December 8, 2022

Toshiba Corporation
Tohoku University Tohoku Medical Megabank Organization
Tohoku University Hospital
National Institute of Information and Communications Technology

Toshiba, ToMMo, Tohoku University Hospital, and NICT Link Quantum Security and Personal Authentication, Successfully Deliver Secure Personalized Healthcare Use Case

Genome data of numerous individuals stored and transmitted in theoretically secure method and utilized only with individual consent

Summary

Toshiba Corporation, Tohoku University Tohoku Medical Megabank Organization (ToMMo), Tohoku University Hospital, and the National Institute of Information and Communications Technology (NICT) have demonstrated the world's first[*1] personalized healthcare[*2] system that stores genome data from many individuals in multiple locations and utilizes them for medical treatment and healthcare using an information theoretically secure method based on the quantum key distribution (QKD) link, the secret sharing system and personal authentication technology. This system is theoretically secure against the threat of store now and decrypt later attacks, prevents data leaks, falsification, and loss of genome data. In this system, data decryption and reconstruction[*3] are performed by using personal authentication and individual consent. The system is expected to contribute to the realization and spread of personalized healthcare.

A part of this work was performed for Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), “Photonics and Quantum Technology for Society 5.0” (Funding agency: QST).

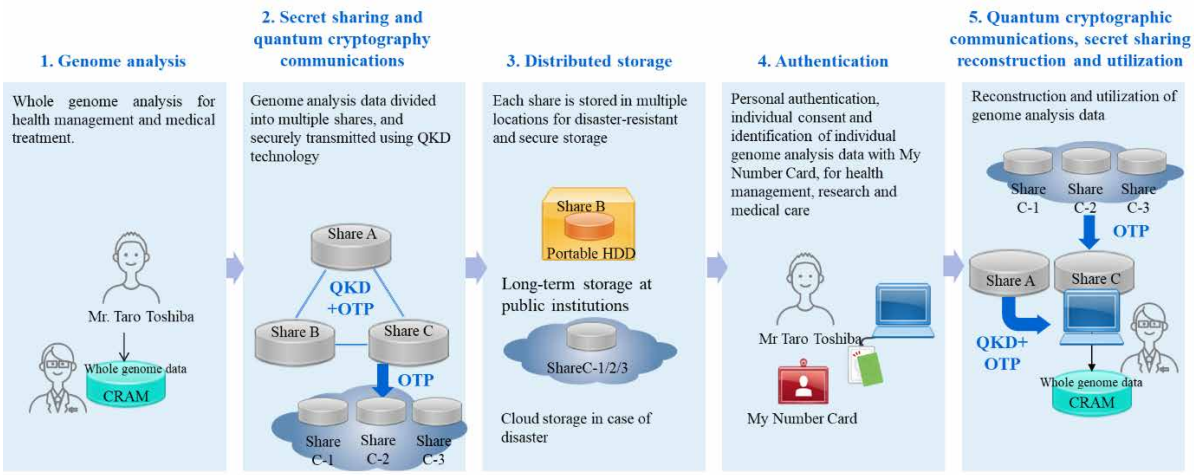

Fig. 1 Entire sequence of personalized healthcare

Development Background

Treatment of patients has long been provided on the basis of identification of symptoms and disease, plus factors like gender and age. However, advances in genome data analysis technology are opening the way to personalized healthcare that combines an individual's genetic information with environmental factors, particularly lifestyle patterns, and then calculates risk of disease, and advises on optimal preventive measures. It is an approach with recognizing the diversity of individuals.

However, personalized healthcare relies on analysis of individual genome data and requires strict security in transmission, storage and utilization of personal health data. In Japan, for example, genome data is recognized as a personal identification code that must be treated and protected as such under the Amended Act on the Protection of Personal Information. If leaked, it can pose a multi-generational risk for families, and has to be protected for a span of centuries. Given this, personalized healthcare must be supported by a framework that can accommodate genome data analysis technology, plus secure data transmission and storage of genome data associated with individual IDs, and data decryption, reconstruction and use must be done only with individual consent.

In July 2021, Toshiba, ToMMo, Tohoku University Hospital and NICT used distributed storage technology based on quantum security technology to successfully demonstrate the world's first experimental backup of large-scale genome analysis data to multiple sites, and their successful reconstruction (announced in August 2021[*4]). This was achieved by combining the QKD link, which is grounded in quantum mechanics and realizes encrypted communications secure against any attempt at wiretapping or decoding, with secret sharing system, which realizes secure data storage by converting original data into multiple distributed fragments (shares) that look like random numbers. However, this backup method focused on bulk transmission and the storage of large-volume data. Managing individual genome data for many people is much more difficult. In addition, the functions of the QKD and secret sharing systems were implemented and carried out independently of one another, as it is difficult to operate them as a unified large-scale system. The next step was clear: to develop a technology for efficient operation of large-scale systems.

Features of This Technology

The four partners have now done this by developing an integrated key management and share control system. They have demonstrated its ability to advance personalized healthcare with a use case that demonstrates secure storage and reconstruction of bulk genomic data, plus storage of genome data on numerous individuals in multiple locations, that can be reconstructed as and when necessary, with personal authentication.

The integrated platform provides unified management and operation of quantum cryptography, secret sharing, and personal authentication. It integrates functions for generating cryptographic keys and random numbers, which are used in large numbers in quantum cryptography and secret sharing, and achieves unified operation of data transmission and storage. As it provides cryptographic keys and random numbers in the same format, and they can be used interchangeably, it realizes highly efficient operation of a large-scale system.

Transmission of large volumes of genomic data requires a large number of quantum encryption keys, but the speed at which they can be generated is limited. Until now, distributed backup systems that store multiple shares at multiple locations have done so at fixed locations. However, the current use case requires distributed storage of large volumes of data on individuals at any given time, which would prevent efficient use of quantum cryptographic keys. The new share control system uses information on remaining quantum encryption keys at each distributed location to determine optimal storage locations, and associates the information with individual IDs. This realizes efficient, secure storage of large volumes of personal data at multiple locations in a large-scale system, and also facilitates data reconstruction and secure use by using personal authentication to identify and reconstruct specific shares.

By linking the distributed data storage technology demonstrated in July 2021 and integrating it with the newly developed key management system and share control system, the four parties have established a personalized healthcare system. Authentication of individuals and personal genome data sharing and reconstruction is based on the My Number Card, an individual ID card issued by the Japanese Government. Genome analysis data cannot be reconstructed at medical centers without the cardholder's consent, preventing information leakage. Demonstrations at ToMMo and Tohoku University Hospital have confirmed the feasibility of this information-safe, practical personalized healthcare system, that can also reconstruct data from shares stored at other sites if data is lost at one site due to a disaster.

Project Structure

Toshiba: Development, construction and operation of QKD system and personalized healthcare system

ToMMo and Tohoku University Hospital: Provision of a verification base, application in individualized healthcare use cases, operation and confirmation of effectiveness

NICT: Development and operation of high-speed one-time pad[*5] (OTP) technology and high-speed secret sharing technology

Future Prospects

Toshiba will continue to advance R&D of QKD technologies, including system demonstrations combined with secret sharing technology, and promote early practical application of QKD technologies in various applications, such as medicine, finance, public administration, and communications infrastructure.

ToMMo and Tohoku University Hospital will continue to promote the use of safe and secure ICT technology to realize future-oriented medicine based on genome information and personalized healthcare.

NICT will continue to engage in research and development of quantum communications technologies such as quantum cryptography and optical quantum control in order to contribute to advanced and fundamental research and development and quantum industry.

*1 December 8 2022, Toshiba Research

*2 Personalized healthcare: health risk management optimized for individuals by analyzing personal genome data together with environmental factors such as lifestyle habit, and calculating the risks of contracting diseases for each individual

*3 Reconstruction: The decryption of data encrypted with quantum cryptography

*4: https://www.global.toshiba/ww/technology/corporate/rdc/rd/topics/21/2108-02.html

*5 One time pad (OTP): An encryption key that is used once and discarded

Appendix: Reference Material

Outline of the Demonstration

1. Details of the Demonstration

(1) Genome analysis data can be transmitted and stored in a theoretically secure manner using QKD technology and secret sharing technology.

(2) The genome analysis data can be reconstructed and utilized only with the consent of the individual, using the individual's My Number Card as the access key.

2. Personalized Health care Sequence Assumed in This Demonstration

(1) Genome analysis: Whole genome analysis of individuals is carried out, on the assumption that the data will be used for lifelong health management and medical treatment.

(2) Secret sharing and quantum cryptography communications: Genome analysis data obtained from analysis are divided into shares and selected an appropriate storage location for each share using secret sharing technology, then the data is securely transmitted to the selected locations using the QKD link.

(3) Distributed storage: Shares are stored at multiple locations for reasons of security and to protect and preserve data in the event of a disaster. A part of the shares is also stored in the cloud, to facilitate data recovery in the event of a large-scale disaster.

(4) Authentication: Each individual will authorize the use of their genome data analysis in relation to health management, research and medical care, and uses their My Number Card to do so.

(5) Quantum cryptographic communication, secret sharing reconstruction and utilization: When an individual consents to use of the genome analysis data, the secret sharing technology aggregates the dispersed shares containing that data at a single location, to which they are transmitted via through cryptographic communications using QKD technology. Once there, the original genome analysis data is reconstructed and utilized.

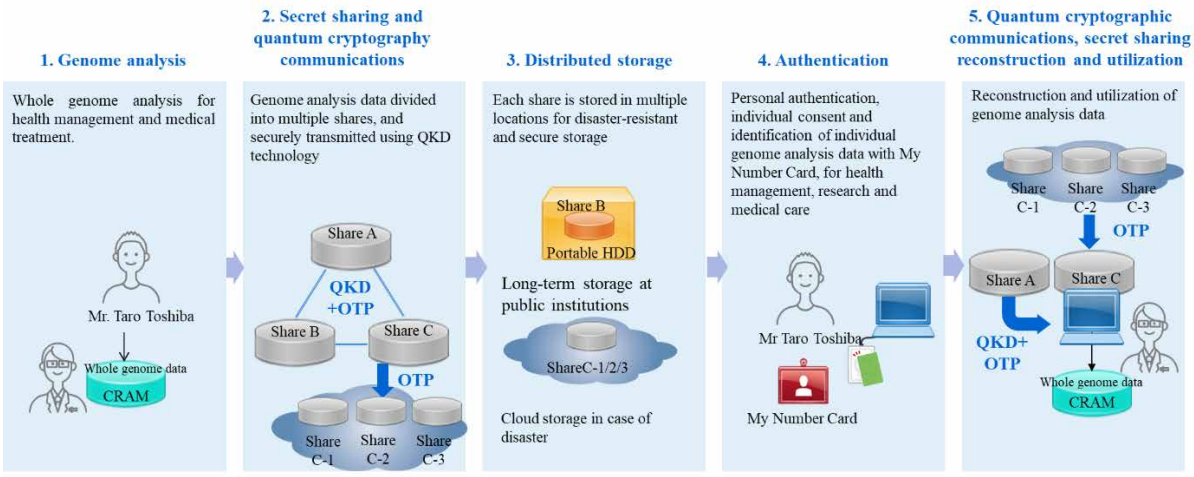

Fig. 1 Entire sequence of personalized healthcare

3. Configuration of the Demonstration System

The demonstration environment consisted of ToMMo (Seiryo-machi, Aoba-ku, Sendai City, Miyagi Prefecture), Toshiba Life Science Analysis Center (LSA, Minami-Yoshinari, Aoba-ku, Sendai City, Miyagi Prefecture), and Tohoku University Hospital (Seiryo-machi, Aoba-ku, Sendai City, Miyagi Prefecture), and three servers on an IP network that function as a cloud formed outside the three sites (three sets). QKD devices and an OTP encryption system developed by Toshiba are installed at the link between ToMMo and LSA and the link between ToMMo and Tohoku University Hospital. The QKD device shares cryptographic keys whose security is guaranteed by quantum mechanics, and the OTP cryptographic communications function using the shared cryptographic keys enables cryptographic communications with theoretically unbreakable security.

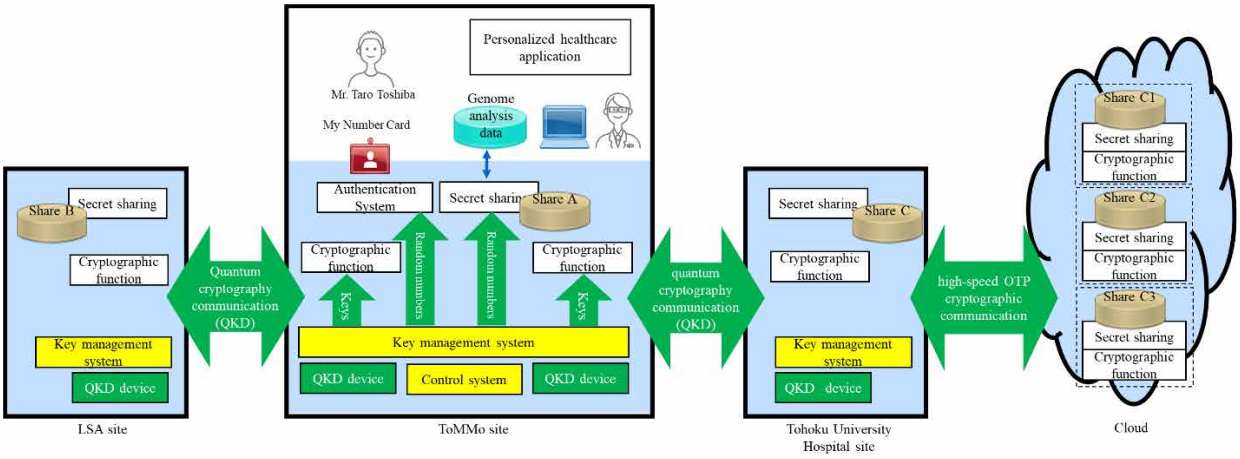

Fig. 2 Configuration of the demonstration system

To facilitate the experiment, cryptographic communications among the three sites and the three servers forming the cloud is done using the high-speed OTP encryption system developed by NICT, which treats random number data shared and stored between the sites as the quantum cryptographic key, instead of a QKD device. Each of the three sites and the three servers on the cloud are equipped with a newly developed integrated key management system and a secret sharing function. The ToMMo site is also equipped with a share control system and a personalized healthcare application for data storage and reference. The integrated key management system unifies provision of cryptographic keys for the cryptographic communications, and random number data to the secret sharing function and authentication system. The share control system determines where shares are stored in the demonstration environment and manages information on the storage location of each share.

4. Demonstration Procedure

(1) Genome analysis data (16 GB) is in the CRAM format (a data format for genome data), which ToMMo uses for genome analysis. ID and authentication data (hash data) for an individual's genome analysis data is stored on that individual's My Number Card. In this demonstration experiment, a contactless IC card conforming to the ISO/IEC 14443 type A standard was used for personal authentication (hereinafter "contactless IC card").

(2) At the ToMMo site, CRAM format genome analysis data is converted into three shares (Share A, Share B, and Share C) by the secret sharing function. The share control system is used to distribute the shares to multiple locations

for storage. The share control system determines where to store the shares based on the system status, including the remaining quantum cryptographic keys for each inter-site link. For example, Share B is encrypted and transmitted to LSA, while Share C is encrypted and transmitted to Tohoku University Hospital, by quantum cryptography communications. The share control system manages the information by associating an individual's ID with the ID of the corresponding share of genome analysis data and its storage locations.

(3) The share (Share C) stored at Tohoku University Hospital is further divided into three shares (Share C1, Share C2, and Share C3) by secret sharing, and these shares are distributed and stored on the three servers in the cloud. Share transmissions between Tohoku University Hospital and the cloud, and between three servers on the cloud are encrypted using high-speed OTP encryption. After storing the shares on three servers on the cloud, the share at Tohoku University Hospital (Share C) is deleted. At this time, the share control system changes the storage location information of each share to the server in the cloud.

(4) A person who presents his or her contactless IC card at ToMMo in order to authorize the use of genome analysis data is identified and authenticated by the authentication system, based on the ID and authentication information stored in the contactless IC card. The share control system is referenced to identify the individual’s shares to be reconstructed.

(5) Two shares (e.g., Share C1 and Share C2) are selected in the cloud, and Share C is reconstructed and transmitted to Tohoku University Hospital using OTP encryption. Share C is then transmitted from Tohoku University Hospital to ToMMo with the QKD system, and ToMMo reconstructs the original genome analysis data from Share C and Share A. After the hash data stored in the contactless IC card is confirmed to be that particular individual’s data, the genome analysis data can be used. The share (Share B) stored in the LSA functions as backup data, and is stored, for example, at a remote public medical institution.

# # #