Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766333

REPUTATION-BASED SCALABLE BLOCKCHAIN SYSTEM USING SHARDING SCHEME FOR HIGH THROUGHPUT LOW LATENCY APPLICATION IN E-HEALTHCHAIN

S.Banupriya [#1] ,Dr.P.Sharmila [#2]
Ph.D Scholar [#1] , Professor[#2]
priyamca126@gmail.com, mavish2009@gmail.com
[#1]PG and Research Department of Computer Science
Navarasam Arts and Science College for Women, Arachalur, Erode.
[#2]Departement of Computer Science,
KPR College of Arts Science and Research, Arasur, Coimbatore.

ABSTRACT—
Blockchain technology plays a vital role in the medical industry for COVID-19 by facilitating communication between various stakeholders. Blockchain-based healthcare systems like eHealthChain, etc., have been developed over the past years to store health records securely. But most systems depend on storing data either directly on the blockchain, which is not scalable, or on off-chain storage systems, which are not completely decentralized. Also, scalability was one of the major issues in these systems, which limits blockchain technology in high-throughput, low-latency applications such as COVID-19 image sharing. Hence, in this article, a Scalable eHealthChain (SeHealthChain) is proposed that introduces a new reputation-based scalable blockchain system via a sharding scheme in the eHealthChain system. First, a novel dual-chain structure including a transaction chain and reputation chain is proposed. A modified Raft-based synchronous consensus scheme is used for the transaction chain, whereas the synchronous Byzantine fault-tolerant consensus is used for the reputation chain. This prevents attacks on both the reputation score and the associated transaction blocks. It supports a high-throughput transaction chain with tolerable generation speed. Additionally, a reputation-based sharding and leader selection technique is proposed to increase the system throughput and security level. Thus, this system enhances the throughput and security level for COVID-19 image transmission. Finally, an extensive experiment shows that the SeHealthChain achieves higher throughput and level of security compared to other sharding-based blockchain systems.

Keywords—Blockchain, Healthcare applications, eHealthChain, Sharding, Reputation

I. INTRODUCTION

Blockchain is an anonymous, decentralized security design in which users within a state's bounds work together to verify the entire system. Users of a decentralized network, such as an IoT network, must be acquainted with the activities of unfamiliar parties [1, 2]. False operations can be found because actions are gathered in a shared blockchain that is accessible to network users [3]. As a result, data confidentiality is guaranteed in a distributed and non-interfering way. They are critical in enabling many important uses, such as industry 4.0, medicine, and

smart homes [4, 5]. In the healthcare industry, it allows users to monitor clinical data or images [6–8].

The pharmacy and medical sectors will use blockchain technology to eradicate fraudulent transactions, allowing for the tracking of all these items. It helps to identify the cause of the falsification. If a therapeutic strategy is developed, it can guarantee the confidentiality of medical data [9]. Furthermore, this information cannot be modified. They use completely dedicated computers to implement this distributed system. Experts can estimate chances for therapies, medicines, and surgeries for many kinds of medical issues using the resources saved by such systems [10]. According to the blockchain in the healthcare industry, distributed ledgers are ideal for keeping and sharing clinical data for particular purposes: (i) the medical industry is made up of several entities (such as patients, doctors, caregivers, and pharmacists), (ii) entities must have more trust in one another than they do now, (iii) without a mediator, credibility and performance rise, (iv) it is critical to reliably monitor physiological bio-signals, and (v) bio-signal information must be consistent throughout the moment to be used for advanced assessment [11].

As the number of IoT devices and associated records has grown, an internet supplier known as an information dealer or data exporter has arisen. Clients are not directly informed about how confidential information is collected, processed, and used because security requirements are hidden within lengthy agreements [12]. Blockchain could be used to store and distribute customer medical records from clinical IoT devices. Patients can gain control over their confidential information by using a blockchain-based Personal Health Information Management System (PHIMS), which allows them to interact with other organizations without the need for an information broker [13]. Conversely, new criteria are required to collect information from the clinical IoT system and disseminate it to the blockchain for preservation to use blockchain solutions for PHIMS implementation. To address this issue, Pawar et al. [14] created eHealthChain, a blockchain-based PHIMS for controlling medical records acquired from clinical IoT devices and other uses.

Although blockchain has been used to keep health records, most depend on keeping data either directly on the blockchain, which is not scalable, or on off-chain storage systems, which are not completely decentralized or interoperable with blockchain [15]. Furthermore, very few studies have been conducted on big datasets such as medical imaging data. Nevertheless, inadequate scalability has limited the broad acceptance of conventional blockchain technology in high-throughput, low-latency applications such as COVID-19 image sharing in the medical industry.

Hence, this article proposes Scalable eHealthChain (SeHealthChain), which adopts a new reputation-based scalable blockchain system via a sharding scheme for increasing the scalability of high-throughput and low-latency applications. It combines the reputation scores with a sharding-based blockchain to explicitly represent the heterogeneity among the evaluators and consider the basis for the incentive strategy. In this system, a novel dual-chain structure is initially proposed, such as a transaction chain and a reputation chain. A modified Raft-based synchronous consensus scheme is used for the transaction chain, whereas the synchronous Byzantine fault-tolerant consensus is used for the reputation chain. This prevents attacks on both the reputation score and the associated transaction blocks. It supports a high-throughput transaction chain with tolerable generation speed. Additionally, a reputation-based

sharding and leader selection technique is proposed to increase the system throughput and security level. Thus, this system enhances the throughput and security level for low-latency healthcare applications like COVID-19 image transmission.

The following sections are organized as follows: Section II discusses the various shard-based blockchain architectures. The suggested model is explained in Section III, and its efficacy is demonstrated in Section IV. Section V provides a summary of the study and suggests some next steps.

II. RELATED WORKS

Kim [16] developed a two-stage cooperative bargaining game technique for a shard-based blockchain consensus model. The overall transactions per period were split for all shards according to the egalitarian bargaining solution. The assigned transactions in all shards were validated by blockchain nodes based on the proportional bargaining solution. But it needs additional factors like reputation to achieve a highly complex and efficient consensus solution. Sohrabi & Tari [17] designed a ZyConChain system, which generates various chains such as parentBlock, sideBlock, and state block. The parentBlock was created based on the Nakamoto consensus algorithm, and the sideBlocks were created according to the Zyzzyva consensus algorithm. A sharding scheme was used to enable nodes to verify the cross-shard transactions. But it needs more storage per node. Mizrahi and Rottenstreich [18] investigated traffic-aware sharding to find memory-light sharding with a low cross-shard rate. An abstract cost was used to model the memory needed to define a function. But it needs other costs, like incentives, to improve efficiency. A few shards process more transactions than others, which leads to a load imbalance among shards.

Hong et al. [19] presented a hierarchical sharding blockchain system called Pyramid, where one shard can store the full records of multiple shards, and transactions can be processed and checked within those shards. However, both the shards were completely isolated, and each node belongs to one shard, which tends to cause a lack of communication between each other, resulting in variances in packing costs.

Amiri et al. [20] developed a scalable permissioned blockchain system called SharPer to increase scalability with deterministic security promises. The nodes were grouped, and all data shards were replicated on the nodes of a group. In this system, the blockchain ledger was created as a directed acyclic graph, and only a view of the ledger was maintained by all groups. Also, the cross-shard consensus was achieved by the decentralized flattened protocols. But the latency was too high to process more transactions because of slow consensus.

Hellings and Sadoghi [21] presented a unifying model called ByShard that reduces the total number of consensuses. Also, centralized orchestration and distributed orchestration were used based on Byzantine primitives to reduce the latency needed to achieve agreement. But it needs to tolerate Byzantine nodes that may attack the system. Hashim et al. [22] applied a transaction-based sharding scheme called MedShard to create shards based on the patient's historical medical records. But the cross-shard transmission was removed that degrading the network efficiency in any shared framework. Zheng et al. [23] presented Meepo for enhancing cross-shard transaction efficiency via the cross-epoch and cross-call. Also, a partial cross-call fusion method was applied to control the multi-state dependency in contract calls. A shadow shard-based recovery was used to enhance the shard's robustness. But the heterogeneity of various shards and the incentive price were required.

Huang et al. [24] developed a cross-shard blockchain protocol called BrokerChain for account-based state sharding. The fine-grained state partition reduced the number of cross-shard transactions in the whole blockchain system. This leads more mutual brokers to seem in the same shard. Also, it impacts the throughput and transaction validation latency. Zheng et al. [25] presented an Aeolus blockchain to execute distributed blockchain transactions. A distributed blockchain transaction design was applied that considers additional variables to split the transaction execution into multiple stages and enable distributed execution. Also, distributed state update sharding was used to reduce the execution period and achieve consensus. But it needs to consider the incentive cost of the shards to improve efficiency.

From the literature, it is observed that the previous schemes rarely defined the incentive cost of the sharding blockchain. In contrast with the previous schemes, the SeHealthChain is a novel incentive strategy suitable for sharding the blockchain, which greatly improves the throughput, latency, and transaction success rates.

III. PROPOSED METHODOLOGY

This section briefly explains the SeHealthChain system. First, it summaries the blockchain-enabled medical systems, 3-level edge-IoT system in healthcare systems, and SeHealthChain. Then, the reputation-based sharding scheme is described briefly for the SeHealthChain system.

3.1 Blockchain-enabled Medical System

Figure 1 shows an example of the blockchain scheme to process healthcare data [14], where different kinds of data archives, including data regarding mobile clinics, life insurance, family medical histories, and doctor's prescriptions, are gathered in the blockchain server on the cloud. Such data can be accessed by licensed medical professionals, patients, and academics with the patient's consensus.

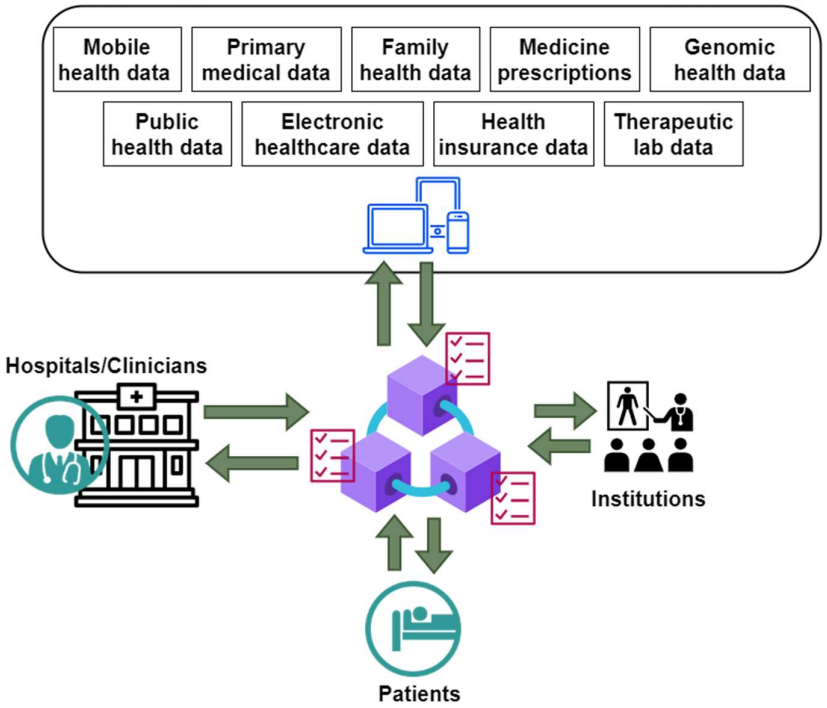

Figure 1. Blockchain-Enabled Medical System

3.1.1 Blockchain-enabled 3-level Edge-IoT System

Figure 2 illustrates the blockchain-edge model for healthcare IoT applications, which has multiple IoT groups, and all of them are linked to the corresponding edge nodes. IoT groups are resource-constrained; so such IoT groups are merged with respective edge nodes through a gateway. As a result, it enables medical system applications to collect, process, store, and retrieve data locally and satisfy the low delay demands for local latency-critical stages.

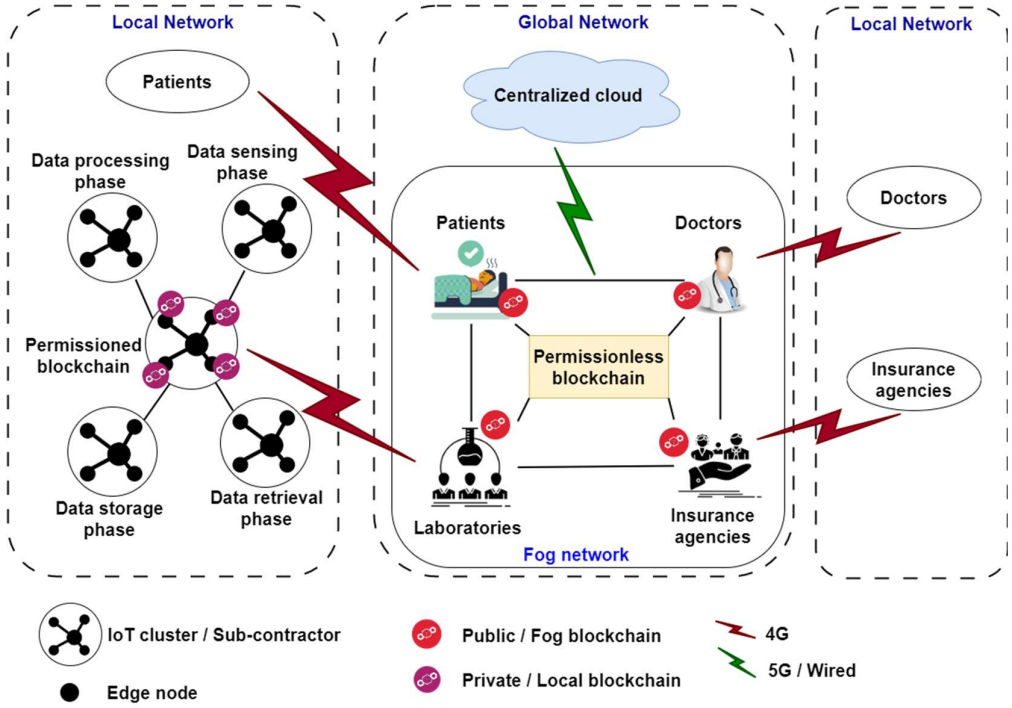

Figure 2. Blockchain-enabled 3-level Edge-IoT System in Medical Applications

Consider lightweight private/permissioned blockchain at the IoT-edge networks which will enable secure and trusted distribution of the desired data among various IoT-edge groups. At the local networks, a blockchain is a useful tool for data analysis and transmission. With smart contracts (stakeholders related to healthcare), a sub-contractor (patients, doctors, and insurance agencies) can verify the data sources and other stakeholders in the chain. The local blockchain also ensures the authentication and access control mechanism at the local network. In case the required service/resource is not available at the local networks, the request is forwarded to the edge networks. Edge networks operate without a connection to the public network or even the closest access network base station. Local edge nodes send requests to edge networks for data processing and executing higher-resource-intensive operations. The edge network is vital in providing elastic resources and services with low-latency access for smart healthcare systems. It can provide a healthcare platform where the provider can generate the data sources and give authority to the user. Various edge nodes will need to share the necessary data of ongoing processes. They assume permissionless or public blockchains and share limited data in the network. The global network, i.e. centralized cloud can provide the highest resource capabilities as compared with the above two networks. It follows the traditional centralized cloud computing approaches that provide a globally available service platform for applications requiring high storage and computational capacity. Transactions occurring between multiple networks and organizations are stored on the blockchain as permanent records. To effectively implement the different stages, each of the networks must operate collaboratively at all stages.

3.2 Scalable eHealthChain System for Medical Applications

The SeHealthChain system is a shard blockchain-enabled PHIMS for collecting, processing, and disseminating patients' health information sensed by clinical IoT tools. It interfaces clinical IoT tools with the shard blockchain storage using a unique interface unit. This interface unit is used to gather information from IoT systems and accumulate them in shards. Also, it retrieves information from the shards and transmits it to the app, which gives a client-friendly interpretation of accumulated information. Figure 3 portrays the structure of the SeHealthChain system that engages 4 major layers: shard blockchain, interface, application, and system layer.

1. Shard blockchain layer: Blockchain is a distributed Hyperledger Fabric that maintains a record of all transactions that are executed in the network. The health information obtained from clinical IoT tools is stored in the shards. This data is only accessed by authorized entities based on the consent of the owner. Consensus is the task of granting and synchronizing ledger files across the network. The shard blockchain layer independently maintains a dual-chain structure, i.e., transaction chain and reputation chain. For the transaction chain, a modified Raft consensus is applied to efficiently generate transaction blocks. For the reputation chain, a Byzantine fault tolerance consensus is used. The hashes of several verified transaction blocks are packed and the reputation scores of each evaluator are computed according to their behaviors into a single reputation block to tradeoff the gains and overheads of Byzantine consensus.
2. Interface layer: It interfaces the application layer with the shard blockchain layer to get the client's medical information using the OAuth 2.0 protocol, which guarantees the patient's authority. It utilizes REST Application Programming Interfaces (APIs) provided by the shards to write health data to the shards.

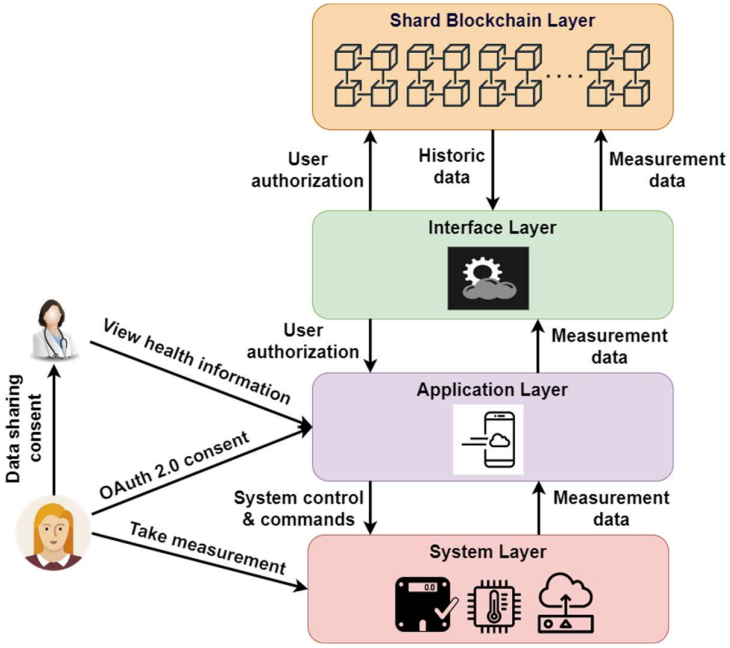

Figure 3. Layered Structure of SeHealthChain System in Medical Applications

- Application layer: It consists of mobile apps which gather data from patients' medical IoT tools. Such apps facilitate information distribution to external parties using OAuth

2.0 protocol. OAuth is a consensus scheme that enables third-party apps to get restricted access to client profiles.

- System layer: It comprises clinical IoT gadgets which are connected to handheld mobile devices using short-range wireless technologies such as Bluetooth. Each patient is in charge of revising personal medical data. The patient can validate other stakeholders who have access to their information.

3.2.1 Overview of Dual-Chain Structure-based Sharding Blockchain Scheme

The SeHealthChain system utilizes a dual-chain structure called a transaction chain and a reputation chain. It comprises four main modules: sharding and leader selection, consensus, reputation method, and synchronization.

- Sharding and leader selection: At the beginning of all epochs, each node is divided into multiple shards. After that, all shards choose their shard leader.
- Consensus: The users transmit transactions (Tx) to the shards that are responsible for the input UTXO (Unspent Transaction Output). After that, such shards execute an intra-shard synchronous consensus to develop the transaction blockchain and the reputation blockchain. For cross-shard transactions, the SeHealthChain system implements an atomic cross-shard protocol.
- Reputation method: For behaviors of each evaluator by the consensus, they can determine and realize a consensus on the reputation scores. The reputation scores improve the earlier two modules by supporting the selection of a high-ability leader in the shard leader selection and balancing several shards to have an equal ratio of active, inactive, honest, and malevolent evaluators. So, every three modules operate together to offer high throughput, secure, and high-incentive blockchain.
- Synchronization: At the termination of all epochs, all shards create a state block to accomplish the transaction and reputation blockchains. The nodes coordinate and fine-tune the stored reputation scores depending on such state blocks from each shard. After that, the system initiates a new epoch.

All these modules are explained below section.

a) Sharding and Leader Selection

In this module, the SeHealthChain system upholds the following features:

1. Uncertainty: It is difficult to forecast the outcomes of the sharding and leader selection.
2. Tradeoff: All shards have an equal overall reputation score, i.e. an equal ratio of active, inactive, honest, and malevolent evaluators. It is complex for malevolent users to take control of one shard.
3. Homogeneity: The outcomes are confirmed by all evaluators locally without high communication overhead.
4. Incentive: For the evaluators, the higher the reputation, the higher the chance of being elected as the shard leader.

If a new epoch e initiates, each evaluator can be shared into multiple clusters. For all clusters, a shard leader can be chosen according to their reputation. Decaying is a crucial aspect of the reputation method, so a sliding window ω is implemented to determine the total reputation

score R^{ω}. The evaluators utilize the random generator (RNG) from the random $seed^{e}$ to create random numbers. This seed can be created by the secure distributed bias-resistant uncertainty creation protocol. R^{rank} is the evaluators' reputation scores ranked in descending order. After that, all evaluators are randomly allocated to a minimum size for preserving the tradeoff feature.

For shard leader selection, the below scheme is adopted to concurrently keep the incentive, uncertainty, and homogeneity features. Initially, the evaluators with reputation scores greater than the average have a probability of being chosen as the leader. All scores of such evaluators can be split into a number, which is randomly created depending on a similar $seed^{e}$. The evaluator with the minimum outcome is chosen as the leader.

Algorithm 1 Sharding and leader selection

Input: A random $seed^{e}$ and the total reputation score $R^{\omega} = r_1^{\omega} + r_2^{\omega} + \cdots + r_n^{\omega}$ over earlier ω epochs;

Output: k shards $C = \{C_1, \ldots, C_k\}$ and k leaders $L = \{l_1, \ldots, l_k\}$

1. **Begin**
2. Initialize $C_i = \emptyset, L_i = \emptyset$ for all $1 \le i \le k$;
3. Set the seed of RNG as $seed^{e}$;
4. $R^{rank} = rank(R^{\omega})$;
5. $\boldsymbol{for}\left(all\ r_i^{rank}\ in\ R^{rank}\right)$ //r_i^{rank} is evaluator Ev_g's score
6. Discover a shard C_t that has the minimum cardinality;
7. $\boldsymbol{if}(several\ C_t\ satisfied\ the\ requirements)$
8. Choose one randomly depending on RNG;
9. $\boldsymbol{end\ if}$
10. Allocate evaluator Ev_g to C_t, $C_t = C_t \cup \{Ev_g\}$;
11. $\boldsymbol{end\ for}$
12. $\boldsymbol{for}(all\ shards\ C_i \in C)$
13. $ra = average\ of\ the\ subset\ of\ R^{\omega}$ that belongs to C_i;
14. $\boldsymbol{for}\left(all\ evaluators\ Ev_j \in C_i\right)$
15. $\boldsymbol{if}\left(ra \le r_j^{\omega}\right)$
16. Create a random float $0 \le y \le 1$ from RNG;
17. $p_{i,j} = {}^{y}/_{r_j^{\omega}}$;
18. $\boldsymbol{else}$
19. $p_{i,j} = +\infty$;
20. $\boldsymbol{end\ if}$
21. $\boldsymbol{end\ for}$
22. $l_i = Ev_j$, where $p_{i,j} = \min\left(p_{i,1}, p_{i,2}, \ldots, p_{i,m}\right)$;
23. $\boldsymbol{end\ for}$
24. **End**

b) Consensus

All Txs can contain a unique identity, a list of input UTXOs, and output UTXOs. The UTXO is the new data from earlier Tx and has the signature. The shard i can only accumulate the Tx

with a particular identity, i.e., the prefix of the identity is equivalent to i. The input UTXOs may come from various shards. So, this system must manage cross-shard transactions. The shards for the input and output UTXOs are denoted by the input shard and the output shard, respectively. A synchronous consensus is introduced to realize high throughput and ½ flexibility in a shard. This consensus creates 2 different chains in all shards: a Raft consensus to create a transaction chain, and a Byzantine fault-tolerant consensus to create the reputation chain.

1. Intra-shard consensus: Initially, the leader creates a transaction chain through Raft consensus. Consider TxL is the transaction list, which orders the transactions so that all evaluators can determine the reputation according to a similar batch of transactions. $TxDc$ is the transaction decisions provided by the user, which has the decisions in a similar order as TxL. $TxDcS$ is the transaction decision set containing each $TxDc$ from shard users. The transmission design of the Raft consensus includes the following steps: (i) the leader can transmit the TxL to each evaluator; (ii) all evaluators can confirm each Tx in the TxL and create a decision (i.e., accept, discard, or unknown) for all Txs. The unknown is utilized if the evaluator cannot manage several transactions because of hardware constraints. After that, the evaluators can transmit the $TxDc$ to the leader; (iii) the leader collects each $TxDc$ to the $TxDcS$; also, the leader can create the Transaction Block (TB) depending on the $TxDcS$. Afterward, the leader can transmit both the $TxDcS$ and the TB to each evaluator; (iv) when the leader introduces a self-promoting or slandering attack by altering the TxL, $TxDcS$ or TB, the evaluators can transmit a warning to all others and then view-shift occurs. After completing this protocol, the evaluators can determine the reputation scores. The Raft consensus can effectively create the TB and is highly appropriate for a high-ability leader. But, it cannot avoid a Byzantine attack. So, such TBs can only be called candidate TBs and can finally realize consensus in the second stage.

   In the second stage, the SeHealthChain creates Reputation Block (RB) through a Byzantine fault-tolerant consensus. The RB has the hashes of TBs to be finally verified and the reputation of each evaluator. In this system, the RapidChain consensus [26] is combined with CoSi [27] to concatenate several signatures into a unified signature. The RapidChain consensus involves four synchronous rounds: (i) plan: a leader plans a hash digest H for the consensus; (ii) repeat: evaluators transmit H received from the leader to each other node with label repeat when the H is right; (iii) awaiting: when an honest evaluator receives distinct hashed H and H', it can transmit H' with label awaiting; (iv) accept: when an authentic evaluator receives $f + 1$ repeat of the same and only H, it can accept and transmit H with a label accept. After the consensus, the authentic evaluator can identify whether the leader is dishonest or not. Similarly, CoSi involves four rounds: announce, commit, challenge, and response. The plan and repeat rounds in the consensus can be utilized for announce and commit rounds in the CoSi. The leader transmits an announcement of the signing of H in the plan round. The evaluators can repeat the leader H and a random secret in the repeat round. In accept round, the leader transmits a collective Schnorr challenge and H with $f + 1$ repeats. After accept round, the authentic evaluator will determine whether the leader is dishonest or not. So,

another round of response is required: evaluators will accept or discard transmitting the cumulative response to the leader and the leader can create the aggregate signature on the RB.

2. Warning in intra-shard consensus: In intra-shard consensus, a warning is used to identify the malevolent behavior of the leader. When the leader transmits an illegal TxL, $TxDcS$ or the TB, which has Tx less than half of the votes, the $f+1$ authentic evaluator can transmit a warning to the others. After the $f+1$ warning is received, a view-shift occurs. Malevolent leaders will also select to eliminate $TxDc$ of an authentic evaluator from $TxDcS$ to reduce its reputation. The victim will transmit a warning with its $TxDc$ to each other evaluator. Every authentic evaluator can add a reputation to the evaluators. When the malevolent leader declines to add the reputation score of the victim, the hash of RB, H can be varied among the malevolent leader and the authentic evaluators.
3. Cross-shard protocol: The Atomix protocol from Omniledger is utilized for cross-shard transactions, i.e., the user can assist in relaying the transactions across the shard with proof-of-acceptance. In this system, the proof-of-acceptance is the aggregate signature on RB.

c). Reputation Method

This reputation method will improve the security level, incentive feature, and throughput of SeHealthChain.

Reputation score computation: At the termination of one intra-shard consensus, the reputation scores within the shard will be computed by each shard member independently and regularly depending on $TxDcS$ and TB. The reputation score r_x of an evaluator x is computed by

$$r_x = \sum_{j=1}^{l} S(j) * T(j) \quad (1)$$

In Eq. (1), l denotes the number of transactions created after the earlier RB, $T(j)$ denotes the value of transaction j to avoid the scenario, where a few evaluators are authentic on small transactions but corrupt on a large transaction, and the scaling factor $S(j)$ is utilized to reward or penalize various behaviors differentially. This system assigns various scaling factors for various behaviors to make the punishment for corrupt behaviors greater than the reward for authentic ones. Also, an illegal Tx being approved is greatly riskier than a legal Tx being quit. So, when an evaluator provides acceptance if many others provide a discard, the evaluator can be punished more than if it provides a discard when the others provide an acceptance. For unknown decisions, the evaluators neither receive nor lose their reputation scores, because it is assumed that such evaluators usually have lower capacities (CPU, bandwidth, and so on).

Reputation blockchain: Once the reputation score is computed, the evaluators use the aggregate signature to create the RB. The RB has the reputation score, the evaluator owns in this epoch, the authenticated TBs, the earlier RB, and the aggregate signature. The consensus guarantees that the authentic evaluators can sign on the RB when the connected TBs and the reputation scores are accurate. Other shards will verify the aggregate signature, and accept it when greater than ½ of the evaluators sign on the block. Additionally, the malevolent evaluators will split the transaction blockchain and the reputation blockchain.

d) Synchronization

In the sharding system, the transactions are divided into multiple shards. So, at the termination of one epoch, the evaluators of each shard must synchronize their reputation blockchain and

transaction blockchain to arrange for the consecutive epoch. To alleviate the high price of transmitting the entire blockchain, the evaluators initiate a round of RapidChain's consensus with CoSi to create the State Block (SB).

The SB has the total reputation scores of the evaluators over the past ω epochs and the UTXO set of the users by this epoch. The UTXO set is the set having each UTXO, which is created in this epoch. As well, the UTXO that belongs to one public key is combined. For instance, when two UTXOs have the values of a and b, respectively, and both belong to the public key PK_x. The evaluators can decline the UTXO that has the value of b and adds the value onto the residual UTXO. The decline principle is to reserve the UTXO with the UTXO with a smaller address. Therefore, an evaluator only requires downloading the SB of the other shard rather than downloading each data in TB and RB.

e) Application of Dual-Chain Structure

As discussed before, the dual-chain structure is used to improve security level, throughput, and incentive strategy. This dual-chain structure comprises the transaction chain and the reputation chain. The SeHealthChain utilizes Raft consensus to create a transaction chain at a fast speed. To avoid Byzantine fault, the mixture of CoSi and RapidChain's consensus is utilized to verify the RB and its appropriate TBs. The attack on TxL, $TxDcS$, and TB can be identified early by the warning. The Byzantine attack can be recognized through the modified RapidChain consensus. So, the TB can only achieve the consensus when the RB that is connected to it has achieved consensus. After the view-shift occurs, each evaluator can decline the unacceptable RB and the connected TBs.

IV. SIMULATION RESULTS

This section describes the simulation setup of the SeHealthChain and illustrates its performance compared with the existing systems: eHealthChain [14], ZyConChain [17], MedShard [22], and BrokerChain [24]. The performance is measured using Python code, which simulates 1000 nodes for the SeHealthChain system in terms of throughput and user-perceived latency. The experiments are carried out on a system equipped with an Intel ® Core TM i5-4210 CPU @ 2.80GHz, 4GB RAM, and a 1TB hard disk running Windows 10 64-bit. The parameters utilized in this simulation are presented in Table 1. In addition to these parameters, the sliding window ω is set to 10, and the size of the TB is 4MB. Not greater than a 1/3 of the nodes are malicious nodes. The scaling factor $S(j)$ in reputation score computation is set to 0.1 for an accurate decision, 0 for unknown, -0.5 for an inaccurate decision, where the transaction must be approved but the evaluator decides to choose discard, and -1 for an inaccurate decision, where the transaction must be discarded but the evaluator wrongly chooses accept.

To simulate the transactions, all evaluators can produce transactions by themselves and transmit them to every other.

Table 1. Simulation Parameters for Blockchain System in Cloud-Fog Applications

Parameters	Global systems	Fog systems	Edge systems	IoT tools
Upstream bandwidth (Mbps)	162	80	33	12.5
Downstream bandwidth (Mbps)	85	37.5	18	6

Storage abilities/RAM (GB)	16	8	4	1
Processing abilities/CPU (Million Instructions Per Second (MIPS))	14100-20500	8200-12600	4030-8050	500-1600
Transmission delay (ms)	145	45	5	1
Blockchain instructions (M)	20	11	5	-
Blockchain processing power (Watts)	20-80	12-40	1.4-20	-

4.1 Throughput

The throughput is defined as the number of transactions the system processes in one second.

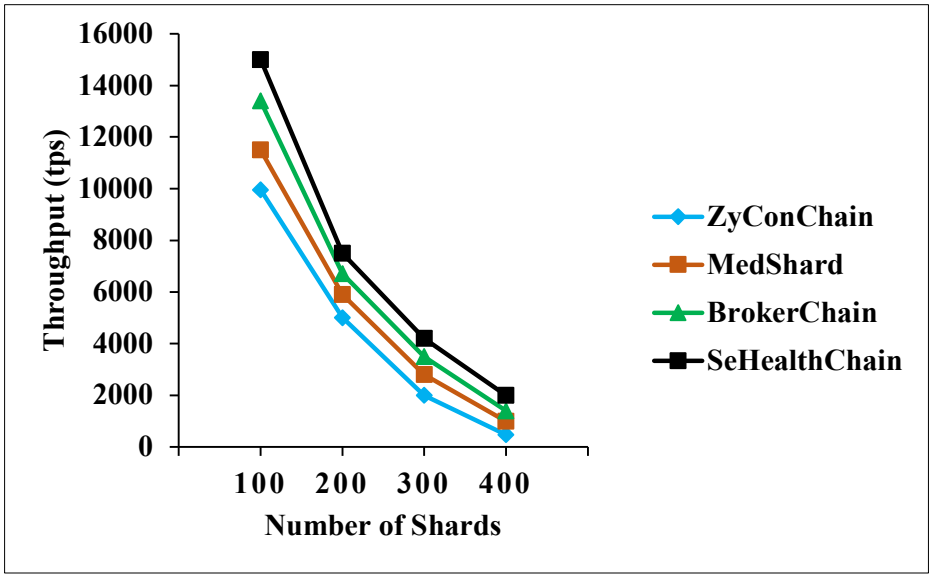

Figure 4. Mean Throughput of the Transactions vs. Number of Shards

The mean throughput results (in transactions per second (tps)) of the proposed and existing shard-blockchain systems for different shard sizes are plotted in Figure 4. The mean throughput for the shard sizes 100, 200, 300, and 400 is 15000 tps, 7500 tps, 4200 tps, and 2000 tps, respectively. On average, the mean throughput of the SeHealthChain while increasing the shard size is enhanced by 64.6%, 35.4%, and 14.8% compared to the ZyConChain, MedShard, and BrokerChain, respectively.

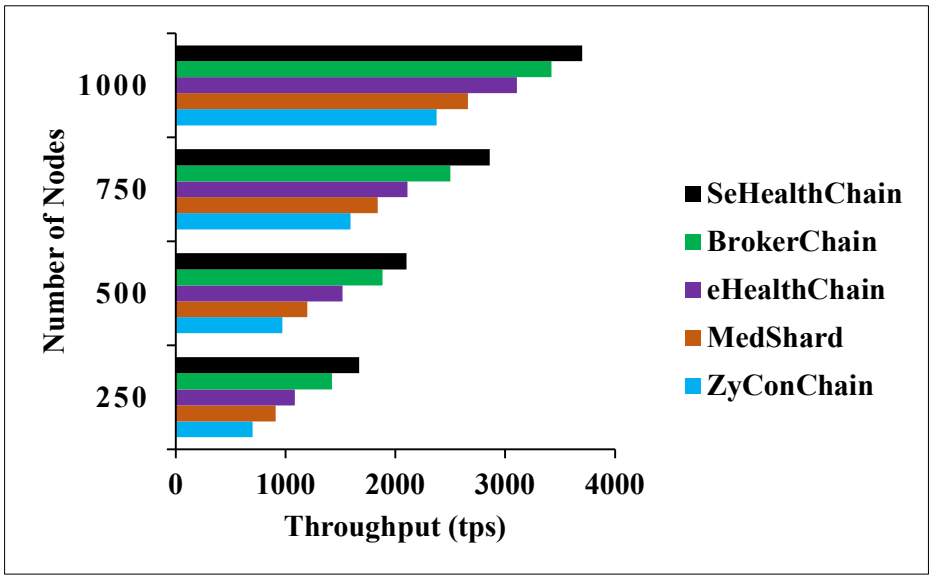

Figure 5. Mean Throughput of the Transactions vs. Number of Nodes

The mean throughput results of the proposed and existing shard-blockchain systems for a varying number of nodes are presented in Figure 5. The mean throughput for the number of nodes 250, 500, 750, and 1000 with the shard size of 200 is 1670 tps, 2100 tps, 2860 tps, and 3700 tps, respectively. On average, the mean throughput of the SeHealthChain system while increasing the number of nodes is increased by 83.4%, 56.3%, 32.1%, and 12%, in contrast with the ZyConChain, MedShard, eHealthChain, and BrokerChain, respectively.

4.2 User-perceived Latency

It is the period that a user transmits Tx to the network until the period that Tx can be verified by any (authentic) node in the system. Figure 6 shows the mean user-perceived latency results (in seconds (s)) of the proposed and existing shard-blockchain systems for different shard sizes. The mean user-perceived latency for the shard sizes 100, 200, 300, and 400 is 19.1s, 55.8s, 82.4s, and 140.6s, respectively.

On average, the mean user-perceived latency of the SeHealthChain while increasing the shard size is reduced by 37.4%, 27.4%, and 16%, in contrast with the ZyConChain, MedShard, and BrokerChain, respectively.

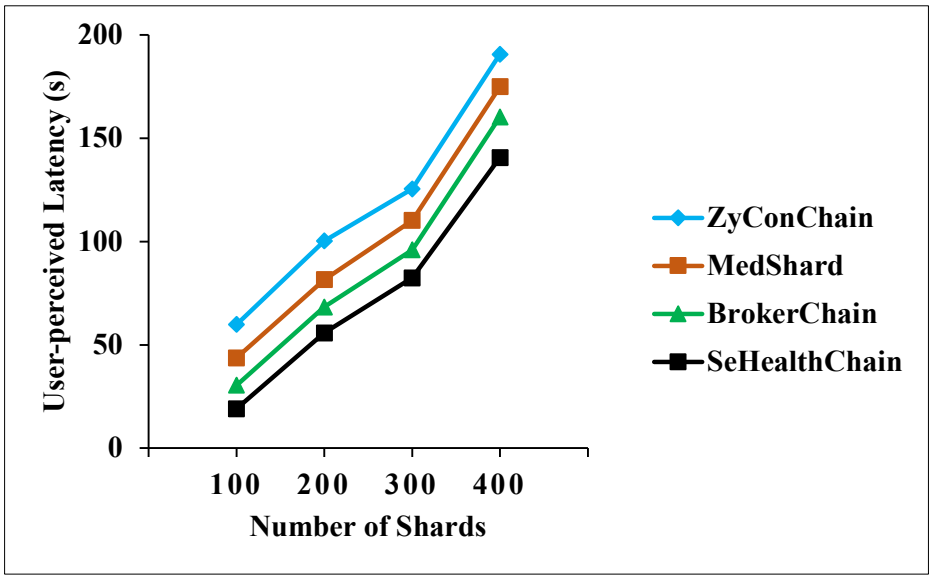

Figure 6. Mean User-perceived Latency of the Transactions vs. Number of Shards

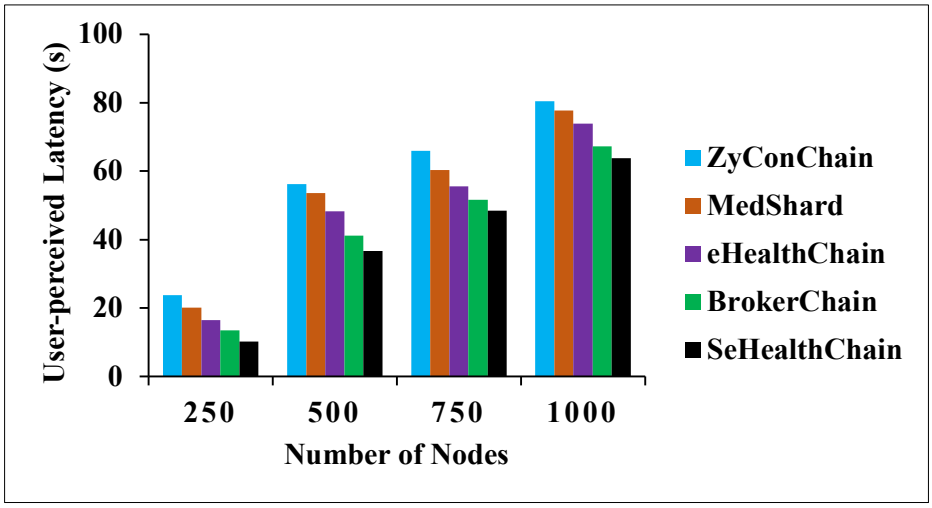

Figure 7. Mean User-perceived Latency of the Transactions vs. Number of Nodes

Figure 7 depicts the mean user-perceived latency of the proposed and existing shard-blockchain systems for a varying number of nodes. The mean user-perceived latency for the number of nodes 250, 500, 750, and 1000 with the shard size of 200 is 10.2s, 36.6s, 48.4s, and 63.8s, respectively. On average, the mean user-perceived latency of the SeHealthChain system while increasing the number of nodes is decreased by 29.7%, 24.9%, 18%, and 8.4% compared to the ZyConChain, MedShard, eHealthChain, and BrokerChain, respectively.

4.3 Security

To analyze the security, three distinct kinds of attack models such as a simple attack, camouflage attack, and observe-act attack with 1000 nodes and a shard size of 200 are tested. The initial two attacks are executed easily. But the performance under an observe-act attack depends on the invader's ability. So, the proposed system is tested with a highly robust invader that can observe the score distribution of each evaluator. In other words, the invaders and authentic evaluators cannot differentiate according to their reputation scores.

The throughput measured for the SeHealthChain under three distinct attack models is plotted in Figure 8. It is observed that the mean throughput for the simple attack is 6741.6 tps, each

invader can have a lower reputation after the 3rd epoch so that they cannot degrade the SeHealthChain efficiency. The mean throughput for the camouflage attack is 6122 tps. After a malicious node becomes leader, the view-shift protocol can be initiated and its reputation score can be removed. So, the malicious node requires a long period (about 10 epochs) to achieve an acceptable reputation score and be chosen as a leader again.

According to this, the SeHealthChain system can run better. The mean throughput for the observe-act attack is 5760 tps, which is worse than the other attack models. These invaders can reduce the SeHealthChain efficiency. In other words, the probability of a malicious leader is about 1/3 in all epochs.

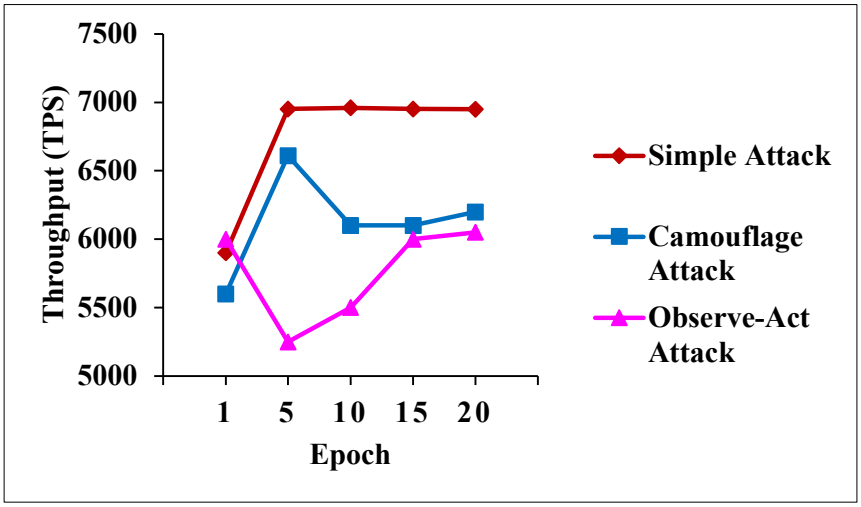

Figure 8. Mean Throughput for Different Attack Models vs. Number of Epochs

From these analyses, it is realized that the SeHealthChain system can increase the performance and security level while encountering a simple attack and a camouflage attack, as well as the same level as the random sharding scheme under an observe-act attack.

V. CONCLUSION

This paper developed the reputation-based secure, rapid, and high incentive blockchain system by sharding called SeHealthChain. This system leveraged reputation scores that characterize the heterogeneity among evaluators and provided the incentive strategy. This system applied a dual-chain structure for sharding blockchain schemes. A modified Raft-based synchronous consensus scheme was adopted for the transaction chain that achieved high throughput. An aggregate signature was applied for the reputation chain that establishes the consensus on the reputation score. Besides, the system throughput and security level for COVID-19 image sharing in medical applications were increased by the reputation-based sharding and leader selection scheme. At last, the simulation proved that the SeHealthChain system has a greater throughput and security level, in contrast with the existing shard blockchain systems in clinical applications.

REFERENCES

[1] Dedeoglu, V., Jurdak, R., Dorri, A., Lunardi, R. C., Michelin, R. A., Zorzo, A. F., & Kanhere, S. S. (2020). Blockchain technologies for IoT. *Advanced Applications of Blockchain Technology*, 55-89.

[2] Uddin, M. A., Stranieri, A., Gondal, I., & Balasubramanian, V. (2021). A survey on the adoption of blockchain in IoT: challenges and solutions. *Blockchain: Research and Applications*, *2*(2), 1-49.

[3] Marbouh, D., Abbasi, T., Maasmi, F., Omar, I. A., Debe, M. S., Salah, K., ... & Ellahham, S. (2020). Blockchain for COVID-19: review, opportunities, and a trusted tracking system. *Arabian journal for science and engineering*, *45*, 9895-9911.

[4] Ratta, P., Kaur, A., Sharma, S., Shabaz, M., & Dhiman, G. (2021). Application of blockchain and internet of things in healthcare and medical sector: applications, challenges, and future perspectives. *Journal of Food Quality*, *2021*, 1-20.

[5] Majeed, U., Khan, L. U., Yaqoob, I., Kazmi, S. A., Salah, K., & Hong, C. S. (2021). Blockchain for IoT-based smart cities: recent advances, requirements, and future challenges. *Journal of Network and Computer Applications*, *181*, 1-33.

[6] Taloba, A. I., Elhadad, A., Rayan, A., Abd El-Aziz, R. M., Salem, M., Alzahrani, A. A., ... & Park, C. (2023). A blockchain-based hybrid platform for multimedia data processing in IoT-Healthcare. *Alexandria Engineering Journal*, *65*, 263-274.

[7] Xie, Y., Zhang, J., Wang, H., Liu, P., Liu, S., Huo, T., ... & Ye, Z. (2021). Applications of blockchain in the medical field: narrative review. *Journal of Medical Internet Research*, *23*(10), 1-18.

[8] Yaqoob, I., Salah, K., Jayaraman, R., & Al-Hammadi, Y. (2021). Blockchain for healthcare data management: opportunities, challenges, and future recommendations. *Neural Computing and Applications*, 1-16.

[9] Chelladurai, U., & Pandian, S. (2022). A novel blockchain based electronic health record automation system for healthcare. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.

[10] Tagde, P., Tagde, S., Bhattacharya, T., Tagde, P., Chopra, H., Akter, R., ... & Rahman, M. H. (2021). Blockchain and artificial intelligence technology in e-Health. *Environmental Science and Pollution Research*, *28*, 52810-52831.

[11] Chen, C. M., Deng, X., Kumar, S., Kumari, S., & Islam, S. H. (2021). Blockchain-based medical data sharing schedule guaranteeing security of individual entities. *Journal of Ambient Intelligence and Humanized Computing*, 1-10.

[12] Jyoti, A., & Chauhan, R. K. (2022). A blockchain and smart contract-based data provenance collection and storing in cloud environment. *Wireless Networks*, *28*(4), 1541-1562.

[13] Pawar, P., Edoh, T. O., Singh, M., & Parolia, N. (2021). Hitching medical IoT devices to blockchain for personal health information management. *Blockchain Technology for IoT Applications*, 191-205.

[14] Pawar, P., Parolia, N., Shinde, S., Edoh, T. O., & Singh, M. (2022). eHealthChain—a blockchain-based personal health information management system. *Annals of Telecommunications*, 1-13.

[15] Sharma, A., Kaur, S., & Singh, M. (2021). A comprehensive review on blockchain and internet of things in healthcare. *Transactions on Emerging Telecommunications Technologies*, *32*(10), 1-53.

[16] Kim, S. (2019). Two-phase cooperative bargaining game approach for shard-based blockchain consensus scheme. *IEEE Access*, *7*, 127772-127780.

[17] Sohrabi, N., & Tari, Z. (2020). ZyConChain: a scalable blockchain for general applications. *IEEE Access*, *8*, 158893-158910.

[18] Mizrahi, A., & Rottenstreich, O. (2020). Blockchain state sharding with space-aware representations. *IEEE Transactions on Network and Service Management*, *18*(2), 1571-1583.

[19] Hong, Z., Guo, S., Li, P., & Chen, W. (2021). Pyramid: a layered sharding blockchain system. In *IEEE Conference on Computer Communications*, pp. 1-10.

[20] Amiri, M. J., Agrawal, D., & El Abbadi, A. (2021). Sharper: sharding permissioned blockchains over network clusters. In *Proceedings of the International Conference on Management of Data*, pp. 76-88.

[21] Hellings, J., & Sadoghi, M. (2021). Byshard: sharding in a byzantine environment. *Proceedings of the VLDB Endowment*, *14*(11), 2230-2243.

[22] Hashim, F., Shuaib, K., & Sallabi, F. (2021). Medshard: Electronic health record sharing using blockchain sharding. *Sustainability*, *13*(11), 1-21.

[23] Zheng, P., Xu, Q., Zheng, Z., Zhou, Z., Yan, Y., & Zhang, H. (2022). Meepo: multiple execution environments per organization in sharded consortium blockchain. *IEEE Journal on Selected Areas in Communications*, *40*(12), 3562-3574.

[24] Huang, H., Peng, X., Zhan, J., Zhang, S., Lin, Y., Zheng, Z., & Guo, S. (2022). Brokerchain: a cross-shard blockchain protocol for account/balance-based state sharding. In *IEEE Conference on Computer Communications*, pp. 1968-1977.

[25] Zheng, P., Xu, Q., Luo, X., Zheng, Z., Zheng, W., Chen, X., ... & Zhang, H. (2022). Aeolus: distributed execution of permissioned blockchain transactions via state sharding. *IEEE Transactions on Industrial Informatics*, *18*(12), 9227-9238.

[26] Zamani, M., Movahedi, M., & Raykova, M. (2018). Rapidchain: scaling blockchain via full sharding. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pp. 931-948.

[27] Kogias, E. K., Jovanovic, P., Gailly, N., Khoffi, I., Gasser, L., & Ford, B. (2016). Enhancing bitcoin security and performance with strong consistency via collective signing. In *25th USENIX Security Symposium*, pp. 279-296.