Privacy Analysis and Comparison of Pandemic Contact Tracing Apps

Yanji Piao[1], and Dongyue Cui[2*]
[1]Information Management and Information System, Yanbian University
Yanji, Jilin 133000 CHINA
[e-mail: piaoyanji@ybu.edu.cn]
[2]Information Management and Information System, Yanbian University
Yanji, Jilin 133000 CHINA
[e-mail: cuidongyue@foxmail.com]
[*]Corresponding author: Dongyue Cui

Received Agust 17, 2021; revised September 30, 2021; accepted October 18, 2021;
published November 30, 2021

Abstract

During the period of epidemic prevention and control, contact tracing systems are developed in many countries, to stop or slow down the progression of COVID-19 contamination. However, the privacy issues involved in the use of contact tracing apps have also attracted people's attention. First, we divide contact tracing techniques into two types: Bluetooth Low Energy (BLE) based and Global Positioning System (GPS) based techniques. In order to clear understand the system structure and its elements, we create data flow diagram (DFD) of each types. Second, we analyze the possible privacy threats contained in various types of contact tracing apps by applying LINDDUN, which is a threat modeling technique for personal information protection. Third, we make a comparison and analysis of various contact tracing techniques from privacy point of view. These studies can facilitate improve tracing and security performance to contact tracing apps through comparisons between different types.

Keywords: Contact tracing apps, COVID-19, LINDDUN, Privacy, Security.

This research was supported by Application Foundation Program of the Science and Technology Development Plan of Yanbian University of China (NO.602021014).

http://doi.org/10.3837/tiis.2021.11.015 **ISSN : 1976-7277**

1. Introduction

Since December 2019, some hospitals in Wuhan City, Hubei Province, China have successively discovered multiple cases of pneumonia of unknown cause, which were confirmed to be acute respiratory infectious diseases caused by the 2019 new corona virus infection [1]. On February 11, 2020, the Director-General of the World Health Organization Tan Desai announced in Geneva, Switzerland that the pneumonia caused by the new corona virus was named "corona virus disease-2019 (COVID-19)" and can be called a pandemic in terms of characteristics [2]. The Center for Systems Science and Engineering of Johns Hopkins University in the United States has produced a "Global New Corona virus Spread Map". The data comes from the World Health Organization, the US Centers for Disease Control and Prevention, Worldometers.info website, BNO news agency, and national governments and health departments. The epidemic is shown to the public through visualization. As of July 6th, 2021, the cumulative number of confirmed cases in China is 118,951, and worldwide is 184,791,441 [3]. COVID-19 has brought us a perceptible impact from politics to economy, from thought to life. One of the measures to deal with the spread of the COVID-19 epidemic is to identify close contacts of infected patients. Many countries have came up with different cryptographic protocols and frameworks and developed different contact tracing applications based on them. However, these applications also have severe implication on user's privacy, such as personal information leakage, mass surveillance and additionally revealing the behavioral patterns of the user [4]. In the early days of using contact tracking applications, cases of exposing patient privacy occurred in various countries, which not only caused fear among citizens but also led to discrimination against patients. Based on this situation, our research attempts to conduct detailed analysis of COVID-19 tracing applications by modeling targeted privacy threats.

The rest of this paper is organized as follows: Section 2 provides related work by researchers and background information about COVID-19 tracing apps and the protocols and frameworks they rely on. While Section 3 describes, LINDDUN, the target threat modeling method we use. Section 4 analyzes the privacy threats of different apps according to different categories of protocols and frameworks, and provides mitigation strategies. Section 5 summarizes the key differences about between privacy and security performance different contact tracing applications after modeling analysis. The paper in then concluded in Section 6.

2. Related Work

On February 24, 2021, MIT Technology Review announced a list of the top ten global technological breakthroughs in 2021, including digital contact tracking used to slow the spread of the new crown virus [5]. Using this technology, health investigators no longer need to rely on the memory of patients to track their whereabouts, which reduces the pressure on disease monitoring. In order to better understand the contact tracking application, we need to classify the existing digital contact tracking protocols and applications. There are two classification principles: technology and system architecture.

2.1 Technology

Current contact tracking solutions mainly use Bluetooth and GPS to determine user's absolute and relative position with others. The following is a brief introduction to Bluetooth, GPS, relative location and absolute location.

2.1.1 BLE and Relative Location

Bluetooth Low Energy (BLE) is a personal area network technology designed and sold by the Bluetooth Technology Alliance. Compared with classic Bluetooth, BLE is able to facilitate data exchanges with minimum power and maintain the standard communication range during the exchange [6].

Relative location refers to the location relative to other nearby individuals by using the device's Bluetooth data [7].

2.1.2 GPS and Absolute Location

Global Positioning System (GPS) is a satellite-based radio navigation system that can provide real-time geographic location and time information to global positioning system receivers anywhere on the earth [8].

Absolute location refers to historical location data based on GPS data, cell phone towers, Wi-Fi routers and data of the third-party service providers [7].

2.2 System Architecture

The relationship between architecture, protocol and application is that the application is designed on the basis of protocol (also including national standard). According to the characteristics of different protocols, such as different server functions and storage methods, it is mainly divided into two different system architectures, namely centralized architecture and decentralized architecture. The following is a brief introduction to these architectures based on BLE and GPS techniques.

2.2.1 Centralized

In a centralized architecture based on BLE technology, a user must first register and bind personal information in a central server. The central server records user registration information and saves Personally Identifiable Information (PII) such as ID and mobile phone numbers in the back-end database. The server generates a privacy-preserving Temporary ID (TempID) according to the user information. The central server authority uses encryption key management tool to generate TempID and then sent to the user's device. The application saves TempID locally. The APP will also receive the TempID broadcast by other users, record and save each observed TempID, corresponding neighbors, occurrence time, and contact duration to the local contact database. The storage days can be adjusted according to the size of the infection window (the default is 14 days, and the configuration is determined by the health authority). Once the user is diagnosed as an infected person, he will upload the encounter data to the central server. The server maps the TempIDs in data to individuals to identify vulnerable contacts.

In a centralized architecture based on GPS technology and assisted with Quick Response (QR) code, a user also must first register and bind personal information in a central server and he also need to authorize the application to collect GPS data and mobile database station data. Similarly, the central server records user registration information and saves PII such as ID and mobile phone numbers in the back-end database. The application uploads information to the central server in real time instead of uploading information until the user is diagnosed as an infected person. The purpose of this is to determine whether the user is a close contact of the confirmed patient by comparing whether the users appear at the same time and the same place.

If the user is infected, the central server will broadcast his anonymous information to all other users. In addition, the user must display his QR code or have his QR code scanned in relatively crowded public places in order to record the time and place information of visiting public places in the central database.

2.2.2 Decentralized

In a decentralized architecture based on BLE technology, a user does not need to 'pre-register' before using it to avoid having PIIs stored on the server. The application generates an initial secret key SK_0 locally, which is globally unique and can be associated with a unique device. The secret key is only saved locally and not uploaded to the server. The application generates the derived key DK_t used on the day according to the initial key SK_0 every day and then creates a pseudonyms or 'chirps' protecting privacy with a very short lifespan. The application stores the derived key DK_t for the most recent days locally, and the number of days can be adjusted according to the size of the infection window (the default is 14 days, and the configuration is determined by the health management agency). It will also receive chirps broadcast by other users, record and save each observed chirp, corresponding neighbors, occurrence time, and contact duration to the local database. Once the user is diagnosed as an infected person, he will upload his DK_ts to the central server. The central server pushes the DK_ts of the infected person to all other devices, and other users evaluate the risk of infection locally.

In a decentralized architecture based on GPS technology, a user does not need to 'pre-register' before using it to avoid having PIIs stored on the server. However, the user must authorize the application to collect GPS data and mobile database station data. The application records information including body temperature, isolation status, infection status, and geographic location locally, usually for 14 days. If the user is infected, he will upload the information to the central server in an anonymous form and other users download the patient's anonymous location data through the central server and compare it with their own travel trajectory locally.

2.3 Classification of Protocols and Applications

We found 48 protocols and applications by searching in the Apple App Store and Google Play Store, and by reviewing previous literature. Then we screened through the official website and shared information on GitHub. Finally, we concluded 28 representative protocols and applications as shown in **Table 1**.

Table 1. Collected data sets

Techno-logy	**System Architecture**	**Country**	**Name**	**Protocal or APPs**	**Detail**
BLE	Centralized	Europe	PEPP-PT NTK[9]	Protocal	
BLE	Centralized	France &Germany	ROBERT[10]	Protocal	
BLE	Centralized	Singapore	Bluetrace[11]	Protocal	
BLE	Centralized	Australia	CovidSafe[12]	APP	Based on Bluetrace
BLE	Centralized	France	TousAntiCovid[13]	APP	Based on ROBERT
BLE	Centralized	Singapore	Trace Together[14]	APP	Based on Bluetrace
BLE	Decentralized	Europe	DP-3T[15]	Protocal	
BLE	Decentralized	Italy	Pronto-C2[16]	Protocal	

BLE	Decentralized	US	Apple&Google alliance[17][18]	Protocal	
BLE	Decentralized	US(MIT)	PACT(EAST-coast)[19]	Protocal	
BLE	Decentralized	US(the University of Washington)	PACT(WEST-coast)[20]	Protocal	
BLE	Decentralized	US&Canda	TCN[21]	Protocal	
BLE	Decentralized	Austria	Stopp Corona[22]	APP	Based on Apple&Google alliance
BLE	Decentralized	Canada	CovidAlert[23]	APP	Based on Apple&Google alliance
BLE	Decentralized	Dutch	PrivateTracer[24]	APP	
BLE	Decentralized	Germany	Corona-Warn[25]	APP	Based on DP-3T&TCN
BLE	Decentralized	Iceland	Rankingn C-19[26]	APP	
BLE	Decentralized	UK	C19X[27]	APP	Based on Apple&Google alliance
BLE	Decentralized	US(Georgia Tech Research Institute)	CoEpi[28]	APP	Based on TCN
BLE	Decentralized	US(the University of Washington)	CovidSafe(UoW)[29]	APP	Based on PACT(WEST-coast)
BLE	Decentralized	US(Virginia's Department of Health)	CovidWise	APP	Based on Apple&Google alliance
GPS	Centralized	China	National Government Service Platform Epidemic Prevention Health Code[30][31][32]	APP	
GPS	Centralized	Columbia	CoronApp[33]	APP	
GPS	Centralized	Malaysia	MySejahtera[34]	APP	
GPS	Centralized	South Korea	New Crown Pneumonia Epidemic Intelligent Management System[35]	APP	
GPS	Decentralized	Germany	Pandoa[36]	APP	
GPS	Decentralized	Israel	Hamagen[37]	APP	
GPS	Decentralized	US(MIT)	PrivateKit: Safe Paths[38]	APP	

2.4 Contact Tracing Solutions

Among the above contact tracing solutions, some claim to protect privacy, while others are considered coercive or used for electronic surveillance. This sparked a debate about the architecture, data management, efficacy, privacy, and security of the contact tracking solution [39]-[44]. These scholars mainly discussed protocol' design decisions from a privacy perspective or contact tracing apps' key attributes, including system architecture, data management, privacy, security performance, proximity estimation, attack vulnerabiliby, usability, inclusivity, content evaluation. Some scholars also considered the ethics of contact tracing applications and the trade-off between privacy and epidemic control [45] [46].

After reading the research results of previous scholars, we have found two main deficiencies: First, from the perspective of the analysis object, the existing research mainly focuses on the analysis of BLE technology-based applications developed in European and American countries and few people have studied GPS technology-based applications [47]. Second, from the perspective of analysis method, the existing analysis on privacy and security is not systematic enough, and few people have analyzed privacy threats by establishing security threat models [48].

Therefore, this article attempts to conduct a security analysis and a comprehensive comparison of COVID-19 contact tracing applications by using LINDDUN, a threat modeling framework.

3. Target Privacy Threat Modeling

Due to space limitations and the focus of this article on the privacy and security analysis and comparison of different contact tracing applications, the analysis focuses on the first three steps of LINDDUN, and the latter three steps are integrated into a threat mitigation strategy (See 3.1 for details). Also due to space limitations, we do not draw the client-server diagrams of the four types of APP, but directly summarize the respective DFD diagrams. This chapter is divided into five sections according to the simplified analysis steps. Four types of applications are introduced in the last four sections, which is convenient for comparative analysis and can reduce redundancy.

3.1 LINDDUN Threat Modeling Framework

Threat modeling involves systematically identifying, eliciting, and analyzing privacy- and/or security-related threats in the context of a specific system [49]. There are various threat modeling frameworks, such as LINDDUN, STRIDE, PASTA, and NIST, which focus on software system privacy, system security, application security, and data-centric risks, respectively.

LINDDUN is the most promising systematic threat modeling framework which uses data flow diagrams as the basis of analysis, and every element in DFD is systematically and thoroughly examined to prevent privacy threats [50].

Specifically, it consists of six steps, defining DFD, mapping privacy threats to DFD elements, identifying threat scenarios, prioritizing threats, eliciting mitigation strategies and selecting corresponding PETS.

The primary contribution of LINDDUN is to provide a systematic methodology for modeling specific privacy threats and to provide a comprehensive catalog of specific privacy threat tree models [51][52]. Then we use LINDDUN to perform target privacy threat modeling and security analysis on four types of applications.

3.2 Creating the Data Flow Diagram (DFD)

The DFD graphically expresses the logical function of the system, the logical flow of data in the system and the logical transformation process, which consists of the following four parts:

external entities, data stores, processes and data flows. The DFDs of four types of applications are shown in **Fig. 1** to **Fig. 4**. **Table 2** shows a detailed description of all DFD elements.

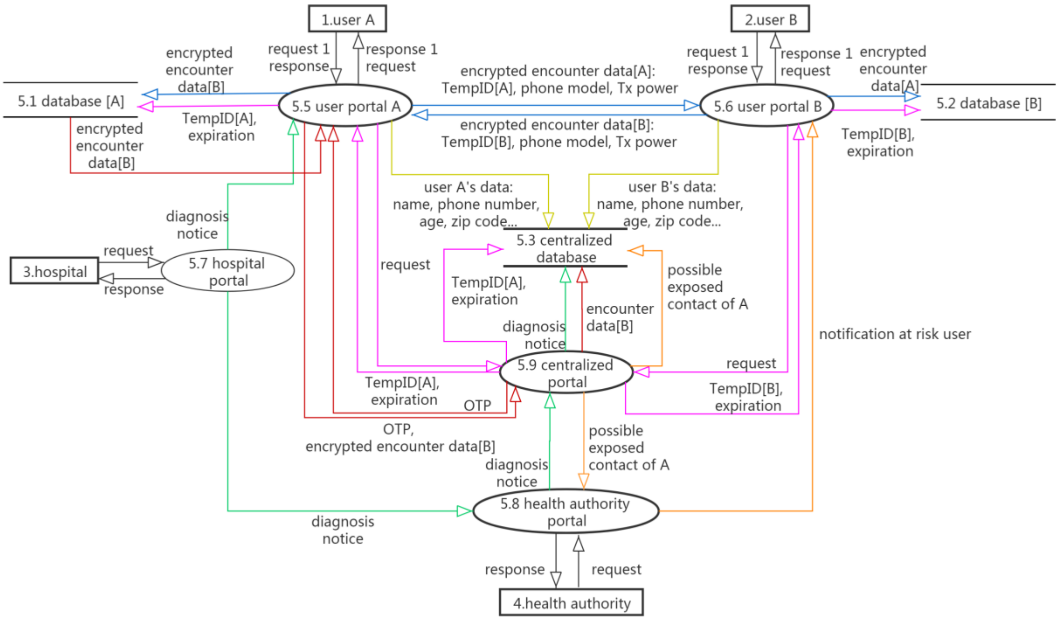

Fig. 1. DFD of centralized APPs based on BLE technology

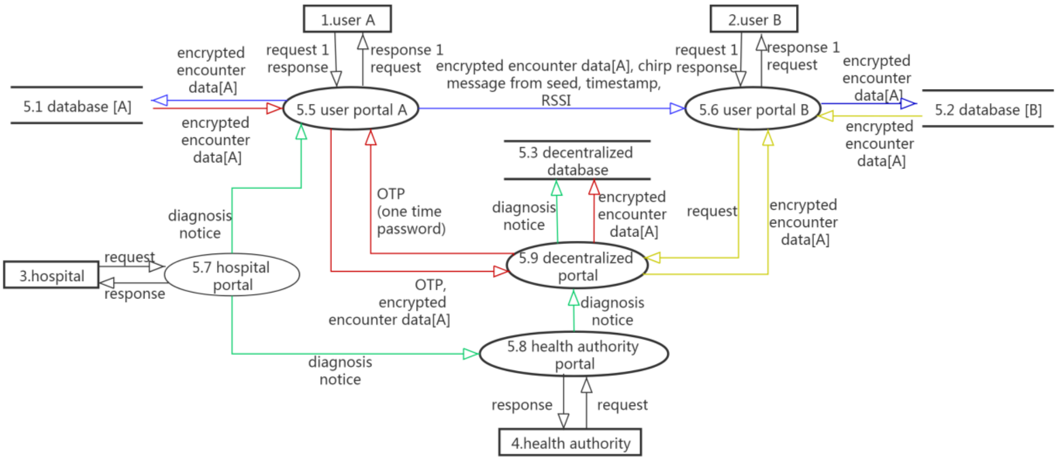

Fig. 2. DFD of decentralized APPs based on BLE technology

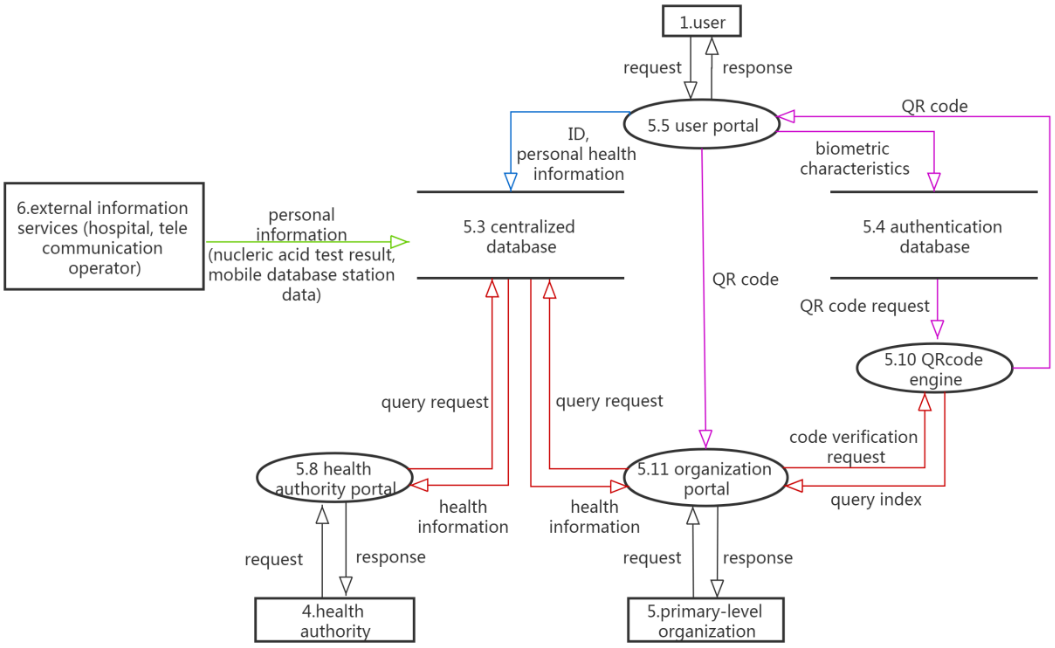

Fig. 3. DFD of centralized APPs based on GPS technology

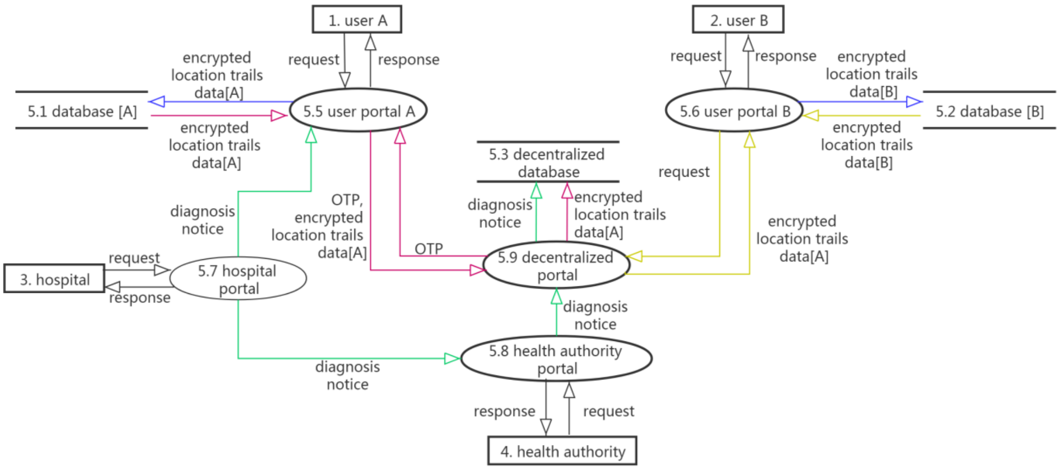

Fig. 4. DFD of decentralized APPs based on GPS technology

Table 2. Detailed description of the DFD elements

Group	Component	Explanation
Entity	user A (1)	A user who installs PHI-code displaying application
	user B (2)	
	hospital (3)	An agency responsible for providing nucleic acid test reports
	health authority (4)	An agency responsible for managing relevant information

	primary-level organization (5)	A user who installs PHI-code scanning application. ex. hospital, railway station, airport, etc
	external information services (6)	Hospital, transport department, telecommunication operator, etc.
Data Store	database[A] (5.1)	Local database on smartphone
	database[B] (5.2)	
	(de)centralized database (5.3)	(De)centralized server database
	authentication database (5.4)	Database for storing identity authentication information
Process	user portal A (5.5)	The user interface front end
	user portal B (5.6)	
	hospital portal (5.7)	The hospital interface front end
	health authority portal (5.8)	The health authority interface front end
	(de)centralized portal (5.9)	The (de)centralized server backend
	QR code engine (5.10)	Generate and verify PHI-code
	organization portal (5.11)	The organization interface front end
...... omit		
Data Flow	health authority - health authority portal (4 - 5.8)	Data flow that requests login
	health authority portal - health authority (5.8 - 4)	Data flow that responses login

3.3 Mapping of Threats to DFD

As shown in Table 3, different DFD element types are subject to different privacy threats. And the relationship is specified in the template provided by LINDDUN. An ‘X’ represents a possible privacy threat type for the corresponding DFD item type.

Table 3. Detailed description of the DFD elements

	L	I	N	D	D	U	N
Entity	X	X				X	
Data Store	X	X	X	X	X		X
Process	X	X	X	X	X		X
Data Flow	X	X	X	X	X		X

Table 4. Contact tracking applications mapping table

	L	I	N	D	D	U	N
Entity		X				X	X
Data Store							X
Process			X				X
Data Flow		X	X				X

In order to simplify the analysis, ignore some unimportant threats and reduce the number of ‘X’, we make some assumptions within a reasonable range. In order to save space, we only show the final mapping table of threatened elements in Table 4. The four DFD elements all face the same threat, that is, policy and consent non-compliance. This threat means that the data subjects do not use personal information in accordance with laws, policies or user agreement. The threat of entity identifiability means that the misactor can identify a user based on a variety of user data. Process and data flow are threated by non-repudiation, which means the user cannot deny having sent a message. The threat of entity unawareness means that the entities that add information are unaware of the consequences of sharing information. The last threat is identifiability of data flow, in short the misleader can abstract user's identity from the data flow. Details are shown in **Table 6**.

Assumptions are as follows:

1. We believe that the back-end of contact tracing applications developed by governments or research institutions are protected against external threats. So internal processes will only be affected by internal threats. Because one process threat is able to represent all of them, we merge process threats and only examine one.

2. In centralized architecture systems, there is non-repudiation threat. For example, in a centralized system based on BLE technology, the unique TempID to each user from central portal and in a decentralized system based on GPS technology, the real-name information from external information services make sure that every action of the user is real-named. So there is irrefutable evidence.

However, in decentralized architecture systems, the situation is just the opposite. Because in these types of systems, the attacker is not able to prove what the user knew, did or said.

3. Detectability is not considered a threat to this particular system. All these system privacy issues are focused on the data itself, rather than the detectability of the data.

4. Non-compliance is a major threat. It is not specific to any part of these systems, but affects the entire system. Therefore, we do not distinguish the various DFD elements of this threat.

5. In centralized architecture systems, identifiability of entitiy (user) is considered a threat, when all entities use systems, they need to login with unique TempID or real-name authentication.

However, in decentralized architecture systems, the situation is just the opposite. In order to further protect privacy, the identity of the entity (user) needs to be hidden in these type of systems.

6. Identifiability threat of data flow only exists in a centralized system based on GPS technology. Because in this system, the transmission of information is not encrypted, so that users cannot hide the connection between identity and information.

7. Linkability of entities (user, hospital, health authority, primary-level organization, external information services) is not considered a threat no matter in which system. When all entities use systems, they need to login with unique TempID or real-name authentication. This is determined by the definition of linkability. In a centralized system, the identity of an entity can be determined by a unique ID, which does not meet the prerequisites for linkability. In a decentralized system, there is no linkability due to information encryption. For the same reason, linkability of data flows is not considered a threat.

8. Linkability and identifiability do not apply to internal processes. Because knowing that two operations belong to the same user does not infringe the patient's privacy. The user's privacy is only infringed if the content of the behavior is leaked (threat of information leakage).

9. In centralized architecture systems, content unawareness only applies to user, hospital and external information services, who input information in the system, as other entities do not input any information.

However, in decentralized architecture systems, entities do not threatened by content unawareness because users with the right to input information know that their encrypted information will only be decrypted on other users' local devices, that is, they understand the consequences of sharing information.

10. Data stores are fully protected in all types of systems and that any attacks are not possible for malicous purpose.

11. Information Disclosure threat of the data flows is not considered because they require a lot of analysis and the extracted information does not match the workload.

12. Information Disclosure threat of the processes is not taken into account because all processes are correctly implemented.

3.4 Identify and document threat scenarios

Table 5. Threat tree of four kinds of APPs

Threat tree			
Non-Compliance (four kinds of APPs)			
1	NC		
	1.1	NC_2: Incorrect or insufficient privacy policies	
		1.1.1	NC_3: Inconsistent/insufficient policy management
Identifiability of Entity (centralized APPs)			
1	I_e		
	1.1	I_e1: identifiable login using untrusted communication	
		1.1.1	I_e2: identifiable log-in used
Non-Repudiation (centralized APPs)			
1	NR_df		
	1.1	NR-df2: no or weak deniable encryption	
		1.1.1	NR-df11: prove data can be decrypted to the plaintext
1	NR_p		
	1.1	NR_p1: process is securely logged	
Unawareness (centralized APPs)			
1	U		
	1.1	U_2: unaware of stored data	
		1.1.1	U_5: unable to review personal information
Identifiability of Data Flow (centralized APPs based on GPS technology)			
1	I_df		
	1.1	I_df2: identifiability of contextual data	
		1.1.1	I_df6: non-anonymous communication traced to entity

Table 6. Misuse cases of threat tree

MUC	Description
MUC01	Threat Tree: NC (Non-Compliance) Summary: The data subjects (primary-level organization, health authority) do not process personal (health) information in compliance with legislations, policies or user agreement. Primary misactor: Insider/outer person/system operator Basic path: Bf1. The misactor does not adhere to national guidelines or laws (e.g. the user's personal (health) information is passed to third parties) Consequence: The user's personal (health) information is spread to others and even spreaded widely in society, causing inconvenience and psychological pressure on user's live. When detected, the reputation of the health authority even the entire government is damaged DFD element(s): process, data flow
MUC02	Threat Tree: Identifiability of Entity Summary: A misactor can identify a user based on a variety of user data.

	Primary misactor: Insider Basic path: Bf1. The misactor runs a series of targeted queries on the user data store to get detailed information Bf2. The misactor can abstract the user's identity from the results of individual requests on account of weak anonymization. In addition, he can link some of the results together to provide identifiable information Consequence: The misactor can get the identity of the user although this should be kept secret DFD element(s): entity
MUC03	Threat Tree: Non-Repudiation Summary: The user cannot deny having sent a message. Primary misactor: Insider/system operator Basic path: Bf1. The misactor has the ability to prove that the data can be decrypted into valid plain text Bf2. The misactor can trace back to the user by a secure log that contains a summary of actions Consequence: The misactor can prove that a user knew, did or said something DFD element(s): process, data flow
MUC04	Threat Tree: Unawareness Summary: The entities that add information are unaware of the consequences of sharing information Primary misactor: Authority Basic path: Bf1. The entities are not aware of the result, so they add information to the system that is able to easily identify them Consequences: When related members search information, it returns identifiable information. The user's privacy is infringed because he believes that his information is completely protected DFD element(s): entity
MUC05	Threat Tree: Identifiability of Data Flow Summary: The misleader abstracts the user's identity from the information stream and associates it with the diagnosis Primary misactor: unskilled insider/skilled outsider Basic path: Bf1. The misactor can intercept the dataflow or access external information services Consequences: The misactor understands which user is diagnosed DFD element(s): data flow

Threat tree is the result of the threats that may occur in LINDDUN threat elements confirmed in the 4.2. **Table 5** reflects the threat tree for four kinds of systems. We map the specific threats of this case to the LINDDUN privacy threat tree catalog [52]. Then we use misuse cases to elaborate on specific threats as shown in **Table 6**. The misuse case shows how threats occur through scenario and result. The misuse case structure is provided in LINDDUN privacy threat modeling [50].

3.5 DREAD for Contact Tracing Apps

We use DREAD [59], a threat rating model, to measure the risk of different threats. DREAD is the abbreviation of five evaluation indicators, they are respectively damage potential, reproducibility, exploitability, affected users and discoverability. Using them as a benchmark, assign scores of 1-3 and calculate the total score. The greater the score, the higher the threat level. **Table 7** shows the risk level of each threat.

Table 7. DREAD for contact tracing apps

Threat	D	R	E	A	D	Sum
MUC01	2	3	3	3	3	14
MUC02	2	3	2	3	2	12
MUC03	1	1	2	1	2	7
MUC04	3	3	3	3	3	15
MUC05	2	3	2	3	2	12

3.6 Threat Mitigation Strategy

According to the analysis of this article, the four systems face five threats in total. We propose mitigation strategies to cope with the privacy threats. **Table 8** represents the privacy requirements and mitigation strategies, also known as privacy enhancing technology (PET) corresponding to the misuse cases. In the issue, application designers receive enhanced guidance on the solution selection process.

Table 8. Mitigation strategy

Misuse Cases	Privacy Requirements	Mitigation Strategies
Policy and consent Non-Compliance (MUC01)	Ensure users aware that they have the right to take legal actions in the event of a violation	Users can sue the developer of the application if users' personal data is not processed in accordance with the content.
Identifiability of entity(MUC02)	Use identity management to ensure unlinkability (from an attacker's point of view) between the partial identities of individuals required by the applications is properly maintained	Employ privacy preserving identity management, e.g. proposed in [53], together with user-controlled identity management system [54] to provide user-controlled linkability of personal data.
Non-Repudiation of process and data flow(MUC03)	Plausible deniability of process and data flow	Privacy preserving authentication, e.g. deniable authentication [55] and off-the-record messaging [56].
Content Unawareness of entity(MUC04)	Information inputters should be aware that they only should only provide the minimum amount of personal information they need	Use feedback tools to raise their privacy awareness.
Identifiability of data flow(MUC05)	User anonymity so that the user is not identified by the content; channel confidentiality	Use anonymity system, such as TOR [57] to communicate within the system.

4. Comparison after Modeling Analysis

After analyzing the target privacy threat modeling, we obtain a comparison table of four contact tracking applications regarding the content and quantity of privacy threats and the difference of data storage. **Table 9** shows the comparison of four contact tracking applications.

Table 9. Comparison of four contact tracking applications

		Centralized APPs based on BLE technology	**Decentralized APPs based on BLE technology**	**Centralized APPs based on GPS technology**	**Decentralized APPs based on GPS technology**
Privacy	MUC01	X	X	X	X
	MUC02	X		X	
	MUC03	X		X	
	MUC04	X		X	
	MUC05			X	
	Number of threats	4	1	5	1
Data Storage	(De)centralized server	X	X	X	X
	Device	X	X		X
Evaluation	Advantage	Normally, the APP can make people who have been in contact with the infected person be notified that they are close contacts without revealing the identity of the contact or the location information where the contact occurred.	The APP can make people who have been in contact with people infected with the new corona virus understand that they are close contacts without revealing the identity of the contact or the location information where the contact occurred.	The combination of GPS location data and base station location data improves the accuracy of judging user mobility; in addition, the combination of QRcode improves the security level of public places and reduces the difficulty of finding close contacts.	The APP can enable people who have been in contact with an infected person to be notified that they are close contacts without revealing the identity of the contact.
	Disadvantage	Although measures to protect privacy have been adopted, the central server still has the ability to identify contacts and other relevant information.	If close contacts conceal their identity and do not perform nucleic acid testing or isolation, it will hinder the efficiency and effectiveness of epidemic prevention and control.	The central server has too much non-encrypted user information, and the privacy threat it faces is the highest among the four types of applications.	GPS technology to determine the absolute position is easily affected by complex environments such as indoors.

5. Conclusion and Future Work

Different countries have different national conditions, and the prevention and control strategies and contact tracing methods and programs adopted are very various. Countries such as China and South Korea, which are greatly influenced by Confucian culture and authoritarian traditions, are more inclined to collectivism and have relatively strict daily life and organization. They need an orderly hard core policy with sufficient overall leadership to fight the plague. Therefore, they opted for applications with a centralized structure based on GPS technology that can grasp more information. However, similar anti-epidemic ideas are difficult to implement in most European countries. Personal privacy has always been the focus of Western politics and the protection of human dignity. Most European and American countries believe that personal information is a part of citizens' personality and human rights, and emphasize the respect and protection of citizens' privacy rights to the greatest extent. The Data Protection Regulation (GDPR) passed by the European Union in 2018 is also known as the most stringent personal privacy protection law in European history. Therefore, they tend to choose applications with a centralized or decentralized structure based on BLE technology. But it is undeniable that in the face of prevention and control, no matter which application we use, we have sacrificed privacy to a certain extent.

One of the main concerns in the use of these apps is related to the security and privacy issues. First, in order to better protect personal privacy, we analyze the privacy threats of each contact tracing techniques. Then, we present a comparison study of contact tracing apps for COVID-19 from privacy point of view. We note that each architecture has its strengths and weaknesses. We hope this paper will help the research community to understand different technological of tracing apps and improve the security performance of tracing apps through comparisons. Through the extensive use and deployment of privacy protection contact tracking technology as a technical tool, information synchronization of the epidemic can be achieved more efficiently, and the purpose of precise prevention and control and normalized prevention and control can be achieved. At the same time, the privacy of citizens is effectively protected. As a future work, we will further analyze information storage issue.

References

[1] D. Zhou, P. Zhang, C. Bao, Y. Zhang, and N. Zhu, "Emerging Understanding of Etiology and Epidemiology of the Novel Coronavirus (COVID-19) infection in Wuhan, China," *Preprints*, Feb. 2020. Article (CrossRef Link)

[2] https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[3] https://coronavirus.jhu.edu/map.html

[4] R. Raskar, I.Schunemann, R. Barbar, K. Vilcans, J. Gray, P. Vepakomma, S. Kapa, A. Nuzzo, R. Gupta, and A. Berke, "Apps gone rogue: maintaining personal privacy in an epidemic,"*arXiv:2003.08567 [cs.CR]*, Mar, 2020.

[5] https://www.technologyreview.com/2021/02/24/1014369/10-breakthrough-technologies-2021

[6] L. Angela, C. Peter, C. Joseph, M. Bassam, H. Thaier, "Security Vulnerabilities in Bluetooth Technology as Used in IoT," *Journal of Sensor and Actuator Networks*, vol. 7, no. 3, 2018. Article (CrossRef Link)

[7] "Maximizing Privacy and Effectiveness in Covid-19 Apps,"OpenMined Community, 2020. [Online]. Available: https://blog.openmined.org/covid-app-privacy-advice/

[8] S. Wang, S.Ding, and L.Xiong, "A New System for Surveillance and Digital Contact Tracing for COVID-19: Spatiotemporal Reporting Over Network and GPS(Preprint)," 2020.

[9] Pan-European Privacy-Preserving Proximity Tracing. [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation.

[10] ROBERT: ROBust and privacy-presERving proximity Tracing Claude Castelluccia.
[11] J.Bay, J. Kek, A.Tan, C. S. Hau, L. Yongquan, J. Tan, and T. A. Quy, "Bluetrace: A privacy-preservingprotocolforcommunity-driven contact tracing across borders,"[Online]. Available: https://bluetrace.io/.
[12] COVIDSafe. [Online]. Available: https://github.com/AU-COVIDSafe
[13] TousAntiCovid. [Online]. Available: https://bonjour.tousanticovid.gouv.fr/index-en.html
[14] Trace Together. [Online]. Available: https://github.com/opentrace-community
[15] DP-3T. [Online]. Available: https://github.com/DP-3T
[16] G. Avitabile, V. Botta, V. Iovino, and I. Visconti, "Towards defeating mass surveillance and sars-cov-2: The pronto-c2 fully decentralized automatic contact tracing system," in *Proc. of Workshop on Secure IT Technologies against COVID-19 (CoronaDef)*, 2020. Article (CrossRef Link)
[17] Apple, "Privacy preserving contact tracing," 2020. [Online]. Available: https://www.apple.com/covid19/contacttracing.
[18] Google, "Exposure notification api," 2020. [Online]. Available: https://www.google.com/covid19/exposurenotifications/
[19] R. L. Rivest, J. Callas, R. Canetti, and et al., "The pact protocol specifications," Technical report, vol. 0.1, April, 2020. [Online]. Available: https://pact.mit.edu/wp-content/uploads/2020/04/The-PACTprotocol-specification-ver-0.1.pdf
[20] J. Chan, S. Gollakota, E. Horvitz, J. Jaeger, S. Kakade, T. Kohno, J. Langford, J. Larson, S. Singanamalla, J. Sunshine et al., "Pact: Privacy sensitive protocols and mechanisms for mobile contact tracing," *arXiv preprint arXiv:2004.03544*, 2020.
[21] TCN Coalition, "TCN protocol for decentralized, privacy-preserving contact tracing,". [Online]. Available: https://github.com/TCNCoalition/TCN
[22] Stopp Corona. [Online]. Available: https://github.com/austrianredcross/stopp-corona-ios
[23] CovidAlert. [Online]. Available: https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert.html
[24] PrivateTracer. [Online]. Available: https://gitlab.com/PrivateTracer
[25] Corona-Warn. [Online]. Available: https://github.com/corona-warn-app
[26] Rankingn C-19. [Online]. Available: https://www.covid.is/app/is
[27] C19X. [Online]. Available: https://github.com/c19x
[28] CoEpi. [Online]. Available: https://github.com/Co-Epi
[29] CovidSafe(UoW), 2020. [Online]. Available: https://covidsafe.cs.washington.edu
[30] National Government Service Platform Epidemic Prevention Health Code. [Online]. Available: http://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=ED391A20971F017E8DBC265ECD66CCCE
[31] National Government Service Platform Epidemic Prevention Health Code. [Online]. Available: http://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=672AF632394BC01A8D07B221C799923E
[32] National Government Service Platform Epidemic Prevention Health Code. [Online]. Available: http://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=09EBF512C9729D09237B646E7DBE1652
[33] CoronApp. [Online]. Available: https://apps.apple.com/cn/app/id1502037648#?platform=iphone
[34] MySejahtera. [Online]. Available: https://gamma.malaysia.gov.my/appdetails/721#tab3
[35] H. Lee, S. Kim, S. Lee, "Evaluation Criteria for COVID-19 Contact Tracing Technology and Security Analysis," *Journal of The Korea Institute of Information Security & Cryptology*, vol. 30, no. 6, pp. 1151–1166, Dec. 2020. Article (CrossRef Link)
[36] Pandoa. [Online]. Available: https://github.com/wirewirewirewire/pandoa
[37] Hamagen. [Online]. Available: https://github.com/MohGovIL/hamagen-reactnative/blob/master/README.md
[38] PrivateKit: Safe Paths. [Online]. Available:http://privatekit.mit.edu/

[39] Fraunhofer AISEC, "PANDEMIC CONTACT TRACING APPS: DP-3T, PEPP-PT NTK,AND ROBERT FROM A PRIVACY PERSPECTIVE," Apr. 2020.

[40] S.O.Blacklow, S.Lisker, M.Y. Ng, U. Sarkar, and C. Lyles, "Usability, inclusivity, and content evaluation of COVID-19 contact tracing apps in the United States," *Journal of the American Medical Informatics Association*, 1982-1989, May. 2021. Article (CrossRef Link)

[41] N. Ahmed, R. Michelin, W.Xue, S.Ruj, and S. Jha, "A Survey of COVID-19 Contact tracing apps." *IEEE Access*, 8, 134577-134601, 2020. Article (CrossRef Link)

[42] M. Hatamian, S. Wairimu, N.Momen, and L. Fritsch, "A privacy and security analysis of early-deployed COVID-19 contact tracing Android apps," *Empirical Software Engineering*, vol. 26, no. 3, 2021. Article (CrossRef Link)

[43] M.L.Messai, and H.Seba, "Short Paper: Privacy Comparison of Contact Tracing Mobile Applications for COVID-19," Oct. 2020.

[44] B.Sowmiya, V. S.Abhijith, S.Sudersan, R. S. J. Sundar, M.Thangavel, and P. Varalakshmi, "A Survey on Security and Privacy Issues in Contact Tracing Application of Covid-19," *SN Computer Science*, vol. 2, no. 3, pp. 1-11, 2021. Article (CrossRef Link)

[45] A.A. Harith, N. A. Muhamad, and R. Griffiths, "Digital Contact Tracing in Combating COVID-19 Pandemic in Malaysia, New Zealand and China," 2021. Article (CrossRef Link)

[46] M.J. Parker, C. Fraser, L.Abeler-Dörner, and D. Bonsall, "Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic," *Journal of Medical Ethics*, vol. 46, no. 7, 2020. Article (CrossRef Link)

[47] H. Cho, D. Ippolito,and Y. W. Yu, "Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs," *arXiv:2003.11511 [cs.CR]*, Mar. 2020.

[48] L.S. Park, J. Singh, J.Malbeuf, and A. A.Mardon, "Evaluation of Effectiveness of Digital GPS Contact Tracing Technology in Response to COVID-19," *The Pacific Journal of Science and Technology*, vol. 21, no. 2, pp. 341-345, Nov. 2020.

[49] A.Gangavarapu, E.Daw, A. Singh, R.Iyer, and R.Raskar, "Target Privacy Threat Modeling for COVID-19 Exposure Notification Systems," *arXiv:2009.13300 [cs.CR]*, Sep. 2020.

[50] L. Sion, K.Wuyts, K.Yskout, D. V. Landuyt, and W.Joosen, "Interaction-Based Privacy Threat Elicitation," in *Proc. of IEEE EuroS&PW*, 2018. Article (CrossRef Link)

[51] K. Wuyts, and W. Joosen, "LINDDUN privacy threat modeling: a tutorial," Department of Computer Science, KU Leuven; Leuven, Belgium, Report CW 685, Jul. 2015.

[52] M. Deng, K. Wuyts, R. Scandariato, B. Prenee, and W. Joosen, "a privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements,"*Requirements Engineering*, vol. 16, pp. 3-32, 2011. Article (CrossRef Link)

[53] K. U.Wuyts, R. U.Scandariato, and W. U. Joosen, "LIND(D)UN privacy threat tree catalog," Department of Computer Science, KU Leuven; Leuven, Belgium, Report CW 675, Sep. 2014. Article (CrossRef Link)

[54] M.Hansen, P. Berlich, J. Camenisch, S. Clauß, A. Pfitzmann, and M. Waidner, "Privacy-enhancing identity management," *Information Security Technical Report*, vol. 9, no. 1, pp. 35–44, 2004. Article (CrossRef Link)

[55] S. Clauß, A. Pfitzmann, M. Hansen, and E.V. Herreweghen, "Privacy-enhancing identity management,"*The IPTS Report*, 67, 8-16, Jan. 2002. Article (CrossRef Link)

[56] M. Naor, "Deniable ring authentication," in *Proc.Crypto*, pp. 481–498, 2002.

[57] N. Borisov, I. Goldberg, and E. Brewer, "Off-the-record communication, or, why not to use pgp," in *Proc. of the 2004 ACM Workshop on Privacy in the Electronic Society, WPES 2004*, Washington, DC, USA, pp. 77–84, 2004.

[58] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proc. of the 13th USENIX Security Symposium*, August 2004.

[59] A. hostack, "Experiences Threat Modeling at Microsoft," in *Proc. of Modeling Security Workshop, Dept. of Computing*, 2008. Article (CrossRef Link)

Yanji Piao received the Ph.D. degree from the Department of Computer Science, Ajou University, Korea, in 2014. She is currently an assistant professor with the Information Management and Information System, Yanbian University, China. Areas of her current interests include information security, intelligent computing, and artificial intelligence.

Dongyue Cui is a master student majoring in technical economy and management in the School of Economics and Management, Yanbian University, China. She received her bachelor degree from Jilin University, China, in 2020. Her research interests include network security and data mining.