2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM) | 978-1-7281-6267-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/HEALTHCOM49281.2021.9398978

Blockchain-empowered Contact Tracing for COVID-19 Using Crypto-spatiotemporal Information

Zheng WEN, Keping YU, Xin QI, Toshio SATO, Yutaka KATSUYAMA, Takuro SATO, Wataru KAMEYAMA
Faculty of Science and Engineering
Waseda University
Tokyo, Japan
robinwen@fuji.waseda.jp

Fumiyuki KATO, Yang CAO, Masatoshi YOSHIKAWA
Department of Social Informatics
Kyoto University
Kyoto, Japan
cao.yang.8x@kyoto-u.ac.jp

Min LUO, Jun HASHIMOTO
EY Japan
Tokyo, Japan
min.luo@jp.ey.com

***Abstract*— The pandemic of the COVID-19 [1] has re-awakened people that viruses are still the greatest threat to human society. Quarantining the patients and tracking close contacts has been used for hundreds of years in the battle between humans and the plague, which are still useful today. In the information society, we can employ information communications technology (ICT) to suppress the spread of epidemics and lower the epidemic curve. By using spatiotemporal information, we can trace the trajectories of patients and their close contacts. However, spatiotemporal information also involves personal privacy, and it has become a topic of concern about whether people's privacy should be sacrificed for epidemic control. In this paper, we propose a close contact tracing solution based on crypto-spatiotemporal information (CSI). First, the solution encrypts spatiotemporal information to protect personal privacy. Then, it uses a blockchain platform to realize the proof of CSI and uses Intel SGX [2] based trusted execution environment [3] to perform close contact judgment. Finally, it can trace close contacts while protecting personal privacy. The evaluation results indicate that the advantages and efficiency of the proposed scheme are significant.**

Keywords-component; COVID-19; close contact tracing; blockchain; crypto-spatiotemporal information(CSI); trusted execution environment (TEE)

I. INTRODUCTION

In the human recorded history, the earliest plague was the Great Plague of Athens, which occurred in 430 BC [4]. Plagues such as the Black Death [4], Influenza [4], and SARS [4] not only brought death and destruction to human civilization, but also drove the survivors to pursue the scientific truths hidden behind the virus, to improve medical technology, and to improve social policies to perpetuate human civilization. In late 2019, the COVID-19 [1] outbreak, which has become a global pandemic, causes a tremendous crisis. COVID-19, as a common enemy of all human beings, has caused tremendous losses to the world, and at the time of writing, the worldwide epidemic of the virus is still spreading.

The information from World Health Organization (WHO) [5] suggests that current evidence indicates that COVID-19 spreads between people through direct, indirect (through contaminated objects or surfaces) or close contact with infected people via mouth and nose secretions. The COVID-19 viruses are released from the mouth or nose when an infected person coughs, sneezes, speaks, or sings.

To avoid contact with these virus droplets, it is important to stay at least 1 meter away from others, clean hands frequently, and cover the mouth with a tissue or bent elbow when sneezing or coughing. When physical distancing (standing one meter or more away) is not possible, wearing a fabric mask is an important measure to protect others [5].

According to WHO recommendations [5], limiting close contact between an infected person and others, rapid identification of patients and close contacts, and isolation in appropriate facilities can be effective in stopping the spread of the virus. By using spatiotemporal information, we can re-establish the path of virus transmission and trace infected persons and close contacts.

Although tracing infection routes could be an effective way to control the spread of viruses, obtaining users' historical trajectories inevitably involves privacy issues. How to leverage the epidemic prevention while protecting privacy is a complicated problem.

Therefore, in this paper, we use crypto-spatiotemporal information (CSI) to trace the infected person and close contacts. The verifiability of the CSI is supported by the blockchain [6] platform. In the proposed system, the CSI is decrypted and judge by the trusted execution environment-based system (TEE) [3], and users do not need to worry about the private problem. This solution enables a highly efficient method for tracing and managing the virus transmission routes and judging close contacts, meanwhile protecting privacy.

II. RELATED WORKS

Virus as a threat to all human beings, researchers, institutions, and companies around the world are making contributions to the fight against the pandemic.

A. Healthcode in China

In China, where the epidemic first struck, a technology called health code [7] is widely used. Through big data

including personal reporting and location tracking, it is possible to identify whether the user has had close contact with the virus and control the citizens' travel through different colors (red, yellow, and green). Health code is more than just color information; it can also be formed by QR code, which can be scanned by prevention agents to verify the user's health status. The health code function is integrated with social communication software, which is commonly used by Chinese people like WeChat [8], QQ [8] and Weibo [8]. It is required to be presented at checkpoints and on public transportation, so that epidemic prevention agents can identify the user's health status and effectively control the spread of viruses.

B. TtraceTogether-Singapore

The Singapore government, which is one of the first countries that use apps for epidemic prevention and control, has provided many valuable experiences. The Singapore government has released TraceTogether [9], which uses Bluetooth to broadcast anonymous messages, where users using the same app can send and receive anonymous IDs. Once a user is judged to be COVID-19 positive, the anonymous ID sent by his/her device is highlighted. The user can look up his/her contact ID history to check if he/she is in close contact with the patient or not.

As the government-operated pandemic prevention and control app, TraceTogether has a great point; however, these IDs are anonymous to the user and only visible to the government. Central-node based systems are concerned with security and privacy.

C. Google Apple API

As the providers of the most mobile phone operating systems in the world, Google and Apple have jointly released the Close Contact Tracking API [10] to health administrations. The API empowers health administrations in each country by broadcasting anonymous IDs via Bluetooth listening to anonymous IDs from other devices, and the infected person can submit their anonymous IDs to health administrations. The API is allowing users to identify if users are close contact with the COVID-19 by inquiring positive IDs.

Although there are many close contact tracing and health status applications available, the COVID-19 epidemic has not significantly controlled. Close contact tracing using anonymous Bluetooth or QR code could be effective in protecting the privacy of citizens, but it is ineffective in countries with low rates of usage. Countries with mandatory use of GPS (Global Positioning System), personal information have better results in epidemic control, but civil rights and privacy concerns remain as shown in Figure 1.

Therefore, this paper provides a close contact tracing solution that utilizes personal trajectory data while protecting privacy by a trusted environment. Details are presented in the following sections.

III. SYSTEM DESIGN

According to the information from the World Health Organization and the U.S. CDC (Centers for Disease Control and Prevention) [5] [19], the COVID-19 is spread primarily between humans in close contact environments by such as droplets. For COVID-19, close contact is defined as any individual who is within 6 feet of an infected person for at least 15 minutes starting from 2 days before illness onset (or, for asymptomatic patients, two days before positive specimen collection) until the time the patient is isolated. By calculating the spatiotemporal information between individuals, we can determine whether the target is in close contact with the virus. In the era of information, most people always carry smartphones, and smartphones can record their spatiotemporal information.

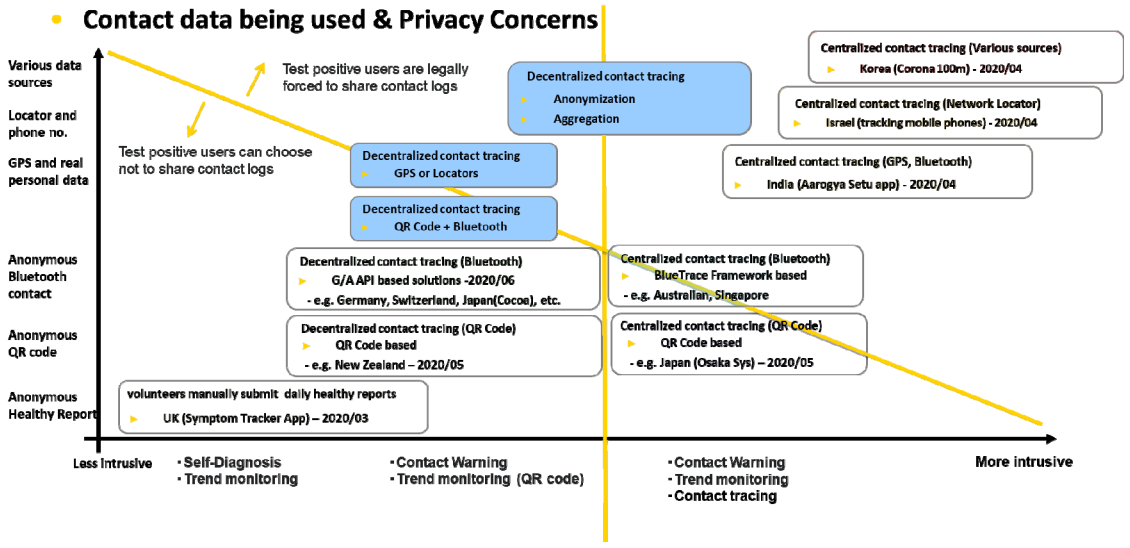

Figure 1. Existing COVID-19 contact tracing solutions.

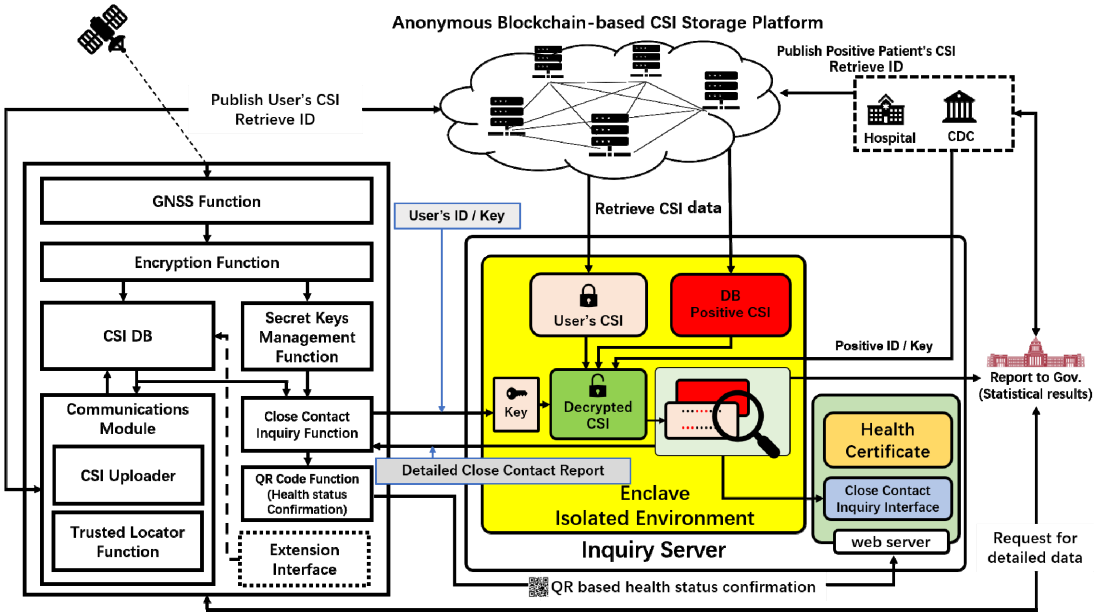

Figure 2. The overview of our system.

As extremely sensitive private data, spatiotemporal information leaks and misuses might have serious consequences. Encrypting personal spatiotemporal information protects privacy, but most of the encrypted data cannot be utilized in identifying close contact. Although homomorphic encryption [11] provides a way to compute encrypted data without decoding, the time and performance cost are inefficient and cannot be utilized by existing devices in the short term. Decoding CSI using secret keys in an insecure environment can be a new security risk.

In our proposed system as shown in Figure 2, we utilize TEE to decode and judge the CSI in a completely isolated execution environment, and it can secure users' privacy. The privacy data inside the TEE is not accessible even to the administrator of the TEE system. Detailed close contact reports is only fed back to the relevant inquirer, and the private information is removed. The inquiry server provides a CSI certificate function for user queries, and reports statistical results to the government to produce early warning of disease cluster, and infection pathways tracing. Medical institutions, such as hospitals and CDC, can also publish CSI to a blockchain platform, providing virus exposure history for the inquiry server.

The last but significant part is the verifiability of CSI. In our proposed system, the trusted app records unaltered GNSS (Global Navigation Satellite System) information; third-party locators are utilized to verify the user's current spatiotemporal information. The CSI is uploaded in a blockchain platform, enabling the CSI to be verifiable, trustworthy, and unalterable.

A. *Crypto-spatiotemporal Information*

As two different reference systems, time and space can be used to observe and describe individuals' movements in space over the time dimension. In computer systems, there are many ways to encode time and geographical location. Common geocode includes address, coordinates, map-code, postal code, geohash [12], etc. For scalability and indexability, we are using Geohash (length of 9) to represent the geographical entity, where Unix timestamp (precision to the second) is used to index time as shown in Figure 3.

Timestamp	Geohash	Crypto-Spatiotemporal information
1577808060	xn775hh8b	1g2KZy34ABF6HElBxxBAkw+A UFXe9G28aA4ODYWmbl12dii4

Figure 3. Crypto-spatiotemporal information

The second level timestamp is sufficient to represent close contact. Geohash is a hierarchical geocoding system that divides space into grids by using z-curves and converts geographic locations into strings. The length of a Geohash string can be used to demarcate the size of an area. Geohash guarantees that the longer a shared prefix between two Geohashes is, the spatially closer they are together. With the Unix timestamp and Geohash, it is easy to retrieve and store spatiotemporal information.

Finally, AES-GCM [13] is applied to encrypt the spatiotemporal information to generate the CSI as shown in Figure 3.

B. *Proof of CSI*

The verifiability of CSI also affects the efficiency of close contact tracing. Inaccurate spatiotemporal information may reduce the accuracy of the tracing system; falsified spatiotemporal information may interfere with the pandemic management. By using trusted third-party locators and trusted application, trusted spatiotemporal information can be obtained.

1) Third-party Locator-based Verification

Third-party locators can generate trusted geographic information for users. It is an open standard that can be implemented by any systems which can identify and locate the user. The locator system can utilize the surveillance system to identify the user's identity and location in a specific place, generating trusted spatiotemporal information. The users' payment activities, such as using the public transportation system, using credit cards, and making digital payments, can also be used to verify the spatiotemporal information. Users can also use public Wi-Fi access points, base stations, and triangulation to locate the user's spatiotemporal location, as shown in Figure 4.

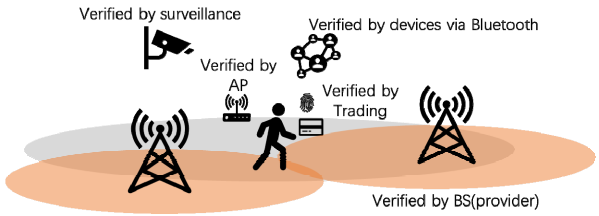

Figure 4. Third-party locators

2) Trusted Application

Both Android and iOS have many functions to protect users' privacy and data security.

As the most effective method, GNSS-capable device-based tracking applications could comprehensively record the user's outdoor trajectories during the pandemic. Secured and trusted technologies (Such as Trusty TEE [14], Secure Enclave [15], etc.) enable the application to obtain trusted GNSS records. In the case of no third-party locators, the recording user's trajectories by the trusted application are critical.

C. Blockchain-based CSI Storage Platform

The spatiotemporal information recorded by third-party verification or trusted apps can be uploaded to the blockchain platform.

The blockchain is decentralized, open, independent, secure and anonymous, and is the best platform for storing verified CSI. The verified CSI is written to the blockchain platform anonymously where the data is synchronized to each peer node, and all information is open and transparent and cannot be falsified. By encoding the data access control logics into a smart contract on the blockchain, it ensures only the trusted locators/applications are allowed to upload data to the blockchain platform, while only the transactions authorized by the CSI's owner can access the CSI. Without the secret key, the CSI cannot be decoded. During the judgment of CSI, the user only needs to submit the transaction ID to the inquiry system, instead of uploading all the CSI data. Figure 5 illustrates the data transaction from and to the blockchain platform.

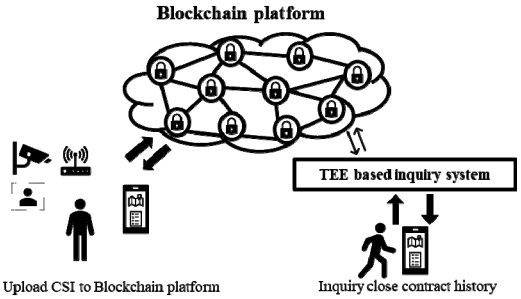

Figure 5. Blockchain-based CSI storage platform

D. The TEE-based Close Contact Inquiry System

By creating enclaves [2] and importing spatiotemporal data into the isolated environment, the Intel SGX-based TEE system can avoid the risk of privacy issues. The system is shown in Figure 6.

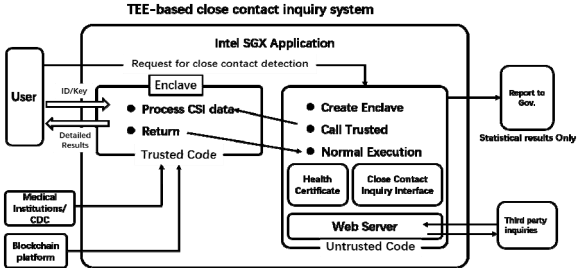

Figure 6. TEE-based close contact inquiry system

The close contact inquiry system can judge the close contacts by comparing spatiotemporal data of patients and users. The CSI cannot be utilized directly and requires the user to provide the key to decrypt the CSI. The privacy data leakage and misuse might cause serious harm to users. Therefore, the users' secret keys are directly inputted into the enclaves through a secured channel, and the operations such as decryption of CSI and close contact judgment are performed inside the enclaves. The detailed close contact report is fed back to the user. Anyone, even the administrator of the TEE system, can only have close contact related statistical results without any private data, and users' private data are not to be disclosed out of the enclave.

As the difference from the existing contact tracing technologies, TEE provides us with the leverage to balance privacy and pandemic surveillance. It realizes to utilize secured privacy data efficiently.

IV. Field Experiment and Evaluation

A. Field Experiment

Limited by the currently ongoing epidemic, we only performed a small-scale field experiment to verify our proposal. The experiment utilized IBM Hyperledger fabric [16] to construct the blockchain platform. The user terminal utilized AES-GCM to encrypt the spatiotemporal information. An Intel SGX-based TEE platform was utilized to judge the close contact. The experiment parameters are shown in TABLE I.

TABLE I. EXPERIMENT PARAMETERS

Experiment location		Nishi-Waseda Campus, Waseda University, Tokyo
TEE environment		HP Z2 SFF G4 Workstation(Intel Xeon E-2174G, 64GB RAM)
Blockchain platform		AWS D2 2 Cores 7GB RAM
Locator	Node	Intel Compute Stick BLKSTK2MV64CC(Intel Core m5-6Y57, 4GB RAM)
	Camera	Logitech C270
Terminals		Android
Close contact conditions		2 meters / 10 mins

The experiment was performed in two scenarios, indoor and outdoor, as shown in Figure 7.

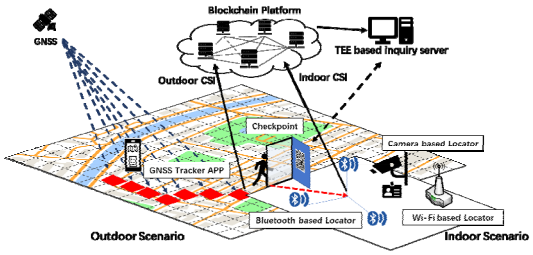

Figure 7. Experiment overview.

1) Outdoor Scenario

We use Android-based GNSS trackers to record users' trajectories. The users' terminals encrypt the GNSS records into CSIs and upload them to the blockchain platform anonymously.

2) Indoor Scenario

The experiment is focused on verifying the locator technology. The positioning of the user is recorded by the device with the functions of face recognition, Bluetooth and Wi-Fi applications. By using the recorded facial information and terminals' MAC addresses, when the experimental participant approaches the locator node, his/her identity can be recognized, and the CSI can be uploaded to the blockchain platform.

When entering a specific room in the experiment, the participants were asked to scan a QR code at the checkpoint and uploaded their CSI information to the TEE-based close contact inquiry system. The TEE-based inquiry system compares the user's CSI with the trajectories of the positive patients, finally feeds the user's current health status back to the terminal. We mark one of the participants as a COVID-19 positive patient. The participant who had long time exposure with a positive patient is reported as a close contact. At the checkpoint, we can detect close contacts in time. Some of the experiment pictures are shown in Figure 8.

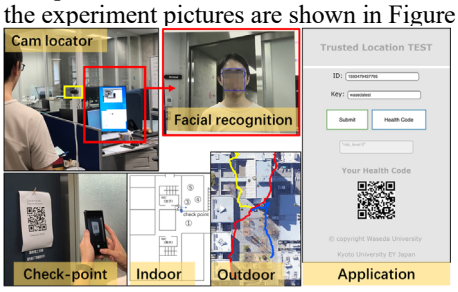

Figure 8. Experiment pictures

By analyzing the raw spatiotemporal information of two participants, we can clearly understand that using CSI can be effective in detecting close contacts. The conditions of close contact can be modified to match more diseases. Figure 9 shows users' trajectories and the detection of close contact.

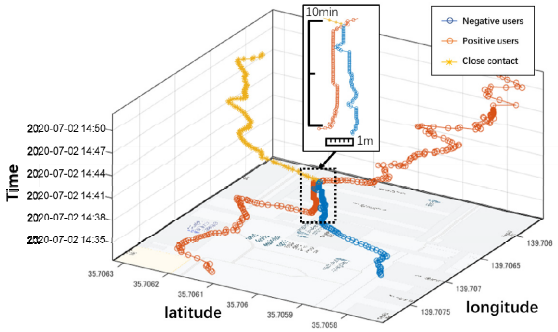

Figure 9. Close contact judgment

B. Evaluation

We evaluate the performance of the system by simulating large-scale accesses, as shown in TABLE II. By using techniques such as GNSS and facial recognition, we can obtain the user's spatiotemporal data without technological difficulties. As a worldwide epidemic, the virus can be defeated only by most people engaging with it. Insufficient users can affect the efficiency of the system. Massive users can also generate many payloads on the system.

TABLE II. DEPLOYMENT REQUIREMENTS

Region	Population	Total Tx/Day	Payload/s	Chains
All JAPAN	130M	58M	2 GB/sec	7
Metropolitan Area	38M	17M	0.63 GB/sec	2
Tokyo	14M	6.24M	0.23 GB/sec	1
Osaka	2.7M	1.2M	0.045 GB/sec	1

In our proposal, one single CSI size is approximately 0.104KB. With five seconds logging interval, each user generates 1.8MB CSI per day. The IBM Hyperledger Fabric-based blockchain platform enables 100TPS (less than 16 nodes) and 100MB per transaction [17]. During epidemics, reducing unnecessary outdoor activities is an effective way to reduce the possibility of infection. It assumes that only one person per household goes out per day, and the amount of transaction to the blockchain platform is about 58M per day in Japan [18]. It also requires at least seven chains to maintain the 58M transaction.

As a privacy-protecting close contact query interface for the public, the performance of the TEE platform is also a significant bottleneck for the system. By simulating the operations of the epidemic data, we obtain the following data as shown in Figure 10, where the Intel SGX enclave initialization time increases with its memory consumption.

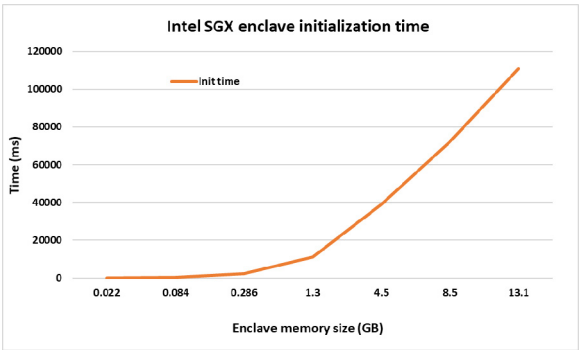

Figure 10. Enclaves initialization time

The memory consumption of the enclave increases with the volume of positive patient data. We evaluate close contact judgment time consumption with four different patient data sets. The client's CSI data volume is set to 4000, 45000, and 100000, which represent different record periods and different CSI logging intervals. When the patient data are 0.1 MB (data size:1.9 MB) and 1M (data size:19 MB), the close contact judgment time is less than 100ms. When the patient data increase to 10 MB (data size: 190 MB) and 100 MB (data size: 1900 MB), the enclave's memory consumption and close contact judgment time are significantly increased. With 100,000 MB client-side CSI data and 100 MB server-side patient data, the close contact judgment time is close to 10 seconds. When the CSI logging interval is 30 seconds, two weeks generate about 40,000 CSI data, and the TEE judgment times are all less than 5 seconds, as shown in Figure 11.

Figure 11. Close contact judgment time

V. CONCLUSION

The proposed system provides a close contact tracking solution that uses spatiotemporal information while protecting privacy. It records and encrypts trusted personal spatiotemporal information and publishes it to the blockchain platform to generate an unalterable verifiable record. The TEE platform utilized to decrypt and match the CSI in an isolated environment significantly secures the user's private data. As a prototype of close contacts, improvements are still needed in many areas. This solution provides a novel approach to the use and protection of personal privacy. It balances between privacy protection and epidemic prevention and control.

ACKNOWLEDGMENT

This research and development work was supported by the MIC/SCOPE #205003002, "Certification System Using Simplified LPWA based Distributed Ledger Technology".

REFERENCES

[1] Coronavirus disease (COVID-19) pandemic. (2020). WHO. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/ (last accessed on 08/24, 2020).

[2] McKeen, Frank, et al. "Intel® software guard extensions (intel® sgx) support for dynamic memory management inside an enclave." Proceedings of the Hardware and Architectural Support for Security and Privacy 2016. 2016. 1-9.

[3] Sabt, Mohamed, Mohammed Achemlal, and Abdelmadjid Bouabdallah. "Trusted execution environment: what it is, and what it is not." 2015 IEEE Trustcom/BigDataSE/ISPA. Vol. 1. IEEE, 2015.

[4] List of epidemics. (2020). Wikipedia. https://en.wikipedia.org/wiki/List_of_epidemics (last accessed on 08/24, 2020).

[5] COVID19 Q&A Detail. (2020). WHO. https://www.who.int/news-room/q-a-detail (last accessed 08/24, 2020)

[6] Swan, Melanie. Blockchain: Blueprint for a new economy. "O'Reilly Media, Inc.", 2015.

[7] Liu, Wei, Xiao-Guang Yue, and Paul B. Tchounwou. "Response to the COVID-19 Epidemic: The Chinese Experience and Implications for Other Countries." (2020): 2304.

[8] Li, Jiawei, et al. "Data mining and content analysis of the Chinese social media platform Weibo during the early COVID-19 outbreak: retrospective observational infoveillance study." JMIR Public Health and Surveillance 6.2 (2020): e18700.

[9] Bay, Jason, et al. "BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders." Government Technology Agency-Singapore, Tech. Rep (2020).

[10] Panzzarino, M. "Apple and Google are launching a joint COVID-19 tracing tool for iOS and Android." Tech Crunch 10 (2020).

[11] Gentry, Craig, and Dan Boneh. A fully homomorphic encryption scheme. Vol. 20. No. 9. Stanford: Stanford university, 2009.

[12] Niemeyer, Gustavo. "Geohash." Retrieved June 6 (2008): 2018.

[13] Salowey, Joseph, Abhijit Choudhury, and David McGrew. "AES Galois Counter Mode (GCM) cipher suites for TLS." Request for Comments 5288 (2008).

[14] Developers, Android. "Trusty tee." (2018).

[15] Han, Byron B., and Craig A. Marciniak. "Enrollment using synthetic fingerprint image and fingerprint sensing systems." U.S. Patent No. 8,913,801. 16 Dec. 2014.

[16] Cachin, Christian. "Architecture of the hyperledger blockchain fabric." Workshop on distributed cryptocurrencies and consensus ledgers. Vol. 310. No. 4. 2016.

[17] Gorenflo, Christian, et al. "Fastfabric: Scaling hyperledger fabric to 20,000 transactions per second." 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). IEEE, 2019.

[18] Population and Households. (n.d.). MIC, Japan. https://www.soumu.go.jp/main_content/000633277.pdf (last accessed on 08/24, 2020)

[19] Coronavirus Disease 2019 (COVID-19) - Transmission. (2020, June 16). Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html (last accessed on 08/24, 2020)