Blockchain-Assisted Secure Fine-Grained Searchable Encryption for a Cloud-Based Healthcare Cyber-Physical System

Mamta, *Member, IEEE*, Brij B. Gupta, *Senior Member, IEEE*, Kuan-Ching Li, *Senior Member, IEEE*, Victor C. M. Leung, *Fellow, IEEE*, Kostas E. Psannis, and Shingo Yamaguchi, *Senior Member, IEEE*

***Abstract*—The concept of sharing of personal health data over cloud storage in a healthcare-cyber physical system has become popular in recent times as it improves access quality. The privacy of health data can only be preserved by keeping it in an encrypted form, but it affects usability and flexibility in terms of effective search. Attribute-based searchable encryption (ABSE) has proven its worth by providing fine-grained searching capabilities in the shared cloud storage. However, it is not practical to apply this scheme to the devices with limited resources and storage capacity because a typical ABSE involves serious computations. In a healthcare cloud-based cyber-physical system (CCPS), the data is often collected by resource-constraint devices; therefore, here also, we cannot directly apply ABSE schemes. In the proposed work, the inherent computational cost of the ABSE scheme is managed by executing the computationally intensive tasks of a typical ABSE scheme on the blockchain network. Thus, it makes the proposed scheme suitable for online storage and retrieval of personal health data in a typical CCPS. With the assistance of blockchain technology, the proposed scheme offers two main benefits. First, it is free from a trusted authority, which makes it genuinely decentralized and free from a single point of failure. Second, it is computationally efficient because the computational load is now distributed among the consensus nodes in the blockchain network. Specifically, the task of initializing the system, which is considered the most computationally intensive, and the task of partial search token generation, which is considered as the most frequent operation, is now the responsibility of the consensus nodes. This eliminates the need of the trusted authority and reduces the burden of data users, respectively. Further, in comparison to existing decentralized fine-grained searchable encryption schemes, the proposed scheme has achieved a significant reduction in storage and computational cost for the secret key associated with users. It has been verified both theoretically and practically in the performance analysis section.**

***Index Terms*—Cloud-based cyber-physical systems (CCPS), data encryption, healthcare information search and retrieval, keyword search, public-key cryptosystems, searchable encryption.**

I. Introduction

CYBER-PHYSICAL systems (CPS) tightly intertwine software and physical components that can have applications in nearly any field we can think of, including healthcare, energy conservation, environment protection, defense, agriculture and many more. CPS tirelessly produce huge silos of data, and to ensure scalability and efficient storage, large companies like Microsoft and Honeywell, etc., are moving towards cloud-based solutions. Cloud-based CPS (CCPS) improves existing CPS functionalities, but at the same time, presents security challenges.

The data related to healthcare and defense may contain sensitive information. Therefore, in CCPS, in addition to applying security policies at the physical level, there is a need to ensure security at the cyber level, i.e., the cloud component. The cloud-based CPS for healthcare offers many benefits like monitoring and controlling patient's health by deploying sensors and actuators in the form of wearable devices. These devices keep track of the essential vitals and may even alert associated medical practitioners in case of critical situations. The data involved in the entire process of monitoring and alerting is highly sensitive, whether it is the location of the patient or the vitals stored by these devices. Therefore, a need for a security mechanism is as essential as providing healthcare services to the patients. Further, the data involved in the continuous monitoring may be massive, and the devices which gather this data are resource constrained. Hence the data is generally stored at a third party cloud server. For this purpose, an obvious solution is to encrypt sensitive data first and then store it to an untrustworthy cloud platform. Encryption indeed ensures security but severely debilitates the accessibility of data. It makes even the most basic operation of searching, a highly challenging task. The searchable encryption (SE) technique is the answer to this problem.

Manuscript received October 15, 2020; revised December 3, 2020; accepted February 13, 2021. Recommended by Associate Editor Laurence T. Yang. *(Corresponding author: Brij B. Gupta.)*

Citation: Mamta, B. B. Gupta, K.-C. Li, V. C. M. Leung, K. E. Psannis, and S. Yamaguchi, "Blockchain-assisted secure fine-grained searchable encryption for a cloud-based healthcare cyber-physical system," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 12, pp. 1877–1890, Dec. 2021.

Mamta is with the Department of Computer Engineering, National Institute of Technology, Haryana 136119, India (e-mail: er.mamta.dabra@gmail.com).

B. B. Gupta is with the Department of Computer Engineering, National Institute of Technology, Haryana 136119, India, and also with the Asia University, Taichung, Taiwan, China, and also with the Macquarie University, Sydney, NSW 2109, Australia (e-mail: gupta.brij@gmail.com).

K.-C. Li is with the Department of Computer Science and Information Engineering, Providence University, Taichung 43301, Taiwan, China (e-mail: kuancli@pu.edu.tw).

V. C. M. Leung is with the Computer Science and Software Engineering, Shenzhen University, Shenzhen 518000, China, and also with the University of British Columbia, Vancouver V6T 1Z4, Canada (e-mail: vleung@ece.ubc.ca).

K. E. Psannis is with the Department of Applied Informatics, University of Macedonia, Thessaloniki 54636, Greece (e-mail: kpsannis@uom.gr).

S. Yamaguchi is with the Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Ube 755-8611, Japan (e-mail: shingo@yamaguchi-u.ac.jp).

Digital Object Identifier 10.1109/JAS.2021.1004003

SE enables the cloud server to search confidential data without revealing any information about the data being searched. Further, in the cloud environment, users from multiple domains interact; therefore, access control must be embedded to enable fine-grained searching functionalities. All the stated features in the encryption scheme including search capability and access control fulfills our requirement for cloud-based CPS for healthcare. However, these features come with an inherent cost, and cannot be directly applied for the resource-constraint devices. There is a need of a mechanism to reduce this cost, and this is the key goal of this paper. To enable fine-grained searching capabilities, we have used attribute-based encryption (ABE) and specifically its ciphertext-policy (CP) variant as it makes it possible for the data owners to implement access rule over the encrypted data. Also, the data owner has full control over his shared data which is the essential requirement in any healthcare system. To reduce the associated cost with the ciphertext-policy ABSE scheme, we have leveraged the blockchain technology, where most of the computational load of the ABSE scheme is delegated to the blockchain network.

A. Research Methodology

This section describes the approach followed for the reduction in computational complexity, starting with the problem formulation of the proposed scheme.

1) Problem Formulation

The ABSE scheme provides fine-grained search capabilities to its users but it comes with an inherent computational overhead. In a typical ABSE scheme the associated storage and the computational cost varies linearly with the number of attributes possessed by a user. In a healthcare CCPS, the devices participating in the network are resource constraint. Thus it is not feasible to directly apply the ABSE schemes to access the encrypted information.

Given the above problem, we aim to construct a keyword search scheme with fine-grained search capabilities that can handle a large number of diverse attributes. Furthermore, it should be lightweight enough to be able to be deployed at devices with resource constraints.

2) Proposed Hypotheses

The proposed hypotheses state that the computational complexity of the existing state-of-the art fine-grained searchable encryption schemes is high and also varies with the number of attributes involved in the system. To use such security solutions for devices with limited resources, there is a need to develop a delegation mechanism which can reduce the computational burden from the entities involved in the system. The proposed scheme aims to reduce the associated cost of a fine-grained searchable encryption scheme with the assistance of blockchain technology.

3) Proposed Solution

To address the above-stated problem, in our view, there can be two possible solutions. First, is to make the associated cost independent of the number of attributes. It has already been the focus of many works, like [1]–[4]. Second, resource-constraint devices should not execute computationally intensive tasks involved in a typical ABSE scheme by themselves and leverage other technology like blockchain technology, which can improve both flexibility and reduce system overhead.

With blockchain technology, the proposed scheme can attain the following characteristic features:

Decentralization: It is the key property of blockchain technology, which means that there is no central control, i.e., there exists no single authority responsible for governing the system. The searchable encryption system developed using the blockchain technology inherits this fundamental property and hence results in a fully decentralized system.

Reduction in computation overhead: In the blockchain-assisted searchable encryption system, the task of system initialization is not the responsibility of a single entity. However, it is handled together by blockchain consensus nodes. Furthermore, the task of search token generation is assisted by consensus nodes to reduce the burden from end-users.

Improves system reliability and free from a single point of failure: In the blockchain-assisted searchable encryption system, there is no need for a central authority, and no master secret key is needed to generate user credentials. Hence, this makes the system more reliable in case one or more nodes fails or becomes malicious.

The remaining of this article is organized as follows: Section II discusses related works. In Section III, the reader is enlightened with the essential background required to understand the construction of the proposed scheme. Further, it presents an introduction to the blockchain and other retrospective techniques which forms the basis of the paper. Section IV gives the system and security definitions and explains the system and security model of the proposed scheme. Section V provides detailed construction of the proposed scheme along with its correctness and security analysis. Section VI finally concludes the paper with possible directions for future work.

II. Related Work

ABSE enables fine-grained searching capabilities in a multi-user environment. Several attribute-based searchable schemes have been developed using either of the two design frameworks (ciphertext-policy or key-policy) [1]–[3]. Moreover, in the area of healthcare, there exist some attribute-based keyword search schemes [5]–[8]. However, most of them need a central authority for management and distribution of secret keys to the cloud users. Consider a scenario where a data owner wants to share his data with a large number of users who possess attributes from different domains. Then the single authority cannot efficiently manage such a large and diverse set of attributes alone. Such a scenario is prevalent in the healthcare system, e.g., in healthcare networks, a data owner, i.e., a patient may want to share his data with users like a researcher, or a doctor or some insurance agent. All of these users belong to an entirely different domain; hence, there is a need for a multi-authority system where each authority is responsible for the management of a disjoint set of attributes from different domains.

One such multi-authority searchable encryption scheme was recently proposed by Miao *et al.* [9]. They achieved decentralization by eliminating the central authority. However, the schemes mentioned above, including the one in [9] comes with tremendous computational complexity inherited from the underlying ABE scheme. To reduce this computational overhead and to achieve decentralization at the same time, the authors leveraged the blockchain technology. They delegated most of the computationally intensive tasks either to the blockchain network or to the cloud server. In the proposed scheme, the system initialization and partial search token generation tasks are handled by the blockchain network, while the cloud server handles the search task. Thus, the users can stay relaxed and get their task completed by the smart blockchain technology. There exist several scenarios in literature where blockchain technology has been used [10]–[13]. In [14], the searchable encryption scheme has been developed by leveraging blockchain technology, but it focuses on the searchable encryption in the symmetric setting. In [12], a searchable encryption scheme has been proposed with the assistance of blockchain technology in the public-key setting, but they did not consider the fine-grained searching capabilities into account.

III. Background

This section gives the necessary information about the bilinear map, hardness assumptions on which the proposed scheme relies, and the structure of the access policy used in the proposed scheme.

A. Bilinear Map and Hardness Assumption

Let G be a source cyclic group of prime order, p, and G_T be a target cyclic group of same order p and g be the generator of the source group, G. Let e be a map between G and $G_T, e : G \times G \to G_T$, it is called a bilinear map if it satisfies the following conditions:

1) $\forall X = g^u \in G_1$ and $Y = g^v \in G_2 : e(g^u, g^v) = e(g,g)^{uv}$ where $u, v \in Z_p$. This is called the bilinearity property.

2) $e(g,g)$ is the generator of the target group, G_T, if g is the generator of G. This property is called non-degeneracy.

3) $\forall X, Y \in G$; $e(X,Y)$ is efficiently computable.

Assumption (q-decisional parallel bilinear Diffie-Hellman exponent (DPBDHE)): Let $e : G \times G \to G_T$ be a bilinear map defined above. Consider another cyclic group Z_p of integers of order again p. Choose $s, a, b_1, b_2, \ldots, b_q \in Z_p$, $U \in G_T$ and compute the tuple D, which consists of the following elements: $g, g^s, \left\{g^{a^i}\right\}_{i\in[2q], i\neq q+1}, \left\{g^{b_j a^i}\right\}_{(i,j)\in[2q,q], i\neq q+1}$,

$$\left\{g^{\frac{s}{b_j}}\right\}_{j\in[q]}, \left\{g^{\frac{sa^i b_j}{b_k}}\right\}_{(i,j,k)\in[q+1,q,q], j\neq k.}$$

The distinguishing probability of the distributions $\left(D, V = e(g,g)^{sa^{q+1}}\right)$ and (D, U) is negligible [15].

B. Access Structure

To construct the proposed searchable encryption scheme, we have used the monotonic access structure, which is defined as: Let the attribute universe be represented by $\mathcal{U}$. A monotonic access structure, $\mathbb{A}$, on $\mathcal{U}$ consists of non-empty subsets of $\mathcal{U}$ such that if a subset belongs to $\mathbb{A}$ then its superset must also be there in $\mathbb{A}$, i.e., $\forall B, C \in \mathbb{A}$, if $B \in \mathbb{A}$ and $B \subseteq C$, then $C \in \mathbb{A}$ [16]. To distribute the secret over the access structure we use a linear secret sharing scheme (LSSS), π, over Z_p, which is defined as follows [16]:

1) The shares of a secret $s \in Z_p$ forms a vector over Z_p.

2) For each access structure defined on the attribute universe, there exists a share generating matrix, $\mathbb{A}$, of order $l \times n$ whose elements are taken from Z_p. Let $\rho : [l] \to \mathcal{U}$ be a function that maps each row of the matrix to the attributes from set $\mathcal{U}$, i.e., $\rho(i)|_{i=1}^{l} \in \mathcal{U}$. To generate the secret shares of s, consider a vector $\overrightarrow{v} = (s, v_2, \ldots, v_m)$ where $v_2, \ldots, v_m \in Z_p$ and compute $\overrightarrow{\lambda} = \mathbb{A}\overrightarrow{v}^T$ where $\lambda_i = \mathbb{A}_i \overrightarrow{v}^T|_{i=1}^{l}$ represents a secret share of s.

LSSS should satisfy the linear reconstruction property which states that if $\exists S \in \mathbb{A}$ such that I represents the set of rows (l) such that $\rho(l) \in S$ then $\exists \{c_i\}_{i\in I} \in Z_p : \sum_{i\in I} c_i \lambda_i = s$, given λ_i represents the valid shares of secret, s.

C. Shamir's Secret Sharing Scheme

It is used to divide a secret into n shares such that the secret can only be constructed if a minimum of k shares is available among them. Therefore, k represents the threshold because after collecting k shares, the secret can be revealed. It is based on polynomial interpolation, given k points, (x_i, y_i), such that no two different points have the same x-axis component, and in a two-dimensional space, there exists a single unique polynomial, q, of degree $k-1$ such that $q(x_i) = y_i \forall i$. Thus, the idea is to construct a polynomial of degree $k-1$, where the data to be shared is placed at the position of a constant term and rest of the coefficients are selected randomly, i.e., $q(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{k-1} x^{k-1} : c_0 = s, c_i|_{i=1}^{k-1} \xleftarrow{\$} Z_p$, where Z_p is the group of integers over which the polynomial is constructed. Now, generate the n points from this polynomial by putting different values of $x_i|_{i=1}^{n}$, i.e., generate (x_i, y_i). The secret can be reconstructed by using any of the k points among the available n points using the Lagrange interpolation method. To find the coefficients of $q(x)$, first, find the Lagrange identities by using these k points as follows:

$$l_j = \prod\nolimits_{\substack{m=1\\ m\neq j}}^{k} \frac{(x - x_m)}{\left(x_j - x_m\right)}. \tag{1}$$

Now, find the polynomial, $q(x) = \sum_{j=1}^{k} y_j l_j$. Among the coefficients computed using this formula, the coefficient of the term, x^0 (constant term), is the secret that we have shared.

D. Consortium Blockchain Platform and its Role in the Proposed Scheme

The consortium blockchain (CB) platform extends the concept of private blockchain where the entire network is managed by a group of organizations instead of a single organization like in a private blockchain platform [14], [17], [18]. A CB platform represents the fine line between public and private blockchain platforms. It adds flexibility in rules to be more like public blockchain. The visibility of the blockchain may be limited only to validators, authorized users

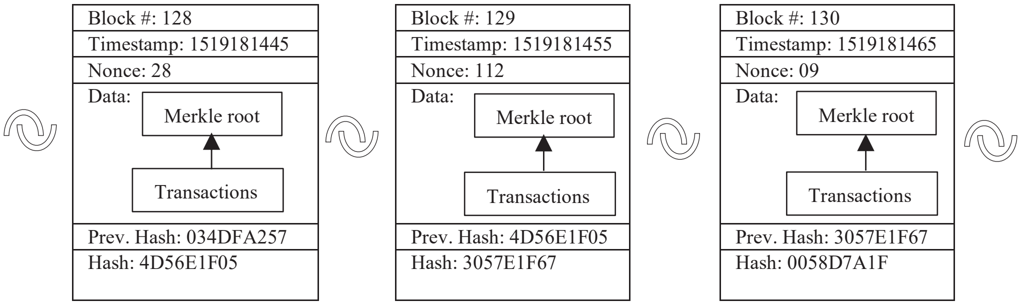


Fig. 1. General structure of a blockchain.

or is visible to everyone, thereby combining the features of both the public and private blockchain platforms. The most noticeable difference between both public and private blockchain platforms is observed during the consensus. Unlike the open system in public blockchain or a completely closed system in a private blockchain, in consortium blockchain, a group of equally powered nodes participate in consensus. The typical structure of the blocks and how these blocks are connected in any blockchain platform is shown below in Fig. 1.

Blockchain, as the name suggests itself that it is a chain of blocks, and this chain is formed and secured with the application of cryptography. Blockchain can also be defined as a sequence of records which is analogous to traditional ledgers with an additional property of immutability. Hence, it is often called an immutable ledger. Now, depending upon the requirement and application area, the responsibility of managing and updating this ledger is given either to everyone in the blockchain network (public blockchain platform, e.g., the Bitcoin network [19]), or to one particular organization (private blockchain platform, e.g., Hyperledger), or to a selected set of users from different organizations with equal power (consortium blockchain platform, e.g., Quorum).

In cloud-based healthcare CPS, the public blockchain platform does not seem suitable because the public blockchain allows access to everyone on the system and the data involved in the healthcare CPS may contain sensitive information which is not for the public use. The private blockchain platform also does not fit in the considered scenario because it is too restrictive, and there is still control by one single authority who can manipulate the data for their benefit. Therefore, the option that is suited well for the present scenario is the consortium blockchain where different entities associated with the healthcare CPS makes a consortium and pre-decides the nodes that will participate in the consensus. By leveraging the benefits of consortium blockchain in healthcare-CCPS, one can impose control as well as can enjoy the decided degree of freedom. The architecture of the consortium blockchain used in the proposed scheme is shown in Fig. 2.

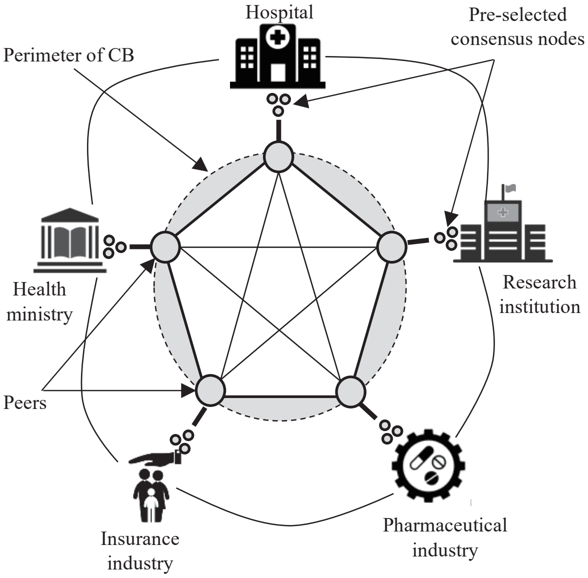


Fig. 2. Architecture of consortium blockchain for healthcare CPS.

The consortium blockchain for healthcare CPS may involve entities like hospitals, research institutions, various state and central health departments managed by the health ministry, and the insurance and pharmaceutical industries, etc. These organizations pre-select some of the nodes as the consensus nodes as per their governance policy. These designated nodes are the one which participate in managing and updating the distributed ledger. In contrast, the other nodes can generate or contribute data. If the number of the consensus nodes is increased, then it results in a more secure and fault-tolerant system. Because now more nodes must agree to reach consensus. Further, the system becomes more scalable in terms of the number of transactions processed. Because in the consortium blockchain consensus nodes are the ones who can contribute a new block in the blockchain. The consensus protocol also requires less computational power because the consensus nodes are chosen by the organizations with a high trust level.

Thus, as the number of consensus node increases, there will be an increase in the number of new blocks, and as a result, more transactions will be included for processing. Lastly, as we increase the number of consensus nodes, the degree of decentralization also increases, which results in a more robust system.

Role of CB in the proposed scheme:

1) System Setup: The consensus nodes initialize the system and generate the global public parameters by using Shamir's secret sharing scheme.

2) User Registration and Generation of Partial Search Token: Any user can register themselves to CB by storing their public key corresponding to their unique global identity. It is also responsible for managing and generating the search

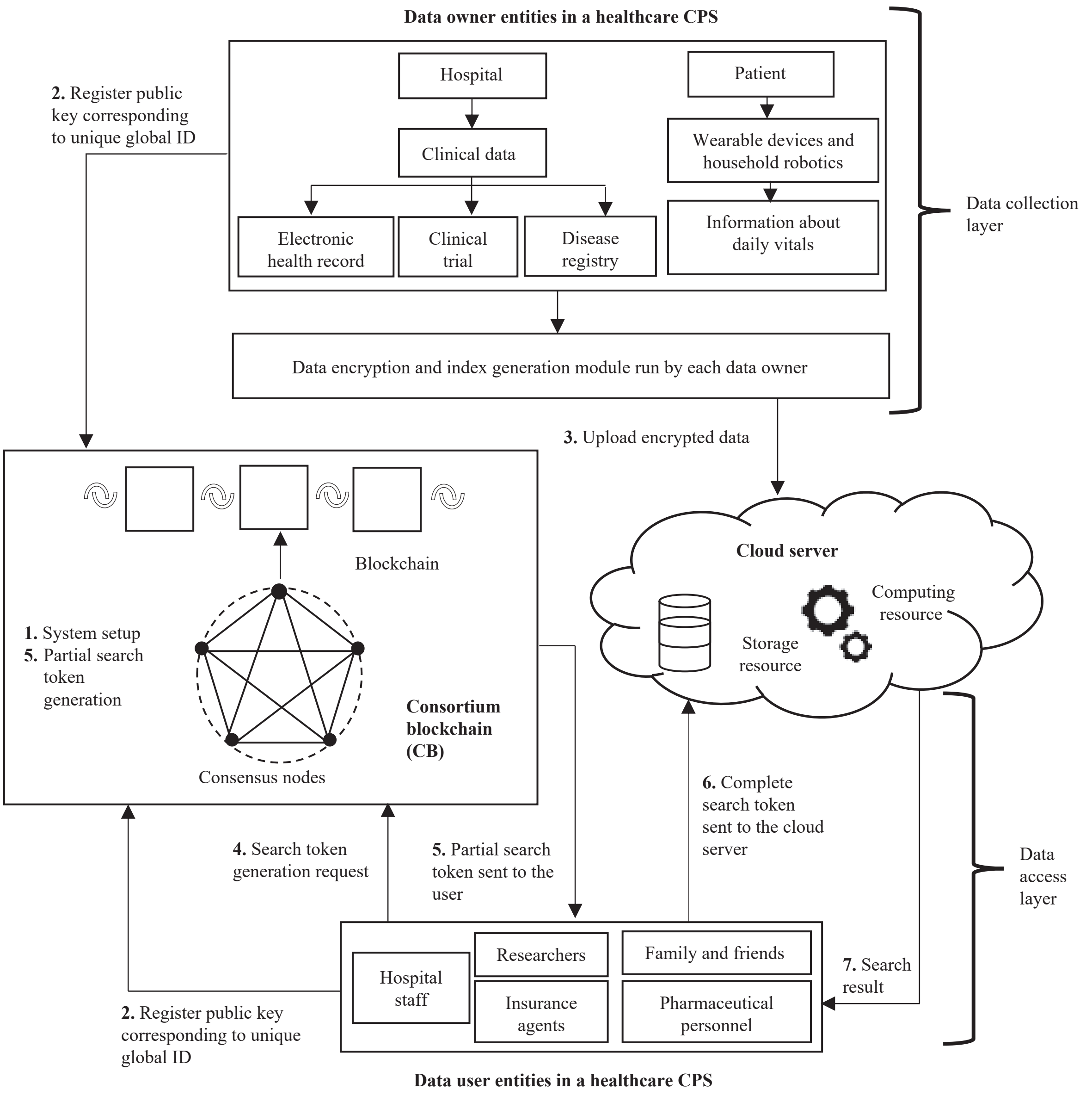

Data owner entities in a healthcare CPS
Hospital
Clinical data
Electronic health record
Clinical trial
Disease registry
Patient
Wearable devices and household robotics
Information about daily vitals
Data encryption and index generation module run by each data owner
Data collection layer
2. Register public key corresponding to unique global ID
3. Upload encrypted data
Blockchain
1. System setup
5. Partial search token generation
Consortium blockchain (CB)
Consensus nodes
Cloud server
Computing resource
Storage resource
6. Complete search token sent to the cloud server
4. Search token generation request
5. Partial search token sent to the user
Data access layer
7. Search result
Hospital staff
Researchers
Insurance agents
Family and friends
Pharmaceutical personnel
2. Register public key corresponding to unique global ID
Data user entities in a healthcare CPS

Fig. 3. Architecture of the proposed scheme for CCPS.

token when some user wants to retrieve encrypted information from the cloud, he/she can reach the CB to generate the partial search token corresponding to the attributes possessed by the user.

3) Free From Keeping the Master Secret: In the CB assisted proposed scheme, there is no single entity that manages the system. Hence, it is free from keeping master secret unlike the existing schemes [11]–[14].

IV. Scheme Preliminaries

This section explains the architecture that provides the general algorithm for the proposed scheme. Further, it gives the security definition along with the game-based security model for the proposed scheme.

A. Architecture of the System

The system is composed of the following four entities, as shown in Fig. 3:

1) Consortium Blockchain (CB): This entity will initialize the system and generate global parameters. Further, it registers the public key of the user corresponding to their unique global ID (GID). When a user wants to perform a search, then he/she can reach the CB along with the attributes possessed by him/her and the CB will generate the partial search token for the user.

2) Data Owner (DO): As the name suggests, DO is the one who owns some data (health data in the considered scenario) and wants to share his/her health data to a third-party cloud server. Before outsourcing both the health data file and the associated data, the keywords are first encrypted. In a healthcare system, the data owner is a patient who encrypts his health data using any standard encryption algorithm and probably the symmetric key algorithm (fast computation). The associated symmetric key used for encrypting data and the keywords contained in that data are encrypted using the *GenIndex* algorithm of the proposed scheme.

3) Data User (DU): DU is the one who wants to access the data stored at the cloud server. He/she will generate the complete search token which is then sent to the cloud server to retrieve the file containing the searched keyword. In a healthcare system, a data user may be a researcher or a doctor, etc.

4) Cloud Server (CS): CS is the one which stores encrypted health data, performs a keyword search on behalf of the data user and returns the search result. A healthcare provider usually manages it in the considered scenario.

The detailed explanation of various steps shown in Fig. 3 is as follows:

1) The consortium blockchain performs system initialization, and the global public parameters are published.

2) The CCPS users (data owner and data user) register themselves on CB by storing their public key generated by them locally, along with their GID.

3) The data owners run the *GenIndex* algorithm to generate the ciphertext and upload it to the cloud server for secure sharing.

4) To get the required data from the cloud server, the DU sends a search token generation request to the CB. The search token request contains the global ID of the user, along with his/her set of attributes.

5) In CB, the set of consensus nodes will verify the user's attributes and generate a partial search token for them using those attributes, which is then sent to the user who requested it.

6) Using the partial search token, DU will generate the complete search token using his/her secret key corresponding to his/her GID. The complete search token generated by DU is then sent to the cloud server.

7) Using the complete search token, CS runs the $Search$ algorithm on behalf of the user, and the search result is then returned to DU.

B. Primary Algorithms

The proposed scheme consists of the following probabilistic polynomial time (PPT) algorithms:

1) $GPP \leftarrow GSetup(1^n)$*:* CB runs this algorithm by taking the security parameter in the unary format as input and generates global public parameters. Let $CN = \{cn_1, cn_2, \ldots, cn_n\}$ be the set of consensus nodes, Att, be the universe of attributes and $\mathcal{W}$ represents the keyword space such that each keyword is a binary string.

2) $[PK_{Gid}, SK_{Gid}] \leftarrow KeyGen(GPP, Gid \in \mathcal{GID})$*:* It is run locally by each CCPS user to get their key pair (public, PK and secret, SK) corresponding to their Gid. PK_{Gid} is sent to the blockchain and SK_{Gid} is kept by the user.

3) $C \leftarrow GenIndex(GPP, (\mathbb{A}, \rho), w \in \mathcal{W})$*:* It is executed by DU to get the searchable index for the keyword, w. As we are using the CP-ABE design framework, this algorithm takes access structure, $\mathbb{A}$, in addition to GPP to encrypt the keyword.

4) $PTK \leftarrow PTok(GPP, Gid \in \mathcal{GID}, \mathrm{S} \subseteq Att)$*:* This algorithm is run by the consensus nodes in the consortium blockchain. It takes Gid and the attributes possessed by the user. CB will first verify that the attributes claimed by the user are genuinely his/her attributes by checking the information associated with his/her Gid. After verification is successful, then CB will output the partial search token for that user.

5) $CTK \leftarrow CTok(GPP, SK_{Gid}, w')$*:* The data user will use this algorithm to generate the complete search token for the keyword, w', using his secret key credentials, SK_{Gid}.

6) $0/1 \leftarrow Search(GPP, C, CTK)$*:* This algorithm is run by the cloud server on behalf of the data user. It will take the ciphertext, C, and the search token, CTK, of the keyword, w, and performs a search for that keyword without decrypting it. If a match is found then it returns 1 else, it returns 0 and if a user is not authorized to perform the search, then it returns $\perp$.

The interaction between different entities involved in the system, along with their role, is shown in Fig. 4. The detail of each step shown in Fig. 4 has already been given in Section IV-A.

Correctness: The proposed scheme is correct if the following condition holds:

$$Search(C, TK) = 1. \tag{2}$$

For the given $GPP \leftarrow GSetup$, $[LPP_\tau, MSK_\tau] \leftarrow AASetup$, $SK_{Gid} \leftarrow KeyGen$, $C \leftarrow GenIndex$ and $TK \leftarrow Trapdoor$.

C. Security Definition and Model

In this section, we will first define the meaning of security against a non-adaptive chosen keyword attack. Then the security of the proposed scheme is modelled as a game between a challenger, C, and an adversary, $\mathcal{A}$, where challenger tries to break the security by leveraging the adversary's output.

1) Security Definition

Indistinguishability against non-adaptive chosen keyword attack (CKA): It states that given the two challenge keyword ciphertexts, $\mathcal{A}$ should not be able to determine which ciphertext is the encryption of a keyword of their choice even if they gain access to the trapdoors for all the keywords except the challenge keywords. In the non-adaptive security definition, an adversary chooses the consensus nodes to be corrupted after looking at the global public parameters and keeping them the same throughout the game, which can corrupt at most $k-1$ nodes. Moreover, the adversary selects the challenge access policy, $\mathbb{A}^*$, which $\mathcal{A}$ plans to attack.

The above-stated security definition is modelled as a game between C and $\mathcal{A}$, with the following phases:

a) Global setup phase: Here, C runs $GSetup(\lambda)$ algorithm to get the global public parameters, which are then sent to $\mathcal{A}$.

b) Initialization phase: After looking at the global parameters, a set of corrupted consensus nodes, $CN_c \subseteq CN : |CN_c| \leq k-1$, and a set of non-corrupted consensus nodes, $CN_n \subseteq CN : (CN_c \cap CN_n) = \emptyset$ are chosen by $\mathcal{A}$. Further, $\mathcal{A}$ selects the challenge access policy, $\mathbb{A}^*$.

c) Query phase: Here, $\mathcal{A}$ can submit the query for the search token with (Gid_i, S_i). The queried trapdoors are stored in a list, L_{tk}, maintained by C.

d) Challenge phase: In this phase, $\mathcal{A}$ is asked to choose keywords, $\{w_0, w_1\}$ of equal length such that $\{w_0, w_1\} \notin L_{tk}$ which are then sent to C. Upon receiving the challenge keywords, C selects a binary bit, $b \in \{0,1\}$ and generates the

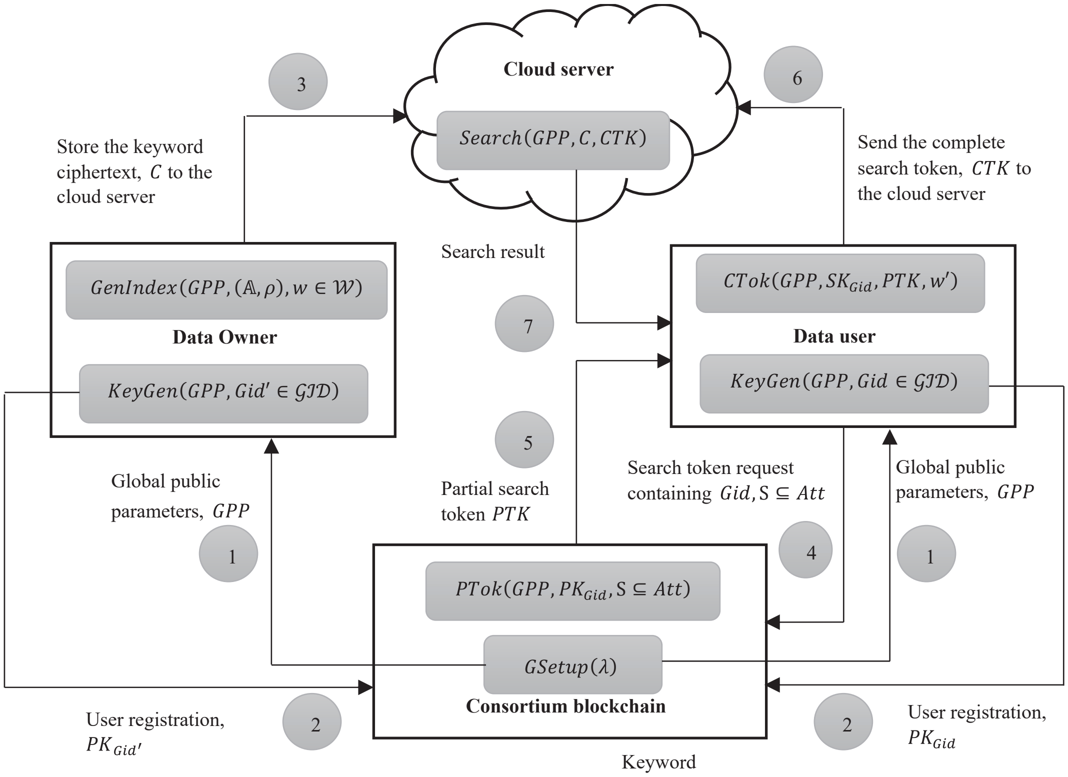

Cloud server
$Search(GPP, C, CTK)$
3
Store the keyword ciphertext, C to the cloud server
6
Send the complete search token, CTK to the cloud server
Search result
$GenIndex(GPP, (\mathbb{A}, \rho), w \in \mathcal{W})$
Data Owner
$KeyGen(GPP, Gid' \in \mathcal{GID})$
7
$CTok(GPP, SK_{Gid}, PTK, w')$
Data user
$KeyGen(GPP, Gid \in \mathcal{GID})$
5
Global public parameters, GPP
Partial search token PTK
Search token request containing $Gid, S \subseteq Att$
Global public parameters, GPP
1
4
1
$PTok(GPP, PK_{Gid}, S \subseteq Att)$
$GSetup(\lambda)$
Consortium blockchain
User registration, $PK_{Gid'}$
2
2
User registration, PK_{Gid}
Keyword

Fig. 4. Interaction Between Different Entities in the Proposed Scheme.

challenge ciphertext, C^*, for the keyword, w_b, by using *GenIndex* algorithm and the challenge access policy, $\mathbb{A}^*$. The resulting C^* is then sent to the adversary, $\mathcal{A}$.

e) Output phase: Here $\mathcal{A}$ outputs its guess b' of b and becomes the winner if $b' = b$ with the advantage as follows:

$$Adv_{\mathcal{A}}^{sCKA}(k) = \left|Pr(\mu = \mu') - \frac{1}{2}\right| = \epsilon. \tag{3}$$

If for all PPT $\mathcal{A}$ this advantage is negligible, then the proposed scheme is said to be CKA secure.

V. The Proposed Scheme

This section discusses the construction of the proposed scheme along with a detailed correctness analysis, security analysis, performance analysis and a comparative analysis with the existing schemes.

A. Detailed Construction

The proposed scheme uses a linear secret sharing scheme (LSSS) and monotonic access structure represented by an LSSS matrix, $\mathbb{A}$. The detail of each algorithm involved in the scheme is given below:

1) $GSetup(1^n)$

The global setup algorithm is run by a set of consensus nodes, $CN = \{cn_1, cn_2, \ldots, cn_n\}$. It selects a cyclic source group, G, of prime order, p, with a generator, g, and generates a bilinear map, $e : G \times G \to G_T$, where G_T is a target cyclic group of prime order, p. Let $\mathcal{GID}$ be the global identity space, Att be the attribute universe, and $\mathcal{W}$ be the keyword space. Select three collision-resistant hash functions, $H_1 : \mathcal{GID} \to G$, $H_2 : \mathcal{U} \to G$, $H_3 : \{0,1\}^* \to Z_p$ where Z_p is a group of integers of prime order, p, and each keyword, $w \in \mathcal{W}$, is assumed to be a binary string. Select two secrets $\{\alpha, \beta\} \leftarrow Z_p$, and calculate the shares of these two secrets, $\{\alpha_i, \beta_i\}_{i=1}^n$ using Shamir's secret sharing scheme (k, n) for n consensus nodes. Using the secret shares, each consensus node computes the two local public parameters, $e(g,g)^{\alpha_i}$ and g^{β_i} and broadcast them to the blockchain network. To get the global public parameters, k or more consensus nodes combine their secret shares using the Lagrange polynomial interpolation method as follows:

$$e(g,g)^{\alpha} = \prod_{i=1}^{k} \left(e(g,g)^{\alpha_i}\right)^{L_i} = e(g,g)^{\sum_{i=1}^{k} \alpha_i L_i} \tag{4}$$

$$g^{\beta} = \prod_{i=1}^{k} \left(g^{\beta_i}\right)^{L_i} = g^{\sum_{i=1}^{k} \beta_i L_i}. \tag{5}$$

Consensus nodes publish the global public parameters, $GPP = \left(e, p, g, H_1, H_2, H_3, e(g,g)^{\alpha}, g^{\beta}\right)$.

2) $KeyGen(GPP, Gid \in \mathcal{GID})$

It is executed locally by each CCPS user to get his/her public and private key pair corresponding to his/her Gid. Randomly select, $\gamma_{Gid} \in Z_p$ and compute $g^{\gamma_{Gid}}$. The user then sends the public key, $PK_{Gid} = g^{\gamma_{Gid}}$ to the blockchain (user registration at blockchain) and keeps the corresponding secret key, $SK_{Gid} = \gamma_{Gid}$ to himself/herself.

3) $GenIndex(GPP, (\mathbb{A}, \rho), w \in \mathcal{W})$

DO run this algorithm to encrypt the keyword, w, under access policy, $\mathbb{A}$, which is a $l \times m$ LSSS matrix with a function $\rho : [l] \to \mathcal{U}$ which maps $\mathbb{A}$'s rows to the attributes. DO select two random vectors, $\vec{v} = (s, v_2, \ldots, v_m)$ and $\vec{u} = (0, u_2, \ldots,$

$u_m) \in Z_p$ and compute shares of the first element of both $\vec{v}$ and $\vec{u}$ as: $\vec{\lambda} = \mathbb{A}\left(\vec{v}\right)^T = \{\lambda_x\}_{\forall x=\rho(l)}$ and $\vec{w} = \mathbb{A}\left(\vec{u}\right)^T = \{w_x\}_{\forall x=\rho(l)}$ where λ_x and w_x represents the secret shares of s and 0, respectively. For each row, $x \in \mathbb{A}[l]$ select a random number, $z_x \in Z_p$, and compute: $C_1 = e(g,g)^s$, $C_{2,x} = e(g,g)^{\lambda_x} e(g,g)^{\alpha z_x}$, $C_{3,x} = g^{w_x} g^{\beta z_x}$, $C_{4,x} = g^{z_x}$, $C_{5,x} = H_2(\rho(l))^{z_x}$ and for some fixed attribute, x', compute $C_{5,x'} = H_2(\rho(l'))^{z_{x'} H_3(w)}$ and outputs a keyword ciphertext, $C = \left(C_1, \{C_i\}_{2 \le i \le 5, \forall x \in \rho(l)}, (\mathbb{A},\rho)\right)$, which is then sent to the CS.

4) $PTok(GPP, PK_{Gid}, S \subseteq Att)$

To extract the needed information DU sends a search token generation request to the consortium blockchain. The request includes the global ID of the user and the attributes possessed by the user. Upon receiving the request, the consensus nodes first verify the claimed attributes and then executes this algorithm to get the partial search token for that user on demand. For each attribute, $x \in S$, each consensus node selects a random number, $r_{x_i} \in Z_p$, and computes: $PTK_{1,i} = g^{\gamma_{Gid} \alpha_i}$, $PTK_{2,i} = H_1(Gid)^{\beta_i}$, $PTK_{3,x,i} = g^{r_{x_i}}$, $PTK_{4,x,i} = H_2(x)^{r_{x_i}}$. The k consensus nodes together combine their partial search token components using the Lagrange polynomial interpolation method, to generate the partial search token as $PTK = [PTK_1, PTK_2, PTK_{3,x}, PTK_{4,x}]$ where $PTK_1 = \prod_{i=1}^{k} (g^{\gamma_{Gid}\alpha_i})^{L_i}$, $PTK_2 = \prod_{i=1}^{k} \left(H_1(Gid)^{\beta_i}\right)^{L_i}$, $PTK_{3,x} = \prod_{i=1}^{k} (g^{r_{x_i}})^{L_i}$, $PTK_{4,x} = \prod_{i=1}^{k} (H_2(x)^{r_{x_i}})^{L_i}$. The generated *PTK* is then sent to the user.

5) $CTok(GPP, SK_{Gid}, PTK, w')$

Upon receiving the partial search token from CB, DU now generates the complete search token for the keyword *w*', by selecting a random number, $t \leftarrow Z_p$, and computing; $CTK_1 = (PTK_1)^{\frac{1}{\gamma_{Gid}}} \cdot PTK_2^t$, $CTK_{2,x} = (PTK_{3,x})^t$, $CTK_{3,x} = (PTK_{4,x})^t$, $CTK_4 = H_1(Gid)^t$. For the same x' as in *GenIndex*, compute $CTK_{3,x'} = (PTK_{4,x})^{tH_3(w)}$. Output trapdoor, $CTK = (CTK_1, CTK_{2,x}, CTK_{3,x}, CTK_4)$. This search token is now sent to CS to enable search on his/her behalf.

6) Search(GPP, C, CTK)

The CS finds a subset of rows $\rho(l)$ of $\mathbb{A}$ such that the span of these rows is $(1,0,\ldots,0)$ and for all such attributes, x, labelled by those rows of $\mathbb{A}$, the cloud server computes, $D_x = C_{2,x} \times e(C_{3,x}, CTK_4) / e(C_{4,x}, CTK_1)$. The cloud server now chooses constants $c_x \in Z_p : \sum_x c_x \mathbb{A} = (1,0,\ldots,0)$, and finds $D'_x = \prod_x (D_x)^{c_x}$ and outputs

$$\begin{cases} 1, & \text{if } D'_x \times e(C_{5,x}, CTK_{2,x}) = C_1 \times e(C_{4,x}, CTK_{3,x}) \\ 0, & \text{Otherwise.} \end{cases} \tag{6}$$

Correctness analysis:

From L.H.S. of Equation (7.3), compute, D_x:

$$= \frac{e(g,g)^{\lambda_x} e(g,g)^{\alpha z_x} e(g, H_1(Gid))^{w_x t} e(g, H_1(Gid))^{\beta t z_x}}{e(g,g)^{\alpha z_x} e(H_1(Gid), g)^{\beta t z_x}}$$

$$= e(g,g)^{\lambda_x} e(g, H_1(Gid))^{w_x t} \tag{7}$$

$$D'_x = \prod_x \left(e(g,g)^{\lambda_x} e(g, H_1(Gid))^{w_x t}\right)^{c_x} = e(g,g)^s. \tag{8}$$

$\lambda_x c_x$ will reconstruct s and $w_x c_x$ will reconstruct 0. λ_x are the shares of secret s, while w_x are the shares of 0.

$$e(C_{5,x}, CTK_{2,x}) = e\left(H_2(\rho(l))^{z_x H_3(w)}, g^{r_x t}\right)$$
$$= e(H_2(x), g)^{z_x H_3(w) r_x t}. \tag{9}$$

Multiply (8) and (9):

$$D'_x \times e(C_{5,x}, T_{2,x}) = e(g,g)^s \times e(H_2(x), g)^{z_x H_3(w) r_x t}. \tag{10}$$

From R.H.S. of (6), compute:

$$C_1 \times e(C_{4,x}, CTK_{3,x}) = e(g,g)^s \times e\left(g^{z_x}, H_2(x)^{r_x t H_3(w)}\right)$$
$$= e(g,g)^s \times e(g, H_2(x))^{z_x H_3(w) r_x t}. \tag{11}$$

From (10) and (11), (6) holds, provided the keyword in both the ciphertext and the trapdoor are same.

B. Security Analysis

Security against collusion attacks: In the proposed scheme, each user possesses a unique global identity, Gid, which binds the search token components of a user together in a consistent manner. If some user wants to collude with another user then the binding fails and thus prevents two users with different Gid to collude [20].

Security against non-adaptive CKA: The following theorem guarantees the security against the non-adaptive CKA:

Theorem 1: The proposed scheme is secure against non-adaptive chosen keyword attack if the decisional parallel bilinear Diffie-Hellman exponent (DPBDHE) problem is hard.

Proof: We will start the proof by assuming that if there exists a probabilistic-polynomial time (PPT) adversary, $\mathcal{A}$, who can break the proposed scheme, then we can construct another probabilistic-polynomial time adversary, C, who can break DPBDHE assumption, given an instance (D, V) of the DPBDHE problem. C simulates the challenger in the security game defined for the proposed scheme as follows:

Global Setup Phase: Given the DPBDHE problem instance (D, U), C publishes global public parameters, $GPP = (e, p, g, G, G_T)$.

Initialization Phase: $\mathcal{A}$ selects a set of corrupted consensus nodes, $CN_c \subseteq CN : |CN_c| \le k-1$, and constructs the challenge LSSS matrix, $\mathbb{A}'$, of size $l^* \times m^*$ using the challenge access structure, $\mathbb{A}^*$. Let $c = |CN_c|$.

Query Phase: In this phase, C first generates the local public parameters for each non-corrupted consensus node, $\tau \in CN_n$ using the approach given in [21]. If the consensus protocol does not involve τ, then C randomly selects $\{\alpha_i, \beta_i\} \leftarrow Z_p$. In another case, if the consensus protocol involves τ, then C will incorporate the DPBDHE instance by choosing $\left\{\alpha'_i, \beta'_i\right\} \leftarrow Z_p$ and setting

$$e(g,g)^{\alpha_i} = e(g,g)^{\alpha'_i} \prod_{x \in X} e\left(g^{b_x a}, g^{a^q}\right)^{\mathbb{A}'_{x,1}} \tag{12}$$

$$g^{\beta_i} = g^{\beta'_i} \prod_{x \in X} \prod_{j=2}^{m'} \left(g^{b_x a^{q+2-j}}\right)^{\mathbb{A}'_{x,j}} \tag{13}$$

where $m' \le q$ and X are the set of attributes in the challenge access structure.

Hash queries $(H_1, H_2,$ *and* $H_3)$*:* To answer H_1 queries, C

maintains a list, L_{Gid}, which contains (Gid, h_{Gid}). If an entry exists in L_{Gid} for the queried Gid, then C returns the corresponding h_{Gid} value. Otherwise, if $Gid \notin Q_{sk}$, then C selects a random value $h_{Gid} \in G$ and adds it to L_{Gid}. If $Gid \in Q_{sk}$, then C considers the following two subcases: 1) If $\rho^*(l^*) \cap S_i = \emptyset$, then C randomly selects, $h_i \in Z_p$ and sets $h_{Gid} = g^{h_i} \times g^a \cdots g^{a^{m'-1}}$. 2) If $\rho^*(l^*) \cap S_i \neq \emptyset$, i.e., there are some rows X' whose associated attributes belong to S_i and $S_i \nvDash \mathbb{A}'$, then $\exists \overrightarrow{d_i} \in Z_p^{1\times m}$ with first term 1 and $\overrightarrow{d_i} \times \mathbb{A}'^{T}_{x\in X'} = 0$. Now, by the construction of $\mathbb{A}'$, it is known that X' will span over the subspace with dimension c. Therefore, $\overrightarrow{d_i}$ is orthogonal to the vectors that have exactly one 1 in the last c positions, which implies $d_{i,j}|_{m^*-c+1\leq j\leq m} = 0$. In this case, C selects a random element, $h_i \in G$ and sets $h_{Gid} = g^{h_i} \times (g^a)^{d_{i,2}} \cdots \left(g^{a^{m'-1}}\right)^{d_{i,m'}}$. To answer H_2 queries, C maintains a list, L_{Att}, which contains (x, h_x). If an entry exists in L_{Att} for the queried attribute x, then C returns the corresponding h_x value. Otherwise, if $x \notin \rho^*(l^*)$, then C randomly selects, $h_x \in G$ and adds it to L_{Att}. If $x \in \rho^*(l^*)$, C selects a random element, $f_x \in G$ and sets $h_x = g^{f_x} \times \prod_{j\in[m']}\left(g^{b_x a^{q+1-j}}\right)^{\mathbb{A}'_{x,j}}$. Let the query to H_3 be w_l, where C responds to $H_3(w_l)$ with a random number from Z_p. Now, using the output of H_1 and H_2 hash oracles, C will answer the partial search token queries for $\{Gid_i, S_i\}$ as follows:

If $x \notin \rho^*(l^*)$, then C knows α_i, β_i and generates the secret key as in the original scheme. If $x \in \rho^*(l^*)$ and $S_i \cap \rho^*(l^*) = \emptyset$, which implies that the attribute is in the challenge access structure, but none of the attributes of user i is present in it, then C uses the output of H_1 and H_2 oracles and sets

$$PTK_1 = (g^{\gamma_{Gid_i}})^{\alpha'} \tag{14}$$

$$PTK_2 = \left(g^{\beta'}\right)^{h_i} \times g^{\sum_{x\in X}\sum_{j=2}^{m'}\sum_{k=2}^{m'} b_x a^{q+1-j+k}\mathbb{A}'_{x,j}} \tag{15}$$

$$PTK_3 = g^{r_x} = \prod_{k\in[m']}\left(g^{a^k}\right)^{-1} \tag{16}$$

$$PTK_4 = (g^{r_x})^{f_x} g^{-\sum_{x\in X}\sum_{j=2}^{m'}\sum_{k=2}^{m'} b_x a^{q+1-j+k}\mathbb{A}'_{x,j}}. \tag{17}$$

Otherwise, $x \in \rho^*(l^*)$ and $S_i \cap \rho^*(l^*) \neq \emptyset$, which means the user has some shares of the challenge access policy. Then for each such $x \in X$, C uses another output of H_1 and H_2 oracles and sets

$$PTK_1 = g^{\gamma_{Gid_i}\alpha'} \tag{18}$$

$$PTK_2 = \left(g^{\beta'}\right)^{h_i} \times g^{\sum_{x\in X}\sum_{j,k=2,j\neq k}^{m'} b_x a^{q+1-j+k}\mathbb{A}'_{x,j} d_{i,k}} \tag{19}$$

$$PTK_3 = g^{r_x} = \prod_{k\in[m']}\left((g^{\gamma_{Gid_i}})^{a^k}\right)^{-d_{i,k}} \tag{20}$$

$$PTK_4 = (g^{r_x})^{f_x} g^{-\sum_{x\in X''}\sum_{j,k=2,j\neq k}^{m'} b_x a^{q+1-j+k}\mathbb{A}'_{x,j} d_{i,k}}. \tag{21}$$

The partial search token is then used by C to answer the search token queries for any keyword.

Challenge Phase: To generate the challenge ciphertext, C chooses a bit, $b \in \{0,1\}$ randomly and encrypts w_b, $C_1 = V$. For other components, C implicitly sets $\overrightarrow{v} = \left(sa^{q+1}, 0, \ldots, 0\right)$ and $\overrightarrow{u} = \left(0, sa^q, \ldots, sa^{q-m'+2}\right)$. If the consensus protocol involves corrupted consensus nodes then $\lambda_{x^*}, w_{x^*} = 0$ and C randomly selects $z_{x^*} \in Z_p$ and computes: $C_{2,x^*} = e(g,g)^{\alpha z_{x^*}}$, $C_{3,x^*} = g^{\beta z_{x^*}}$, $C_{4,x^*} = g^{z_{x^*}}$, $C_{5,x^*} = H_2(\rho(x^*))^{z_{x^*}}$. Otherwise, set $\lambda_{x^*} = sa^{q+1} \times \mathbb{A}'_{x^*,1}$ and $w_{x^*} = \sum_{j=2}^{m'} sa^{q-j+2} \times \mathbb{A}'_{x^*,1}$, then for each of row in the access matrix, C implicitly sets $z_{x^*} = s/b_{x^*}$ and computes

$$C_{2,x^*} = \prod_{x\in X\setminus\{x^*\}} e\left(g, g^{sb_x a^{q+1}/b_{x^*}}\right)^{\mathbb{A}'_{x^*,1}} \tag{22}$$

$$C_{3,x^*} = \prod_{x\in X\setminus\{x^*\}} \prod_{j=2}^{m'}\left(g^{sb_x a^{q+2-j}/b_{x^*}}\right)^{\mathbb{A}'_{x^*,j}} \tag{23}$$

$$C_{4,x^*} = g^{s/b_{x^*}} \tag{24}$$

$$C_{5,x^*} = \prod_{x\in X''} \prod_{j=2}^{m'}\left(g^{sb_x a^{q+1-j}/b_{x^*}}\right)^{\mathbb{A}'_{x^*,j}}. \tag{25}$$

Output Phase: If the guess of the adversary is correct, i.e., $b' = b$, then C correctly outputs the challenge term $V = e(g,g)^{sa^{q+1}}$, otherwise C outputs a random element of G_T.

■

C. Performance Analysis

In this section, we will first present the analysis of the proposed scheme comparatively with exiting schemes in the literature and then provide the implementation details of the proposed scheme to validate the theoretical claims.

1) Comparative Analysis

The comparative analysis of the proposed scheme is done in terms of the features, the storage cost, and the computational cost with the similar CP-ABE based searchable encryption schemes in the literature. The notation used throughout the comparative analysis is given in Table I.

Table II compares the basic features of the proposed scheme with related schemes.

2) Asymptotic Complexity Analysis

In this section, we will compare the storage and the computational cost of the existing CP-ABE based similar searchable encryption schemes in the literature with the proposed scheme, as shown below in Table III. Here, we will not compare every CP-ABE based searchable encryption scheme, but we will focus only those CP-ABE based searchable encryption schemes which have multi-authority support to manage a diverse set of users. The storage cost is computed in terms of the number and size of the group elements involved. In contrast, the computational cost is computed in terms of the number and types of operations involved in the output of each algorithm.

As shown in Table III, the storage and computational cost of the key generation algorithm is constant in the proposed scheme, unlike other similar schemes where it varies with the number of attributes. Therefore, the proposed scheme reduces

TABLE I
NOTATION AND DESCRIPTION

Notation	Meaning
$\|G\|, \|G_T\|$	Size of the source group (Symmetric bilinear map)
$\|G_1\|, \|G_2\|$	Size of source groups (Asymmetric bilinear map)
$\|G_T\|$	Size of the target group
ξ, N	Number of attributes associated with users and access policy, respectively
k	Number of keywords
E, E_T	Exponent operation in the source and the target group, respectively
P, H	Bilinear pairing and hash operation, respectively
$S_{Z_p}, M_{Z_p}, D_{Z_p}$	Modular subtraction, multiplication and division in the group Z_p
$\|CN\|$	Total number of consensus nodes
CP-ABE	Ciphertext-policy attribute-based encryption
KP-ABE	Key-policy attribute-based encryption
MBF	Monotone Boolean function
MSEDDH	Multi-sequence of exponents decisional Diffie-Hellman
DBDH	Decisional bilinear Diffie-Hellman
DPBDHE	Decisional parallel bilinear Diffie-Hellman exponent
ABDHE	Augmented decisional bilinear Diffie–Hellman exponent assumption
LSSS	Linear secret sharing scheme
FS, SS	Full security and selective security resp.

TABLE II
COMPARISON OF BASIC FEATURES OF THE PROPOSED DECENTRALIZED SEARCHABLE ENCRYPTION SCHEME WITH EXISTING SCHEMES

Ref.	BaseTech.	AccessStructure	Authorities	Security model & Assumption
[1]	KP-ABE, CP-ABE	Tree	Single	SS, Decision Linear
[2]	CP-ABE	AND	Single	SS, DBDH
[7]	CP-ABE	Tree	Single	SS, DBDH
[8]	CP-ABE	AND	Single	SS, MSEDDH
[9]	CP-ABE	MBF	Multiple	SS, DPBDHE
[22]	CP-ABE	MBF	Multiple	ABDHE and DBDH
Proposed scheme	CP-ABE	MBF	Multiple consensus nodes	SS, DPBDHE

the burden from attribute authorities to manage the secret key of users. Instead, it is managed by the consensus nodes, and that too is done only when the consensus nodes receive a search token generation request from the end-users. For remaining algorithms, storage and computational cost is dependent upon the number of attributes. However, the proposed scheme is genuinely decentralized and more secure due to the application of blockchain technology. Further, the search token needs not be stored by the user; it can be computed on-demand with consensus nodes. Consensus nodes generate a partial search token, upon which the user will perform some exponent and multiplication operations (minimal cost operations) to compute the complete search token. Therefore, we can say that the cumbersome process of computing search tokens corresponding to the user's attributes is delegated to the blockchain. Furthermore, system initialization is also managed by the consensus nodes in the blockchain. Therefore, the proposed scheme is genuinely decentralized and is free from a single point of failures.

3) Experimental Analysis

Experimental setup: To evaluate the performance of the proposed scheme, the implementation is done on a 64-bit Windows 10 system with an Intel Core i3 processor 2.00 GHz and 4 GB RAM, in JAVA using the Netbeans-8.1 IDE and java pairing-based cryptography library (JPBC) [23]. To create the symmetric bilinear map, we used a Type A pairing constructed on the elliptic curve, $y^2 = x^3 + x$ over a field F_q, where $q = 3 mod 4$ is a prime. In this pairing both G_1 and G_2 are the group of points from $E\left(F_q\right)$ and hence it is called a symmetric pairing.

Parameter setting: The size of the base field is set to be 512-bit which offers a security equivalent to 1024-bit DLOG [23] and the order, p of source group G and target group G_T is set to be 160-bit. The efficiency of the proposed scheme relies on the operations performed by the consensus nodes in the blockchain network and the remaining operations by the end users and the cloud server. The performance of the blockchain operations depends entirely upon the consensus mechanism

TABLE III
COMPARISON OF STORAGE AND COMPUTATIONAL COST OF THE PROPOSED DECENTRALIZED SEARCHABLE ENCRYPTION SCHEME WITH EXISTING SCHEMES

Ref.	Algorithm	Storage cost	Computational cost
[22]	$KeyGen$	$\xi\|G\|$	$2\xi E+\xi H$
	$GenIndex$	Offline: $(2N+1)\|G_T\|+5N\|G\|$ Online: $1\|G_T\|+4N\left\|Z_p\right\|$	Offline: $(3N+2)E_T+7NE$ Online: $3NM_{Z_p}+ND_{Z_p}+1E_T$
	$GenTok$	$(\xi+3)\|G\|+1\left\|Z_p\right\|$	$(\xi+3)E+(k+1)H$
	$Search$		$2NP+(2N+1)E+NE_T$
[9]	$KeyGen$	$(3\xi+4)\|G\|+3\left\|Z_p\right\|$	$(3\xi+8)E$
	$GenIndex$	$(3N+3)\|G\|+2\|G_T\|+N\left\|Z_p\right\|$	$(4N+3)E+2E_T+P$
	$GenTok$	$2\|G\|+1\left\|Z_p\right\|$	$(2\xi+2)E$
	$Search$		$2P+NM_{Z_p}$
Proposed scheme	$KeyGen$	$1\left\|Z_p\right\|$	$1E$
	$GenIndex$	$3N\|G\|+(N+1)\|G_T\|$	$4NE+(2N+1)E_T+(N+1)H$
	$PTok$ by blockchain	$(2\xi+2)\|G\|$	$(2\xi+2)E+(\xi+1)H$
	$CTok$ by user	No need to store	$(2\xi+2)E$
	$Search$		$4NP+NE_T$

TABLE IV
AVERAGE EXECUTION TIME (AET) (SECOND) OF THE ALGORITHMS OF THE PROPOSED SCHEME WHERE THE NUMBER OF ATTRIBUTES IN ATTRIBUTE UNIVERSE, ACCESS POLICY AND THE SET ξ IS KEPT THE SAME

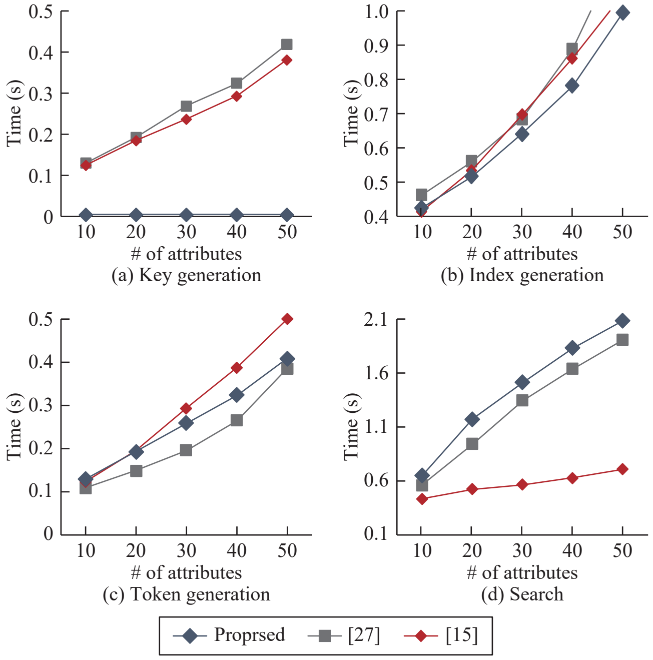
Proposed scheme (AET (s))		Number of attributes				
		10	20	30	40	50
Algorithms	$KeyGen$	0.004948404	0.004698573	0.004915196	0.004712041	0.004631696
	$GenIndex$	0.424942489	0.517242993	0.641555095	0.782029384	0.995143897
	$CTok$	0.130210341	0.193235129	0.259217566	0.324164851	0.408873766
	$Search$	0.676173615	1.226010711	1.583026596	1.919774288	2.183988908

and the computational power of the consensus nodes, which is not the focus of this paper. In this paper, we mainly focus on the regular operations performed by the end users and how blockchain technology assisted in reducing burden from these end users and the cloud, and compared that with existing multi-authority searchable encryption schemes [9], [22].

Procedure: To demonstrate performance, we have varied the number of attributes in the attribute universe, and the access policy in the set ξ from 10 to 50 with a step length of 10. In each step, the experiment has been executed 20 times to find the average time taken by each algorithm which is listed below in Table IV.

For the better demonstration of the results, we plotted the time complexity of each algorithm against the number of attributes comparatively, as shown in Fig. 5. As mentioned in the asymptotic complexity analysis section, we will compare the time complexity of the proposed scheme with existing similar decentralized CP-ABE based searchable encryption schemes.

Fig. 5(a) denotes the time taken by the key generation algorithm of the proposed scheme and other related schemes [9], [22]. The key generation time for the proposed scheme is constant and does not vary with the number of attributes, while in [9] and [22] a linear graph can be observed. In the



(a) Key generation

(b) Index generation

(c) Token generation

(d) Search

Fig. 5. Average execution time taken by (a) *KeyGen*, (b) *GenIndex*, (c) *CTok*, and (d) *Search* algorithms.

proposed scheme, the attributes are not embedded while generating the secret key for the user. At the same time, it is done directly during the search token generation by the consensus nodes. Here, the user will generate a public and private key pair for himself/herself corresponding to his/her unique global identity and sends the public key to blockchain network to register himself/herself to the blockchain network. Therefore, the proposed scheme eliminates the need of a central authority for the management of users attributes. It is done by the consensus nodes when a search token generation request is received from the user.

Fig. 5(b) shows the time taken by the index generation algorithm. As it can be observed from the Fig. 5(b), the index generation time of the schemes under consideration is almost the same and varies linearly with the number of attributes. Fig. 5(c) presents the time taken by the search token generation algorithm at the user's end. It also increases linearly with the number of attributes for all the three schemes. Fig. 5(d) shows the time taken by the search algorithm executed by the cloud server. It also varies linearly with the number of attributes, and the proposed scheme has higher time complexity as compared to [9] and [22]. The time complexity of [9] is least among the three since it only has two pairing operations, and thus the time taken by the search algorithm is almost constant. However, this algorithm is executed by the cloud server, which is assumed to have plenty of resources, and will not affect the performance of the scheme. Hence, the experimental results validate the theoretical claims made in the asymptotic complexity analysis section.

VI. Conclusion and Future Directions

In this paper, we have proposed a truly decentralized, robust and computationally efficient ABSE scheme for healthcare CCPS with the assistance of consortium blockchain. The devices involved in a typical healthcare CCPS are resource constrained. Further, the users in a healthcare CCPS may belong from a variety of domains, and to manage such a large number of diverse users, a single authority is not sufficient.

Keeping that in mind, we have proposed a scheme where computationally intensive tasks are delegated to the consensus nodes in the blockchain network to increase efficiency and to avoid single point failure. The consensus nodes are responsible for initializing the system and generating partial search tokens for users. This helps in reducing the computational burden from data users and also eliminates the need for a global central authority which sets up the system. Further, consensus nodes in the blockchain network handle the user's attributes, unlike the central trusted authority in similar ABSE schemes. Therefore, a large number of attributes from diverse domains can efficiently be handled by the proposed scheme. The proposed scheme addresses both of the aforementioned constraints in a healthcare CCPS, and thus works well in this scenario.

As a part of future work, one can work in two directions. First, one can impart efforts to incorporate a proper reward mechanism like the bitcoin network. In the proposed scheme, if we add a proper mechanism to reward users financially by means of a token, which can be spent by the users either in availing healthcare services or can be converted to other cryptocurrencies in return for sharing their data with government and healthcare agencies. By doing this, we can inspire users to utilize the system for the benefit of both the healthcare agencies and themselves. Second, one can work on adding scheme-specific features like verifiability of the search results and handle the event of user revocation with the assistance of the blockchain network to increase the correctness, robustness and dynamicity of the system.

References

[1] Q. Zheng, S. Xu, and G. Ateniese, "VABKS: Verifiable attribute-based keyword search over outsourced encrypted data," in *Proc. IEEE Infocom*, 2014, pp. 522–530.

[2] W. Sun, S. Yu, W. Lou, Y. T. Hou, and H. Li, "Protecting your right: Attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud," in *Proc. IEEE Infocom Conf. Computer Communications*, 2014, pp. 226–234.

[3] Mamta and B. B. Gupta, "An efficient KP design framework of attribute-based searchable encryption for user level revocation in cloud," *Concurr. Comput. Pract. Exp.*, p. e5291.

[4] Mamta and B. B. Gupta, "Secure fine-grained keyword search with efficient user revocation and traitor tracing in the cloud," *J. Organ. End User Comput.*, vol. 32, no. 4, pp. 112–137, 2020.

[5] Y. Miao, J. Ma, X. Liu, F. Wei, Z. Liu, and X. A. Wang, "m 2-ABKS: Attribute-based multi-keyword search over encrypted personal health records in multi-owner setting," *J. Med. Syst.*, vol. 40, no. 11, p. 246, 2016.

[6] C. Guo, R. Zhuang, Y. Jie, Y. Ren, T. Wu, and K.-K. R. Choo, "Fine-grained database field search using attribute-based encryption for e-healthcare clouds," *J. Med. Syst.*, vol. 40, no. 11, p. 235, 2016.

[7] Z. Chen, F. Zhang, P. Zhang, *et al.*, "Verifiable keyword search for secure big data-based mobile healthcare networks with fine-grained authorization control," *Futur. Gener. Comput.* Syst., vol. 87, pp. 712–724, 2018. DOI: 10.1016/j.future.2017.10.022.

[8] Mamta and B. B. Gupta, "An attribute-based keyword search for m-health networks," *J. Comput. Virol. Hacking Tech.*, pp. 1–16, 2020.

[9] Y. Miao, R. Deng, X. Liu, K.-K. R. Choo, H. Wu, and H. Li, "Multi-authority attribute-based keyword search over encrypted cloud data," *IEEE Trans. Dependable Secur. Comput.*, 2019. DOI: 10.1109/TDSC.2019.2935044

[10] L. Chen, W.-K. Lee, C.-C. Chang, K.-K. R. Choo, and N. Zhang, "Blockchain based searchable encryption for electronic health record sharing," *Futur. Gener. Comput. Syst.*, vol. 95, pp. 420–429, 2019.

[11] S. Cao, G. Zhang, P. Liu, X. Zhang, and F. Neri, "Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain," *Inf. Sci.* (*Ny*)., vol. 485, pp. 427–440, 2019.

[12] Y. Wang, A. Zhang, P. Zhang, and H. Wang, "Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain," *IEEE Access*, vol. 7, pp. 136704–136719, 2019.

[13] R. Guo, H. Shi, Q. Zhao, and D. Zheng, "Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems," *IEEE Access*, vol. 6, pp. 11676–11686, 2018.

[14] C. Esposito, M. Ficco, and B. B. Gupta, "Blockchain-based authentication and authorization for smart city applications," *Inf. Process. Manag.*, vol. 58, no. 2, p. 102468, 2021. DOI: 10.1016/j.ipm. 2020.102468.

[15] B. Waters, "Ciphertext-policy attribute-based encryption: An expressive, efficient, and provably secure realization," in *Int. Workshop on Public Key Cryptography*, 2011, pp. 53–70.

[16] A. Beimel and ע. למיב, *Secure Schemes for Secret Sharing and Key Distribution*. Technion-Israel Institute of Technology, Faculty of Computer Science, 1996.

[17] C. L. Stergiou, K. E. Psannis, and B. B. Gupta, "IoT-based big data secure management in the fog over a 6G wireless network," *IEEE Internet Things J.*, 2020. DOI: 10.1109/JIOT.2020.3033131

[18] D. Li, L. Deng, B. B. Gupta, H. Wang, and C. Choi, "A novel CNN based security guaranteed image watermarking generation scenario for smart city applications," *Inf. Sci.* (*Ny*)., vol. 479, pp. 432–447, 2019.

[19] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system (White Paper), " [Online]. Available: https//bitcoin.org/bitcoin.pdf, 2008.

[20] A. Lewko and B. Waters, "Decentralizing attribute-based encryption," in *Proc. Annual. Int. Conf. Theory and Applications of Cryptographic Techniques*, 2011, pp. 568–588.

[21] Y. Rouselakis and B. Waters, "Efficient statically-secure large-universe multi-authority attribute-based encryption," in *Proc. Int. Conf. Financial Cryptography and Data Security*, 2015, pp. 315–332.

[22] Q. Xu, C. Tan, W. Zhu, Y. Xiao, Z. Fan, and F. Cheng, "Decentralized attribute-based conjunctive keyword search scheme with online/offline encryption and outsource decryption for cloud computing," *Futur. Gener. Comput. Syst.*, 2019. DOI: 10.1016/j.future.2019.02.067

[23] A. De Caro and V. Iovino, "jPBC: Java pairing based cryptography, " in *Proc. IEEE Symp. Computers and Communications*, 2011, pp. 850–855.

Mamta received the Ph.D. in applied cryptography from the Department of Computer Engineering at National Institute of Technology, India, in October 2020. Her research interests include public-key cryptographic primitives, searchable encryption, information security, cloud computing and its extensions, and blockchain technology. She has published various research articles with various reputed publishers, like IEEE, Springer, Wiley, and Taylor and Francis group.

Brij B. Gupta (SM'17) received the Ph.D. degree from Indian Institute of Technology Roorkee, India, in the area of information and cyber security. He published more than 250 research papers in international journals and conferences of high repute. His biography was selected and published in the 30th Edition of Marquis Who's Who in the World, 2012. Dr. Gupta also received Young Faculty Research Fellowship Award from MeitY, government of India in 2017. He is also working as Principal Investigator of various R&D projects. He is serving as Associate Editor of *IEEE Access*, Associate Editor of IJICS, Inderscience and Executive Editor of IJITCA, Inderscience, respectively. He was also Visiting Researcher with Yamaguchi University, Japan, with Deakin University, Australia and with Swinburne University of Technology, Australia during 2015, 2017, and 2018, respectively. Additionally, he was Visiting Professor with Temple University, USA and Staffordshire University, UK during June 2019 and July 2019, respectively. At present, Dr. Gupta is working as Assistant Professor in the Department of Computer Engineering, National Institute of Technology, India. His research interest includes information security, cyber security, cloud computing, web security, intrusion detection and phishing.

Kuan-Ching Li is a Distinguished Professor in the Department of Computer Science and Information Engineering (CSIE) at Providence University, China, where he also serves as the Director of the High-Performance Computing and Networking Center. He received the Licenciatura in mathematics, and M.S. and Ph.D. degrees in electrical engineering from the University of Sao Paulo (USP), Brazil, in 1994, 1996, and 2001, respectively. He published more than 200 scientific papers and articles and is the Co-author or Co-editor of more than 25 books published by Taylor & Francis, Springer, and McGraw-Hill. He is the Editor in Chief of the journal *Connection Science* and serves as an Associate Editor for several leading journals. Also, he has been actively involved in many major conferences and workshops in program/general/steering conference chairman positions and has organized numerous conferences related to computational science and engineering. He is a Fellow of IET and a Senior Member of the IEEE. His research interests include parallel and distributed computing, big data, and emerging technologies.

Victor C. M. Leung is currently a Distinguished Professor of computer science and software engineering with Shenzhen University, China. He is also an Emeritus Professor of electrical and computer engineering and the Director of the Laboratory for Wireless Networks and Mobile Systems, The University of British Columbia (UBC), Canada. His research is in the broad areas of wireless networks and mobile systems. He has published widely in archival journals and refereed conference proceedings in these areas; several of his papers have won Best Paper Awards. He is a Fellow of the Royal Society of Canada, Canadian Academy of Engineering, and Engineering Institute of Canada. He was a recipient of the 1977 APEBC Gold Medal, Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarships from 1977 to 1981, a 2012 UBC Killam Research Prize, IEEE Vancouver Section Centennial Award, the 2017 Canadian Award for Telecommunications Research, and the 2018 IEEE TCGCC Distinguished Technical Achievement Recognition Award. He has coauthored articles that won the 2017 IEEE ComSoc Fred W. Ellersick Prize, the 2017 IEEE Systems Journal Best Paper Award, the 2018 IEEE CSIM Best Journal Paper Award, and the 2019 IEEE TCGCC Best Journal Paper Award. He is named in the current Clarivate Analytics list of "Highly Cited Researchers." He has provided leadership to the Technical Program Committees and Organizing Committees of numerous international conferences. He is serving on the Editorial Boards of the *IEEE Transactions on Green Communications and Networking*, the *IEEE Transactions on Cloud Computing*, *IEEE Network*, and *IEEE Access*. He has previously served on the editorial boards of the IEEE Journal on Selected Areas in Communications Wireless Communications Series and Series on Green Communications and Networking, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Vehicular Technology*, the *IEEE Transactions on Computers*, and the *IEEE Wireless Communications Letters*.

Kostas E. Psannis is currently an Associate Professor in communications systems and networking at the Department of Applied Informatics, School of Information Sciences, University of Macedonia, Greece, Director of Mobility2net Research & Development & Consulting JP-EU Lab and Member of the EU-JAPAN Centre for Industrial Cooperation. Dr. Psannis received the degree in physics (Department of Physics, founded in 1928), Faculty of Sciences, from Aristotle University of Thessaloniki (AUTH, founded in 1925), Greece, and the Ph.D. degree from the School of Engineering and Design, Department of Electronic and Computer Engineering of Brunel University, UK. From 2001 to 2002 he was awarded the British Chevening scholarship. The Chevening Scholarships are the UK government's global scholarship programme, funded by the Foreign and Commonwealth Office (FCO) and partner organisations. The programme makes awards to outstanding scholars with leadership potential from around the world to study at universities in the UK. Dr. Psannis' research spans a wide range of digital media communications, media coding/synchronization and transport over a variety of networks, both from the theoretical as well as the practical points of view. His recent work has been directed toward the demanding digital signals and systems problems arising from the various areas of ubiquitous big data/media and communications. This work is supported by research grants and contracts from various government organisations. Dr. Psannis has participated in joint research works funded by Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science (JSPS), KAKENHI Grant, The Telecommunications Advancement Foundation, International Information Science Foundation, as a Principal Investigator and Visiting Consultant Professor in Nagoya Institute of Technology, Japan. Dr. Psannis was invited to speak on the EU-Japan Co-ordinated Call Preparatory meeting, Green & Content Centric Networking (CCN), organized by European Commission (EC) and National Institute of Information and Communications Technology (NICT)/ Ministry of Internal Affairs and Communications (MIC), Japan (in the context of the upcoming ICT Work Programme 2013) and International Telecommunication Union (ITU-founded in 1865), SG13 meeting on DAN/CCN, Berlin, July 2012, amongst other invited speakers. Konstantinos received a joint-research Award from the Institute of Electronics, Information and Communication Engineers, Japan, Technical Committee on Communication Quality, July 2009 and joint-research Encouraging Prize from the IEICE Technical Committee on Communication Systems (CS), July 2011. Dr. Psannis has more than 60 publications in international scientific journals and more than 70 publications in international conferences. His published works has more than 2100 citations (h-index 24, i10-index 41). Dr. Psannis supervises a post-doc student and seven Ph.D. students. Prof. Psannis is serving as an Associate Editor for *IEEE Access* and *IEEE Communications Letters*. He is Lead Associate Editor for the Special Issue on Roadmap to 5G: Rising to the challenge, *IEEE Access*, 2019. He is a Guest Editor for the Special Issue on Compressive Sensing-Based IoT Applications, *Sensors*, 2020. He is a Guest Editor for the Special Issue on Advances in Baseband Signal Processing, *Circuit Designs, and Communications, Information*, 2020. He is a Lead Guest Editor for the Special Issue on Artificial Intelligence for Cloud Based Big Data Analytics, *Big Data Research*, 2020. He is TPC Co-Chair at the International Conference on Computer Communications and the Internet (ICCCI 2020), Nagoya Institute of Technology Japan, ICCCI to be held in 2020 June 26–29 at Nagoya, Japan, and Conference Chair at the World Symposium on Communications Engineering (WSCE 2020- http://wsce.org/) to be held at University of Macedonia, Thessaloniki, Greece, October 9–11, 2020.

Shingo Yamaguchi (SM'14) is a Professor in the Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Japan. He received the B.E., M.E., and D.E. degrees from Yamaguchi University, Japan, in 1992, 1994, and 2002. He was a Visiting Scholar at the University of Illinois at Chicago, US, in 2007. He is currently the Director of Information and Data Science Education Center, Yamaguchi University. His research interests include AI, IoT, big data, and cybersecurity. He was the Executive Conference Chair of IEEE ICCE 2021. He is a Member of the Board of Governors of IEEE Consumer Technology Society. He is also the Editor-in-Chief of *IEEE Consumer Electronics Magazine*. He is a Senior Member of IEEE and IEICE.