Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology

Hanne Æsøy Nes

Privacy in the Norwegian Automatic Contact Tracing App Smittestopp

Master's thesis in Communication Technology
Supervisor: Colin Boyd
June 2021

NTNU
Kunnskap for en bedre verden

Hanne Æsøy Nes

Privacy in the Norwegian Automatic Contact Tracing App Smittestopp

Master's thesis in Communication Technology
Supervisor: Colin Boyd
June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology

Title:	Privacy in the Norwegian Automatic Contact Tracing App *Smittestopp*
Student:	Hanne Æsøy Nes

Problem description:

During the current COVID-19 pandemic, contact tracing is an important measure to help break chains of transmission and contain the spread of the coronavirus. To assist manual contact tracing, many countries have implemented Automatic Contract (ACT) systems that use mobile devices to track close encounters. In Norway, the original implementation of such a system, the app Smittestopp, was banned because it collected a lot of information about its users and stored the data centrally to use for analysis. To take the privacy concerns into account, the next version of Smittestopp is based on the Google and Apple Exposure Notification (GAEN) system that is decentralised and does not store any data centrally.

This master thesis will look at the privacy of the new version of the Norwegian Smittestopp. It will consider the privacy measures and concerns of ACT systems based on GAEN especially with regards to Smittestopp.

Date approved:	2021-03-15
Supervisor:	Colin Boyd, IIK

Abstract

During the ongoing COVID-19 pandemic, Automatic Contact Tracing (ACT) systems have been used to more effectively trace possible infected individuals, but the systems have also brought along new concerns, especially related to privacy. In this thesis, different ACT systems based on Bluetooth Low Energy (BLE) are explored, focusing on the Google and Apple Exposure Notification (GAEN) based systems. GAEN was created as a joint effort between Google and Apple to create a more privacy-preserving decentralised ACT system that addresses issues using BLE for contact tracing. Norway uses the implementation in their new ACT system, Smittestopp. Different privacy attacks have been successful on different apps based on GAEN. The main focus of the thesis is on attacks that try to identify infected individuals or identify movement patterns and create social graphs. Integrity and other concerns are also explored. The implications of the attacks vary. Contact tracing is not privacy-preserving in nature, but the extent and distributed nature of the data could make it easier for outsiders to access information on COVID-19 diagnosed individuals.

Sammendrag

I løpet av den pågående COVID-19-pandemien har automatiske smittesporingssystem blitt brukt for å spore mulige smittede individer mer effektivt. Disse systemene har også ført til nye bekymringer, spesielt knyttet til personvern. I denne oppgaven blir forskjellige smittesporingssystem basert på Bluetooth Low Energy (BLE) utforsket, med fokus på system som baserer seg en løsning kalt Google and Apple Exposure Notification (GAEN). GAEN ble opprettet som en felles innsats mellom Google og Apple for å forbedre personvernet med et desentralisert smittesporingssystem som adresserer problemer med å bruke BLE til smittesporing. Norge bruker implementeringen i sitt nye system, Smittestopp. Ulike personvernangrep har vært vellykkede på forskjellige apper basert på GAEN. Denne oppgavens hovedfokus er på angrep som prøver å identifisere smittede individer eller identifisere bevegelsesmønstre og lage sosiale grafer basert på data. Bekymringer knyttet til integritet og andre bekymringer blir også utforsket. Implikasjonene av angrepene varierer. Smittesporing i seg selv er ikke personvernbevarende, men omfanget og den distribuerte kvaliteten til dataen kan gjøre det lettere for utenforstående å få tilgang til informasjon om COVID-19-diagnostiserte individer.

Preface

This master thesis is submitted as the last work of the five-year Master of Science (MSc) in Communication Technology (IIK) at Norwegian University of Science and Technology (NTNU).

I want to thank my supervisor, Professor Colin Boyd, for his guidance and insights. Without his feedback, advice, and support, the work on this thesis would have been much more difficult.

The subject of contact tracing is important, but especially now with the COVID-19 pandemic. It has been fascinating to read about how contact tracing is done both manually and technologically, and I have learned a lot throughout the process.

These five years would not have been the same without all the people I have met. I am especially grateful for my classmates and the study and social environment we have created together. My time in Trondheim will not be forgotten.

I would also like to thank my family and friends for the continuous support over the years.

Hanne Æsøy Nes
Trondheim, 2021

Contents

Appendices

List of Acronyms

ACT Automatic Contact Tracing.

AEMD Associated Encrypted Metadata.

AES Advanced Encryption Standard.

BLE Bluetooth Low Energy.

DP-3T Decentralised Privacy-Preserving Proximity Tracing.

DPIA Data Protection Impact Assessment.

ENE Exposure Notifications Express.

FHI Norwegian Institute of Public Health.

GAEN Google and Apple Exposure Notification.

GDPR General Data Protection Regulation.

GPS Global Positioning System.

HA Health Authority.

JWT JSON Web Token.

MAC Message Authentication Code.

MSIS Norwegian Surveillance System for Communicable Disease.

NHN Norwegian Health Network.

NTNU Norwegian University of Science and Technology.

PEPP-PT Pan-European Privacy-Preserving Proximity Tracing.

ROBERT ROBust and privacy-presERving proximity Tracing.

RPI Rolling Proximity Identifier.

SHA Secure Hash Algorithms.

TEK Temporary Exposure Key.

WHO World Health Organisation.

Chapter 1 Introduction

In December of 2019, at a market in Wuhan in China, a new SARS virus began to spread between humans [Tay21]. Since then, the World Health Organisation (WHO) has declared the virus later named COVID-19, a global pandemic. Severe measures that limit people's freedom have been used to contain the spread of the virus and prevent hospitals from reaching capacity. The measures have affected the world economy, and people are anxious to go back to normal. Different companies have developed vaccines at record speed, but high demand makes the vaccination process slow, and one can only guess when the pandemic will end.

COVID-19 spreads through bodily fluids. This means that it is crucial to isolate infected people and figure out who have been in contact with a contagious person to prevent more people from becoming infected. Therefore, *contact tracing*, the act of establishing who could be close encounters of an infected person, is a vital tool in a virus outbreak. However, manual contact tracing, where health officials directly contact those who test positive for COVID-19, takes many resources, and it sometimes misses possible encounters.

A pandemic is of global concern, and research communities worldwide have looked for better or new solutions to the new problems. Since contact tracing is of high importance, it is not an exception to this. Applications for mobile devices that register encounters between people have been developed. Different types of solutions now exist, and they all go under the *Automatic Contact Tracing (ACT)* umbrella. Using an ACT system makes it easier and possibly more effective to identify encounters that might otherwise be forgotten or overlooked.

With new solutions, new issues often follow. Of particular interest in this thesis, some ACT implementations have been criticised for lack of privacy. These applications collect more data than necessary for contact tracing. In Norway, the first iteration of an ACT system was removed due to privacy concerns. It collected a lot of data that was supposed to be used for research. The second version is based on a widely used framework, the Google and Apple Exposure Notification (GAEN) framework. It is used in many countries and seen as privacy-preserving

In this thesis, the following research questions will be investigated:

RQ1: Is the privacy of infected users conserved in Smittestopp and other GAEN based systems?
RQ2: Can ACT systems based on GAEN be used as surveillance systems to monitor user's movement patterns and social circles?

To look into these questions, alternatives to GAEN are studied, and a literature research to find flaws and privacy concerns has been conducted. Most of the literature used investigates GAEN or systems implemented in other countries than Norway. However, since the systems share many similarities, the research is relevant also for Norway's Smittestopp.

The first chapter, Chapter 2, introduces background information that will be the basis for the rest of the thesis. The concepts include the Norwegian way to handle manual contact tracing and the technical aspects used for automatic contact tracing.

In Chapter 3, a survey of existing ACT systems is conducted. The GAEN system is explored in detail, among some other solutions that have been implemented. All the systems discussed are based on the Bluetooth Low Energy (BLE) technology.

Chapter 4 explains how Norway's two versions of an ACT system operate. The app that is in use now, Smittestopp, is explored in more detail than the first app, Smittestopp v1. Smittestopp is based on GAEN, but the Norwegian government controls much of the system, and those parts are in focus in the chapter.

In Chapter 5, privacy attacks that could be an issue in GAEN based systems are introduced. Integrity and other concerns are also investigated. The privacy attacks are divided into attacks that try to identify COVID-19 positive individuals, and identify movement patterns and create social graphs.

How the attacks explained in Chapter 5 can be a risk in Smittestopp, and the overall impact of the attacks, are then discussed in Chapter 6.

Chapter 2 Background

In this chapter, the terminology and concepts used in the rest of the thesis are introduced.

2.1 Global Response to the COVID-19 Pandemic

COVID-19 took the world off guard, and across the world, strict measures have been put in place to try to contain the spread so that hospitals do not reach capacity. One of the main tactics has been to prevent people from meeting too many other people by restricting and monitoring movement patterns. New familiar terms include “social distancing”, the act of maintaining greater physical distance from others than usual [MW21c], and “lockdowns” where people in a city or even a country have to stay home, only necessary shops are open, and activities outside the house are restricted [MW21b].

International cooperation has also been vital. For instance, the global cooperation, COVAX, works toward faster development of vaccines for COVID-19 and fair access for every country [WHO20].

2.2 Norway’s Response to the COVID-19 Pandemic

On the 12th of March 2020, Norway’s Prime Minister Erna Solberg announced that the country would shut down and introduced the most invasive legislation since World War II [Hel20a]. Since then, Norway has enforced both stricter and less strict countermeasures against the spread of the virus.

The Norwegian government’s strategy against the pandemic, as decided in May of 2020 [Hel20b], defines six important measures to slow the spread of the COVID-19. They are the following:

1. Hygiene measures such as frequent hand washing and general cleaning.

2. Early detection and isolation of infected individuals.
3. Detection and quarantine of close contacts of the confirmed infected.
4. Reduction of the number of travellers to and from areas with high infection rates.
5. Reduction of the contact frequency in the population.
6. Extensive protective measures at nursing homes and hospitals, and otherwise for members of high-risk groups.

At the core of Norway's response to the pandemic is the Norwegian Institute of Public Health (FHI). It is a government agency under the Ministry of Health and Care Services, and its task is to produce, summarise and communicate knowledge with regards to public health [FHI]. One of their competence areas is infectious disease control. FHI is also responsible for the data handling of Norway's central health register, Norwegian Surveillance System for Communicable Disease (MSIS). MSIS is regulated under the health register law and contributes to the surveillance of infectious diseases in Norway [FHI17].

2.3 Contact Tracing

Contact tracing is the identification, notification, and monitoring of individuals that might have been in contact with someone who is confirmed infected by an infectious disease [MW21a]. The goal of contact tracing is to break future chains of infection by locating possible sources of infection. Hopefully, this is achievable before there is a large outbreak of the disease in question.

Both testing and contact tracing are central to uphold measures 2 and 3 of the Norwegian COVID-19 strategy mentioned in Section 2.2. In this thesis, the terminology introduced by FHI [FHI20c] will be used to describe the different actors in a contact tracing situation:

- *Index case:* a person with a confirmed COVID-19 infection that triggers contact tracing.
- *Close contact:* a person that has been in contact with an index case and is therefore at risk of infection.
- *Contact tracer:* a person that works with tracing close contacts of the index cases.

In Norway, a close contact is defined as someone who has been within two metres of an index case for more than 15 minutes or someone who has been in physical contact with the index case or their bodily fluids [FHI20a]. These encounters must also occur when the index case is infectious, which is considered two days before the first symptom or, in the case of no symptoms, two days before a positive test result, and until the symptoms have passed, or ten days after exposure. Other factors that can decide if a person is at high risk of infection [FHI20c] are:

- what way people are together (e.g., face-to-face),
- if the encounter was inside or outside,
- if the symptoms are more likely to spread fluids (e.g., coughing),
- the age of the infected (children are less likely to infect others),
- if any activity that led to heavier breathing occurred,
- and if the encounter occurred within the most infectious period of the COVID-19 infection (usually sometime between two days before the first symptom and three days after).

These are many deciding factors, and in the end, it is up to the contact tracer to decide who is a close contact.

The contact tracers consist of the chief municipal physician (kommuneoverlege) and its staff, and the general practitioners of the municipalities [FHI20c]. When an index case is found through testing, the chief municipal physician is notified. Next, a contact tracer contacts the index case to inform them of what to do next and to get a list of contacts. The tracers then evaluate the list, and the contacts defined as close are contacted and asked to quarantine and get tested. The close contacts should stay quarantined for ten days after the encounter unless they test negative on day seven. If any of the tests come back positive, the close contact triggers a new round of contact tracing and is now an index case. This illustrates the close link between contact tracing and testing, and for contact tracing to be efficient, it should be easy to test possible cases of COVID-19.

Based on the information given by the infected cases, a contact tracer can also consider further extending the tracing by involving other municipalities or by alerting the media of a possible outbreak somewhere, for instance. All identified cases and contacts are stored and reported to FHI.

2.4 Mobile Devices

Worldwide, there are around 3.8 billion unique smartphone users. This makes up approximately 48.41% of the population [Tur21]. In Norway in 2020, the number of people that owned a smart mobile device was around 96% according to an annual survey about media usage done by Statistics Norway [Sta21].

The most common operating systems are those based on Google's Android, making up 71.9% of the number of devices, followed by Apple's iOS making up 27.33% [Sta20]. In Norway, iOS devices make up 59.56% and Android devices 39.92%.

Mobile devices usually have many different types of technologies that can be used for different purposes by developers. In the following subsections, some of the technologies used in ACT systems are introduced.

2.4.1 BLE

Bluetooth is a short-range wireless technology that operates in the 2.4GHz frequency band [Spo18]. In addition to the classical Bluetooth, a less power-consuming option, Bluetooth Low Energy (BLE) exists. All devices that support Bluetooth 4.0 or newer versions support BLE. Classical Bluetooth is used for audio streaming, while BLE is used for data transmission, location services, and to create device mesh networks for, for instance, IoT. The location services include technologies such as broadcasting beacons or indoor navigation. Mobile devices can use both Classical and BLE at the same time by using a time-sharing mechanism.

In Apple's iOS-devices, the possibility to run background tasks is limited. This includes BLE services, and for apps created by external developers, the usage of BLE only works when the app runs in the foreground, thereby using a lot of battery. The functionality is restricted to prevent app developers from collecting too much data without the user knowing. For instance, before the restrictions were in place, some companies placed beacon transmitters at locations so that their apps could register if users were at the location and this was used for advertising purposes [Wel19]. The technology is therefore also used to create targeted advertisements. It is easier for Android developers to use the technology in the background but still limited.

For broadcasting beacons, there are two main ways. One is to broadcast to and collect from all nearby devices that are listening. Another is to connect through a handshake before data is transmitted. The range of BLE is specified to be up to 100 metres in Bluetooth 4.0 and up to 400 metres in Bluetooth 5, but in reality, it is much shorter [Spo18]. Due to factors such as device receivers, transmitters, surroundings, and antennas, the range is usually around 10 metres at best. This also makes the range much shorter indoors than outdoors.

2.4.2 GPS

Global Positioning System (GPS) is a navigation system that provides positioning services anywhere on Earth. The US government created the technology, and it consists of satellites that orbit the Earth and transmit current position and time [US 20]. User devices have receivers that collect these signals and calculate their three-dimensional position and time based on them. GPS requires line-of-sight to the satellites and does not work as well indoors because of it. Devices that use GPS continuously search for satellites to connect to, and if none is available, the device tries to connect to all, which increases the battery consumption of applications reliant on GPS [Lia18].

2.5 Cryptography Background

Three concepts that are often used to evaluate the information security of something are confidentiality, integrity, and availability. In short, confidentiality is about keeping unauthorised people from accessing data they are not supposed to access. Integrity is about keeping the data from being altered or deleted by unauthorised people. Finally, availability is about maintaining access to those that should have access and preventing unauthorised people from disrupting this access.

Privacy

Privacy can be placed under confidentiality, and concerns keeping the personal data, data that can identify an individual or the data subject, safe. The entity that uses the data subject's data is called the data controller. The entity that collects the data on behalf of the data controller is called the data processor.

In Europe, the privacy and security law, General Data Protection Regulation (GDPR) defines principles the data controller should follow to give the users a higher degree of privacy. These are the following [Wol21]:

1. *Lawfulness, fairness and transparency:* all processing of data must be lawful, fair, and transparent to the person the data belongs to.
2. *Purpose limitation:* data must only be processed for legitimate and explicitly stated purposes.
3. *Data minimisation:* only the most essential data should be collected.
4. *Accuracy:* the data must be kept accurate and up to date.
5. *Storage limitation:* data must not be stored for longer than what is necessary for the stated purpose.
6. *Integrity and confidentiality:* data must be processed in a way that confidentiality and integrity are ensured.
7. *Accountability:* the data processor is responsible for keeping all of these principles and be open about how.

For instance, if data is processed outside of the data controllers control, it would decrease the user's privacy as the principle about integrity and confidentiality is not upheld. In this thesis, these seven principles are kept in mind when privacy risks are estimated.

Anonymity

The anonymity of users can help improve the privacy of users. In data collection and processing cases, anonymity can be achieved by ensuring that the information is processed so that individuals are non-identifiable. For instance, it should not be possible to pick out the data of one individual in a data set. Creating social graphs as defined in Section 2.6.3 or deducing movement patterns reduces the anonymity set and therefore contributes to identifying individuals.

Encryption

Data is often encrypted to provide confidentiality to data. This means that the data is encoded so that it cannot be understood without reversing the encoding. Two types of encryption are symmetric and asymmetric encryption. In symmetric encryption, a shared key is used, and in asymmetric encryption, the keys are different but mathematically related, a public key for encryption and a private key for decryption. The keys must be kept safe and secret in all types of encryption.

The most common encryption standard is the Advanced Encryption Standard (AES) [Nat01], which is a symmetric block cipher. A block size of 128-bits can be used, and the keys can be of different sizes. For example, AES-256 means that a 256 bit-sized key is used to encrypt the data. In an ACT system, encryption is used to keep data that is transferred between devices anonymous and random, thereby increasing the users' privacy.

A key derivation function is used to derive a key used for encryption. In GAEN, a HMAC-based key derivation function (HKDF) is used. It uses a key, a salt, some information, and the output length to create a new key, and a hash function must be chosen. A hash function is an algorithm that takes a message and outputs a bit array of fixed length. They are one-way, meaning that it is not possible to get the original message from the output. One example is the Secure Hash Algorithms (SHA) family [Nat15]. SHA-256 is used in the HKDF used in GAEN to encrypt data.

Authentication

Authentication, to confirm that the message is from the sender it claims to be from and that the message is not altered, is also often important. A Message Authentication Code (MAC) can be used to provide integrity and authentication and is, for instance, used to validate data sent between two entities that share a secret key. A secret shared key is often used. For HMACs, a cryptographic hash function and a secret key are used.

For instance, in Smittestopp, a Chaum Pedersen zero-knowledge proof [CP92] is used by the verification server when anonymous tokens are used. A zero-knowledge proof is a way to prove that an entity knows a secret without exposing the secret to the verification entity. It is used to prove that the private key was used to sign the verification request.

2.6 Automatic Contact Tracing (ACT)

Manual contact tracing can be time-consuming and resource-demanding, and it does not scale well. Besides, many encounters that should be classified as close are difficult to trace. For instance, if an index case has taken any public transportation while being infectious, these contacts will not be known to the index case and, therefore, difficult to identify for the contact tracers. Worst case, this leads to unidentified chains of infection.

Consequently, many countries have implemented Automatic Contact Tracing (ACT) systems. These systems take advantage of the fact that so many citizens own a mobile device, and use this to record encounters that might otherwise be missed or forgotten. The goal is to improve the efficiency of contact tracing.

The use of smartphones also allows for the usage of the technologies they offer. Both GPS and BLE are used, along with, for instance, QR codes that a user can scan with the device camera to register their presence at a location. In this thesis, systems based mainly on BLE are examined. GPS based systems consume a lot of battery and cannot estimate the distance between people to the precision that contact tracing needs, especially not indoors. Systems created that use GPS also collect more data and are often seen as less privacy-preserving.

Systems that use BLE use it to look for nearby devices to register encounters. However, since both Android and iOS devices face issues with background-running of BLE, apps that base themselves on BLE must implement workarounds or encourage users to run the app in the foreground.

Global cooperation is also no exception when it comes to ACT systems. Many new protocols or systems are open-source so that other countries can use the systems or take ideas from them for their implementations. In addition, researchers from different countries have worked together to create solutions. For example, the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) protocol was created as a cooperation between scientists in Switzerland, Belgium, and Italy [PEP20]. PEPP-PT met controversy due to its centralised approach, and some of the researchers broke from it to create the Decentralised Privacy-Preserving Proximity Tracing (DP-3T) protocol. It was created by researchers from the same countries as well as the Netherlands, UK, and Portugal [TPH+20]. Both of the protocols will be explored in more detail in Chapter 3.

Generally, an ACT system consists of the following actors:

- *Users:* members of a country with a mobile device.
- *Health Authority (HA):* usually in charge of testing and manual contact tracing. In an ACT system, the HA is also, most often, the responsible owner of the system and in charge of verifying infections.

The main components are usually:

- *Backend:* consists of servers and data centres that receive and distribute data on who is infected.
- *App:* collects and transmits data. The data could consist of BLE beacons, GPS location data, QR codes, or different combinations.
- *Verification solution:* verifies claims of infection by checking if a positive test result exists. The verification is often outside of the contact tracing protocols and implemented by the HA but is still an important part of an ACT system.

Other concepts that are often included:

- *Risk score:* an estimate of the risk that a user is infected based on collected data and criteria defined by the government such as the criteria mentioned in Section 2.3.
- *Encounter metadata:* consists of the variables used to calculate risk scores and includes, for instance, the signal strength of the BLE signal, duration of the encounter, and other variables.
- *Pseudonym:* is an anonymous, randomised identifier broadcast to be able to identify encounters later.
- *Exposure notification*: the notification sent to the user to notify them that they are at risk of infection.

From a technical perspective, there are two main types of contact tracing systems, *centralised* and *decentralised*. The main difference between the two is where the risk score is calculated as shown in Figure 2.1 and Figure 2.2. All steps until number four are the same in these figures, but the two approaches differ at the server and in the following steps.

2.6.1 Centralised ACT Systems

In the centralised systems, risk scores are calculated centrally, usually by the HA. An overview can be seen in Figure 2.1. The identities of the index case and their close contacts are often revealed to the HA. In many solutions, the users register using their phone number, and if they are later at risk of infection by having been near an index case, they are contacted directly through their number. Some systems

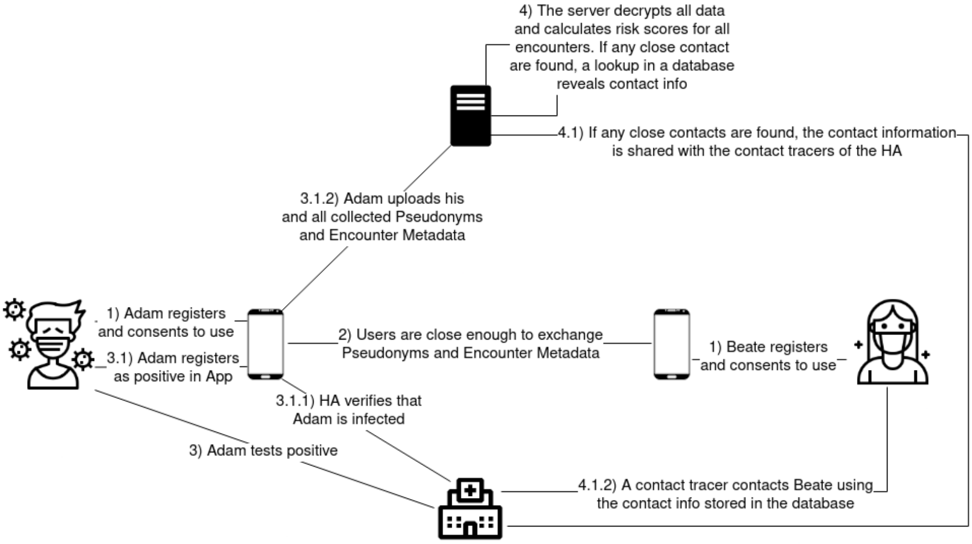

4) The server decrypts all data and calculates risk scores for all encounters. If any close contact are found, a lookup in a database reveals contact info
4.1) If any close contacts are found, the contact information is shared with the contact tracers of the HA
3.1.2) Adam uploads his and all collected Pseudonyms and Encounter Metadata
1) Adam registers and consents to use
3.1) Adam registers as positive in App
2) Users are close enough to exchange Pseudonyms and Encounter Metadata
1) Beate registers and consents to use
3.1.1) HA verifies that Adam is infected
3) Adam tests positive
4.1.2) A contact tracer contacts Beate using the contact info stored in the database

Figure 2.1: Flow of a general *centralised* ACT system

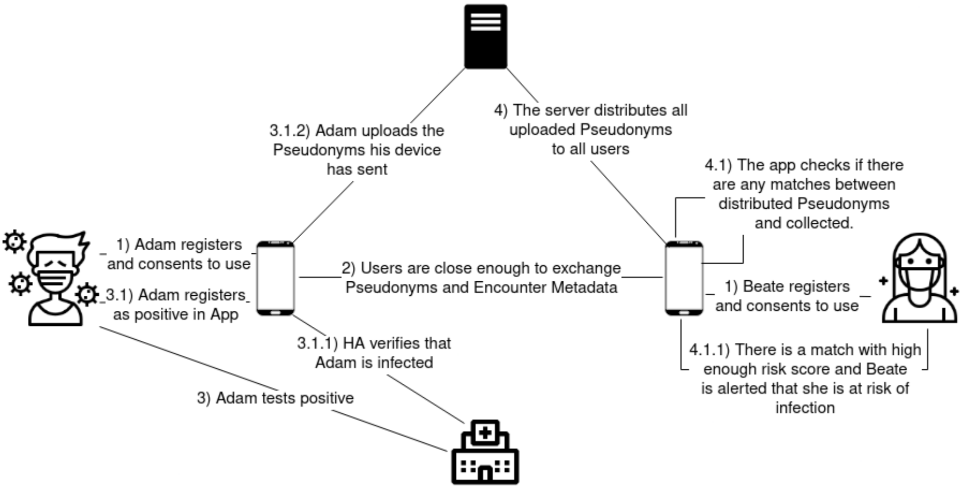

3.1.2) Adam uploads the Pseudonyms his device has sent
4) The server distributes all uploaded Pseudonyms to all users
4.1) The app checks if there are any matches between distributed Pseudonyms and collected.
1) Adam registers and consents to use
3.1) Adam registers as positive in App
2) Users are close enough to exchange Pseudonyms and Encounter Metadata
1) Beate registers and consents to use
3.1.1) HA verifies that Adam is infected
4.1.1) There is a match with high enough risk score and Beate is alerted that she is at risk of infection
3) Adam tests positive

Figure 2.2: Flow of a general *decentralised* ACT system

also regularly collect all user data regardless of infection and store it centrally. In other centralised systems, the computation of the risk score is done on encrypted data not to reveal the user's identity to the HA.

The HA must be a trusted third party. The decrypted data stored centrally gives much information about the users. The data is often controlled by the HAs, and the data could be used to create social graphs. This way, the HA can learn much about the population. The graphs can, for instance, be used to find individuals that are tied to a person of interest.

Moreover, if the keys are lost, a lot of data can be obtained by other actors. The central servers could therefore be of high interest to external attackers. Even when all computation is done on encrypted data, access to the data can create valuable social graphs. On the other hand, it is difficult for individual users to find out anything about each other. The identity of the index case is therefore well preserved between individual users.

2.6.2 Decentralised ACT Systems

In a decentralised system, the risk score is calculated locally at the user's device, and only the user will know if they are at risk of infection. An overview can be seen in Figure 2.2. The amount of data the government controls is limited, making it easier for users to trust and adopt the system. The HA usually only redistributes data without storing data for long or decrypting anything. It must be trusted not to alter data or store data longer than stated, but it does not control as much information as the servers in the centralised systems do.

At the same time, because of the distributed solution, more data is at the individual users' devices, and if any vulnerabilities exist, more people can exploit them. That means that it is easier for individual users to, for instance, try to identify infected individuals.

2.6.3 Privacy Concerns in ACT Systems

Two privacy concerns related to ACT systems that will be discussed in this thesis are the possibility to identify who is infected and to trace who users have been with and where they have been. A positive COVID-19 diagnosis is health data and is therefore sensitive and essential to keep confidential. Identifying an individual that has tested positive is, therefore, a breach of the user's privacy.

Information on whom an individual knows and where they are can be used to identify an individual and create graphs that can show movement patterns and social connections in the population. Social graphs can be used for deanonymisation as research has shown [SMB17], and an individual's movement patterns can be used to try to deanonymise the individual where a person has been is personal information.

2.6.4 Integrity Concerns in ACT Systems

If a large number of false exposure notifications exist in an ACT systems, the system's value would decrease. To send false exposure notifications could be done to prevent a specific person or group of people of attending a specific event, prevent an event, or keep a group of people or a person at home. Another effect of false exposure notifications could be to create mistrust in the government. Since the HA is usually in charge of the system, a non-functioning ACT system would reflect poorly on the government.

Chapter 3 ACT systems

In this chapter, some examples of Automatic Contact Tracing (ACT) systems are described. There are other systems and variants of the ones described in use today that will not be discussed. Firstly, centralised systems will be explored, followed by decentralised systems. At the end of this chapter, an overview of the systems is presented.

3.1 Centralised ACT Systems

As mentioned in Section 2.6, centralised systems usually calculate risk scores centrally. Some variations of this type of system reveal the user's identity to the HA, and some do not.

3.1.1 BlueTrace

One of the first nationwide deployments of an ACT system was a centralised system, *TraceTogether* in Singapore [BKT+20]. This system is based on a protocol called BlueTrace, which uses BLE to register encounters. An open-source version called OpenTrace also exists. Devices connect through handshakes and can either possess the role of Peripheral or Central, and devices usually alternate between the two. Peripherals advertise, and Centrals look for these advertisements to connect to the Peripherals to exchange a collection of data through reads and writes performed by the Central device. The data transferred are called Encounter Messages and are UTF-8 encoded JSON messages that contain the device's pseudonym and the metadata required to calculate risk scores. Scanning for and advertising these messages are done at different cycles, scanning at 15-20% of the time and advertising at 90-100% to conserve resources but also to ensure that devices register each other. In addition, devices that have communicated blacklist each other to ensure even distribution of devices noted.

Pseudonyms

The pseudonyms, called TempIDs in BlueTrace, are a part of the Encounter Messages. When the user registers with their phone number, the backend generates a unique, randomised UserID stored with the phone number at a central database. It is also possible to implement BlueTrace without requiring a phone number by automating the exposure notifications, but the standard is to use phone numbers. The TempIDs are created at the user device as shown in Figure 3.1. They are rotated every 15 minutes.

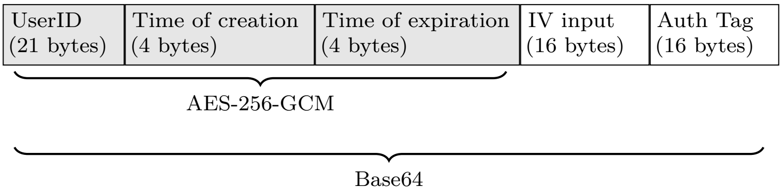


Figure 3.1: TempID (pseudonym) in BlueTrace

Metadata

In addition, the Encounter Messages contain metadata required to calculate risk scores. The messages consist of the Peripheral Device's TempID and device model, a code saying something about which HA is in charge of the system, and the BlueTrace protocol version. The Central Devices write the same back but with their own TempID and device model and an indicator of the signal strength of the message they received from the Peripheral. The messages are stored locally until a user is infected or at risk of infection.

In Case of Confirmed Infection

When a user registers as infected, the user will be asked to share the stored Encounter Messages with the Health Authority. A PIN or verification code is issued by the HA and sent with the relevant encounter messages to verify infection. The HA uses the code to verify the upload. Only the HA can decrypt the data collected at the server. When they do, the collected TempIDs will reveal the UserIDs associated with the phone numbers of the users that have been in contact with the index case.

Risk Score Calculation

When the HA has decrypted the information, the contact tracers look at the information collected. The signal strength, exposure time, and distance to the index case are used to see if the collected TempIDs should be notified. In Singapore, the information gathered by the app is compared to the information the index case gives on the phone to the contact tracers. The individuals seen as at risk are contacted directly.

BLE Limitations

Due to the limitation on Bluetooth running in the background of especially iOS devices, iOS users are encouraged to run the app in the foreground. This leads to higher power consumption, and if a user forgets to open the app before leaving their home, no contact tracing occurs. BlueTrace was used in Australia's CovidSafe until December 2020, but because of these problems, it has been replaced with another protocol called Herald Bluetooth Protocol [Aus20]. It improves Bluetooth performance for iOS devices by, for instance, using Android devices as a data sharer of iOS background beacons.

3.1.2 Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT)

Another centralised approach is the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) protocol [PEP20] created by researchers from different countries of Europe. It works similarly to BlueTrace and uses BLE, but the devices do not connect through a handshake. Instead, they continuously advertise and collect beacons. The pseudonyms are called Temporary IDs and are generated pseudo-randomly and changed frequently. The beacon is encrypted information that allows mapping to a persistent pseudonym of the device that temporary pseudonyms are derived from. The app collects a list of Temporary IDs of other users, which are uploaded to a central server in case of infection. Here, the risk score is calculated, and the Temporary IDs of high enough risk are decrypted, and the user is contacted by the HA. In PEPP-PT, the HA issues a one-time use verification token on verified infection, and the user can choose to upload their encounters.

PEPP-PT received criticism due to its centralised approach [Bus20]. An open letter published by scientists and researchers from more than 25 countries stated that governments could use such technology for surveillance. France and Germany created apps based PEPP-PT but both have since changed their systems due to criticism and limited uptake. The website *pepp-pt.org* has been abandoned, and it is not easy to find detailed information about the protocol.

3.1.3 ROBust and privacy-presERving proximity Tracing (ROBERT)

The ROBust and privacy-presERving proximity Tracing (ROBERT) protocol is a centralised protocol created as a proposal for PEPP-PT [Inr20]. The main difference is that the infected individual's identity and encounters are not revealed to the HA despite this being a centralised approach. The protocol uses BLE beacons to register encounters. For each device registered, the following data is stored at the server: an authentication key used to authenticate messages from the user, an encryption key used to encrypt information sent from the server to the specific user, the permanent user ID, a flag that says if the user has received an exposure notification, a time field that keeps track of if the user has asked for their status, and a when an encounter with an index case occurred. All time variables are stored as epochs from the time the system was created. To avoid automatic registrations, proof-of-work systems such as CAPTCHA are used.

Pseudonyms

When a user registers, they get a permanent ID and a set of pseudonyms, called Ephemeral Bluetooth Identifiers (EBIDs), from the central server that can only be linked together by the user or the HA. These are stored at the server, but the server should not know to whom these belong. The EBIDs that are given to a user at registration and later at a set interval are generated for each epoch i using the permanent ID and the server key.

Metadata

In addition to the EBIDs, ROBERT uses an encrypted country code that is, among others, created using the specific users EBID and a “federation key” that is shared among all servers in Europe. The HA uses the encrypted country code, and the HA can only decrypt it at the backend. The messages sent between devices can be seen in Figure 3.2. The MAC of A is the HMAC-SHA256 of the first three fields and the user’s authentication key.

Encrypted country code	EBID of broadcasting user	Time	MAC of A

Figure 3.2: Messages broadcasted in ROBERT

In Case of Confirmed Infection

Verification of infection is not specified in ROBERT and is up to the HA of a country to implement. If a user is confirmed infected, he or she uploads collected pseudonyms. Users ask if they are at risk of infection by submitting one of their pseudonyms to the HA. The HA then checks if the permanent ID is flagged as at risk. Since the pseudonyms are not linked to identity, the HA will not know whom they flag as at risk. Furthermore, the server is not supposed to store whether or not the flags occurred from the same user.

Risk Score Calculation

The risk score is calculated based on how many times a user's pseudonyms have been in contact with an infected user and how many pseudonyms are flagged as "exposed". Other parameters such as duration can also be added. If the score is higher than a set threshold, the user will be notified. How exactly the score is calculated and which values are used is up to the HAs.

3.2 Decentralised ACT Systems

In decentralised systems, most of the data is stored at the users' devices, and the central server works more or less like a bulletin board, distributing information to the users. Some examples of decentralised systems are presented in this section.

3.2.1 Decentralised Privacy-Preserving Proximity Tracing (DP-3T)

One example of a decentralised approach to contact tracing is the Decentralised Privacy-Preserving Proximity Tracing (DP-3T) protocol [TPH+20]. It was created as a response to the controversy that PEPP-PT created. The protocol contains three different variants so that developers can choose, one that is low-cost, one called unlinkable, and one hybrid. All three protocols are based on broadcasting of BLE beacons where the pseudonyms are called Ephemeral IDs, EphID. The EphIDs are stored locally together with an indication of when the beacon was received and measurements to calculate the risk score later. A backend server that is trusted not to change or remove data distributes information.

All apps also create dummy traffic to the server to protect the identity of the infected individuals that upload their information. Users ask the server for information regularly, and upon receiving, the EphIDs are reconstructed and compared to collected EphIDs. Based on the collected measurements associated with the EphID, a risk score is calculated locally.

Pseudonyms

In the simplest variant, the low-cost one, the derivation of the pseudonyms is based on a key that rotates daily. The secret day key ($\mathrm{SK_t}$) is computed by using a hash function on the previous $\mathrm{SK_t}$. What hash function is used is not specified, and therefore one can assume the choice is the HAs. The pseudonyms are then derived from this and changed at a specific time interval that can be different based on the needs of a country. The pseudonyms are generated at the beginning of the day and are 16-bytes. The order of which the generated pseudonyms are used is random. The IDs and $\mathrm{SK_t}$s are stored for 14 days.

Metadata

When the pseudonyms are collected, an exposure measurement, for instance, signal attenuation, and the day the beacon was received are stored at the user's phone. This constitutes the metadata used for risk score calculations. To save space, multiples of metadata for the same pseudonym are stored together.

In Case of Confirmed Infection

When a user tests positive, the user can upload relevant $\mathrm{SK_t}$s that were used when the user was contagious. The app also picks a brand new $\mathrm{SK_t}$ and deletes the uploaded ones. Using the $\mathrm{SK_t}$s, other users re-compute all pseudonyms and checks if any matches exist between the generated ones and the ones collected. For each match, the risk score is calculated based on the exposure measurement and time.

Risk Score Calculation

Each device calculates the risk score and checks if the score is above a threshold determined by the HA. How the score is calculated can vary between systems. For example, in Switzerland, a per-day score is calculated based on the exposure measurements of all matches that day.

The Other DP-3T Variants

Unlinkable DP-3T, is an extension of the simple variant that is supposed to make it more difficult for adversaries to link pseudonyms of infected users. Users can also decide not to upload specific pseudonyms used at specific times. In this case, the time intervals (i) have a fixed starting point that is shared among all users. The pseudonyms are derived using a 32-byte seed different from other pseudonyms. This means that the keys are more difficult to link. The smartphone stores the exposure measurement and the day but in a hashed string for each observed beacon. This differs from the low-cost version, where the pseudonym is stored in its raw form.

If a user test positive, the app uploads i and $\mathrm{seed_i}$ and excludes those the user wishes to exclude. The server creates a Cuckoo filter of each pair of time and seed every two hours. The users check if the filter contains any of their collected hashes. If so, the time and exposure measurement is received. The point of the filter is to decrease the chance of false positives while hiding the pseudonyms.

The last variant, the hybrid, combines the two above. Random seeds are generated for each time window, and they are used similar to how the low-cost design generates pseudonyms for all time intervals within it. Seeds are only uploaded if they are relevant to exposure estimation for others. By having a shorter time window, this design offers more protection against linking pseudonyms.

3.2.2 Google and Apple Exposure Notification (GAEN)

In April of 2020, Google and Apple announced that they would join forces to create an interface to provide optimised access to Bluetooth / BLE and make it easier to transmit beacons in the background [GA20]. Since then, the interface has become more than just an interface and can today be used as a functioning ACT system by itself. The HA can decide to implement parts of the system themselves, or use something called Exposure Notifications Express (ENE) [Goo]. ENE makes it easier for HAs to implement contact tracing systems by simply providing a configuration file to decide risk parameters and similar to Google and Apple, and they do the rest. If the HA wants to do more itself, it needs to create an app, a server backend and a verification solution. Google and Apple provide a reference framework for these parts. No complete source code of the parts Google and Apple control is available to the public, except for some sample and partial code for both the iOS and Android solutions that were released in July of 2020.

The main difference between GAEN and other similar decentralised approaches is that the functionality is implemented on the operating system layer and not the application layer. As mentioned, both Apple and Google restrict how app developers can use the Bluetooth technology, but with GAEN, approved app developers can use BLE in the background for exposure notification. The applications must meet set privacy requirements. All Android devices running Android 6.0 (Marshmallow) and all iOS devices running iOS 12.5 and higher support exposure notifications. In addition to allowing BLE in the background, the fact that it runs on the operating system layer makes it possible to hide data from apps on the application layer. Moreover, it ensures that the app will work on all, and between all, devices that run new enough versions of the operating system. It also makes cooperation across borders where countries use the framework easier.

In the standard, responsibilities of the system are shared the following way. In Figure 3.3, blue corresponds to GAEN's responsibilities, the red to the app's responsibilities, and the bolded black to the server's.

- *GAEN:* all key generation, derivation, and exchanging of RPIs. This data is stored locally on the phone within GAEN. The Diagnosis Keys are also used to derive RPIs and check for matches in GAEN. A rough exposure score is calculated based on the matches and their metadata.
- *App:* communicates between GAEN and the server. When data that will be used to calculate risk score is distributed, the app provides the to the GAEN framework. GAEN then provides exposure information that the app evaluates and decides whether or not to notify the user. When a user registers as infected, the app can also add an infectiousness score to each key, related to when the user is deemed most infectious.

- *Server:* collects and distributes keys of infected users, and calculates and provides exposure information. The server's role can be seen in the bolded black step in Figure 3.3.

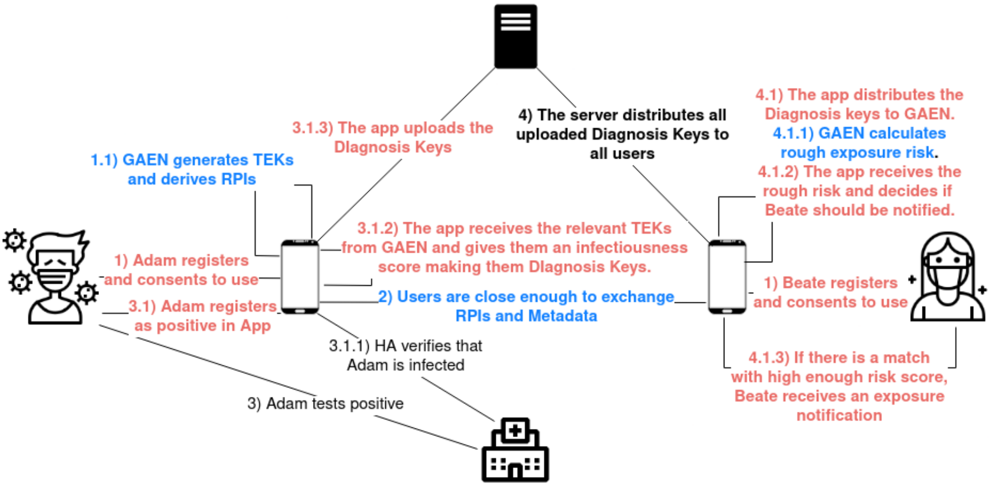

3.1.3) The app uploads the Diagnosis Keys
4) The server distributes all uploaded Diagnosis Keys to all users
4.1) The app distributes the Diagnosis keys to GAEN.
4.1.1) GAEN calculates rough exposure risk.
4.1.2) The app receives the rough risk and decides if Beate should be notified.
1.1) GAEN generates TEKs and derives RPIs
3.1.2) The app receives the relevant TEKs from GAEN and gives them an infectiousness score making them Diagnosis Keys.
1) Adam registers and consents to use
3.1) Adam registers as positive in App
2) Users are close enough to exchange RPIs and Metadata
1) Beate registers and consents to use
3.1.1) HA verifies that Adam is infected
3) Adam tests positive
4.1.3) If there is a match with high enough risk score, Beate receives an exposure notification

Figure 3.3: Flow of a GAEN based ACT system

ACT systems based on GAEN are often similar since much of the system depends on Google and Apple's framework. The solutions of different countries might differ in server infrastructure, how the app looks and what it offers of functionality, and if and how verification of infection is implemented. GAEN supports interoperability between countries, and a server called European Federation Gateway Service exchanges data between countries in EU and EEA [FHI21a].

Pseudonyms

The BLE beacons transmitted between devices with an app based on GAEN consists of a Bluetooth pseudorandom identifier, or Rolling Proximity Identifier (RPI), and encrypted metadata [GA20]. The RPIs are derived from a Temporary Exposure Key (TEK) that is changed at a set interval called the TEKRollingPeriod. The TEKs are randomly generated numbers of 16-bytes. They are stored together with their interval number for 14 days. The interval is 24 hours in GAEN.

The RPIs are derived as shown in Figure 3.4. For the key derivation function that derives the RPI key, SHA-256 is used. AES-128 is used to create a 16-byte RPI. The RPI is rotated ever 10-12 minutes, called an ENIntervalNumber that starts from the TEKRollingPeriod.

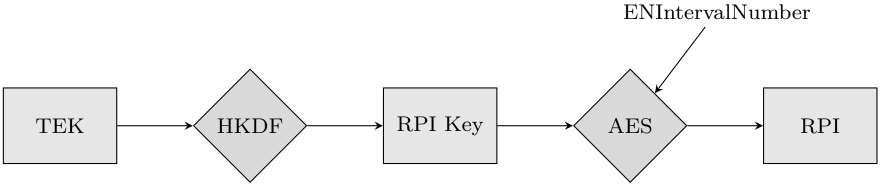


Figure 3.4: How RPIs are derived in GAEN.

Metadata

The metadata is called Associated Encrypted Metadata (AEMD), and it is encrypted in a similar way to the RPIs. AEMDincludes variables used to calculate risk scores, for instance, the Bluetooth signal strength. The metadata is derived as shown in Figure 3.5. AES-CTR is AES-128 in counter mode that outputs data of the same length as the input without padding, which allows the metadata to be smaller in size and not 128-bits. The RPI is used as the initialisation vector.

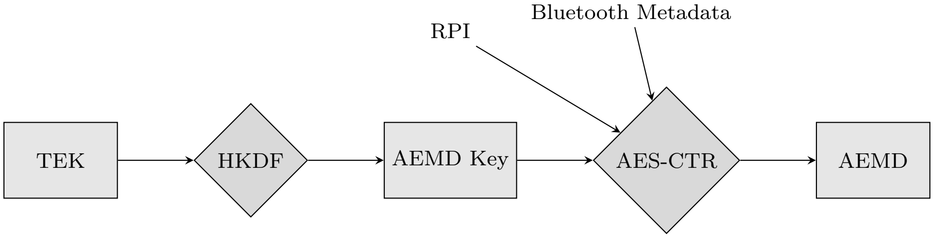


Figure 3.5: How the metadata is encrypted in GAEN.

In Case of Confirmed Infection

When a user tests positive and wants to upload their data, the TEKs of relevance and the interval number of that TEK are uploaded to the central servers. This data is then distributed to all devices from the central server. The uploaded TEKs are uploaded together with an infectiousness score set by the app and are from now on referred to as Diagnosis Keys.

To upload the data, the user should first verify infection. The solution to this varies among the countries' systems. In Google and Apple's developer guide, a recommendation can be found [Goo20c].

Risk Score Calculation

To detect risk of infection, new Diagnosis Keys are retrieved by the app from the backend. The keys are then given to GAEN which uses this to calculate Exposure

Windows that each represents one TEK. The Exposure Windows are calculated by deriving all RPIs based on the retrieved TEKs and checking for matches [Goo20a]. The HA can change parameters for how the Diagnosis Keys should be translated into exposure window data. One TEK can have more than one Exposure Window. Based on the Exposure Windows, GAEN can calculate the risk score, or the HA's app can retrieve the windows and do it themselves.

The Exposure Window defines the signal attenuation, the duration of contact in epoch time in milliseconds, the infectiousness score, a report type, and a list of the times RPIs based on that TEK is observed. Attenuation is calculated by subtracting the received power from the transmit power. The infectiousness score of a TEK is computed based on days since the symptoms started, or if no start date available, the test date. The report type corresponds to which type of test it is and is defaulted to "CONFIRMED_TEST". If the type is "RECURSIVE" it may be dropped since this is reserved for future use, and "REVOKED" does not lead to exposures.

If the HA wants GAEN to handle the risk score calculations, GAEN calculates a weighted duration for each Exposure Window by weighting the duration data for each encounter of the TEK. In addition, a score is computed. In the reference [Goo20e], the score is computed as follows in Equation (3.1). The different values are weighted based on risk of infection. High infectiousness will be weighted more than low infectiousness, for instance.

$$\begin{aligned} \text{RiskScore} = \text{reportTypeWeight}[\text{TEK.reportType}] \, * \\ \text{infectiousnessWeight}[\text{infectiousness}] \, * \, \text{weightedDuration} \end{aligned} \tag{3.1}$$

The Exposure Windows that have a score higher than a set threshold are summarised, and the one with highest value is stored separately. These will then allow the app to decide if the user should receive an exposure notification. The HA can choose to do the risk scoring mostly themselves by receiving the Exposure Windows before the score is calculated and use their own equations and weights to fit the country's requirements.

Terms and Conditions for HA

To develop a system with the GAEN framework, the owner has to be a government public health authority as seen in the terms and conditions [Goo20d][App20]. The system must also be used exclusively for contact tracing of COVID-19. In addition, it is not allowed for any app or HA to collect any identifiable information such as phone number or similar. Apps can also not ask for location permissions or other types of permissions. This is to try to prevent usage of the framework for purposes other than COVID-19 contact tracing.

The app can also not send data other than the Diagnosis Keys of consenting infected users to the central server. This means that centralised apps cannot use the framework. Apps that do not meet the set requirements are not allowed access to the API. They will therefore have more difficulty running BLE in the background. Technically, once given access, Google and Apple will not know if the HAs follow restrictions without checking in physically.

3.3 Comparisons of Solutions

Besides the decentralised or centralised nature, the differences are not that big between the solutions described in this chapter. BlueTrace distinguishes itself by using BLE handshakes while the rest use broadcasting to all. All the BLE-based systems struggle with background running of BLE and have to find workarounds for that, except for GAEN.

GAEN's implementation is similar to the low-cost version of DP-3T. Due to the BLE issues, the systems that exist based on DP-3T today, for instance, SwissCovid in Switzerland, often leverages GAEN. Therefore, DP-3T does not differ that much from GAEN. The main differences are that GAEN does not use dummy traffic and the filter of the Unlinkable and Hybrid versions of DP-3T.

The pseudonyms that are broadcast are created differently, but the general idea is the same. For instance, pseudonyms are generated from a time specific key. In BlueTrace, PEPP-PT, and ROBERT, a permanent key is stored centrally, and in all but ROBERT, this key is linked to the user's identity. In ROBERT, the pseudonyms are also generated centrally. The pseudonyms and time specific keys are generated locally at the user's device and not linked to identity in the decentralised solutions.

All the metadata relies on an estimation of Bluetooth signal strength and some indication of time and duration. In addition, the days the user is most infectious is usually accounted for. DP-3T only stores rough metadata from the start, only storing day and duration, while GAEN keeps most time data but at the operating system level.

Not all systems have implemented verification of infection solutions. Those that have often rely on the use of a token or a code. The time specific keys are used to check for matches, and then the metadata is used to calculate the risk score. Based on a set threshold, the users get an exposure notification.

It is good that there are choices to chose from based on the needs of a country. As seen in Table 3.1, of these solutions, GAEN is the most used. There are many reasons why this could be. Firstly, the fact that it is the one that works best with BLE. Also, it is easier to implement as many apps exist, which also allows for more effortless interoperation between countries. Because many centralised systems are criticised, the decentralised approach could also be a factor.

Name	**Centralised/ Decentralised**	**HA knows if at risk**	**Tech used**	**Countries**
BlueTrace/ OpenTrace	Centralised	Yes	BLE	Singapore, Australia (backwards compatible), Fiji, Nepal, Morocco
PEPP-PT	Centralised	Yes	BLE	Abandoned
ROBERT	Centralised	No	BLE	France
DP-3T	Decentralised	No	BLE	Switzerland, Estonia, Belgium
GAEN	Decentralised	No	BLE	Austria, Barbados, Belgium, Bermuda, Brazil, Canada, Cook Islands, Croatia, Cyprus, Czech Republic, Denmark, Ecuador, England and Wales, Estonia, Finland, Germany, Gibraltar, Iceland, Italy, Japan, Jersey, Kazakhstan, Latvia, Lithuania, Malta, Mongolia, Netherlands, New Zealand, Northern Ireland, Norway, Panama, Poland, Portugal, Republic of Ireland, Saudi Arabia, Scotland, Slovenia, South Africa, Spain, Switzerland, Taiwan, Uruguay, some states of the USA

Table 3.1: Overview of the ACT systems discussed in this chapter.

Chapter 4 How Smittestopp Works

The two versions of Smittestopp are examined in more detail in this chapter. The first iteration of Smittestopp will be referred to as Smittestopp v1, and the one currently in use as simply Smittestopp. Smittestopp v1 went beyond the contact tracing scope and was highly focused on data gathering for research. The new and in-use Smittestopp was developed with much inspiration from the Danish GAENbased Smitte|stop and is also based on GAEN.

4.1 Smittestopp v1

When the subject of an ACT system in Norway was introduced, FHI employed Simula Research Laboratory, an information and communication technology research organisation owned by the Norwegian government, to develop it. Smittestopp v1 used both GPS and Bluetooth, and collected data from every device with the app is uploaded automatically to a server at specific intervals. This data is used as both a tool for research and as a contact tracing aid. The data was encrypted and stored centrally for 30 days, and close encounters were supposed to be automatically notified. The user registered using their phone number for identification.

4.1.1 Contact Tracing in Smittestopp v1

The contact tracing part of Smittestopp v1 used the GPS location data to see where an index case had been and check if anyone had been nearby to see if they could be at risk [LLB+20]. BLE is used to estimate the distance between phones, and device IDs and signal strength are encrypted and stored locally on the user devices. The specifics were never figured out as Simula wanted to test the system before actually using it. When the collected location and BLE data is sent to the central server, it is deleted from the user devices. If the server does not hear from a user for more than a week, all data belonging to that specific user is deleted from the servers.

When a person tests positive for COVID-19 and has a phone number linked registered in the app, data from the server is fetched, and two graphs are produced, one with only GPS data and one with both Bluetooth and GPS. Nodes in both graphs represent users. In the BLE and GPS graph, edges represent contacts between two users and contain duration and a risk score. In the GPS-only graph, the edges represent trajectories. To account for the issues with running Bluetooth in the background of iOS devices, close encounters of the index case's close encounters are also of relevance.

Users can see their own data through *helsenorge.no*, and who has accessed it. If a user gets a notification stating that they are at risk of infection, the user can also check if this is legitimate by cross-checking with their data.

4.1.2 Smittestopp v1 as a Research Tool

Smittestopp v1 was meant to be a tool for both research and contact tracing. The research part of the system uses the graphs created as explained in Section 4.1 to look into how COVID-19 restrictions change movement patterns in the Norwegian population. The main goals for this research were to identify which government decisions have the highest effect on social distancing, give input to and further develop FHI's models for epidemics, and how diseases spread. When the app was first put into use, only the research part of the system was up and running to validate the system's usefulness and calibrate how the risk score is calculated to avoid false positives.

The collected data was anonymised and split into data sets of aggregated data. The data sets were aggregated at three different geographic levels and time definitions, and every user was placed in an age category. The geographic levels were regions with more than 200 users, more than 2000 users, and more than 50 000 users. The time definitions were one-hour intervals, 3-hour intervals, and days. The goal was to find a balance between geographic level and time to achieve adequate anonymity. However, when the app was released, this balance was not in place yet, meaning that the data was not adequately anonymised. The data sets can have two granularities, either with more details or less.

4.1.3 Criticism of Smittestopp v1

The app received criticism for its extensive collection of data and lack of transparency in developing within Norway [Gun20]. Amnesty International listed it as one of the most dangerous contact tracing apps for privacy [Amn20]. Claudio Guarnieri, Head of Amnesty International's Security Lab, stated the apps listed could be used as "highly invasive surveillance tools, which go far beyond what is justified in efforts to tackle COVID-19".

The Norwegian Data Protection Authority or Datatilsynet banned the app from collecting personal data on the 12th of June 2020 [Dat20]. The decision was based on the fact that the principle of data minimisation was not upheld. In addition, the benefit of the app was low, especially since no contact tracing had yet started and the anonymisation solution was not finished. They also criticised that users could not choose to opt out of one of the two parts of the system, meaning that the app's purpose was not adequately limited. Based on Datatilsynet's decision and all the criticism, FHI chose to scrap the system and start from scratch.

4.2 Smittestopp Based on GAEN

For the second iteration of Smittestopp, it was decided to use GAEN and focus more on privacy. Since GAEN should not be used if more data than BLE is collected, the research part of the app was scrapped altogether. Smittestopp v2 is made by a team consisting of employees from FHI, from Netcompany, a Danish Consultancy firm, and from the Norwegian Health Network (NHN) that operates *helsenorge.no*.

Smittestopp is divided into the following parts [FHI20d]:

- The *backend* which is hosted on Netcompany's data centres outside of Copenhagen. Here the keys of infected users are stored for 14 days, access tokens are verified, and the keys uploaded are distributed to user devices.
- The *app* checks if the user is exposed by calculating the risk score based on the exposure data it receives from GAEN.
- The *verification* part is a web service that distributes access tokens if a verified user (verified through ID-porten) has tested positive within the last two days and checks if uploaded access tokens are legitimate

4.2.1 The App and GAEN

In the app, the user initiates contact tracing by consenting to the terms and conditions. Only then GAEN can start creating TEKs and RPIs to broadcast, as well as start collecting nearby RPIs. The users should be over 16 years old of age, but there is no way to verify this at that point. This is because children under 16 in Norway need consent from their parents on issues regarding health and welfare [Hel18].

4.2.2 Verification

The Norwegian solution uses the MSIS health register and the common Norwegian identification portal, ID-porten, to verify that the person who wants to upload their keys are infected [FHI20d]. To register as infected in the app, the infected user must first agree to a new set of terms and conditions. The new agreement is done to ensure that the users understand what this registration entails. Then the user will be asked to log in through ID-porten.

Suppose a positive COVID-19 test result taken within the last two weeks exists in the MSIS database, the user is at least 16 years old, and the user has not verified their infection more than three times the last 24 hours. In that case, the user's app receives the date of the test and an access token that allows the user to upload the keys to the central server. The application also asks the user to insert the date the symptoms started to combine the keys with an infectiousness score. This combination makes a Diagnosis Key. If no date is inserted, the date of the test is used. The Diagnosis Keys are then uploaded to the central server. If there is no positive test result or the user is under 16 years old, the user gets a corresponding error, and no keys are uploaded. The flow is seen in Figure 4.1.

Some data is stored at the verification solution to check if users have not verified their infection more than three times in the last 24 hours. Each user gets a unique pseudonym from ID-porten, which should not be traced back to the user. Together with the number of times the pseudonym is used for verification, this pseudonym is stored for 24 hours. If users can upload their TEKs from many different devices, false notifications of infection could occur. However, users can verify infection three times every 24 hours to upload keys from a work device and a personal device.

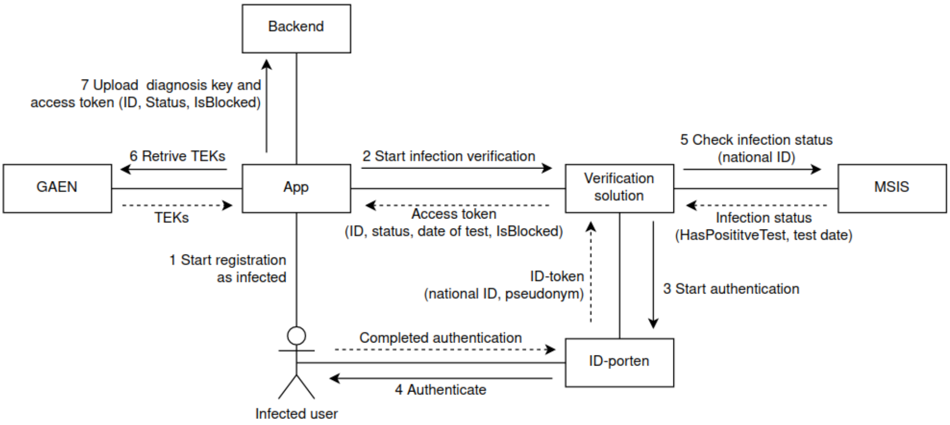


Figure 4.1: Verification in Smittestopp. Adapted from [Smi20]

Every verification is seen as users querying to access their records, and MSIS logs the verifications as personal queries to ensure traceability in where data from MSIS is sent.

Anonymous Tokens

The verification system, as explained above, has been criticised. For instance, if the verification and backend work together, it could be possible to link the ID-token to the user [Arv20]. This creates a theoretical possibility to make social graphs if FHI and NHN worked together with MSIS. Therefore, a new way to verify infection has been suggested, which is now implemented and backward compatible with the old way of verifying infection [Smi20]. Silde, Strand, and Moe, who created the new solution, won Datatilsynet's built-in privacy award of 2020 [Dat21].

Anonymous tokens are inspired by the Privacy Pass protocol [MSS20]. Privacy pass was created to make anonymous browsing easier by reducing the number of necessary internet challenges such as CAPTCHAs [DGS$^+$18]. Instead of having to do a new challenge every time, when a user does one challenge, they receive several anonymous tokens that can be used next time. This makes it easier to move around anonymously on the internet.

In the scope of contact tracing, anonymous tracing is used for verification of infection. When a user wants to verify their infection, the app uploads a message consisting of a hash and some randomness. MSIS validates the request and gives the date of the result. The message is then signed by FHI. When the user gets the signed message back, the application removes the randomness and sends the new signed message with the original message to the backend. The backend can then check if everything is correct. If the message is stored, the backend can ensure that the same token cannot be used more than once.

In the new implementation of the verification system, the following is the sequence of creating tokens in Smittestopp using point G on elliptic curve E [MSS20]. The numbers correspond to the numbers in Figure 4.2.

1. Public ($K = kG$) and private keys (k) are generated by FHI and private key k is distributed to the backend and verification system.
2. When a user wants to report as infected, the app generates a request $P = rT$, based on a random r and a new point T. T is generated using a random seed t and a hash function. The masked point P is then sent to the token generator.
3. The verification service then generates response $Q = kP$. A Chaum Pedersen zero-knowledge proof is used to prove that k was used to sign P. Q is sent back together with (c, z) used for zero-knowledge proof of correctness.
4. The app then checks the proof and removes r as T is masked by both r and k. Then the tuple (t, W) where $W = kT$ is generated and sent to the backend.
5. The backend generated T from t, and checks that W is generated from t by computing kt. t is stored so that no token can be used twice.

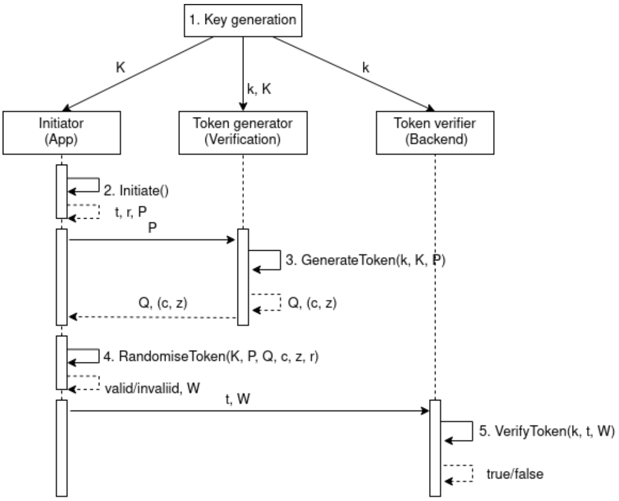


Figure 4.2: Anonymous tokens. Adapted from [MSS20]

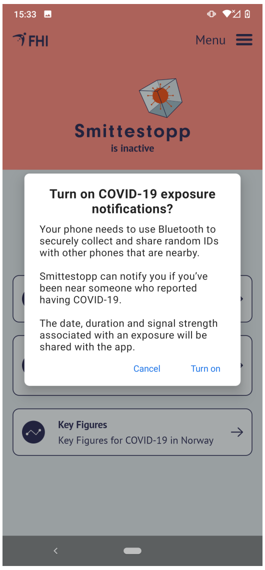
In addition to this, the tokens have public metadata that contains a timestamp, and the keys used are rotated regularly [SS21]. This is to prevent users from posting delayed keys used while the user had recovered from the disease.

4.3 Smittestopp In Action

When users download Smittestopp, they have to click through four pages with information about the app, scroll through and read the terms and conditions, and consent to it. Then, users have to respond to two prompts asking to turn the functionality on, one of which can be seen in Figure 4.3. Only then is the app activated as in Figure 4.4.

In device settings, the user can turn on and off exposure notifications, delete collected IDs and see how many times the app has sent the collected Diagnosis Keys to check for exposures and if any were found. This can be seen in the Figures 4.5, 4.6 and 4.7. In Figure 4.7, the number of keys is the total number of new Diagnosis Keys provided to GAEN from the app. None of these resulted in a match in the Figure, but if a user got a notification, it would be possible to check in these settings how many matches occurred. It is unclear if the number corresponds to the number of RPIs or the number of Diagnosis Keys.

To test Smittestopp, an experiment was run for this thesis. An Ubertooth One was set up at Gløshaugen, a university campus of NTNU, Trondheim. The Ubertooth One is a device that can be used for Bluetooth experiments that allow for capturing



15:33

FHI Menu

Smittestopp
is inactive

Turn on COVID-19 exposure notifications?

Your phone needs to use Bluetooth to securely collect and share random IDs with other phones that are nearby.

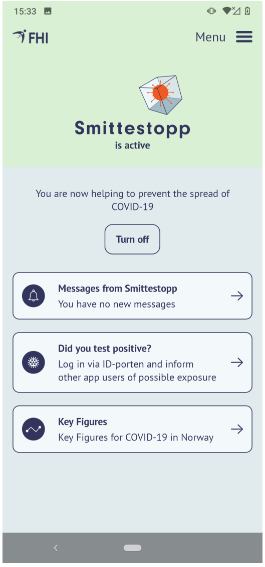
Smittestopp can notify you if you've been near someone who reported having COVID-19.

The date, duration and signal strength associated with an exposure will be shared with the app.

Cancel Turn on

Key Figures
Key Figures for COVID-19 in Norway

Figure 4.3: Activating Smittestopp



15:33

FHI Menu

Smittestopp
is active

You are now helping to prevent the spread of COVID-19

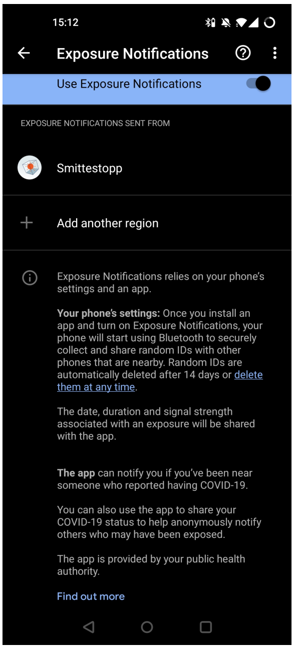
Turn off

Messages from Smittestopp
You have no new messages

Did you test positive?
Log in via ID-porten and inform other app users of possible exposure

Key Figures
Key Figures for COVID-19 in Norway

Figure 4.4: Active Smittestopp



15:12

Exposure Notifications

Use Exposure Notifications

EXPOSURE NOTIFICATIONS SENT FROM

Smittestopp

Add another region

Exposure Notifications relies on your phone's settings and an app.

Your phone's settings: Once you install an app and turn on Exposure Notifications, your phone will start using Bluetooth to securely collect and share random IDs with other phones that are nearby. Random IDs are automatically deleted after 14 days or delete them at any time.

The date, duration and signal strength associated with an exposure will be shared with the app.

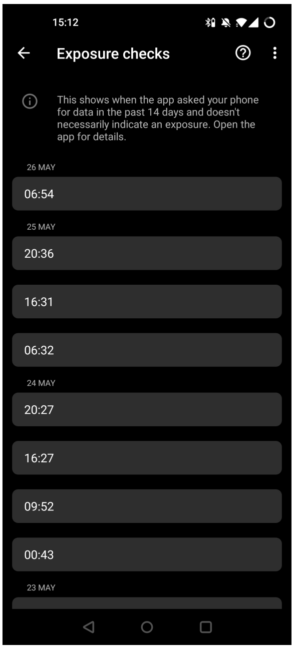
The app can notify you if you've been near someone who reported having COVID-19.

You can also use the app to share your COVID-19 status to help anonymously notify others who may have been exposed.

The app is provided by your public health authority.

Find out more

Figure 4.5: GAEN in phone settings



15:12

Exposure checks

This shows when the app asked your phone for data in the past 14 days and doesn't necessarily indicate an exposure. Open the app for details.

26 MAY

06:54

25 MAY

20:36

16:31

06:32

24 MAY

20:27

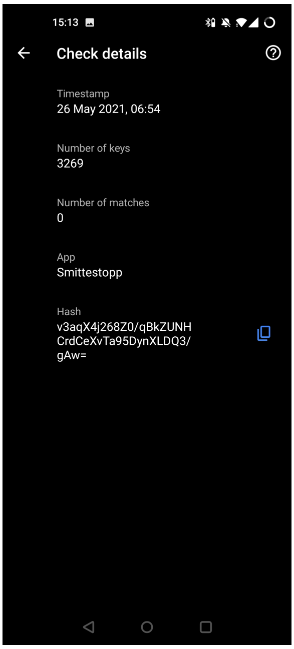
16:27

09:52

00:43

23 MAY

Figure 4.6: Exposure checks



15:13

Check details

Timestamp
26 May 2021, 06:54

Number of keys
3269

Number of matches
0

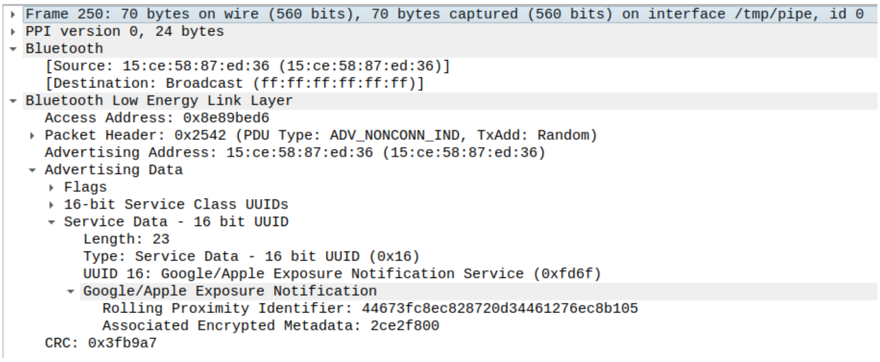
App
Smittestopp

Hash
v3aqX4j268Z0/qBkZUNHCrdCeXvTa95DynXLDQ3/gAw=

Figure 4.7: An example of an exposure check

BLE directly into Wireshark [Gre]. It collects BLE beacons at a larger range than a mobile device.

The data was filtered using three different filters. The `bluetooth.gaen` filter, which is only available in Wireshark 3.4.0 and later. It filters out GAEN beacons which contain an `AEMD` (for the metadata) and an `RPI` field that makes it easy to filter on both. The two other filters, `!_ws.malformed` and `!btle.crc.incorrect` were used to remove malformed packets and packets with checksum errors. There were a lot of these, which could be because of inaccuracy in the Ubertooth device.



```
▸ Frame 250: 70 bytes on wire (560 bits), 70 bytes captured (560 bits) on interface /tmp/pipe, id 0
▸ PPI version 0, 24 bytes
▾ Bluetooth
    [Source: 15:ce:58:87:ed:36 (15:ce:58:87:ed:36)]
    [Destination: Broadcast (ff:ff:ff:ff:ff:ff)]
▾ Bluetooth Low Energy Link Layer
    Access Address: 0x8e89bed6
  ▸ Packet Header: 0x2542 (PDU Type: ADV_NONCONN_IND, TxAdd: Random)
    Advertising Address: 15:ce:58:87:ed:36 (15:ce:58:87:ed:36)
  ▾ Advertising Data
    ▸ Flags
    ▸ 16-bit Service Class UUIDs
    ▾ Service Data - 16 bit UUID
        Length: 23
        Type: Service Data - 16 bit UUID (0x16)
        UUID 16: Google/Apple Exposure Notification Service (0xfd6f)
      ▾ Google/Apple Exposure Notification
          Rolling Proximity Identifier: 44673fc8ec828720d34461276ec8b105
          Associated Encrypted Metadata: 2ce2f800
    CRC: 0x3fb9a7
```

Figure 4.8: Example of a Smittestopp beacon.

An example of a captured GAEN beacon can be seen in Figure 4.8. Here it is possible to see the RPI and metadata fields. The numbers are the result of the outputs of the process shown in Figure 3.4 and 3.5. The RPI is as expected 128-bytes, and the metadata is smaller.

Chapter 5
Security Analysis of GAEN

GAEN was created to be more privacy-preserving and solve problems regarding the usage of BLE in background-running applications. Many countries have adopted systems based on the GAEN framework, and the framework and Google and Apple have to be trusted. However, some privacy and other concerns exist. Concerns not related to privacy are also discussed.

5.1 Confidentiality

For confidentiality, the biggest concern is the privacy of the users. Attacks on the privacy of ACT systems usually have one of two primary goals: finding out who is infected or establishing social graphs and movement patterns of people. The attacks have been found through a literature study, and the most relevant attacks are explained in this section. The main source of inspiration is Vaudenay and Vuagnoux's research on the GAEN application in Switzerland, SwissCovid [VV20]. Some of the attacks have been tested by Baumgärtner et al. [BDF^{+}20]. The research is not conducted on Smittestopp, but as the fundamental principles in a GAEN based system are the same, the same issues are likely to be found in other countries' solutions.

In Table 5.1 an overview of the discussed privacy attacks can be seen. The last attack, the "time identification attack" is not explained in this section because it is based on an integrity attack, and it will be investigated in Section 5.2.

Short name	**Description**	**Goal**	**Difficulty**	**Impact**
Educated guess	Use encounter information to try to find the infected	Identify individual	Low	Low
Paparazzi attack	Collect only one target's RPIon a device	Identify individual	Low	Low
Extended paparazzi attack	Collect RPIs while noting other data	Identify individuals	Medium	Medium
Strategic devices	Place devices at strategic places to collect location data	Movement patterns	High	High
Little Thumb	Collect and link RPIs that rotate out of synch	Movement patterns	Medium	Low
Using Graphs to Deanonymise Individuals	Trying to find information on individuals from the graphs	Identify individuals	High	High
Time Identification Attack	Moving user into future or past to trigger the use of the same RPI	Identify individuals	High	High

Table 5.1: Overview of confidentiality attacks

5.1.1 Identifying Infected Individuals

If the RPI reveals a user's identity, it would be easy to determine who is infected. The RPIs are derived from the TEKs that are generated as random numbers and are not supposed to be linked to users.

Educated Guess

In the "educated guess" and the "extended paparazzi" attack, the attacker needs to collect the RPIs and Diagnosis Keys outside of GAEN. This is because the details of encounters are kept at the operating system layer in GAEN, and only coarse encounter information such as day and duration is revealed to the app. As mentioned in Section 3.2.2, it is not allowed for the HA to create applications that collect more information than BLE beacons of nearby devices and distributed Diagnosis Keys. Nevertheless, nothing stops a user from noting, for instance, the location of their device when a beacon is collected outside of the official app.

When a user consents to upload their keys when infected, the keys are supposed to reveal which RPIs can be derived from the same key. Therefore, one Diagnosis Key links 24 hours worth of RPIs. This means that each device will know how many RPI rotations they have been with the index case based on how many of the RPIs they collected that are derived from the same Diagnosis Key. This information alone is not easy to link to identity. However, if users are social distancing, they might have a better idea of who they have been with for an extended time, or at least whom they have not. If the specific time of the encounter is revealed, the narrowing down might be more straightforward. In GAEN the specific time is included in the metadata that is decrypted at the operating system layer, but it is not revealed to the application.

The "educated guess" attack base itself a lot on guessing, and an attacker will not know entirely if they guess correctly. Most likely, the target is known to the attacker, and other sources of information can therefore be just as telling. For instance, the information given from manual contact tracers could indicate time without the data from the app.

Paparazzi Attack

It could be possible to figure out who is infected by restricting the anonymity set and targeting a user without needing to collect data outside of GAEN. One attack is referred to as the "paparazzi attack" [VV20]. Suppose an attacker ensures that only BLE beacons from one target are collected on a device. If the target is later infected and uploads their keys, the device that collected only the target's beacons will be notified. Since the device only has one option, the attacker will know that the target is infected.

This sort of attack is difficult to mitigate, and it is not illegal. The consequences can be severe for the targeted individual, but the attack is not very scalable as it would require a large number of devices to track many people. It is also challenging to ensure that only one person's RPIs are recorded, and more RPIs will likely be collected. One way to record the RPIs of only one person is to turn on the exposure notifications when no other people are around. Since no data is collected outside of GAEN, it is also necessary to be around the person for more than 15 minutes in order to trigger the exposure notification.

Extended Paparazzi Attack

Suppose that RPIs are collected outside of GAEN while noting the time, location, and who is nearby. This information is then stored associated with the specific RPI. If the Diagnosis Keys then derive any of the collected RPIs, the other information linked to the RPI could deanonymise the infected individual. If a device is set up solely for this purpose, collecting BLE beacons and registering people arriving (for instance, using surveillance cameras), it could be possible to at least narrow down who could be infected if said device receives an exposure notification. The attack will be referred to as the "extended paparazzi" attack. It is more scalable but not the most accurate, and the approach described here needs an extent of guessing to be successful. However, with enough time and money, it would be relatively easy to set up, as demonstrated by the small experiment used to test Smittestopp in Section 4.3.

With this approach, if an attacker could isolate the beacons of a target and note them separately, one collected and stored RPI would be enough to know that an individual that has been in contact with the device is infected if they were infectious at that time. Isolating targets could be done by checking which RPI most likely belongs to the person by looking at the Bluetooth signal strength and the surveillance video footage or by isolating targets with the collecting device by, for example, leaving them alone in an office. Since data is collected outside of GAEN and the Diagnosis Keys can be collected too, only one RPI per target is necessary. A file over interesting targets could be made, and if distributed Diagnosis Keys derive any of them, the attacker will know who is infected. Thus, more people can be affected by this approach, and it is more scalable. However, as in the basic paparazzi attack, it will require some extent of guessing.

5.1.2 Using Information to Monitor Movement Patterns or Create Social Graphs

If it is possible to link one user's RPIs together and location data is known, it could reveal the movement patterns of users. This can also be used to create social graphs if combined with other user's data. Additionally, even if the identity cannot be directly derived from the RPI or the observation of the RPI, an individual's movement pattern could reveal a lot about the individual's identity.

Strategic Placement of Tracking Devices

The most basic way that RPIs are linked is through the distributed Diagnosis Keys. Using additional information will make the possibility to link 24 hours worth of infected RPIs more valuable. If devices are placed strategically across a city and collect data, movement patterns of infected individuals could be easy to obtain. This attack is shown by Baumgärtner et al. [BDF+20]. Strategic places could be at major access roads to cities, industrial areas where many people work, access roads to suburbs or other areas where many people live, or at public transport stations. This can lead to systems that show movement patterns of all infected individuals within the monitored area. If many infected individuals are caught in the system, this could also create social graphs.

Little Thumb Attack

To track individuals that are not reported as infected, other ways to link RPIs are necessary. Since GAEN operates at the operating system layer, key scheduling is defined at this layer. That means that for GAEN based systems, it should be easier to rotate RPIs and operating system based variables, such as the Bluetooth Device Address, simultaneously. The Bluetooth Device Address is a unique address similar to Media Access Control (MAC) addresses. This adds security as it decreases the chance to link consecutive RPIs if the RPI and the Bluetooth Device Address rotate out of sync. The Bluetooth Device Address is changed every 10-12 minutes as the RPIs. However, if they are not rotated simultaneously, it could be another way to link RPIs. Vaudenay and Vuagnox [VV20] found that some rotating was out of sync in some phones. This can lead to a "Little Thumb" attack.

An example is shown in Table 5.2. Here, a listener that sees all three columns will see that two Bluetooth Device Address', BD_Addr1 and BD_Addr2, have the same RPI, RPI1. This indicates that the two addresses belong to the same device, and therefore that RPI2 also belongs to BD_Addr1.

Bluetooth Device Address	**RPI**
BD_Addr1	RPI1
BD_Addr2	RPI1
BD_Addr2	RPI2

Table 5.2: High level example of the Little Thumb bug

If this bug is widespread enough and often occurs, linking RPIs without the keys is trivial. Google and Apple are aware of the bug but cannot do much about it except ensuring that it rarely occurs. Android lacks the functionality to notify applications that the Bluetooth Device Address is changing or has changed [Goo20b]. Therefore when a new RPI is created, the advertising has to be stopped and restarted. There is, therefore, a possibility that the Bluetooth Device Address has already changed before the new RPI is generated.

Using Graphs to Deanonymise Individuals

The graphs created by these devices can show important information about individuals. For example, the strategically placed devices could reveal where an individual work, lives, exercises, and much more based on their placement. That information alone could reduce the anonymity set and narrow down the list of possible infected people. In addition, if the devices combine the collected data with, for instance, surveillance footage as in Section 5.1.1, the system could be further extended to trace individuals.

Using Linking of RPIs to Create Surveillance Systems

With many enough strategically placed devices, it could also be possible to follow a person's movements for the duration of the RPI rotation. Combined with trying to guess whom a specific RPI belongs to, this can be used to trace a person within the area with devices by looking at where the same RPI shows up next. This sort of tracking would require many devices as the range of BLE is limited. The same can be achieved if many app users collect data outside of GAEN on their devices and upload the data to a shared server.

5.2 Integrity

When it comes to integrity, the main issue is if it is possible to create false exposure notifications. Such attacks could be made to prevent a specific person from attending events, prevent many people from attending events, or create mistrust in the system and the government. The integrity concerns will not be discussed in detail. Most of the attacks are from Vaudenay and Vuagnoux's research [VV20], and from research they have conducted together with Iovino [IVV21]. In Table 5.3 an overview of the discussed attacks can be seen.

Short name	**Description**	**Goal**	**Difficulty**	**Impact**
Simple Targeted Attack	Collect and link RPIs that rotate out of synch	Send a notification to a specific person	Low	Low
Lazy student attack	Rebroadcast RPIs that are likely to be reported	Prevent event or limit group's movement	Medium	High
Wormhole attack	Rebroadcast many RPIs that are likely to be reported to many places	Prevent movement or event. Mistrust	High	High
Simulated GAEN	Rebroadcast many RPIs that are likely to be reported to many places	Prevent movement or event. Mistrust	High	High
Inverse-sybil attack	Many users pretend to be the same user	Prevent movement	High	High
Time attack	Move device back in time to trigger old notification	Prevent movement	Medium	High

Table 5.3: Overview of integrity attacks

Simple Targeted Attack

A simple targeted attack can be done by using a diagnosed user's device before they upload their Diagnosis Keys, get close to a target and upload the keys after the target has collected enough RPIs to be notified. Another solution is to get a diagnosed person to verify infection on a device that the target has been close to. This will create an exposure notification at the target's device and can be used to prevent someone from going somewhere. A black market where Diagnosis Keys are sold before they are uploaded to the server can also be created for this. Then RPIs can be derived from the keys and transmitted to more people before the keys are uploaded. In addition, devices only download Diagnosis Keys a few times per day, so an attacker could be able to derive RPIs from distributed keys and transmit them to the target before they download them.

If the user adds an antenna to their device to increase the range of the broadcast, this type of attack could involve more people. That would require changing the way the signal strength looks so that it seems that the device with the antenna was within two metres.

Simple Replay Attack / Lazy Student Attack

Another simple attack is to collect RPIs from a place with high infection rates and rebroadcast them elsewhere. This is a type of replay attack, and Vaudenay and Vuagnox refer to this as the 'lazy student attack' as it could be used to try to put a class in quarantine to avoid having to sit an exam [VV20]. In this case, the metadata must be changed to occur as a match at the user's device. In GAEN, an encounter is verified by comparing the metadata and checking if the match was sent and received within the same time interval. However, the metadata is not authenticated and could therefore be malleable.

Wormhole Attack

Baumgärtner et al. [BDF^{+}20] showed a large scale attack of the above mentioned kind. Here they used a wormhole attack to re-transmit RPIs to many places. The metadata used to calculate risk scores is possible to alter since the data is not authenticated.

Simulated GAEN

Using devices that broadcast fake RPIs derived from fake TEKs and later reporting the TEKs, could lead to many people getting a notification. A device can also here transmit with high power but fake the metadata to seem closer at broadcast time.

Inverse-Sybil Attack

Another attack is referred to as the Inverse-Sybil attack [ACK^{+}21]. In this type of attack, many users pretend to be the same user by using the same TEKs to derive RPIs. If one of the participants then tests positive, many devices receive an exposure notification without actually having been exposed since they have been in contact with RPIs from the same Diagnosis Keys as uploaded.

Time Attacks

If it is possible to trick devices into thinking that they are in the past, it could be possible to trick devices into thinking TEKs were valid when they should not be. For example, that could allow an attacker to transmit old RPIs from already reported TEKs. Then, when the device is back in present time, the Diagnosis Keys corresponding to those RPIs would result in a notification.

The most straightforward time corruption attack is to set the time back physically with access to the device. It is also possible to use a rogue server and ARP-spoofing to get the NTP (Network Time Protcol) to synchronise the clocks of devices connected to the same WiFi. A rogue base station that can send fake NITZ (Network Identity and Time Zone) messages, usually used to tell a phone that they have changed time zones, can also be used.

The time must be changed for 15 minutes since that is how long the duration should be before a notification is sent in most countries, but Iovino et al. suggest speeding up the process by moving time so that GAEN thinks that a signal was missed and sends a new one faster.

Iovino et al. [IVV21] tested many of these time attacks in different implementations of GAEN apps, including the Danish Smitte|stop, which was developed by the same company as the Norwegian app, Netcompany. To mitigate these types of time attacks, the devices' clock should be more secure.

Using Time Attacks to Identify Individuals

Iovino et al. [IVV21] also point out a way to figure out to whom a device belongs by using the possibility to move a device in time. The exploit that is used to do this is that new TEKs are generated when the date is modified or at midnight. If a TEK has already been generated for the day, it will be reused. That can be exploited by sending the device back to a specific time when the RPIs used are known and checking if the RPIs show up. If so, the device is present. This could be used to get the users to announce themselves every time they arrive somewhere.

In addition, if an attacker sends the target into the future, the attacker can note the RPIs that will be used at the actual time in the future. This can be used to see if a user attends an event when the date comes.

5.3 Google and Apple

Google and Apple control all key generation, derivation, and the first metadata handling. The TEKs are to be generated randomly and not be possible to link to the user, and the metadata should not be stored. Since there is no available source code, users have to trust that Google and Apple oblige by this.

Google and Apple have made themselves indispensable, especially for iOS devices [Hoe20]. It is challenging to create an ACT system based on BLE without following Google and Apple's terms. As mentioned before, countries that use DP-3T also use GAEN. Therefore, DP-3T based systems are reduced to additional measures on top of the framework. For instance, DP-3T is used in Switzerland's SwissCovid, but it also uses GAEN. This means that Google and Apple still set the premises and decides much of the functionality. GAEN is similar to DP-3T, but it takes away the implementation of a large part of the system from the developers of DP-3T's control.

Some solutions use BLE without Google and Apple, but that requires additional measures. TraceTogether and COVIDSafe are examples of successful apps that have tried to create workarounds to the BLE issues. They are centralised, which makes it easier to find issues and test configurations since the data can be collected centrally and cross-checked with the information contact tracers collect. This way, the system can be tweaked, and the BLE issues can be met with better solutions. Decentralised variants might struggle more to work around the issues and might have to compromise with Google and Apple as SwissCovid and DP-3T has done.

In addition, it might make countries that do not want to use GAEN more likely to choose a GPS based solution. For instance, the background running issues were one of the reasons why Smittestopp v1 used GPS as well as BLE. It also sets the premise on what type of BLE based contact tracing should be used and might hinder other solutions from being created. This means that other BLE based implementations that could work better or be more privacy-preserving might be missed.

5.3.1 Operating System Layer

The fact that GAEN works at the operating system layer means that a user cannot simply uninstall the contact tracing app to remove the functionality [Hoe20]. The fact that intrusive systems such as this should only be used while necessary is a vital privacy measure. In GAEN systems, a user cannot simply uninstall an app as the code is within the operating system. All newer smartphones have the necessary functionality coded into the operating system. With ENE the users do not even need an app. They only need to consent within settings. Users can, therefore, turn the functionality off but not remove it.

5.3.2 Access to Data

In April 2021, Reardon of AppCensus found that GAEN logs important exposure notification data in system logs of Android devices [Rea21]. These logs are available to many pre-installed, privileged apps. Both broadcast and collected RPIs were logged, together with the Bluetooth Device Address of the collected RPI. The apps in question are apps that hardware developers have pre-installed and have the READ_LOGS permission.

The data logged also allowed the apps with access to identify the individual in question as the apps could access phone numbers and similar. In the log, Reardon found a message stating that no exposure was found, which could indicate that if an exposure risk existed, that would be logged as well. If not, the lack of ‘no exposure’ message could indicate exposure. The information could also be used to create social graphs since it could be collected from many devices by apps with privilege. Google has since stated that the issue is fixed [NTB21].

5.4 The HAs

The terms and conditions should hinder anyone but HAs from creating apps that use the GAEN framework. Google and Apple must therefore be trusted not to open this functionality up to anyone else or use it for other things themselves. In addition, the HA has to agree on not collecting other data and not storing the data centrally. However, there is nothing technically stopping HAs from using GAEN to create centralised surveillance systems, as noted by Hoepman [Hoe20]. It is mainly based on trust. GAEN must trust the HA, and GAEN is trusted to follow up and check that restrictions are followed.

The HA must be trusted not to create an app that sends the result of the exposure checks back to the server together with an identifier of the user. If information is collected this way, the HA could also target groups of people by using a specific TEK and if the TEK results in an exposure, send the result back to the server. Then the HA will know that the person has been in contact with that TEK. This can, for instance, be used in police investigations to uncover crime networks, or it can be used to target specific ethnic groups.

In addition, how the HA defines the data broadcast is essential for what is legal or not. If the RPIs are not defined as personal information, it will not be illegal to set up the surveillance systems previously explained [VV20]. Many citizens could be tracked this way in areas with high infection rates and a high adoption rate of the app. If it is possible to find other ways than the “Little Thumb” bug to link RPIs outside of the Diagnosis Keys, or if that bug is widespread enough, all other users could also be possible to track.

Since the functionality is at the operating system layer, the GAEN framework could be extended when not needed in a pandemic situation. For instance, the collection attacks mentioned in previous sections can be made after a pandemic to see who is at a place. Governments can use this to target specific groups of people.

In January this year, Singapore confirmed that their TraceTogether based on BlueTrace would allow law enforcers to use the data to aid criminal investigations[Yu21]. This was announced after approximately 78% of the population had already downloaded the app as it was made mandatory in September 2020. Moreover, the Australian CovidSafe has had an instance where an Australian spy agency collected COVID-19 contact tracing data from the app. However, according to the Australian government, this was not on purpose, and the data would not be decrypted and used for anything [Whi20]. Both of these systems use a centralised approach where more data is stored centrally, and the data is, therefore, easier to exploit. Nevertheless, it shows how BLE contact tracing data can be misused for purposes other than contact tracing.

Chapter 6 Practical Consequences in Smittestopp

The work after Smittestopp v1 was focused on increasing the privacy of the solution. FHI conducted a Data Protection Impact Assessment (DPIA) [Knu20a] and a risk assessment [Knu20b] before releasing the system. The overall risk of using Smittestopp is seen as "acceptable" based on these papers.

Threat actors that might be interested in information about the Norwegian population could include foreign intelligence. In the risk assessment [Knu20b], China and Russia are pointed out as especially interested in Norwegian interests. As seen in recent events, also US Intelligence could be interested, and since the servers are located in Denmark and US intelligence have cooperated with Denmark, this could be of concern [GSDL21].

FHI concludes in the DPIA [Knu20a] that the Diagnosis Keys and RPIs are personal information. This could mean that attacks that try to deanonymise users using the RPI or attacks that collect many RPIs are considered illegal in Norway. However, the data is broadcast and picked up by many devices other than devices with the app anyway, and it will not be easy to prevent and discover attacks.

In this chapter, the research questions stated in the introduction will be explored and, among some other concerns, discussed in the context of Smittestopp.

6.1 RQ1: Is the privacy of infected users conserved in Smittestopp and other GAEN based systems?

For the first research question, this thesis has investigated privacy attacks that try to identify infected users. The attacks explained in Section 5.1.1, "educated guess", "paparazzi", and "extended paparazzi", are not the most difficult to do, but the consequences are smaller. The attacks are not entirely accurate as they rely on an extent of guessing. Furthermore, the attacks are challenging to scale, maybe except the "extended paparazzi" attack, and they mostly affect the individual. As mentioned, other factors outside of the ACT system could be used to figure out who is COVID-19 positive so that issue will persist even without an app.

These attacks can be mitigated if apps outside of GAEN cannot use the Diagnosis Keys to derive RPIs. For instance, when pointing out the same issue in DP-3T, Vaudenay [Vau20] suggests encrypting the secret day keys with a rotating key shared among the apps. The secret day keys correspond to the TEKs in GAEN. Since GAEN and DP-3T are similar, this solution is possible for GAEN as well, but as Vaudenay points out, it would require infrastructure that is difficult to deploy.

The "time attack" exploits an integrity concern to deanonymise users. It might be more challenging to do than the previously mentioned attacks, but it is more scalable within a network. As mentioned, it could be mitigated by securing the device's clock, but that is difficult for the developers of Smittestopp to do anything about.

The anonymous tokens introduced in Smittestopp remove the possibility that the backend and verification server cooperate to figure out who is infected. However, people who use older versions of the app will not have this functionality. It was also noted in the paper explaining the implementation that a dishonest verification service could use special keys to track users [SS21].

Another source of concern could be the traffic between ID-porten and the application in the Norwegian Smittestopp. If an eavesdropper listens in on the traffic, they could possibly identify who is trying to verify their infection and, therefore, who is infected. Telecommunication companies could be able to identify a lot of COVID-19 positive individuals that way. In DP-3T they use dummy traffic to avoid this.

For the backend, FHI sees the risk that people with access to the servers will modify the Diagnosis Keys as low [Knu20b]. This is because few people have access, and they are trusted. However, it could be a concern that Danish intelligence has cooperated with the US to spy on Norwegian politicians if they had decided to include the traffic to and from the servers. The data could then possibly be used to identify individuals who upload their keys to the server.

6.2 RQ2: Can ACT systems based on GAEN be used as surveillance systems to monitor user's movement patterns and social circles?

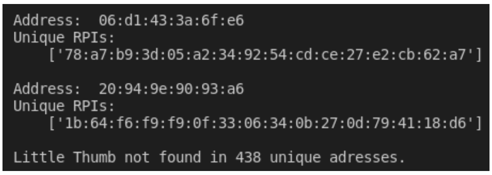
For the second research question, attacks that try to monitor users or create social graphs from contact tracing data are investigated. The attacks explained in Section 5.1.2 that try to do this require time and resources to complete. The devices collecting RPIs could be like the Ubertooth One used in the test of Smittestopp in Chapter

4, but more likely other small Bluetooth receivers could be used. Therefore, the setup can be easy, but to place them without noticing might be more difficult, time-consuming, and, depending on the devices, expensive.

In small towns in Norway, it could be easier to place devices in a valuable way as the geographic area is smaller. Furthermore, fewer people would reduce the anonymity set, and, therefore, it could be easier to use the data to deanonymise individuals. However, as the easiest way to link RPIs is through the Diagnosis Keys, individuals that can be traced are individuals who have tested positive for COVID-19 and uploaded their keys. Therefore, the infection rates have to be high for the attack to give much value.

In the unlinkable version of DP-3T, the pseudonyms are derived from a new seed every time, meaning the pseudonyms cannot be linked even when a user uploads the seeds. This would make it more difficult to find movement patterns and social graphs between infected individuals. However, it would require more data to be transferred to and from the backend server. In the low-cost version of DP-3T and in GAEN, only one Diagnosis Key per day is necessary to upload, while in this approach, it would require one seed for every pseudonym.

The "Little Thumb" attack could make it possible to link RPIs of healthy people if it is widespread enough. To test if the bug could be found in Smittestopp, the experiment in Section 4.3 was extended. Beacons were captured for between 2 to 6 hours for 7 days resulting in a captured amount of 1590 unique addresses. The capture files were converted to JSON files using Tshark and run through the code found in Appendix A.1. An example of the output of this script can be seen in Figure 6.1. No evidence of the bug was found on any of the days the experiment was run.



```
Address:  06:d1:43:3a:6f:e6
Unique RPIs:
    ['78:a7:b9:3d:05:a2:34:92:54:cd:ce:27:e2:cb:62:a7']

Address:  20:94:9e:90:93:a6
Unique RPIs:
    ['1b:64:f6:f9:f9:0f:33:06:34:0b:27:0d:79:41:18:d6']

Little Thumb not found in 438 unique adresses.
```

Figure 6.1: Example of output of script that looks for Little Thumb bug

However, the experiment had some limitations. When running the Ubertooth software, it frequently crashed, which seemed to be a bug in the firmware. Every session, therefore, resulted in many capture files separated by some seconds. This means that data that could have shown the bug was lost every time. A bash script was used to merge the files using `mergecap`. Also, many packets were malformed or had checksum errors, resulting in packets of interest being lost. Therefore the bug could have occurred but not been noticed.

In addition, the captures occurred with fewer students visiting Campus due to the pandemic. The range of BLE beacons is also not that high in reality, as mentioned in section 2.4.1. There were two devices with Smittestopp installed close to the receiver and at least three people in the same room with the app active. Other students walk past and study nearby, but most of the beacons were most likely from the same devices. This limits the diversity of devices and decreases the amount of data collected.

The implementation of anonymous tokens increased the privacy of Norwegian users of Smittestopp, but the implementation could be attacked by a dishonest verification service [SS21]. The verification service could give special keys that could be used to track individuals. This attack is detectable if the users share their public keys to ensure that they are consistent.

6.3 Other Factors

Overall, the confidentiality risk is seen as small in the DPIA conducted by FHI since little sensitive information is stored, none of which is stored centrally until someone chooses to upload their Diagnosis Keys. Furthermore, since the server is supposed to distribute the Diagnosis Keys, the confidentiality risk here is also seen as small.

In addition to confidentiality concerns, the following is also of interest when evaluating the Norwegian Smittestopp.

6.3.1 False Reports of Infection

The integrity risk is seen as small in the DPIA and it is seen as unlikely that anyone can manipulate the data, and few people have direct access to it.

The consequences of many false reports could be significant. Threat actors, such as foreign intelligence, could want to influence the Norwegian population's movement patterns and activity level. Politicians and other people in positions of power could be especially vulnerable. Actors could try to hinder people from voting, for instance, or it could simply be used to weaken the trust of the Norwegian population in the government and democracy. In the context of a pandemic, a lack of trust in the government could mean that the population could stop listening to the government when new measures to contain the virus are put in place, leading to higher infection rates and hospitals reaching capacity.

In Smittestopp, another integrity concern is that a user can upload keys from three different devices every day. It could lead to a variation of the "lazy student attack" if, for instance, three students ask a verified COVID-19 positive individual to register as infected on their devices. This would require that the infected person got physical access to log into ID-porten on their devices. In that case, there is a chance that all students in their class with the app receive an exposure notification, even if the infected person never attended the class. Since the number of people who are notified is small, the consequences are most likely to be small, but again it could reduce the trust in the app for groups of people affected by it.

6.3.2 Google and Apple

Many of the attacks exploit GAEN, and it is Google and Apple's job to ensure that they cannot be misused. The developers of Smittestopp will therefore not be able to make much of a difference. This takes away control from the developers and the HA. For instance, the issue that Google logged exposure notification data in the systems logged was, according to Reardon, communicated to GAEN before the issue was released. However, a solution was not prioritised until the issue got media coverage 60 days later [Rea21]. The same issue could possibly have been prioritised earlier by the developers of Smittestopp.

The main problem with the fact that Google and Apple are in charge might be transparency. Since only limited source code is released, there is no way for others to quality check the code or come with ideas for improvement. For example, Smittestopp's open-source code allowed Slide and Strand to suggest the improvement of adding anonymous tokens, and one could wonder which improvements could be added to GAEN if that was also available.

ACT systems are hopefully only necessary for a limited time. To ensure purpose limitation, the functionality should be removed after the pandemic is over. It could still be available for a possible future epidemic, but it is not at any point to collect the data without the pandemic. That could increase the chance that the data could be misused. However, it makes it harder for attackers while the system is necessary to obtain data, which increases integrity and confidentiality.

6.3.3 Usefulness

It is not easy to estimate how effective a decentralised ACT system is. Since no data is collected, there is no data to base assumptions on whether or not the app is actually discovering exposures that the manual contact tracers cannot. However, such data minimisation is positive from a privacy perspective. BLE is not an exact technology either, and as previously mentioned, the signal is affected by environmental factors.

In the risk assessment of Smittestopp [Knu20b], FHI concludes that three factors decide if an ACT system based on GAEN will be effective. The technology must be accurate, enough people must download it, and it is more effective if the infection rates are high. In addition to these, the test capacity must be adequate.

For the technology to be accurate, there must be a small number of false positives, and the application must be able to note most of the contacts of a device. Before releasing the application, FHI conducted controlled tests to find variables that will sense the presence of most nearby devices. The study was released in May of 2021 and compared two different configurations for risk scoring and found that one of them identified 80% of close contacts and 34% of contacts further away [MMM^{+}21]. The values used for the most successful configuration were 58/68 dBm. They also found that iOS devices collected less than Android. The scenarios tested were scenarios that contact tracers would struggle to find, such as in a queue or public transportation. The study was used to optimise the values used in the app. On the 1st of June 2021, they rolled out a new version of Smittestopp, where the rate is updated to 93% [FHI21b].

It is difficult to estimate the number of false reports due to the decentralised approach. FHI knows how many people verify their infection through the app. However, they do not know how many receive exposure notifications or if they are sent on the correct basis. The tests used to estimate effectiveness indicate that the number of exposure notifications should be approximately accurate. It is also not possible to know if users take the notification seriously and follows the suggested approach.

As of the 8th of June 2021, only 19% of the Norwegian population has downloaded the app [FHI20b]. It is not certain that all 19% keep the app activated. For each index case, approximately one in five of their close contacts have the app. This means that even though the accuracy is precise, the actual effect is smaller. If 20% of index cases have the app, and 20% of their close contacts have it, only 4% of contacts are actually picked up. For this thesis, the numbers released by FHI have been monitored, and as can be seen in Appendix A.2, the percentage of positive cases that register it in the app averages to 4.63%.

To ensure high adoption, FHI points out that it is essential that the citizens of Norway know the app and how to use it and that they trust the government and app developers. The government has sent out a text message to every citizen over the age of 16 asking them to download the app, but it has still only been downloaded by 19% [FHI20b]. Maybe this is because the trust in an ACT system was damaged by the first iteration of Smittestopp and the amount of negative media coverage it got. It could also be because the Norwegian government is afraid to do anything wrong this time and therefore wants to ensure that downloading the app is entirely up to the individual and do not want to force it on the population.

In addition, some problems have been noted by users of the app. For instance, in February, users of Smittestopp reported a delay in time between a positive test result was received and the possibility to verify infection [Krü21]. For one user, it took three days before the user could register as infected in the app. The point of ACT systems is to be able to notify more people faster than manual contact tracing, and issues such as this will reduce the usefulness of the system. FHI stated that it was a delay in MSIS that caused the issue, and it was a known and fixed flaw. Nevertheless, similar issues could reduce the trust the user has in the system and reduce usefulness.

As mentioned in section 2.3, in manual contact tracing, it is up to a human contact tracer to decide whether or not a contact of an index case is or is not a close contact. An ACT system will lose this human decision aspect and could be stricter or less strict than a person. If a system is too strict, false reports of infection could increase, but if is not strict enough, manual contact tracers could probably find more contacts than the system can. In addition, the guidance the contact tracers give a close contact could vary. Some will be asked to quarantine, some to stay home and get tested. The experiments FHI have conducted to find the best configuration of Smittestopp, try to find the balance of a human contact tracer, but it is not easy to know if this balance is met.

6.4 Why Does It Matter?

Google and Apple already collect a lot of information on their users. It is therefore easy to think that it is acceptable that they collect more data. However, the data collected using GAEN is a new type of data that can be used for different purposes than data already collected. It is not the same as location data, which does not necessarily reveal who are together. Due to the low accuracy of GPS indoors, BLE is more precise and reliable in this regard. In GPS, one could look to be at the same location but, for instance, be at different floors of a building. The RPIs collected can say something about who a person is close together with, when, and how often. This is valuable information and not previously available with the accuracy that BLE offers.

One could also think that it is not that important if people know that a user is positive for COVID, but it could have consequences for the individual. The test result could be used to discriminate people. This has been seen in smaller communities in Norway. For instance, during an outbreak of COVID in Hammerfest, a list of names of infected people was sent around, which made the threshold to get tested higher[KG21]. This means that releasing names could hinder contact tracing as well as being a burden on the infected people as people will not check if they are infected to avoid public humiliation. The infection status is also medical information and should therefore be treated as confidential data.

In addition, since such a high percentage of the world owns either an Apple iOS or Google Android device, and GAEN functionality exists them, not only in countries that have implemented a system, it could allow for mass surveillance of almost everyone if misused.

Chapter 7 Conclusion

This thesis has examined different ACT solutions focusing on GAEN and the privacy risk it could entail. The Norwegian ACT application Smittestopp has been studied in more detail. The main privacy risks of the GAEN framework, Smittestopp, and other systems based on GAEN have been investigated through literature research. In Chapter 5, different attacks and concerns were identified. They were discussed and put into the context of the Norwegian Smittestopp in Chapter 6.

The first research question aimed to investigate if the identity of a confirmed COVID-19 positive person is preserved in Smittestopp and similar GAEN based systems. With the decentralised approach of GAEN, more data is located at the individual users' devices, and many attacks are challenging to mitigate. However, the impact is relatively small, as most of the attacks target individuals. It can be a concern for an infected individual to have their positive test leaked, but information from manual contact tracing can lead to the same conclusions. The extended attacks where the identity of more people can be revealed have a higher impact, but they are also more challenging to do successfully. In Smittestopp, improvements to the verification solution have increased the users' privacy.

For the second research question, the possibility to use data from GAEN to identify movement patterns or create social graphs was investigated. The decentralised approach of GAEN makes this more difficult as the data is distributed among many devices. However, an attacker could use many devices of their own to collect data. The data is supposed to be distributed, and collection is trivial, but setting up devices and using the information is more complicated. Linking of RPIs is necessary for an attacker to be successful. The primary way to do this is to link RPIs of infected individuals through the distributed Diagnosis Keys. This makes it easier to use collected data to identify movement patterns of infected individuals.

Other concerns also exist. Google, Apple and the HA can use data they possess or functionality they have control over to find out more about their users. It is also vital that the purpose of ACT systems is limited to containing the spread of a virus and not extended or used after the threat is over. It is also challenging to determine if the risk is worth it since there is no way to know if all those that are at risk are notified, if cases overlooked by manual contact tracers are found, or if the false report rate is low. The tests conducted by FHI implies that the configuration of Smittestopp is effective, but the adoption rate is still only 19% as of the 8th of June 2021.

To summarise, Smittestopp is a more privacy-preserving and functional contact tracing app than Smittestopp v1 was. However, the concerns and limitations discussed could be misused by both external and internal actors, and the new data collected is important to keep safe. On the other hand, manual contact tracing is not privacy-preserving itself as the identity of the exposed individuals is disclosed to the contact tracers. Therefore, one could argue that the fact that privacy concerns exist in ACT systems is logical. Be that as it may, the problem lies in the amount of data collected and that people can misuse it outside of the HA.

More people are getting vaccinated, but the chances are that the COVID-19 pandemic will be around for a while longer. Therefore, the systems explored in this thesis will stay relevant at least for a while longer, and also if a new virus starts to spread. However, it is essential that even though the ACT systems could be of use later, their usage is limited to when they are needed. Future work for this thesis could be to test if the attacks explored are feasible in Smittestopp and if the collected data can be exploited in other ways.

References

[ACK$^+$21] Benedikt Auerbach, Suvradip Chakraborty, Karen Klein, Guillermo Pascual-Perez, Krzysztof Pietrzak, Michael Walter, and Michelle Yeo. *Inverse-Sybil Attacks in Automated Contact Tracing.* In Kenneth G. Paterson, editor, *Topics in Cryptology – CT-RSA 2021*, pages 399–421, Cham, 2021. Springer International Publishing.

[Amn20] Amnesty. *Bahrain, Kuwait and Norway contact tracing apps among most dangerous for privacy.* [Online]. Available: https://www.amnesty.org/en/latest/news/2020/06/bahrain-kuwait-norway-contact-tracing-apps-danger-for-privacy/, 2020. Last visited: 12-02-2020.

[App20] Apple. *Exposure Notification APIs Addendum.* [Online]. Available: https://developer.apple.com/contact/request/download/Exposure_Notification_Addendum.pdf, 2020. Last visited: 16-02-2020.

[Arv20] Eivind Arvesen. *A Final Post About Smittestopp.* [Online]. Available: https://www.eivindarvesen.com/blog/2021/01/02/a-final-post-about-smittestopp, 2020. Last visited: 11-03-2020.

[Aus20] Australian Government. *Technology behind COVIDSafe.* [Online]. Available: https://www.covidsafe.gov.au/technology.html, 2020. Last visited: 12-02-2021.

[BDF$^+$20] Lars Baumgärtner, Alexandra Dmitrienko, Bernd Freisleben, Alexander Gruler, Jonas Höchst, Joshua Kühlberg, Mira Mezini, Markus Miettinen, Anel Muhamedagic, Thien Duc Nguyen, et al. *Mind the gap: Security & privacy risks of contact tracing apps. arXiv preprint arXiv:2006.05914*, 2020.

[BKT$^+$20] Jason Bay, Joel Kek, Alvin Tan, Chai Sheng Hau, Lai Yongquan, Janice Tan, and Tang Anh Quy. *BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders.* [Online]. Available: https://bluetrace.io/, April 2020.

[Bus20] Douglas Busvine. *Rift opens over European coronavirus contact tracing apps.* [Online]. Available: https://www.reuters.com/article/us-health-coronavirus-europe-tech-idUSKBN2221U0, 2020. Last visited: 14-04-2021.

[CP92] David Chaum and Torben P. Pedersen. *Wallet Databases with Observers.* In *Proceedings of the 12th Annual International Cryptology Conference on Advances* *in Cryptology*, CRYPTO '92, page 89–105, Berlin, Heidelberg, 1992. Springer-Verlag.

[Dat20] Datatilsynet. *Vedtak om midlertidig forbud mot å behandle personopplysninger - appen Smittestopp.* [Online]. Available: https://www.datatilsynet.no/contentassets/ae1905a8b88d4d869f1e059b60be35fd/Vedtak-om-midlertidig-forbud-mot-a-behandle-personopplysninger.pdf, July 2020.

[Dat21] Datatilsynet. *Pris for innebygd personvern til Anonyme Tokens.* [Online]. Available: https://www.datatilsynet.no/aktuelt/aktuelle-nyheter-2021/pris-for-innebygd-personvern-til-anonyme-tokens/, 2021. Last visited: 03-05-2021.

[DGS$^+$18] Alex Davidson, Ian Goldberg, Nick Sullivan, George Tankersley, and Filippo Valsorda. *Privacy Pass: Bypassing Internet Challenges Anonymously. Proceedings on Privacy Enhancing Technologies*, 2018:164–180, June 2018.

[FHI] FHI. *Om Folkehelseinstituttet.* [Online]. Available: https://www.fhi.no/om/fhi. Last visited: 28-02-2021.

[FHI17] FHI. *Om Meldingssystem for smittsomme sykdommer (MSIS).* [Online]. Available: https://www.fhi.no/hn/helseregistre-og-registre/msis/meldesystemet-for-smittsomme-sykdommer/, 2017. Last visited: 28-02-2021.

[FHI20a] FHI. *Definisjoner av tilfelle og nærkontakt.* [Online]. Available: https://www.fhi.no/nettpub/coronavirus/testing-og-oppfolging-av-smittede/definisjoner-av-mistenkte-og-bekreftede-tilfeller-med-koronavirus-coronavir/?term=&h=1, February 2020. Last visited: 01-03-2021.

[FHI20b] FHI. *Nøkkeltall fra Smittestopp.* [Online]. Available: https://www.fhi.no/om/smittestopp/nokkeltall-fra-smittestopp/#table-container-20976238, December 2020. Last visited: 09-06-2021.

[FHI20c] FHI. *Smittesporing.* [Online]. Available: https://www.fhi.no/nettpub/coronavirus/testing-og-oppfolging-av-smittede/smittesporing/, 2020. Last visited: 11-03-2021.

[FHI20d] FHI. *Teknologien bak Smittestopp.* [Online]. Available: https://www.fhi.no/om/smittestopp/teknologien-bak-smittestopp, 2020. Last visited: 06-03-2021.

[FHI21a] FHI. *Nå virker Smittestopp på tvers av landegrenser.* [Online]. Available: https://www.fhi.no/nyheter/2021/na-virker-smittestopp-pa-tvers-av-landegrenser/, February 2021. Last visited: 03-06-2021.

[FHI21b] FHI. *Oppdatert smittestopp-app oppdager nesten alle nærkontaktene .* [Online]. Available: https://www.fhi.no/nyheter/2021/oppdatert-smittestopp-app-oppdager-nesten-alle-narkontaktene/, June 2021. Last visited: 03-06-2021.

[GA20] Google and Apple. *Exposure Notifications API.* [Online]. Available: https://developers.google.com/android/exposure-notifications/exposure-notifications-api, 2020. Last visited: 20-05-2021.

[Goo] Google. *Exposure Notifications Express overview.* [Online]. Available: https://developers.google.com/android/exposure-notifications/en-express. Last visited: 03-04-2021.

[Goo20a] Google. *Define meaningful exposures.* [Online]. Available: https://developers.google.com/android/exposure-notifications/meaningful-exposures#scoring-options, 2020. Last visited: 20-05-2021.

[Goo20b] Google. *Exposure Notifications API.* https://github.com/google/exposure-notifications-internals/blob/main/README.md#ble-mac-and-rpi-rotation, 2020. Last visited: 17-04-2021.

[Goo20c] Google. *Exposure Notifications verification server.* [Online]. Available: https://developers.google.com/android/exposure-notifications/verification-system, 2020. Last visited: 04-05-2021.

[Goo20d] Google. *Google COVID-19 Exposure Notifications Service Additional Terms.* [Online]. Available: https://blog.google/documents/72/Exposure_Notifications_Service_Additional_Terms.pdf, 2020. Last visited: 15-03-2021.

[Goo20e] Google. *nearby.exposurenotification Reference.* [Online]. Available: https://developers.google.com/android/reference/com/google/android/gms/nearby/exposurenotification/package-summary, 2020. Last visited: 14-04-2021.

[Gre] Great Scott Gadgets. *Ubertooth.* [Online]. Available: https://github.com/greatscottgadgets/ubertooth. Last visited: 04-06-2021.

[GSDL21] Martin Gundersen, Øyvind Bye Skille, Olav Døvik, and Henrik Lied. *Kilder til Danmarks Radio: USA har spionert på norske politikere fra Danmark.* [Online]. Available: https://nrkbeta.no/2021/05/30/kilder-til-danmarks-radio-usa-har-spionert-pa-norske-politikere-fra-danmark/, May 2021. Last visited: 02-06-2021.

[Gun20] Martin Gundersen. *– Jeg vil egentlig ikke anbefale noen å bruke appen slik det er nå.* [Online]. Available: https://nrkbeta.no/2020/04/02/advarer-mot-a-installere-fhis-korona-app/, April 2020. Last visited: 03-03-2021.

[Hel18] Helsetilsynet. *Barns selvbestemmelsesrett.* [Online]. Available: https://www.helsetilsynet.no/rettigheter-klagemuligheter/helse--og-omsorgstjenester/sarskilte-rettigheter-for-barn-og-unge-til-helsetjenester/barns-selvbestemmelsesrett/, 2018. Last visited: 01-04-2021.

[Hel20a] Statsministerens kontor Helse- og omsorgsdepartementet. *Omfattende tiltak for å bekjempe koronaviruset.* [Online]. Available: https://www.regjeringen.no/no/aktuelt/nye-tiltak/id2693327/, Mar 2020. Last visited: 24-04-2021.

[Hel20b] Helse og omsorgsdepartementet. *Langsiktig strategi for håndteringen av covid-19-pandemien.* [Online]. Available: https://www.regjeringen.no/no/dokumenter/langsiktig-strategi-for-handteringen-av-covid-19-pandemien/id2791715/, 2020.

[Hoe20] Jaap-Henk Hoepman. *A Critique of the Google Apple Exposure Notification (GAEN) Framework. CoRR*, abs/2012.05097, 2020.

[Inr20] Inria, PRIVATICS Team. *ROBERT: ROBust and privacy-presERving proximity Tracing.* [Online]. Available: https://github.com/ROBERT-proximity-tracing/documents/blob/master/ROBERT-summary-EN.pdf, April 2020.

[IVV21] Vincenzo Iovino, Serge Vaudenay, and Martin Vuagnoux. *On the Effectiveness of Time Travel to Inject COVID-19 Alerts.* In Kenneth G. Paterson, editor, *Topics in Cryptology – CT-RSA 2021*, pages 422–443, Cham, 2021. Springer International Publishing.

[KG21] Kristina Kalinina and Julie Kristin Karlsen Groseth. *Smittelister florerer på nett i Finnmark: – Jeg er forbannet!* [Online]. Available: https://www.nrk.no/tromsogfinnmark/politiet-advarer-mot-korona-lister-i-finnmark-etter-et-smitteutbrudd-i-hammerfest-1.15506904, May 2021. Last visited: 04-06-2021.

[Knu20a] Gun Peggy Strømstad Knudsen. *Personvernkonsekvensvurdering for registre og systemer.* Technical report, FHI, December 2020.

[Knu20b] Gun Peggy Strømstad Knudsen. *Risikovurdering av ny løsning for Smittestopp.* Technical report, FHI, December 2020.

[Krü21] Louise Krüger. *Tok tre dager å få lagt inn positivt svar i Smittestopp.* [Online]. Available: https://www.vg.no/nyheter/innenriks/i/AlpGPM/tok-tre-dager-aa-faa-lagt-inn-positivt-svar-i-smittestopp?fbclid=IwAR3ax3eaFMZqDLc0rujdI2s5yjspHQTnGD6s8go8UJ2jKsrnwHwrsK9fpbA, February 2021. Last visited: 29-05-2021.

[Lia18] Shannon Liao. *Why GPS-dependent apps deplete your smartphone battery.* [Online]. Available: https://www.theverge.com/2018/8/17/17630872/smartphone-battery-gps-location-services, August 2018. Last visited: 28-02-2021.

[LLB$^+$20] Jeanine Lilleng, Odd Rune Lykkebø, Bjørn Borud, Øyvind Indrebø, Eivind Andreas Arvesen, Aleksander Slater, and Elina Sande Heimark. *Endelig rapport for kildekodegjennomgang av løsning for digital smittesporing av koronaviruset.* Technical report, Simula Research Laboratory, May 2020.

[MMM$^+$21] Hinta Meijerink, Elisabeth H. Madslien, Camilla Mauroy, Mia Karoline Johansen, Sindre Møgster Braaten, Christine Ursin Steen Lunde, Trude Margrete Arnesen, Siri Laura Feruglio, and Karin Nygård. *The first GAEN-based COVID-19 contact tracing app in Norway identifies 80% of close contacts in "real life" scenarios.* Cold Spring Harbor Laboratory Press, 2021.

[MSS20] Henrik Walker Moe, Tjerand Slide, and Martin Strand. *Anonymous Tokens.* [Online]. Available: https://github.com/HenrikWM/anonymous-tokens/wiki, 2020. Last visited: 11-03-2020.

[MW21a] Merriam-Webster. *Contact Tracing.* In *Merriam-Webster.com dictionary.* Merriam-Webster, 2021.

[MW21b] Merriam-Webster. *Lockdown.* In *Merriam-Webster.com dictionary.* Merriam-Webster, 2021.

[MW21c] Merriam-Webster. *Social distancing.* In *Merriam-Webster.com dictionary.* Merriam-Webster, 2021.

[Nat01] National Institute of Standards and Technology. *Advanced Encryption Standard (AES).* Technical Report Federal Information Processing Standards Publications (FIPS PUBS) 197, November 26, 2001, U.S. Department of Commerce, Washington, D.C., 201.

[Nat15] National Institute of Standards and Technology. *Secure Hash Standard (SHS).* Technical Report Federal Information Processing Standards Publications (FIPS PUBS) 180-4, Change Notice 4 August 04, 2015, U.S. Department of Commerce, Washington, D.C., 2015.

[NTB21] NTB. *Sårbarhet i Android-versjonen av Smittestopp appen er rettet.* [Online]. Available: https://www.digi.no/artikler/sarbarhet-i-android-versjonen-av-smittestopp-appen-er-rettet/509932?fbclid=IwAR1gZ8EuBqATJbQqQhJQ46EfQBD2VdqdxtSGt9LudFzrCzXd0R16YRgbrl4, May 2021. Last visited: 10-05-2021.

[PEP20] PEPP-PT Team. *PEPP-PT Documentation.* [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation, 2020. Last visited: 12-04-2021.

[Rea21] Joel Reardon. *Why Google Should Stop Logging Contact-Tracing Data.* [Online]. Available: https://blog.appcensus.io/2021/04/27/why-google-should-stop-logging-contact-tracing-data/, April 2021. Last visited: 10-05-2021.

[SMB17] Ji Shouling, Prateek Mittal, and Raheem Beyah. *Graph Data Anonymization, De-Anonymization Attacks, and De-Anonymizability Quantification: A Survey. IEEE Communications Surveys Tutorials*, 19(2):1305–1326, 2017.

[Smi20] Smittestopp Development Team. *Smittestopp Documentation.* [Online]. Available: https://github.com/folkehelseinstituttet/Fhi.Smittestopp.Documentation, 2020. Last visited: 12-05-2021.

[Spo18] Jon Gunnar Sponås. *Things You Should Know About Bluetooth Range.* [Online]. Available: https://blog.nordicsemi.com/getconnected/things-you-should-know-about-bluetooth-range, February 2018. Last visited: 29-02-2021.

[SS21] Tjerand Silde and Martin Strand. *Anonymous Tokens with Public Metadata and Applications to Private Contact Tracing.* Cryptology ePrint Archive, Report 2021/203, 2021. https://eprint.iacr.org/2021/203.

[Sta20] Stat Counter. *Mobile Operating System Market Share Worldwide - March 2021.* [Online]. Available: https://gs.statcounter.com/os-market-share/mobile/worldwide, 2020. Last visited: 10-03-2021.

[Sta21] Statistisk sentralbyrå. *Norsk mediebarometer 2020.* [Online]. Available: https://www.ssb.no/kultur-og-fritid/artikler-og-publikasjoner/_attachment/452390?_ts=17912355278, April 2021.

[Tay21] Derrick Bryson Taylor. *A Timeline of the Coronavirus Pandemic.* [Online]. Available: https://www.nytimes.com/article/coronavirus-timeline.html, March 2021. Last visited: 10-03-2021.

[TPH$^+$20] Carmela Troncoso, Mathias Payer, Jean-Pierre Hubaux, Marcel Salathé, James R. Larus, Wouter Lueks, Theresa Stadler, Apostolos Pyrgelis, Daniele Antonioli, Ludovic Barman, Sylvain Chatel, Kenneth G. Paterson, Srdjan Capkun, David A. Basin, Jan Beutel, Dennis Jackson, Marc Roeschlin, Patrick Leu, Bart Preneel, Nigel P. Smart, Aysajan Abidin, Seda Gurses, Michael Veale, Cas Cremers, Michael Backes, Nils Ole Tippenhauer, Reuben Binns, Ciro Cattuto, Alain Barrat, Dario Fiore, Manuel Barbosa, Rui Oliveira, and José Pereira. *Decentralized Privacy-Preserving Proximity Tracing. IEEE Data Eng. Bull.*, 43(2):36–66, 2020.

[Tur21] Ash Turner. *How Many Smartphones Are In The World?* [Online]. Available: https://www.bankmycell.com/blog/how-many-phones-are-in-the-world, 2021. Last visited: 10-03-2021.

[US 20] US Government. *The Global Positioning System.* [Online]. Available: https://www.gps.gov/systems/gps/, 2020. Last visited: 10-03-2021.

[Vau20] Serge Vaudenay. Analysis of dp3t. Cryptology ePrint Archive, Report 2020/399, 2020. https://eprint.iacr.org/2020/399.

[VV20] Serge Vaudenay and Martin Vuagnoux. *The Dark Side of SwissCovid.* [Online]. Available: https://lasec.epfl.ch/people/vaudenay/swisscovid.html, 2020. Last visited: 05-06-2021.

[Wel19] Chris Welch. *Here's why so many apps are asking to use Bluetooth on iOS 13.* [Online]. Available: https://www.theverge.com/2019/9/19/20867286/ios-13-bluetooth-permission-privacy-feature-apps, September 2019. Last visited: 28-04-2021.

[Whi20] Zack Whittaker. *Australia's spy agencies caught collecting COVID-19 app data.* [Online]. Available: https://techcrunch.com/2020/11/24/australia-spy-agencies-covid-19-app-data/, November 2020. Last visited: 15-05-2021.

[WHO20] WHO. *COVAX.* [Online]. Available: https://www.who.int/initiatives/act-accelerator/covax, 2020. Last visited: 08-03-2021.

[Wol21] Ben Wolford. *What is GDPR, the EU's new data protection law?* [Online]. Available: https://gdpr.eu/what-is-gdpr/, 2021. Last visited: 08-05-2021.

[Yu21] Eileen Yu. *Singapore police can access COVID-19 contact tracing data for criminal investigations.* [Online]. Available: https://www.zdnet.com/article/singapore-police-can-access-covid-19-contact-tracing-data-for-criminal-investigations/, January 2021. Last visited: 15-05-2021.

Appendix

Appendix A

A.1 Little Thumb Bug Code

```
import json

## GET JSON ##
with open('mf-20_04.json', 'r') as f:
    entireCap = f.read()

cap = json.loads(entireCap)

## CONVERT JSON TO DICTIONARY
gaen = []
for i in range(len(cap)):
    if "btcommon.eir_ad.advertising_data" in cap[i]["_source"]["layers"
        ]["btle"]:
        if "bluetooth.gaen" in cap[i]["_source"]["layers"]["btle"]["
            btcommon.eir_ad.advertising_data"]["btcommon.eir_ad.entry"
            ]:
            tempdict={}
            tempdict["adr"] = cap[i]["_source"]["layers"]["btle"]["btle
                .advertising_address"]
            tempdict["rpi"] = cap[i]["_source"]["layers"]["btle"]["
                btcommon.eir_ad.advertising_data"]["btcommon.eir_ad.
                entry"]["bluetooth.gaen"]
            gaen.append(tempdict)


## DICTIONARY SYNTAX
# gaen = [{
#   "adr": "advertising_address",
#   "rpi": {"bluetooth.gaen.rpi": "rpi", "bluetooth.gaen.aemd": "
    metadata"}
# }]
#
# allRPIsbyAdr = {
#   "adr": [list of RPIs]
# }
```

```
## CREATE DICTIONARY OF ADDRESSES WITH ONLY THEIR RPIs
allRPIsbyAdr = {}
uniqueRPIsbyAdr = {}

# MAKE ALL ADDRESSES A KEY
for e in range(len(gaen)):
    allRPIsbyAdr[gaen[e]["adr"]] = []
    uniqueRPIsbyAdr[gaen[e]["adr"]] = []

# PLACE ALL RPIs IN CORRESPONDING ADDRESS' LIST
for e in range(len(gaen)):
    allRPIsbyAdr[gaen[e]["adr"]].append(gaen[e]["rpi"]["bluetooth.gaen.
        rpi"])

# FILTER OUT DUPLICATES
for adr in allRPIsbyAdr:
    uniqueRPIsbyAdr[adr].append(list(set(allRPIsbyAdr[adr])))

thumbbugcheck = 0
for adr1 in uniqueRPIsbyAdr:
    for adr2 in uniqueRPIsbyAdr:
        if uniqueRPIsbyAdr[adr1] != uniqueRPIsbyAdr[adr2] and
            uniqueRPIsbyAdr[adr1][0] == uniqueRPIsbyAdr[adr2][0]:
            print(adr1, adr2, "have the same RPI. Possible Thumb bug!\n
                ")
            thumbbugcheck = 1


## PRINT ##
for adr in uniqueRPIsbyAdr:
    print("\nAddress: ", adr)
    print("RPIs:")
    for rpi in uniqueRPIsbyAdr[adr]:
        print("    ", rpi)
        if len(uniqueRPIsbyAdr[adr]) > 1:
            print("THUMB BUG?\n")
            thumbbugcheck = 1

if thumbbugcheck == 1:
    print("\nPossible Little Thumb found!\n")
else:
    print("\nLittle Thumb not found in", len(uniqueRPIsbyAdr), "unique
        adresses.\n")
```

A.2 Key Numbers from Norwegian Smittestopp

The following pages include numbers from FHI regarding the number of downloads of Smittestopp [FHI20b], as well as a percentage of number of actual infected vs. registered as infected.

COVID-19 NORWAY - SMITTESTOPP

Updated: (08/06/2021)

Sources:

FHI https://www.fhi.no/om/smittestopp/nokkeltall-fra-smittestopp/#table-container-20976238

VG https://www.vg.no/spesial/corona/

Population https://www.ssb.no/befolkning/faktaside/befolkningen

Assumption: That infected Smittestopp users register positive in app on same day as they get their test result

Population (4rd quarter 2020): 5391369

Dato	Positives reported in app - FHI	Positives registered in MSIS - VG	% of positives recorded in app	App downloads (total) - FHI	% of population downloaded app
21/12/2020	23	427	5.39%	117700	2.18%
22/12/2020	18	602	2.99%	158700	2.94%
23/12/2020	12	522	2.30%	174300	3.23%
24/12/2020	19	476	3.99%	183800	3.41%
25/12/2020	10	317	3.15%	194700	3.61%
26/12/2020	8	428	1.87%	207200	3.84%
27/12/2020	14	381	3.67%	217400	4.03%
28/12/2020	21	525	4.00%	231200	4.29%
29/12/2020	22	696	3.16%	241700	4.48%
30/12/2020	26	719	3.62%	249000	4.62%
31/12/2020	13	557	2.33%	255800	4.74%
01/01/2021	17	379	4.49%	259000	4.80%
02/01/2021	11	329	3.34%	262900	4.88%
03/01/2021	36	447	8.05%	345000	6.40%
04/01/2021	17	526	3.23%	366300	6.79%
05/01/2021	25	929	2.69%	374600	6.95%
06/01/2021	35	803	4.36%	402000	7.46%
07/01/2021	31	825	3.76%	408000	7.57%
08/01/2021	21	683	3.07%	412600	7.65%
09/01/2021	15	443	3.39%	415900	7.71%
10/01/2021	19	554	3.43%	419600	7.78%
11/01/2021	8	431	1.86%	423000	7.85%
12/01/2021	18	715	2.52%	425200	7.89%
13/01/2021	29	668	4.34%	435300	8.07%
14/01/2021	17	456	3.73%	438000	8.12%
15/01/2021	15	466	3.22%	439800	8.16%
16/01/2021	8	243	3.29%	441100	8.18%
17/01/2021	10	206	4.85%	460800	8.55%
18/01/2021	5	383	1.31%	465200	8.63%
19/01/2021	7	422	1.66%	468200	8.68%
20/01/2021	10	431	2.32%	471800	8.75%
21/01/2021	16	372	4.30%	474300	8.80%
22/01/2021	11	306	3.59%	477500	8.86%
23/01/2021	8	238	3.36%	488100	9.05%
24/01/2021	27	279	9.68%	543400	10.08%
25/01/2021	13	233	5.58%	562000	10.42%
26/01/2021	40	279	14.34%	658300	12.21%
27/01/2021	19	367	5.18%	675700	12.53%
28/01/2021	20	168	11.90%	682400	12.66%
29/01/2021	8	299	2.68%	686600	12.74%
30/01/2021	19	218	8.72%	690400	12.81%
31/01/2021	7	173	4.05%	691400	12.82%
01/02/2021	13	296	4.39%	692100	12.84%

02/02/2021	14	291	4.81%	694600	12.88%
03/02/2021	30	286	10.49%	752700	13.96%
04/02/2021	33	373	8.85%	807300	14.97%
05/02/2021	28	271	10.33%	833500	15.46%
06/02/2021	14	170	8.24%	836500	15.52%
07/02/2021	2	119	1.68%	846100	15.69%
08/02/2021	22	346	6.36%	850900	15.78%
09/02/2021	10	231	4.33%	853000	15.82%
10/02/2021	10	210	4.76%	854700	15.85%
11/02/2021	15	264	5.68%	856300	15.88%
12/02/2021	14	426	3.29%	857800	15.91%
13/02/2021	19	149	12.75%	858200	15.92%
14/02/2021	12	117	10.26%	859600	15.94%
15/02/2021	16	353	4.53%	861100	15.97%
16/02/2021	20	286	6.99%	862400	16.00%
17/02/2021	10	358	2.79%	864300	16.03%
18/02/2021	21	326	6.44%	864700	16.04%
19/02/2021	19	283	6.71%	866200	16.07%
20/02/2021	16	208	7.69%	867600	16.09%
21/02/2021	1	217	0.46%	867800	16.10%
22/02/2021	7	226	3.10%	869200	16.12%
23/02/2021	26	390	6.67%	870700	16.15%
24/02/2021	13	341	3.81%	871100	16.16%
25/02/2021	29	551	5.26%	872600	16.19%
26/02/2021	25	524	4.77%	873300	16.20%
27/02/2021	18	262	6.87%	873800	16.21%
28/02/2021	16	180	8.89%	878300	16.29%
01/03/2021	29	729	3.98%	878900	16.30%
02/03/2021	34	499	6.81%	882900	16.38%
03/03/2021	32	689	4.64%	884600	16.41%
04/03/2021	18	570	3.16%	886200	16.44%
05/03/2021	28	690	4.06%	887600	16.46%
06/03/2021	17	363	4.68%	889000	16.49%
07/03/2021	19	396	4.80%	890400	16.52%
08/03/2021	30	828	3.62%	893500	16.57%
09/03/2021	42	714	5.88%	899200	16.68%
10/03/2021	37	685	5.40%	902200	16.73%
11/03/2021	50	874	5.72%	904100	16.77%
12/03/2021	48	908	5.29%	923700	17.13%
13/03/2021	33	838	3.94%	924600	17.15%
14/03/2021	31	656	4.73%	929500	17.24%
15/03/2021	49	866	5.66%	932500	17.30%
16/03/2021	58	1150	5.04%	933500	17.31%
17/03/2021	52	1064	4.89%	963600	17.87%
18/03/2021	66	1034	6.38%	966400	17.92%
19/03/2021	52	989	5.26%	967900	17.95%
20/03/2021	56	820	6.83%	969300	17.98%
21/03/2021	39	577	6.76%	970800	18.01%
22/03/2021	40	1096	3.65%	972400	18.04%
23/03/2021	47	1085	4.33%	973900	18.06%
24/03/2021	47	984	4.78%	975600	18.10%
25/03/2021	39	831	4.69%	976600	18.11%
26/03/2021	43	1072	4.01%	979000	18.16%
27/03/2021	29	462	6.28%	980400	18.18%

28/03/2021	41	681	6.02%	980700	18.19%
29/03/2021	21	989	2.12%	982000	18.21%
30/03/2021	35	1069	3.27%	982200	18.22%
31/03/2021	39	871	4.48%	983600	18.24%
01/04/2021	18	691	2.60%	983800	18.25%
02/04/2021	33	640	5.16%	983900	18.25%
03/04/2021	33	692	4.77%	985100	18.27%
04/04/2021	23	574	4.01%	985300	18.28%
05/04/2021	36	573	6.28%	986500	18.30%
06/04/2021	31	900	3.44%	986700	18.30%
07/04/2021	26	933	2.79%	987800	18.32%
08/04/2021	31	877	3.53%	987800	18.32%
09/04/2021	35	760	4.61%	988300	18.33%
10/04/2021	20	481	4.16%	989500	18.35%
11/04/2021	16	421	3.80%	989600	18.36%
12/04/2021	19	648	2.93%	989800	18.36%
13/04/2021	27	739	3.65%	995200	18.46%
14/04/2021	27	598	4.52%	995600	18.47%
15/04/2021	20	620	3.23%	996900	18.49%
16/04/2021	26	503	5.17%	997200	18.50%
17/04/2021	16	417	3.84%	998300	18.52%
18/04/2021	19	366	5.19%	998400	18.52%
19/04/2021	12	518	2.32%	998600	18.52%
20/04/2021	22	559	3.94%	998800	18.53%
21/04/2021	17	550	3.09%	999000	18.53%
22/04/2021	20	444	4.50%	1000400	18.56%
23/04/2021	9	480	1.88%	1000700	18.56%
24/04/2021	13	329	3.95%	1000800	18.56%
25/04/2021	9	222	4.05%	1002000	18.59%
26/04/2021	18	550	3.27%	1003300	18.61%
27/04/2021	20	524	3.82%	1003005	18.60%
28/04/2021	9	470	1.91%	1004700	18.64%
29/04/2021	18	385	4.68%	1004800	18.64%
30/04/2021	6	429	1.40%	1005000	18.64%
01/05/2021	16	289	5.54%	1006100	18.66%
02/05/2021	12	210	5.71%	1006200	18.66%
03/05/2021	15	483	3.11%	1007400	18.69%
04/05/2021	22	484	4.55%	1007600	18.69%
05/05/2021	29	469	6.18%	1008700	18.71%
06/05/2021	29	506	5.73%	1008900	18.71%
07/05/2021	17	407	4.18%	1010000	18.73%
08/05/2021	10	316	3.16%	1010100	18.74%
09/05/2021	7	309	2.27%	1011200	18.76%
10/05/2021	22	581	3.79%	1011300	18.76%
11/05/2021	19	477	3.98%	1011400	18.76%
12/05/2021	22	500	4.40%	1012600	18.78%
13/05/2021	11	320	3.44%	1012700	18.78%
14/05/2021	11	425	2.59%	1012900	18.79%
15/05/2021	22	312	7.05%	1013000	18.79%
16/05/2021	16	247	6.48%	1014100	18.81%
17/05/2021	8	201	3.98%	1014200	18.81%
18/05/2021	16	314	5.10%	1014300	18.81%
19/05/2021	18	478	3.77%	1014400	18.82%
20/05/2021	13	505	2.57%	1015600	18.84%

21/05/2021	18	714	2.52%	1015800	18.84%
22/05/2021	26	452	5.75%	1015800	18.84%
23/05/2021	39	451	8.65%	1016000	18.84%
24/05/2021	16	238	6.72%	1017300	18.87%
25/05/2021	16	511	3.13%	1018700	18.90%
26/05/2021	16	499	3.21%	1020500	18.93%
27/05/2021	17	409	4.16%	1021700	18.95%
28/05/2021	20	339	5.90%	1021900	18.95%
29/05/2021	14	245	5.71%	1022000	18.96%
30/05/2021	9	169	5.33%	1022100	18.96%
31/05/2021	11	292	3.77%	1022300	18.96%
01/06/2021	16	460	3.48%	1022500	18.97%
02/06/2021	19	353	5.38%	1024900	19.01%
03/06/2021	16	289	5.54%	1025200	19.02%
04/06/2021	9	303	2.97%	1025400	19.02%
05/06/2021	8	168	4.76%	1025500	19.02%
06/06/2021	3	128	2.34%	1025600	19.02%
07/06/2021	5	249	2.01%	1026700	19.04%
08/06/2021	9	188	4.79%	1027000	19.05%
		Average	**4.63%**		

NTNU

Kunnskap for en bedre verden