-notifications-internals/blob/main/en-risks-and-mitigations-faq.md, Oct 2020.

A Tracking Pedestrians in Close Proximity

We evaluated the feasibility of the attack when pedestrians walk in close proximity. In our experiment, two short-distance pedestrians, each carrying a smartphone, passed in front of an attack device. The distance between the pedestrians was set approximately 2 m. The

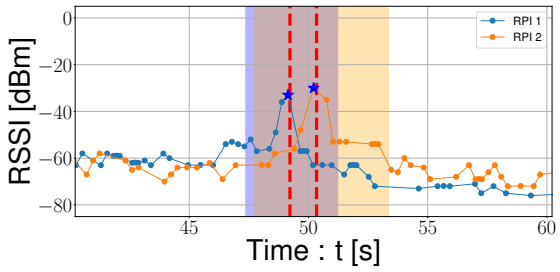


Fig. 20. Time–signal strength graph.

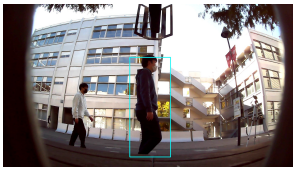


Fig. 21. Captured images corresponding to $t_{\max}(i) = 49.14$ s (RPI1)

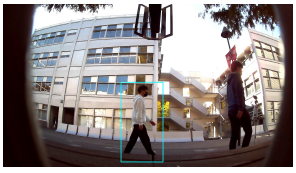


Fig. 22. Captured images corresponding to $t_{\max}(i) = 50.21$ s (RPI2)

first pedestrian walked with an iPhone 8 and the second pedestrian walked with a Huawei P20 Lite. The other conditions were the same as those in the experiments described in Section 7.1. The measured time–signal strength graph is shown in Figure 20. In the figure, the areas shown in purple and orange indicate the time from when the target transmits RPI1/RPI2 frames to the camera's angle of view to when it frames out. The red dotted line indicates the time at which the target arrived at the front of the camera. For RPI1 and RPI2, the BLE signal strength had a maximum value with a clear peak at $t_{\max}(i) = 49.14$ s and $t_{\max}(i) = 50.21$ s, respectively. The images taken at the time of maximum signal intensity for each RPI are shown in Figures 21 and 22. The difference between the time when each pedestrian reached the front of the camera and the maximum signal strength time was 0.08 s and 0.14 s, respectively. These results demonstrate that the RPI can be linked to each pedestrian even when the pedestrians are walking a short distance.

B Simulating the BLE RX Power

In section 5, to simulate the realistic BLE RX power, we introduced additive white Gaussian noise and loss corrections to represent the field environments; we set the additive Gaussian noise as $\epsilon \sim \mathcal{N}(0, 3^2)$ and the loss correction term as $\eta = -12$. The procedure by which these values were determined is as follows. First, field and

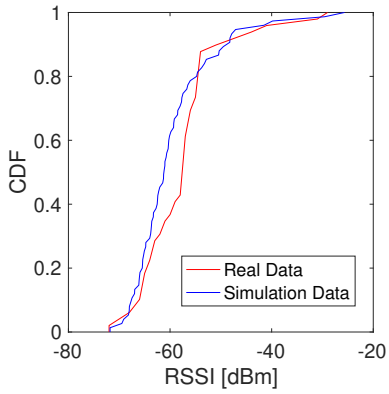


Fig. 23. Empirical Distributions for the received BLE signal strength: real-world experiment vs. simulation.

simulation experiments were conducted under identical conditions. In the field experiment, the target walked 2 m from the attack device with the iPhone XR in his hand on the side of the attack device. In the simulation experiment, the same conditions were set and the signal strength was calculated using Equation 2. In both the experiments, the strength of the received BLE signal was recorded.

Second, by comparing the difference in the received signals obtained from the two experiments, the parameters of the normal distribution assumed as noise and the size of the correction term were determined.

In the following, we show that the corrections determined above are reasonable. The time series data of the signal strength for 10 s before and after the time when the received signal strength reached the maximum value was extracted. The empirical distribution function of the extracted data is shown in Figure 23. We can see that the two distributions are very close. This result shows that the correction of the received intensity in the simulation reproduces the characteristics of the field experiment very well.

C Impact of Street Width

We set the street width as $w = 3$ m in the experiments presented in Sections 5.2 and 5.3. In this section, we evaluate the attack success rate when the street widths were set to $w = 3$, 6, and 9 m. The number of installed attack devices was set to one. The other conditions were the same as those described in Section 5.2. The results are shown in Figure 24. Although increasing the street width leads to a decrease in the attack success rate,

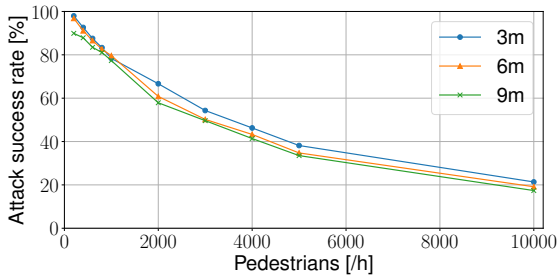


Fig. 24. N (# pedestrians/h) vs. attack success rate (%) for the street width $w = 3, 6, 9$ m.

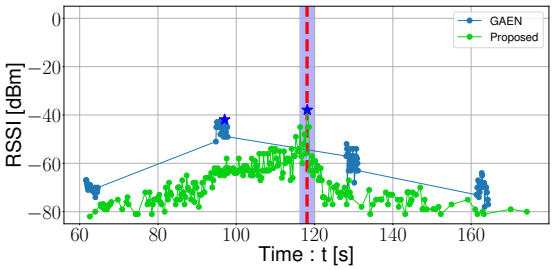


Fig. 25. Time-signal strength graph.

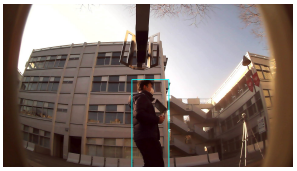


Fig. 26. Captured image at the maximum signal strength (smartphone).

Fig. 27. Captured image at the maximum signal strength (PoC).

overall, the street width does not have a significant impact on the attack success rate, implying that a linking attack is feasible for streets of various widths.

D Parameters for Assessing the I-RPI Overhead

In Section 6.1, the overhead introduced by the I-RPI mechanism was evaluated. To calculate the overhead, we used the number of collected TEKs and received RPIs per time window (15 min). The rationale for deriving these numbers is described below. First, the maximum number of TEKs was equal to the cumulative number of people with positive COVID-19 test results over the 14 days. As an example of the number of TEKs, we adopted 1,192,452, which is the cumulative number of infected individuals observed from February 6 to February 19 in Japan. Next, we measured the number of RPIs received in one time window at three locations at Shibuya station, one of the most congested areas in Japan: the connecting corridor, the sidewalk near the intersection, and the underpass. The number of measured RPIs at each location was 71, 16, and 118. Based on these results, the maximum number of received RPIs in a time window was determined to be 118.

E Empirical Evaluation of the Intermittent Signal Transmission

We developed a PoC implementation of the intermittent transmission scheme and demonstrated its effectiveness in mitigating attacks through field experiments. Because the current iOS and Android smartphone devices cannot configure the BLE signal transmission frequency to $T = 30$ s, as proposed in this study, we implemented the scheme on a PC with a BLE dongle. For comparison, we used an iPhone with GAEN installed. Table 7 summarizes the experimental setup.

In the experiment, a pedestrian passed in front of the attack device while holding the two devices. Figure 25 shows a graph of time–signal strength. For each signal strength, a star was plotted at the maximum value. From the figure, we can see that the resolutions of the signal strength graphs generated by the smartphone and the PoC implementation were high and low, respectively, as expected. Figure 26 and Figure 27 are images taken at the time when the respective signal strength was at its maximum. It is clearly shown that the attack was successful for the smartphone and unsuccessful for the PoC implementation. The difference between the time when the target reaches the front of the camera and the maximum time of the BLE signal strength transmitted by the PoC implementation is 21.1 s, which indicates that the attacker fails to link the target. These results demonstrate that the proposed method is an effective approach for attack mitigation.

Table 7. Sender device's setup.

Equipment	Model
Smartphone	iPhone 13 Pro (iOS 15.1, COCOA 1.4.1)
PoC Implementation	Panasonic CF-AX2 (Ubuntu 20.04 LTS)
BLE Dongle	BSBT4D100BK

Table 8. N (# pedestrians/h) vs. attack success rate (%) (Antenna angle : 45°). The values in parentheses are standard deviations.

200	96.4 (3.3)	1.8 (1.6)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
400	88.6 (2.2)	4.9 (0.9)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
800	82.8 (1.3)	7.1 (0.6)	0.1 (0.2)	0.0 (0.0)	0.0 (0.0)
1K	77.9 (0.8)	8.2 (0.8)	0.4 (0.3)	0.0 (0.0)	0.0 (0.0)
2K	60.9 (2.3)	14.4 (0.6)	1.0 (0.3)	0.1 (0.1)	0.0 (0.0)
3K	51.1 (0.9)	15.4 (0.8)	1.9 (0.3)	0.1 (0.1)	0.0 (0.0)
5K	36.3 (0.9)	17.3 (0.4)	3.1 (0.2)	0.4 (0.1)	0.0 (0.0)
10K	18.3 (0.8)	14.4 (0.7)	4.4 (0.1)	0.8 (0.1)	0.1 (0.0)

F Simulation Study of the Directions of Antenna

We present the simulations that also aim to study the attack success rate when the directional antennas are installed at an angle of 45° from the walking direction of the pedestrians. The other experimental conditions were the same as those described in Section 5.2. The number of pedestrians per hour N ranges from 200 to 10,000. Table 8 lists the results. The attack success rate for $N = 1,000$ was 78% and that for $N = 5,000$ was 36%. The attack success rate is reduced by up to six points compared to the actual experimental results in Section 5.2. However, the results show that the attack success rate remains high.

G Impact of Various BLE Signal Strengths

We evaluated the attack success rate in the condition where there are devices with different BLE transmission powers, by means of a simulation study. In the simulation study, we assumed that the target people could be divided into two groups, Group A and B, where each person in Group A owned an iPhone XR and each person in Group B owned a Nexus 6. The values in Table 9 were used as the parameters for the smartphones carried by the targets. The other experimental conditions were the same as those described in Section 5.2. The number of pedestrians per hour N ranges from 200 to 10,000. Table 10 lists the experimental results. The attack success rate for $N = 1,000$ was 79% and that for $N = 5,000$ was 40%. The attack success rate is comparable to the results in Section 5.2. The experiment shows that the attack is feasible even when devices with different BLE transmit strengths are combined.

Table 9. Parameters used in the simulation (Mix Conditon) .

	Description	Value
G_{ta}	Absolute gain of the Group A target's BLE transmission antenna	-4.9 [dBi]
G_{tb}	Absolute gain of the Group B target's BLE transmission antenna	-3.0 [dBi]
P_{ta}	Group A Target's BLE transmission power	16 [dBm]
P_{tb}	Group B Target's BLE transmission power	6.57 [dBm]

Table 10. N (# pedestrians/h) vs. attack success rate (%) (Mix Conditon). The values in parentheses are standard deviations.

N	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m \geq 5$
200	94.7 (1.8)	2.4 (0.7)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
400	91.4 (3.1)	4.1 (1.4)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
800	86.0 (2.0)	6.0 (0.5)	0.2 (0.2)	0.0 (0.0)	0.0 (0.0)
1K	78.9 (1.7)	8.8 (0.6)	0.3 (0.2)	0.0 (0.0)	0.0 (0.0)
2K	66.1 (2.6)	12.5 (0.6)	1.1 (0.2)	0.1 (0.1)	0.0 (0.0)
3K	55.8 (2.6)	15.9 (0.7)	1.4 (0.0)	0.1 (0.1)	0.0 (0.0)
5K	39.9 (1.7)	18.0 (1.0)	2.8 (0.2)	0.2 (0.2)	0.0 (0.1)
10K	19.9 (0.7)	16.1 (0.4)	4.4 (0.1)	0.9 (0.1)	0.1 (0.0)

H Further Discussion

H.1 Other Potential Targets of the BLE-based Linking Attack

We note that the target of our BLE-based linking attack is not limited to DCTs such as the GAEN framework. This attack can also be a threat to other BLE-based apps, such as fitness trackers or Bluetooth earphones. Becker et al. [63] found that MAC address randomization does not work on Fitbit, a popular wireless-enabled wearable fitness tracker product. In addition, Das et al. [64] found that all six fitness trackers examined use fixed MAC addresses. They further performed a measurement study in the wild and reported that 89% of fitness trackers they observed used fixed MAC addresses. Using these facts, an attacker can link the fixed BLE MAC addresses advertised by fitness trackers with the face photos of the owners. A promising scheme for mitigating such threats is MAC address randomization; we expect such schemes to be widely implemented by applicable manufacturers.

H.2 Positioning of Linking Attack

GAEN framework implements a mechanism by which the identifier transmitted by the smartphone changes every 10–20 min to prevent an attacker from tracking the user over a long period of time. The goal of a linking

attack is not to track users over time but to identify the target's infection status. We also note that the threat of linking attacks has been pointed out from the design stage in DCT frameworks such as the DP3-T and GAEN frameworks [16, 65]. During the design phase, it was believed that the attack required multiple contacts with the target and that it was therefore not feasible. Our extensive experiments revealed that it is feasible to perform a linking attack.