Practical Privacy Measures in Blockchains

Omar S. Saleh, Osman Ghazali and Norbik Bashah Idris

Abstract The Blockchain Technology has recently been a hot topic proposed in many industries such as Financial, Healthcare, Business, E-Government, Education, etc. The Blockchain can simply be defined as a distributed database or public ledger that contains records of all digital transactions/events that have transpired amongst the parties involved. The technology itself is comprised of other more fundamental knowledge namely: cryptography, distributed system, network and game theory. Thus at the more basic level, the blockchain components include functions such as hash, asymmetric cryptography, digital signatures, peer-to-peer network protocols and some elements of a "proof of correctness/work" resulting from a game-like setup. Against a backdrop of such a mixture of functions, "privacy" has emerged to be one of the new challenges in any Blockchain implementation. This research aims to investigate the techniques that can be used to successfully manage privacy in the blockchains. The work has identified the requirements and analyzed the techniques that can be used. Finally, the work was also extended to an analysis on the performance evaluation of blockchains in managing privacy albeit focusing on a specific blockchain—the Hyperledger fabric platform.

Keywords DLT (Distributed Ledger Technology) · Blockchain · Cryptography · Hash · Privacy · Zero-knowledge proofs · Peer-to-peer · Hyperledger

O. S. Saleh (✉)
Studies, Planning and Follow-Up Directorate, Ministry of Higher Education and Scientific Research, Baghdad, Iraq
e-mail: omar_saad@ahsgs.uum.edu.my

O. S. Saleh · O. Ghazali
School of Computing, University Utara Malaysia, Kedah, Malaysia
e-mail: osman@uum.edu.my

N. B. Idris
Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia
e-mail: norbik@iium.edu.my

M. A. Khan et al. (eds.), *Decentralised Internet of Things*, Studies in Big Data 71,
https://doi.org/10.1007/978-3-030-38677-1_2

1 Introduction

A Blockchain has been defined by [1] as a distributed database or public ledger that contains records of all digital transactions/events that have transpired amongst the parties involved. Each transaction is verified through agreement amongst the majority of parties in the system. Once a record is made, the information cannot be erased. Thus, the blockchain serves as an irrefutable record of every single transaction that has been made, thereby allowing the participating parties to know for sure that a digital event has occurred. Each transaction is contained in a block with several blocks being linked to each other linearly and chronologically in the form of a chain.

Authors [2] define the blockchain as continuous and unchangeable chains of data whereby different transactions get stored in the form of timestamped blocks. Blockchain is also known as distributed ledger technology [3]. Each copy of the Blockchain software (node) is able to store the complete copy of the ledger, write new entities to its ledger upon a given consensus among the connected nodes, broadcast transactions and regularly check its copy of ledger if its identical to the ledgers across most connected nodes [1]. The ledger is a combination of connected blocks. The block contains the transactions that took place and is chained with the previous block forming a chain hence the name blockchain. The transactions are compressed and anchored in the block using a Merkle tree model. The header of each block in the chain includes the hash of its content and the hash of all information in the previous block. In the blockchain, cryptography technology takes a significant place. It ensures the confidentiality of user data and transactions to ensure data consistency and provide all possible security. Privacy is a challenge in the blockchain due to the fact of public nature of the network. The transactions of blockchain are public; hence, it is possible to trace and extract the physical identities of the users by data mining. Privacy threats would arise from the transactions and network environment. Hence, this research describes the infrastructure of the blockchain, characteristics of the blockchain, design principles of blockchain, and the working process of the blockchain. This work also analyzes the privacy problems that blockchain still has and introduces the existing measures to these problems.

2 Blockchain Architecture

Blockchain consists of five layers, and each layer involves specific components. These layers are data layer, consensus layer, contract layer, network layer, and application layer.

1. Data layer involves several components such as data block, chain structure, timestamp, hash function, Merkle tree, and digital signature.
2. Consensus layer involves the consensus mechanisms which help the nodes to reach consensus [4].

3. The Contract layer mainly includes the smart contract and other codes which are the control logic of the decentralized application. The smart contract is a computer code embedded into the blockchain, and it comprises a set of rules [5].
4. Network layer involves the data transmission protocols and verification mechanisms [6]. A flat topology is the way of nodes connected in the blockchain, which means there is no trusted node or central node.
5. The application layer involves the applications. The typical applications of blockchain are Bitcoin, Ethereum, and Hyperledger [7]. Figure 1 shows the architecture of the blockchain and Fig. 2 shows the architecture of the blockchain in Bitcoin, Ethereum, and Hyperledger.

Bitcoin, Ethereum, and Hyperledger are the most three dominant blockchains. Many commonalities are in the overall architecture, but they are different in the

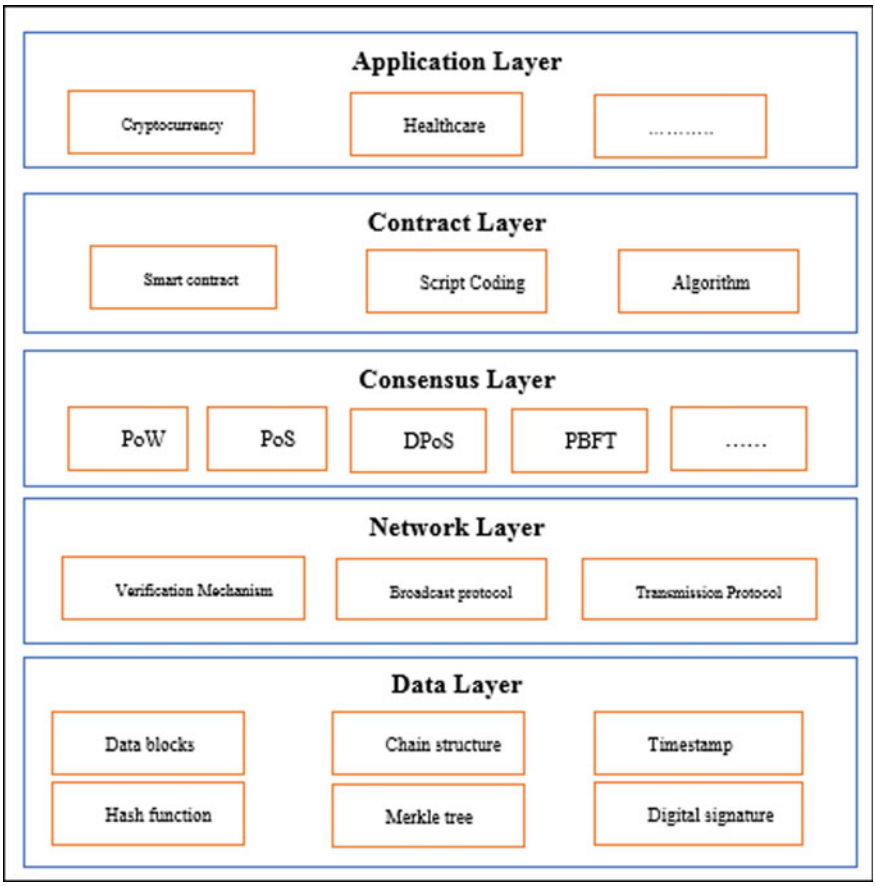

Fig. 1 Blockchain architecture

Layers	Bitcoin	Ethereum	Hyperledger
Application	Bitcoin Trading	Ethereum trading	Enterprise Applications
Contract	Script	Solidity/Script EVM	Go/Java Docket
Consensus	Pow	PoW	PBFT/SBFT
Network	TCP-based P2P	TCP-based P2P	HTTP/2-based P2P
Data	Merkle tree	Merkle Patricia Tree	Merkel Bockt tree

Fig. 2 Blockchain architecture among Bitcoin, Ethereum and Hyperledger

implementation, including data structure, consensus mechanism, smart contract, network, and application. The data structure adopted in Bitcoin is Merkle tree and in Ethereum is Merkle Patricia Tree, and Hyperledger is Merkle bockt tree. Proof of Work (PoW) is the consensus mechanism used in Bitcoin. Proof of Work (PoW) and proof of stake (PoS) are consensus mechanisms used in Ethereum. Practical Byzantine fault tolerance (PBFT) and Speculative Byzantine fault tolerance (SBFT) are consensus mechanisms used in Hyperledger [8–10]. A TCP protocol is used in both Bitcoin and Ethereum while HTTP/2 protocol is used in Hyperledger (Fig. 3).

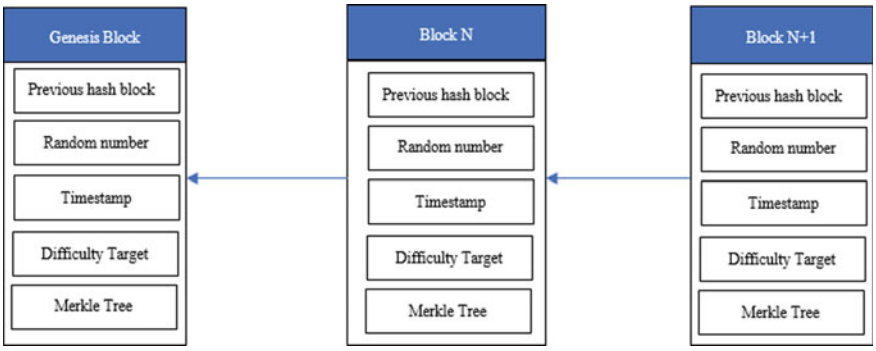

Fig. 3 Blockchain structure

2.1 Blockchain Structure

Blockchain is a chain of connected blocks. Each block is the collection of valid transactions. Any node in the blockchain-based system can start a transaction and broadcast to all nodes present in the network. Nodes in the network validate the transaction using the old transactions and then block added to the existing blockchain. Each block in the blockchain has two components which are block header and list of transactions [11, 12]. Block header contains Metadata of the block. Every block in the blockchain inherits from the previous block. The block is constructed using mining statistics. This measure has to be enough complicated to make them tamper-proof and based on the following formula:

$$H_k = \text{Hash}(H_{k-1}|T| \,|\text{noncel})$$

where T and nonce can be obtained by solving the consensus mechanism. The hash of the current block can be calculated by the Hash of previous block Hash value, transaction root Hash value. Block is also containing the Merkle Tree Root.

2.2 Blockchain Working Process

The Blockchain working process is simply described by four steps [13], as indicated below:

1. The sending node records new data and generates the necessary hash and broadcast that to the network;
2. The receiving node checks the message's hashes and the content if the message is correct, then it will be stored to its block; this process is generally done through what is called a "proof effort", e.g., proof of work (PoW) or proof of stake (PoS) or other model depending on the type of blockchain in use. However, the dominant ones are PoW and PoS [14].

When the majority of the nodes store the block, and it builds on it and moves to the next one. Blockchain working process is shown in Fig. 4.

With the PKI being at the forefront of the blockchain's architecture; the consensus is what makes it all work together. The consensus is what allows the different nodes to agree to the policies in the network. In the case of Bitcoin and many other blockchain technologies like Ethereum (as to the date of writing), the consensus is called the Proof of Work; it is worth noting here that there are many types of consensuses such as Proof of Stake [15].

Proof of Work (PoW)—as highlighted is the software algorithm that maintains both the safety as well as the transparency on the blockchain. In the case of Bitcoin, It uses SHA—256 hash functions to operate. Miner nodes operating on the blockchain consider 10 min worth of bitcoin-based. Blockchain activity and then encode those

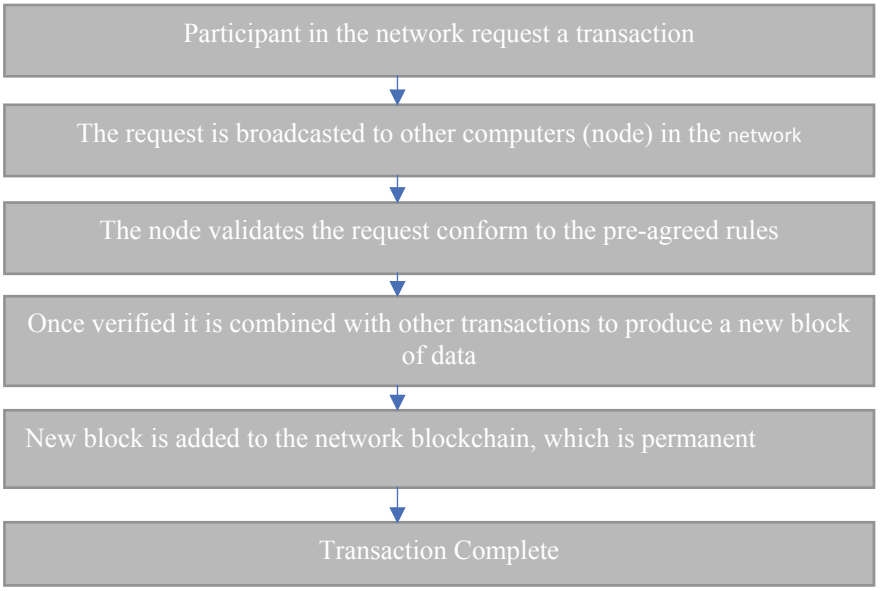

Fig. 4 Blockchain structure

transactions in the form of block. Mathematical solutions are then applied to create a hash value for that block. By doing this, the mining node has not only recorded all the transactions occurring during that time but has also indicated to the more extensive Blockchain network that the work required to hash the transactions has been performed. In return, the Blockchain network provides the mining node with a fixed amount of bitcoins [16].

Making the 'hash' is a crucial process in the proof of work. Here, the researchers [17] said that a proof of work incorporates the same operational principles as a Completely Automated Public Turing test to tell Computers and Humans Apart or CAPTCHA where any potential users have to go through and pass a test in order to access services. The proof of work leverages on the computational power of the Blockchain network, preventing any attempt to interfere or tamper with the blocks.

Proof of Stake (PoS)—This model came to life to tackle the drawbacks of PoW. As we have seen with the PoW the more hashing power you have the likely you will be rewarded; This made for what is called in the crypto community a "whale" where an entity owns most of the power of coins. The entity with the most potent machine will get to control the network. That is why PoS was proposed. In the simplest form PoS is when the users in the Blockchain network buy tokens (the asset-specific to the system like bitcoins in the case of the Bitcoin network) that permit them to conduct transactions involving decision making events and the future of the Blockchain network [18]. Users use the tokens to indicate that they have sufficient balances to participate. The authors [19] said that this is more environmentally friendly than PoW functions as they do not require as much computational power as mining operations.

Hence the fees for transacting using PoS are lower. However, the PoS mechanism has been criticized as being excessively centralized. While [20] pointed out that PoS permits broader user participation amongst users, is less susceptible to centralization.

2.3 Blockchain Types

Blockchain technologies divided into three types: (1) Public Blockchain(permissionless); (2) Consortium Blockchains and (3) private Blockchain(permissioned) [13, 14].

2.3.1 Public Blockchain

A Blockchain was designed to securely cut out the middleman in any exchange of asset scenario. It does this by setting up a block of peer-to-peer transactions. Each transaction is verified and synced with every node affiliated with the blockchain before it is written to the system. Until this has occurred, the next transaction cannot move forward. Anyone with a computer and internet connection can set up as a node that can sync with the entire blockchain history. Although this redundancy has it's advantages such as making public blockchains extremely secure; it is also a contributing reason to making then slow and somewhat wasteful [21]. A public blockchain has several benefits, such as:

1. Every transaction is public, and users can maintain individual anonymity;
2. Provides decentralization and becomes an excellent advantage for situations where a network needs to be decentralized;
3. Full transparency of the ledger; and
4. Faster, secure, and less expensive than the accounting systems and methods used today in the finance industry, however, the costs are higher, and speeds are slower than on a private chain. This also means that Public Blockchains allow any person with an internet connection to participate in the verification of transactions process and set themselves as a node [22].

2.3.2 Private Blockchain

Private blockchain lets the middleman back in, to a certain extent. The user writes and verifies each transaction allowing for greater efficiencies, and significant speed on private blockchains [14]. However, here is the main argument.

The company can choose who has read access to their blockchain's transactions, allowing for greater privacy than a public blockchain. A private blockchain is a better fit for more traditional business. This means that the private chains allow one party to have full control, and they will select a few nodes that are predetermined.

2.3.3 Consortium Blockchain

A Consortium blockchain is partially private. Consortium blockchain platforms have many of the same advantages as a private blockchain but operate under the leadership of a group instead of a single entity. This platform would be great for organizational collaboration. A consortium chains provide many of the same benefits of the private blockchain (efficiency and transaction privacy, etc.), without consolidating power with only one party; Hyperledger is an example [13]. In summary, every Blockchain network has different rules regarding what kind of assets it trades, and under which conditions trading takes place. Those rules encoded into its software called the consensus. The node in the Blockchain network is every device running the Blockchain software and connected to the network [1, 23].

As mentioned earlier, blockchain classified into three options which are public blockchain, private blockchain, and consortium blockchain.

2.4 Blockchain Characteristics

The blockchain technology made through different existing technologies such as cryptography, mathematics, algorithms, and distributed consensus algorithms [22, 24, 25]. As such, the Blockchain has six key characteristics [13].

1. *Decentralization*. It means that the Blockchain does not have to rely on a single centralized node which functions as a master node. Each node can record, store, and update the ledger. Together they form the blockchain community of peer-to-peer nodes.
2. *Transparency.* The block's data recorded by each node and distributed among other connected nodes are visible to each node, thus creating openness among connected nodes.
3. *Open Source*. Most Blockchain systems are open to anyone, allowing anybody to modify the code and technology in ways that best suit their needs. However, this does not mean that anyone can edit a running blockchain solution. Making any modification to a working solution means connected nodes agree to accept the change, and it is only valid when connected nodes adopt the change.
4. *Autonomy*. As there are connected nodes, any changes happen once the majority of nodes accept the change. It enforces good deeds from different nodes making changes or intervention useless; other nodes will easily detect any attempt at making any change.
5. *Immutable*. The records will be preserved forever, and cannot be changed unless someone can take control of more than 51% nodes at the same time (i.e., a simple majority).
6. *Anonymity*. Data is hashed and shared; being hashed makes transactions somewhat anonymous.

3 Privacy Requirements for Blockchain

In the context of Blockchain Technology, privacy and confidentiality mean that the data written to the blockchain and the identities of the parties involved are protected [26]. Privacy and confidentiality in the blockchain still a challenge and open issue [27, 28]. This claim also mentioned by several authors such as [19, 29, 30]. Privacy was pointed out as a problem in the original paper of Bitcoin conducted by [22] as there were no ways to protect the privacy of users. With a public blockchain, the information stored in public ledger. And the transaction contains various information such as the ID of the previous transaction, timestamp, participants address, trade values, and signature of its sender [31]. Hence, there is a possibility of tracing the transaction to extract the users' physical identities or other additional information by data mining [16]. Privacy in the blockchain divided into two types which are (1) privacy of information and (2) privacy of the party [19]. Privacy of information is related to the content of the message posted to the blockchain. The party may wish to hide the content of the message from network members. While the privacy of the party is related to the identity of the party, who will be involved in the transaction in the Blockchain [19]. The privacy requirements in the blockchain are studied by several researchers [11, 19, 32, 33]. Hence, to protect privacy, the following requirements should be considered.

The content of the transactions should be only known to their partakers;

- Transaction details are not visible to unauthorized third parties and the world at large unless one of the counterparties has chosen to reveal that information; and
- Transaction details cannot be collected, analyzed, or matched with "off-Blockchain" metadata to reveal any information about counterparties or transaction details. By this, our definition encompasses the use of graph analysis, pattern matching, and machine learning to construct a profile of a counterparty based on the activities associated in the ledger [27, 28, 34]. The blockchain needs to satisfy several requirements to protect privacy [35], and as follows:

    1. The links between transactions should not be visible or discoverable.
    2. The content of transactions is only known to their partakers [16].
    3. The private or permissioned blockchain could set an access control policy. It gives complete transparency of the blockchain data is not a problem.

The privacy requirements should be considered on two factors [16, 36, 37] and as follows:

1. Identity Privacy: which means intractability between the transaction scripts and the real identities of their partakers, as well as the transactional relationships between users. Even if users apply random addresses (or pseudonyms) when acting in the blockchain, they can only provide limited identity privacy.
2. Transaction Privacy: it means that specified users can only access the transaction contents. Transaction privacy is the primary concern in the public blockchain. In the next section, the security and privacy issues in the blockchain explored.

Various measures and techniques that can be used to achieve the security and the privacy at each layer of blockchain investigated as well.

4 Security and Privacy Issues in Blockchain

Since all transactions in public blockchain are visible and open in the network, so the blockchain is mainly vulnerable to leakage of transactional privacy [11]. The critical evaluation parameter in any blockchain is how well the conditions of security and privacy meet the requirements of blockchain. Hence, analyzing the security and privacy issues of blockchain become a valuable research area. Security is defined based on three main components which are confidentiality, integrity, and availability. In blockchain context, privacy means limits the access to the information through a set of rules. Integrity means the information is accurate and trustworthy, and availability means that the information is grantees to be accessed by authorized people. Privacy is defined based on two components data privacy and user privacy [38].

As mentioned earlier, blockchain architecture includes several layers. Thus identifying the challenges that occur in each layer would be very important to be taken into consideration. Encryption measure can be used to achieve confidentiality in three layers which are a smart contract, network, and data layer [38]. Two measures can be used to achieve integrity, which is the Message Authentication Code (MAC) and Signature Scheme. MAC used for achieving the integrity in three layers of blockchain, which are a smart contract, network, and data layer.

Signature Scheme can be used to achieve integrity in both transaction and consensus layers [38]. Availability made by various measures such as consensus, access control, and protocols. Data privacy-preserving computation measure adopted for achieving the data privacy and user privacy at the smart contract layer.

Access control measure used for achieving the data privacy, while blind signature and ring signature used for achieving the anonymity (user privacy) at the consensus layer. Zero-Knowledge Proofs and Mixing measures used for achieving both data privacy and user privacy at the transaction layer. Access control measure used for achieving the data privacy at the data layer, while IP Anonymity measure used for achieving the privacy of user at the network layer [38]. Table 1 summarizes the cryptographic measures that which adopted for achieving security and privacy of information subjected to the blockchain layers.

5 Privacy and Security Measures Used in Blockchain

In this section, we provide a detailed discussion on a selection of techniques that can be leveraged to enhance the security and privacy of existing and future blockchain systems.

Table 1 Cryptographic measures subject to each layer

Blockchain layer	Confidentiality	Integrity	Availability	Data privacy	User privacy
Smart contract	Encryption	MAC	–	Data privacy preserving computation	Identity privacy preserving computation
Consensus	–	Signature scheme	Consensus	Access control	Blind or ring signature
Transaction	–	Signature scheme	Access structure of transactions	Zero-knowledge proofs, mixing techniques	Zero-knowledge proofs
Network	Encryption	MAC	Protocols	–	IP anonymity
Data	Encryption	MAC	Access control	Access control	–

5.1 Mixing

Mixing measures was proposed by Chaum [39]. It aims to hide the identity of users as well as the content of the communication. The architecture of mixing service is clearly stated in Fig. 5.

Here the explanation of mixing by an example. Assume that we have two entities sender and receiver and a message M is prepared to be delivered at address R. The message will be encrypted with the receiver public key KR and appending the address R. Then the intermediary's public key KI is encrypted with the result and based on the following formula:

$$K_I(r_0,\ K_R(r_1,\ M),\ R) \rightarrow K_R\left(r_{1,}M\right),\ R$$

where r_0 and r_1 random numbers which ensure that no message is transferred more than once s we mentioned in the first place, Bitcoin's blockchain doesn't guarantee

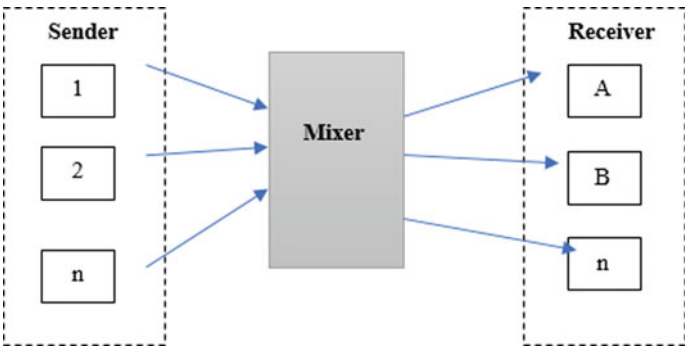

Fig. 5 Mixing service architecture

obscurity for users: transactions use onymous addresses and may be verified publicly, so anyone will relate a user's dealing with her alternative transactions by an easy analysis of addresses she utilized in creating bitcoin exchanges. Additional seriously, once the address of the dealing is coupled to the real-world identity of a user, it's going to cause the outflow of all her transactions. Thus, admixture services (or tumblers) was designed to forestall users' addresses from being coupled. Mixing, literally, it's a random exchange of user's coins with alternative users' coins, as a result, for the observer, their possession of coins is obfuscated. However, these mixing services don't offer protection from coin thieving [40].

Mixcoin was planned by Bonneau et al. in 2014, that provides anonymous payment in Bitcoin and bitcoin-like cryptocurrencies. To defend against passive adversaries, Mixcoin extends the namelessness set to permit all users to combine coins at the same time. To defend against active adversaries, Mixcoin provides namelessness the same as ancient communication mixes. Additionally, Mixcoin uses associate degree responsibleness mechanism to observed stealing, and it shows that users can use Mixcoin rationally while not stealing bitcoins by orienting incentives [41].

5.2 Anonymous Signatures

This section will dive in discussing the two most important and typical anonymous signature schemes which are group signature and ring signature.

Group signature is a cryptography theme which firstly proposed by [42]. Given a bunch, any of its members will sign a message for the whole cluster anonymously by exploitation her personal secret key, and any member with the cluster's public key will check and validate the generated signature and ensure that the signature of some group member is employed to sign the message. The method of signature verification reveals nothing regarding verity identity of the signer except the members of the cluster. Cluster signature encompasses a group manager who manages adding group members, handling the event of disputes, together with revealing the first signer. Within the blockchain system, we have a tendency to conjointly would like a licensed entity to form and revoke the cluster and dynamically add new members to the group and delete/revoke the membership of some participants from the group. Since the group signature needs a bunch manager to line up the group, the cluster signature is appropriate for syndicate blockchain.

Ring signature was proposed by [43] which shown in Fig. 6. It can succeed anonymous through linguistic communication by any member of cluster users. The term "ring signature" originates from the signature algorithmic program that uses the ring-like structure. The ring signature is anonymous if it's troublesome to work out that member of the cluster uses his/her key to sign the message. Ring signatures take issue from cluster signatures in 2 principal ways: 1st, during a ring signature theme, the $64,000 identity of the signer can not be discovered within the event of a dispute, since there's no cluster manager during a ring signature. Second, any users will group A "ring" by themselves while not further setup. Thus, the ring signature

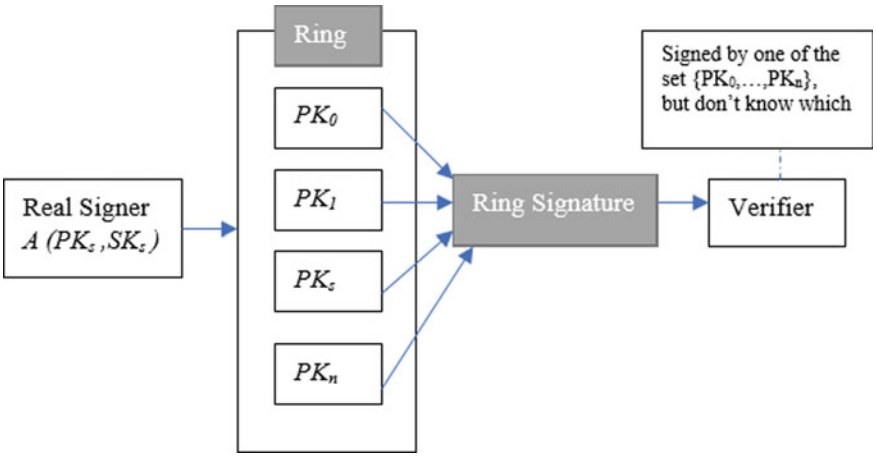

Fig. 6 Ring signature process

is applicable to the general public blockchain. One in every one of the standard applications of a hoop signature is CryptoNote [44]. It adopts ring signature to cover the association between the sender's addresses of transactions. Additional exactly, CryptoNote constructs the sender's public key with many alternative keys, so it is not possible to spot World Health Organization truly sent (signed) the dealing. Thanks to the utilization of ring signature, if the quantity of ring members is n, then the likelihood that AN opponent might with success guess a true sender of dealing is 1/n.

Here is the working process of ring signature measure [43]. User A selects a set of users ($User_0$, $User_1$, $User_n$) and creates a ring. Each user has a pubic key from standard signature scheme such as ECDSA. User A signs a message with his/her private key (SKs) and all the public keys (PK_0, PK_1, PK_n) of the members in the ring. The verifier knows that the message signed by one from the set but does know the real singer. Hence, the ring signature provides anonymity for the signer.

5.3 Homomorphic Encryption (HE)

Homomorphic cryptography (HE) is a powerful cryptography. It will perform sure kinds of computations directly on ciphertext and make sure that the operations performed on the encrypted knowledge, once decrypting the computed results, can generate identical results to those performed by constant operations on the plaintext. There are many parts of homomorphic cryptosystems [45, 46]. One will use homomorphic cryptography techniques to store knowledge over the blockchain with no vital changes within the blockchain properties. This ensures that the information on the blockchain is encrypted, addressing the privacy issues related to public

blockchains. The employment of homomorphic cryptography technique offers privacy protection and permits prepared access to encrypted knowledge over public blockchain for auditing and different functions, like managing worker expenses. Ethereum sensible contracts give homomorphic cryptography on knowledge hold on in blockchain for larger management and privacy.

The working process of HE will be explained by the following scenario.

A and *B* are two main parameters. *A* has secret values (x_1, x_2,..., x_n) and *B* has a function $F(.)$ In order for A and B to calculate $F(x_1,\ldots,x_n)$ together without leaking the secret values.

$E(.)/D(.)$ are the set of the homomorphic encryption system. *A* can send encrypted inputs $E(x_1),\ldots,E(x_n)$*to B*. After that, a normal computation on the encrypted text should be performed and the result send to *A*. *A* will get $f(x_1,\ldots,x_n)$ after the decryption. The privacy of blockchain will be protected in the implementation of homomorphic cryptographic including both the Pedersen commitment scheme and Paillier cryptosystem [16].

5.4 Attribute-Based Encryption (ABE)

Attribute-based encryption (ABE) is a cryptographic method. The conception of attribute-based cryptography was planned in 2005 [47] with one authority. Since then, variety of extensions are planned to the baseline ABE, as well as ABE with multiple authorities to come up with users' personal keys together [48, 49]. ABE schemes that support impulsive predicates [50, 50]. Attribute-based cryptography is incredibly powerful nonetheless few applications thus far deploy it thanks to the shortage of understanding of each core ideas and economical implementation. ABE has not nonetheless been deployed in any type on a blockchain for the data processing thus far. In 2011, a localized ABE theme was planned [49] to use ABE on a blockchain. For instance, on a blockchain, permissions may well be delineated by the possession of access tokens. All nodes within the network, that have an exact token issued to them, are going to be granted access to the special rights and privileges related to the token. The token provides a method of following who has bound attributes Associate in Nursingd such tracking ought to be worn out an recursive and consistent fashion by the approved entity that distributes the token. Tokens are often viewed as badges that represent attributes or qualifications and may be used as non-transferable quantifiers of name or attributes.

5.5 Secure Multi-party Computation

The multi-party computation (MPC) model defines a multi-party protocol to permit them to hold out some computation together over their non-public knowledge inputs while not violating their input privacy, such that associate degree soul learns nothing

concerning the input of an authentic party however the output of the joint computation. The primary large-scale preparation of MPC was in 2008 for associate degree actual auction downside in Denmark [51]. In recent years, MPC has been utilized in blockchain systems to shield users' privacy [52]. Designed and enforced secure multiparty computation protocols on the Bitcoin system in 2014. They made protocols for secure multiparty lotteries with none sure authority. Their protocols are ready to guarantee fairness for honest users despite however dishonest ones behave. If a user violates or interferes with the protocol then she becomes a loser and her bitcoins are transported to the honest users. A decentralized SMP computation platform, referred to as Enigma, is projected in 2015 by [53]. By exploitation a sophisticated version of SMP computation, Enigma employs a verifiable secret sharing theme to ensure the privacy of its process model. Also, Enigma encodes shared secret knowledge employing a changed distributed hash table for economical storage. Moreover, it leverages associate degree external blockchain as a corruption-resistant recording of events and also the regulator of the peer-to-peer network for identity management and access management. Like the Bitcoin system, Enigma provides autonomous management and protection of private knowledge whereas eliminating the requirement and dependency of a sure third party.

As mentioned earlier, MPC is a cryptographic protocol that used for emulating a trusted party. MPC would be very benefit to be used in system with no trusted parties. MPC has two main goals which are the correctness and security [54]. Its working can be shown based on the following example.

Let say P_1, P_2,…,P_n are mutually suspicious. Each party has a secret input (x_1, x_2, x_n). Hence, the joint function y can be computed as follows:

$$y = f(x_1, x_2, \ldots, x_n)$$

The correctness goal will be achieved through that everyone can computes $y = f(x_1, x_2, \ldots, x_n)$ and the security goal will be achieved through that nothing but the output is revealed.

5.6 Non-interactive Zero-Knowledge (NIZK) Proof

Zero-knowledge proofs are another cryptographic technology that has powerful privacy-preserving which was proposed in the early 1980s [55, 56]. The concept of this protocol was proposed by Blum. The goal of NIZK is to preserve the privacy. It aims to provide the prove the correctness of data without leaking addition information. A zero-knowledge protocol can be explained by an example.

Let say we have two parties the prover *P* and the verifier *V*. The prover has a statement which he wants to prove to verifier. Hence, the working process of zero-knowledge proof will be based on the following steps:

1. The prover will compute a proof for the statement and then send the proof to the verifier.
2. The verifier will choose a question and send it to the prover.
3. The prover then calculates the answer for the question and send it to the verifier. Hence, the verifier by using the answer will check whether Prover really knows the statement. Figure shows an interactive zero-knowledge protocol between prover and verifier.

NIZK proof system can be represented according the following formulas [16].

(P, V) are the prover and verifier, the NIZK proof system for language L when $L \subseteq NP$ with k where k is a security parameter and if it meets two properties which are completeness and soundness.

The completeness occurs for any input $x \in L$ and its witness w and polynomial $p(.)$

$$P_r[V(R, x, P(R, x, w)) = 1] \geq 1 - 1/P(x)$$

The soundness occurs for any input x $\notin$ L and algorithms P^* and polynomial $p(.)$

$$P_r\left[V\left(R, x, P^*(R, x,)\right) = 1\right] < 1 - 1/P(x)$$

5.7 The Trusted Execution Environment (TEE) Based Smart Contracts

TEE is an execution atmosphere if it provides a totally isolated environment for application execution, that effectively prevents different software package applications and operative system(s) from meddling with and learning the state of the applying running in it. The Intel software package Guard eXtensions (SGX) maybe a representative technology to implement a TEE. As an example, Ekiden [18] may be an SGX primarily based resolution for confidentiality-preserving sensible contracts. Ekiden separates computation from the accord. It performs sensible contract computation in TEEs on calculating nodes off-chain, then uses a foreign attestation protocol to validate the execution correctness of calculating nodes on-chain.

Intel® SGX measure is a hardware-based solution which provides data protection. It is a platform with built-in CPU instructions that permit the access to the data. Access to the data will be denied or disabled if the code is altered or tampered.

5.8 Discussion

Three main points have to be taken into consideration in order to achieve security and privacy in the blockchain.

(1) No single technology may be a cure for the security and privacy of Blockchain. (2) There's no technology that has no defects or is ideal altogether aspects. Once we add new technology to a posh system, it perpetually causes alternative issues or new form(s) of attacks and (3) there's perpetually a trade-off between security-privacy and potency. We must always advocate those techniques that improve the protection and privacy of blockchain It is clearly stated in the previous sections that blockchain technology classified into three types: (1) public blockchain, Private blockchain and Consortium blockchain. Privacy and confidentiality are still a challenge in the blockchain. Hence, this research comes into address the privacy issues and analyzed the solutions that would be used in order to preserve and maintain the privacy. This research addresses several measures that can be used for preserving the privacy in blockchain and as shown in Table 2.

It is very important to maintain the privacy at all levels such as data level, transaction level and network level. Privacy can be maintained very well in permissioned blockchain because of that (1) running private blockchain is easier that public blockchain, (2) easy to change the rules and revert the transactions, (3) the validators are known so any risk of a 51% attack arising from some miner collusion does not apply, (4) transactions are cheaper and since they only need to be verified by a few nodes that can be trusted to have very high processing power, and do not need to be verified by ten thousand laptops, (5) Nodes can be trusted to be very well-connected and (6) private blockchain can provide a greater level of, well, privacy because the read permissions are restricted. Privacy in private blockchain can be maintained using different measures and it is based on the platform. Hyperledger fabric is a private permissioned blockchain. From all the above privacy techniques which mentioned earlier, Hyperledger uses attribute-based encryption of data in which can restrict data to a user based on user's attribute. Hyperledger fabric can also zero knowledge proof in which verifier can verify issuer without getting access to issuer data. As mentioned above that in Hyperledger fabric we can restrict access between user of same organizations we can create private data side DB in which only those user will have access which are linked to the transaction. Hyperledger fabric ledger consist of world state (Database) and Transaction log (Blockchain), there is public ledger for all permissioned participants but in public ledger there are only hashes not actual data. World state is maintained so reading data doesn't involve traversing the entire blockchain. Each peer can recreate the world state from the transaction log. Private data and Attribute Based Encryption together give enough flexibility to model a non-trivial business process without revealing confidential information. Hyperledger fabric adds certain layers of privacy for user's data, First, it gives access to only permissioned user of the network, then can secure user's actions using attribute based encryption. Moreover, it uses zero knowledge proof where authentication is required between user without giving data of one user to another. In order to maintain the privacy at

Table 2 Privacy measures in blockchain

Techniques	Application	Advantages	Disadvantages
Mixing	MixCoin	It is very helpful in preventing the users' addresses from being linked	There might risk of leakage of user privacy due to the centralized services
Group signature	JUCIX	The ability to hide the signer identity among a group of users	Trusted third party is needed to act as a manager
Ring signature	CryptoNote, Ethereum	The ability to hide the signer identity among a group of users and hence trusted the third party is not required	The signer identity can not be revealed in the event of a dispute
ABE	None	Data confidentiality and fine-grained access control can be achieved simultaneously	Need to resolve the issuance and revocation of attribute certificate in the distributed environment
HE	Ethereum	Privacy-preserving can be achieved by performing a computation on the ciphertext	The computational efficiency of the complex function is very low
SMPC	Engima	The ability to carry out some computation through multi-party without violating the input privacy	Efficiency is low in the complex functions
NIZK	Zcash	Users can prove their balance easily without reveling the account balance	Less efficient
TEE based solutions	Ekiden, Enigma	The privacy of smart contract can be protected by running them in TEE	Need to resolve the attacks on SGX

data level, end to end encryption can be used while preserving the privacy at network level could be through network level configurations as well as using private channels. Preserving the privacy at truncation level could be through implementing access control measures. In case of public blockchain, and based on the analyzing we made for the existing measures and techniques, different encryption schemes for both identity privacy and transaction privacy can be used and based on the use case. Ethereum adopts a couple of measures for maintaining the privacy which are Zero Knowledge-based and Mixers.

From what has been presented, it can conclude that the measures of privacy can be successfully employed based on the blockchain platform and the use case.

6 Performance Evaluation of Blockchains

Ethereum and Hyperledger are the most dominant blockchain platforms. Smart contract and Crypto-currency are an example of the application in Ethereum. The smart contract is executed in Ethereum Virtual Machine (EVM) using Solidity Language and Ethash (PoW) consensus algorithm. In Hyperledger, the smart contract is executed using Dockers and Golang and Java languages, and Practical Byzantine Fault Tolerance (PBFT) consensus algorithm [14]. Authors [57] conducted a performance analysis of Ethereum and Hyperledger.

Hyperledger Fabric consistently performs better than Ethereum both in term of throughput and latency. This research will dive in more details regarding the measures of the private security mechanism of Hyperledger Fabric. The privacy protection measures of Hyperledger Fabric will be divided into four measures:

1. Using symmetric cryptography and zero-knowledge proof. This is for several reasons such as separating the transaction data from on-chain records and protecting privacy from the underlying algorithm.
2. Using the digital certificate management service. This is for guarantees the legitimacy of the organization on the blockchain.
3. Using the design of multi-channel. This is for separating the information between different channels.
4. Privacy data collection. This is for satisfying the need for the isolation of privacy data between different organizations within the same channel.

The channel and privacy data collection are the most typical methods. The channel is dedicated to allowing the data on the channel to be isolated separately and to the blockchain privacy protection. The ledger is shared by the peer on the same channel and the recognition of the channel is needed to be obtained by the transaction peer before it can join the channel and transact with others. The private data collection (PDC) is a group of organizations that are permitted to store private data on a channel. The data stored contains the private data and the hash value of the private data [58].

In the Hyperledger Fabric, the processing of privacy data is divided into two scenarios: new channels are required once the whole dealings and ledger should be unbroken strictly confidential to the skin members of the channel; when the transaction info and ledger have to be compelled to be shared among some organizations, a number of them are going to be able to see all the dealings information, alternative organizations have to be compelled to recognize the prevalence of this dealings to verify the genuineness of the transaction, a non-public information assortment ought to be established during this case. Additionally, as a result of non-public information is propagated through peer-to-peer instead of block, the

privacy information assortment is employed once the dealings data should be confidential for the sorting service peer. The blockchain dealings method involving the privacy information assortment is as follows [59, 60].

Blockchain transaction process involving the privacy data collection based on several steps and as follows:

1. The offer request is submitted by the client application to call the chain code function to the endorsement peer of the private data set authorization and through the provisional domain, the private data is sent.
2. The transaction is simulated by the endorsement peer and the private data is stored in a local temporary repository in the peer. The gossip protocol is used by the endorsement peer in order to disseminate the private data to the authorized peer.
3. The public data is returned by the endorsement peer including the hash value of the private data key-value pair.
4. The transaction is submitted to the sorting service peer by the client application and then distributing the sorting result to each block.
5. The authorized peer can use the collection policy when submitting a block in order to determine if it is authorized to view private data.

7 Conclusion

In this research, we identified and discussed the cryptographic algorithms used in the blockchain. In addition to that, we identified the privacy requirements for the blockchain as well as the privacy techniques that can be used to manage the privacy in the blockchain. Our findings show that there are many techniques which can be used for enhancing the privacy in the blockchains. Finally, we discussed the privacy process in the Hyperledger platform and the measures that can be used for ensuring its privacy.

Acknowledgements Authors would like to sincerely thank Universiti Utara Malaysia (UUM), International Islamic University Malaysia (IIUM), Malaysia and Ministry of Higher Education, Iraq for supporting this research.

References

1. Grech, A. and Camilleri, A. F.: Blockchain in Education. In: Inamorato dos Santos, A. (ed.) EUR 28778 EN (2017). https://www.doi.org/10.2760/60649
2. Ackerman, A.,Chang, A., Diakun-Thibault, N., Forni, L., Landa, F., Mayo, J., van Riezen, R.: Blockchain and Health IT: Algorithms, Privacy and Data (August 8, 2016). Project PharmOrchard of MIT's Experimental Learning "MIT FinTech: Future Commerce.", White Paper August 2016. Available at SSRN: https://ssrn.com/abstract=3209023

3. Duan, Z, Mao, H., Chen, Z., Bai, X., Hu, K., Talpin, J.-P.: Formal modeling and verification of blockchain system, vol. 86, pp. 231–235 (2018)
4. Wu, J., Tran, N.K.: Application of blockchain technology in sustainable energy systems: an overview. Sustain **10**(9), 1–22 (2018)
5. Cui, G., Shi, K., Qin, Y., Liu, L., Qi, B., Li, B.: Application of block chain in multi-level demand response reliable mechanism. In: 2017 3rd International Conference on Information Management (ICIM), pp. 337–341 (2017)
6. Fukumitsu, M., Hasegawa, S., Iwazaki, J., Sakai, M., Takahashi, D.: A proposal of a secure P2P-type storage scheme by using the secret sharing and the blockchain. In: Proceedings of the International Conference on Advanced Information Networking and Applications (AINA), pp. 803–810 (2017)
7. Yuan, Y., Wang, F.Y.: Towards blockchain-based intelligent transportation systems. In: IEEE International Conference on Intelligent Transportation Systems (ITSC), pp. 2663–2668 (2016)
8. Zheng, Z., Xie, S., Dai, H.N., Wang, H.: Blockchain challenges and opportunities: a survey. Work Pap.–2016, December 2016
9. Baliga, A.: Understanding blockchain consensus models. Whitepaper, April, pp. 1–14 (2017)
10. Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H.: An overview of blockchain technology: architecture, consensus, and future trends. In: Proceeding of 2017 IEEE 6th International Congress on Big Data (BigData Congress), pp. 557–564 (2017)
11. Prashanth Joshi, A., Han, M., Wang, Y.: A survey on security and privacy issues of blockchain technology. Math. Found. Comput **1**(2), 121–147 (2018)
12. Le, T., Mutka, M.W.: Capchain: a privacy preserving access control framework based on blockchain for pervasive environments. In: Proceedings of 2018 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 57–64 (2018)
13. Lin, I.-C., Liao, T.-C.: A survey of blockchain security issues and challenges. Int. J. Netw. Secur. **1919**(55), 653–659 (2017)
14. Dinh, T.T.A., Wang, J., Chen, G., Liu, R., Ooi, B.C., Tan, K.-L.: Blockbench: a framework for analyzing private blockchains. In: Proceedings of the 2017 ACM International Conference on Management of Data. ACM (2017)
15. Fabian, B., Ermakova, T., Krah, J., Lando, E., Ahrary, N.: Adoption of security and privacy measures in bitcoin–stated and actual behavior (2018). Available at SSRN:https://ssrn.com/abstract=3184130
16. Feng, Q., He, D., Zeadally, S., Khan, M.K., Kumar, N.: A survey on privacy protection in blockchain system. J. Netw. Comput. Appl. **126**, 45–58 (2019)
17. Duan, B., Zhong, Y., Liu, D.: Education application of blockchain technology: learning outcome and meta-diploma. In: Proceedings of the International Conference on Parallel and Distributed Systems (ICPADS), December 2017, pp. 814–817 (2018)
18. Cheng, R., Zhang, F., Kos, J., He, W., Hynes, N., Johnson, N., ... & Song, D.: Ekiden: A platform for confidentiality-preserving, trustworthy, and performant smart contracts. In 2019 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 185–200. IEEE (2019, June)
19. Axon, L., Goldsmith, M., Creese, S.: Privacy requirements in cybersecurity applications of blockchain, vol. 111, 1st edn. Elsevier (2018)
20. Ruffing, T., Moreno-sanchez, P., Kate, A.: CoinShuffle: practical decentralized coin mixing for bitcoin—bookmetrix analysis. In: European Symposium on Research in Computer Security (ESORICS), vol. 8713, pp. 1–15 (2014)
21. Chen, J., Yao, S., Yuan, Q., He, K., Ji, S., Du, R.: CertChain: public and efficient certificate audit based on blockchain for TLS connections. In: Proceedings of the IEEE INFOCOM, April 2018, pp. 2060–2068 (2018)
22. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system, p. 9. Www.Bitcoin.Org (2008)
23. Turkanovic, M., Holbl, M., Kosic, K., Hericko, M., Kamisalic, A.: EduCTX: a blockchain-based higher education credit platform. IEEE Access **6**, 1–20 (2018)
24. Gervais, A., Karame, G.O., Wüst, K., Ritzdorf, H.: On the security and performance of proof of work blockchains Vasileios Glykantzis Srdjaň Capkun. Bitcoin.org (2017)

25. Garay, J.A.: The bitcoin backbone protocol : analysis and applications the bitcoin backbone protocol : analysis and applications, June 2017, pp. 1–44 (2015)
26. Yang, D., Gavigan, J., Hearn, Z.W.: Survey of confidentiality and privacy preserving technologies for blockchains, pp. 1–32 (2016)
27. Stuart, P.: Confidentiality in Private Blockchain (August 8, 2016). Project "Kadena: Kuro - Private Blockchain.", White Paper August 2016. Available at SSRN:https://www.kadena.io/
28. Chang, P., Yang, C., Yang, C., Hwang, M.: An academic transcript system embedded with blockchains (2018)
29. Ouaddah, A., Elkalam, A.A., Ouahman, A.A.: Europe and MENA Cooperation Advances in Information and Communication Technologies, vol. 520, pp. 523–533. Springer, Cham (2017)
30. Ikeda, K.: Security and privacy of blockchain and quantum computation, 1st ed., vol. 111. Elsevier (2018)
31. Bhowmik, D., Feng, T.: The multimedia blockchain: a distributed and tamper-proof media transaction framework. In: International Conference on Digital Signal Processing (DSP), 2017 August, November 2017
32. Fan, K., Ren, Y., Wang, Y., Li, H., Yang, Y.: Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G. IET Commun. **12**(5), 527–532 (2018)
33. Colloquium, J.N., Zrt, B.E.: Blockchain: solving the privacy and research availability tradeoff for EHR data. In: IEEE 30th Jubilee Neumann Colloquium, pp. 135–140 (2017)
34. Ali, A., Afzal, M.M.: Confidentiality in blockchain. Int. J. Eng. Sci. Invent. **7**(1), 50–52 (2018)
35. Wang, R., He, J., Liu, C., Li, Q., Tsai, W.T., Deng, E.: A privacy-aware PKI system based on permissioned blockchains. In: Proceedings of IEEE International Conference on Software Engineering and Service Science (ICSESS) November 2018, pp. 928–931 (2019)
36. Chen, Y., Xie, H., Lv, K., Wei, S., Hu, C.: DEPLEST: a blockchain-based privacy-preserving distributed database toward user behaviors in social networks. Inf. Sci. (NY) **501**, 100–117 (2019)
37. Casino, F., Dasaklis, T.K., Patsakis, C.: A systematic literature review of blockchain-based applications: current status, classification and open issues. Telematics Inform **36**, 55–81 (2018)
38. Raikwar, M., Gligoroski, D., Kralevska, K.: SoK of used cryptography in blockchain (2019)
39. Chaum, D.: Untraceable electronic mail, return addresses and digital pseudonyms. In Secure Electronic Voting, pp. 211–219. Springer, Boston, MA (2003)
40. Zhang, R., Xue, R., Liu, L.: Security and privacy on blockchain, **1**(1) (2019)
41. Bonneau, J., Narayanan, A., Miller, A., Clark, J., Kroll, J. A., & Felten, E. W. (2014, March). Mixcoin: Anonymity for Bitcoin with accountable mixes. In International Conference on Financial Cryptography and Data Security, pp. 486–504. Springer, Berlin, Heidelberg
42. Chaum, D., Van Heyst, E.: Group signatures. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 547, No. iii, pp. 257–265. LNCS (1991)
43. Wood, G., et al.: How to leak a secret. J. Br. Blockchain Assoc., vol. 2018, November 25, 2016, p. Github site to create pdf, 2016
44. Van Saberhagen, N.: CryptoNote v 2.0. Self-published, pp. 1–20 (2013)
45. Logarithms, D.: A public key cryptosystem and a signature based on discrete logarithms, vol. I, pp. 10–18 (1976)
46. Abidin, A.S.Z., Yusuff, R.M., Bakar, N.A., Awi, M.A., Zulkifli, N., Muslimen, R.: Public-key cryptosystems based on composite degree residuosity classes. In: Lecture Notes in Electrical Engineering (LNEE), vol. 130, pp. 285–299 (2013)
47. Sahai, A., Waters, B.: Fuzzy identity-based encryption BT. In: Advances in Cryptology (EUROCRYPT 2005), vol. 3494, Chapter 27, p. 557 (2005)
48. Chase, M.: Multi-authority attribute based encryption. In: Proceedings of the 4th Conference Theory Cryptography, vol. 4392, pp. 515–534 (2007)
49. Lewko, A., Waters, B.: Decentralizing attribute-based encryption, vol. 2, No. subaward 641, pp. 568–588 (2011)

50. Garg, S., Gentry, C., Halevi, S., Sahai, A., Waters, B.: Attribute-based encryption for circuits from multilinear maps. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8043, PART 2, pp. 479–499. LNCS (2013)
51. Bogetoft, P., et al: Secure multiparty computation goes live. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 5628, pp. 325–343. LNCS (2009)
52. Andrychowicz, M., Dziembowski, S., Malinowski, D., Mazurek, Ł.: Secure multiparty computations on bitcoin. In: Proceedings of the IEEE Symposium on Security and Privacy, pp. 443–458 (2014)
53. Srichaiyo, T., Hjertén, S.: Enigma: decentralized computation platform with guaranteed privacy. J. Liq. Chromatogr **12**(5), 809–825 (2015)
54. Benhamouda, F., Halevi, S., Halevi, T.: Supporting private data on Hyperledger fabric with secure multiparty computation. IBM J. Res. Dev. **63**(2), 1–8 (2019)
55. Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM J. Comput. **18**(1), 186–208 (2005)
56. Essaf, F.: Privacy protection issues in blockchain technology, pp. 124–131 (2019)
57. Pongnumkul, S., Siripanpornchana, C., Thajchayapong, S.: Performance analysis of private blockchain platforms in varying workloads. In: 2017 26th International Conference on Computer Communication and Networks (ICCCN), pp. 1–6. IEEE (July, 2017)
58. Ma, C., Kong, X., Lan, Q., Zhou, Z.: The privacy protection mechanism of Hyperledger fabric and its application in supply chain finance. Cybersecurity **2**(1), 15 (2019)
59. Androulaki, E., et al.: Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the Thirteenth EuroSys Conference, No. 1. ACM (2018)
60. Vukolić, M.: Rethinking permissioned blockchains. In: Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts (BCC), pp. 3–7 (2017)