Blockchain technology the identity management and authentication service disruptor

International Journal on
Advanced Science
Engineering
Information Technology

Vol.8 (2018) No. 4-2
ISSN: 2088-5334

Blockchain Technology the Identity Management and Authentication Service Disruptor: A Survey

Shu Yun Lim[1], Pascal Tankam Fotsing[1], Abdullah Almasri[1], Omar Musa[1], Miss Laiha Mat Kiah[2], Tan Fong Ang[2], Reza Ismail[3]

[1] Faculty of Business and Technology, Unitar International University, 3-01A, Level 2, Tierra Crest, Jalan SS6/3, Kelana Jaya, 47301 Petaling Jaya, Selangor, Malaysia
E-mail: lim_sy@unitar.my, mc170104052@student.unitar.my, abdullah@unitar.my, omarm@unitar.my

[2]Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
Email: misslaiha@um.edu.my, angtf@um.edu.my

[3]SysCode Sdn Bhd, D-13A-02, Menara SuezCap 1, KL Gateway, No.2, Jln Kerinchi, Gerbang Kerinchi Lestari, 59200 Kuala Lumpur, Malaysia
E-mail: reza@syscode.asia

***Abstract*— The Internet today lacks an identity protocol for identifying people and organizations. As a result, service providers needed to build and maintain their own databases of user information. This solution is costly to the service providers, inefficient as much of the information is duplicated across different providers, difficult to secure as evidenced by recent large-scale personal data breaches around the world, and cumbersome to the users who need to remember different sets of credentials for different services. Furthermore, personal information could be collected for data mining, profiling and exploitation without users' knowledge or consent. The ideal solution would be self-sovereign identity, a new form of identity management that is owned and controlled entirely by each individual user. This solution would include the individual's consolidated digital identity as well as their set of verified attributes that have been cryptographically signed by various trusted issuers. The individual provides proof of identity and membership by sharing relevant parts of their identity with the service providers. Consent for access may also be revoked hence giving the individual full control over its own data. This survey critically investigates different blockchain based identity management and authentication frameworks. A summary of the state-of-the-art blockchain based identity management and authentication solutions from year 2014 to 2018 is presented. The paper concludes with the open issues, main challenges and directions highlighted for future work in this area. In a nutshell, the discovery of this new mechanism disrupted the existing identity management and authentication solutions and by providing a more promising secure platform.**

***Keywords*— blockchain; authentication; identity management; distributed ledger technology; ethereum; hyperledger.**

I. Introduction

We all heard about Bitcoin [1], Ether [2] and other cryptocurrencies, which enables people to anonymously perform secure and trustworthy payments and transactions. In the heart of those cryptocurrencies there is a blockchain [3]; a decentralized database which records all transactions since their beginning. The entire network as opposed to a central entity such as a bank or government is continuously verifying the integrity of it. This way, users do not have to trust a central entity, but security is guaranteed by the strength and computing power of the entire network participating in the blockchain.

Authentication as a process of determining whether someone or something is, in fact, who or what it is declaring to be, is the key component of any trustworthy online system which handles sensitive data or transactions. Whether these systems are Internet of Things (IoT), industrial Internet, social networking or payment gateway system, the main aspect of those systems is the authentication process. The process of authentication is very visible to users. It directly influences their perception of trust. An ideal authentication process should be efficient, reliable and able to verify data credentials while protecting user’s privacy.

The identification ecosystem of the past decades is complex and full of middlemen. Service providers have invested billions in system and infrastructure to be compliant with data security regulations. As of today, they are still facing challenges in managing user's identity, authenticating and authorizing users. Every day online users are tasked with providing identity, entering credentials for online and cloud services that they access. These has generated huge volumes of user data with service providers and user private data is stored and left to the discretion of service providers [4] [5]. Last year in Malaysia, there was a massive data leak involving 46.2 million mobile users [6]. Early this year, personal details of over 220,000 organ donors and their next-of-kin had been leaked through government official databases in Malaysia. Their personal details, identity card numbers, addresses and mobile phone numbers may have fallen into the wrong hands [7].

Users who subscribed to multiple online services will have to store passwords in all the servers for authentication and hence authentication data are replicated and withheld in multiple servers. These redundant actions of exchanging authenticating data may lead to an exploit of the authentication mechanism. These vulnerabilities have caused user to suffer from identify theft and data breaches. This server-centric identity management model has deficiencies. From the service providers standpoint, managing and authenticating users is becoming inevitably complicated. Passwords and personal identity information is traditionally stored in a centralized server which makes it possible for hackers to achieve their malicious goals by stealing, misusing or manipulating these data. Therefore, service providers are required to create stronger mechanisms, by adding multiple factors authentication for access and stronger encryption, which further complicates the system [8].

Besides server-centric identity management, federated identity management [9] is adopted currently where organizations allow users to use the same single identity on different online services. This comes in the form of single sign on or Facebook Login, Google ID etc. Although identity federation gives a degree of portability to a centralised identity, but the power remains with the identity provider. The impact of federated identity on user privacy is more profound. Identity providers have access to the information stored by subscribers for authentication purpose and this presents a privacy issue. It is difficult for users to make sure the proper Service Level Agreement (SLA) rules are enforced since there is a lack of transparency that allows the users to monitor their own information. This can be seen in recent Facebook and Cambridge Data Analytica dispute over alleged harvesting and use of personal data [10].

Know-Your-Customer (KYC) compliance obligations for financial institutions are costly and time-consuming. Global financial institutions are burdened by the need to both collect and protect data at the same time. The current personal data ecosystem is archaic, fragmented and inefficient hence a new authentication and identity management framework is needed. Self-sovereign identity management [11] and a decentralized solution with Distributed Ledger Technology (DLT) is required to address these challenges. The blockchain and DLT is undeniably an ingenious invention for nowadays Internet systems, since many people including developers do not understand what the technology is about, the blockchain technology remains one of the most underestimated technologies of the time. In section II, the background of identity management and authentication mechanism is explained. Blockchain technology and a comparative review between Ethereum and Hyperledger blockchain is discussed in section III. An overview of related research works is presented in section IV. Section V and VI concluded the paper with open issues, main challenges and directions of the future blockchain and distributed ledger technology.

A. Overview of Identity Management

Identity management refers to broad administrative area and standards that create, maintain and the de-provision of user account. Sound identity management and governance are needed to manage identities for online services. Identity management is required to simplify the user provisioning process. Enabling new users to get access to online services and de-provisioning users to ensure that only the rightful users have access to services and data.

1) Independent IDM

Majority of Internet identities are centralised [12]. The user credentials are owned and managed by a single entity. But these independent identity repository model has deficiencies. Users do not own their identity record and it can be revoked or misused by the identity provider.

2) Federated IDM

Federated identity management systems [9] [13] can provide authentication and authorization capabilities across organizational and system boundaries. It requires agreements that an identity at one provider is recognized by other providers and contractual agreements on data ownership. User account is managed independently by identity provider and no enterprise directory integration is required. This lower the security risk as credential are not replicated but propagated on demand. This approach is relatively more complex to implement and requires proper agreement and trust relationship between online services.

3) Self-sovereign IDM

Self-sovereign identity is the concept that users should be able to control their own digital identity. People and businesses can store their own identity data on their own devices and provide their identity to those who need to validate it, without relying on a central repository of identity data. Since it is independent from any individual silo, it gives user full control, security and full portability of their data. Sovrin foundation [14] describes self-sovereign identity as an Internet for identity where no one owns it, everyone can use it, anyone can improve it.

B. Overview of authentication process

The establishment of a secure channel permits to exchange sensitive data providing trustworthy, confidentiality and integrity service on the exchanged data. To provide these services, companies will setup an authentication process based on user registration data. Those data will be stored in a server either locally or remotely and to be used whenever they need to identify user. To overcome the issue of identification, companies developed many types of authentication mechanism based on either something you know (such as password) or something you have (such as smart card) or something you are (such as a user profiling, fingerprint or other biometric method).

1) *Password authentication*

Password authentication is simple and easy to use, but it must have a certain level of complication and regular renewal to keep the security. It is an authentication technology with well-known weaknesses in the sense that even if the correct username and password combination is provided; it is still difficult to prove that the request is from the rightful owner and subjected to shoulder surfing attack [15]. Users frequently reuse their passwords when authenticating to various online services. In view of the weak password practices, this brings high security risks to the user account information. Nonetheless password authentication is still the most frequently used authentication technology with more than 90% of transactions [16].

2) *Trusted Platform Module based authentication*

Trusted Platform Module (TPM) is a hardware-based security module that uses secure crypto processor that can store cryptographic keys that protect information. A variant of it is, Mobile Trusted Module (MTM) [17] is a proposed standard by Trusted Computing Group a consortium (TCG) founded by AMD, Hewlett-Packard, IBM, Intel, Microsoft. It is mainly applied to authenticate terminals from telecommunications. However, it is being considered as a online authentication method with Subscriber Identity Module (SIM) due to the generalization of smartphones. User devices can utilise unique hardcoded keys to perform software authentication, encryption, and decryption.

3) *Trusted Third Party authentication*

Employing Trusted Third Party (TTP) services within the cloud leads to the establishment of the necessary trust level and provides ideal solutions to preserve the confidentiality, integrity and authenticity of data and communication. PKI (Public key infrastructure) team up with TTP provides technically sound and legally acceptable means to implement strong authentication and authorization. PKI is an authentication means using public-key cryptography. It enables users to authenticate the other party based on the certificate without shared secret information. One example of TTP authentication in cloud is Single-Sign-On (SSO) [18]. When a user gets authentication from a site, it can go through to other sites with assertion and no authentication process is required. However, the existence of a trusted third party as an authentication server or certification authority is becoming security and fault intolerance bottlenecks for the system.

4) *Multifactor authentication*

Multi-factor authentication [8] [19] [20] [21] ensures that a user is who they claim to be by combining a few means of authentication. The more factors used to determine a person’s identity, the greater the trust of authenticity. ID, password, biometrics [22] [23] [24], certificate are used traditionally for single factor authentication. With the emerging of mobile network, second factor authentication takes the form of SMS, e-mail, and telephony OTPs, PUSH Notifications, and mobile OATH Tokens. Even though it is rather effective for closed communities such as enterprise cloud, these methods are too costly, inconvenient, and logistically difficult especially for the distribution, administration, management and support in the cloud.

5) *Implicit authentication*

This approach uses observations of user behaviour for authentication and it is well suited for mobile devices since they can collect a rich set of users’ information, such as location, motion, communication, and their usage of applications. A number of profiling techniques have been studied to provide a suitable service for user and personal profile information in mobile cloud environment [25] [26] [27]. But to date, a formal model of this approach has yet to be provided and limited device resources are the technical constraints to overcome. Studies on intelligent mobile authentication service are still inadequate.

6) *Blockchain authentication*

The immutable blockchain ledger verifies and ensures that the users, transactions, messages are legitimate. Blockchain authentication [28] is done by smart contracts which are written and deployed to blockchain. A smart contract generator can be programmed through a Smart Contract Authentication (SCA) layer to activate and execute every time an authentication is required by either party and self-govern itself within a predefined scope of actions. The need for a third party to authenticate transactions is eliminated. Costs can be reduced while security and privacy are greatly enhanced. Effort of hijacking the authentication process would be much greater in the distributed environment.

II. Material and Method

This section highlights some basic concept of what a blockchain is and why it can be the best alternative to manage our data credentials and authentication process.

A. *The Block*

A block is referring to files where data pertaining to blockchain network is permanently stored. A block is like pages of a ledger or an account book. Each time a block is completed, it gives way to other block. Data stored in blocks cannot be altered. The genesis block, genesis.Json, is the first block of a blockchain.

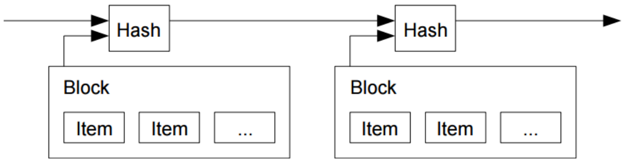

Fig. 1 Structured connections of Blockchain’s blocks [1]

B. *Hashing*

Each block contains a record of transaction and is cryptographically hashed. A hash function takes in input value and creates an output value deterministic of the input value. Every input has a determined output. The process of applying the hash function to any data is called hashing and the output is called the hash value or simply the hash. One critical characteristic of a secure hash function is that it is only one way. This means that given the hash, it is impossible to determine what the input was. Hashing is extensively used with Blockchains. For example, a process of hashing public keys derives addresses on a Blockchain. An Ethereum account is computed by hashing a public key with keccak-256.

C. Blockchain

A blockchain [29] is a chain of blocks of valid transactions. Each block includes the hash to the prior block in the blockchain. It uses a peer-to-peer network, which means every node in the network is connected to every other in the network. After the transaction is verified, it is broadcasted to the network and is added to everyone copy of the blockchain.

Advantages of the blockchain technology includes:

- Immutability: nothing on the blockchain can change. Any confirmed transaction cannot be altered.
- Permanence: A public blockchain will act as a public ledger, data will be accessible if the blockchain remains active.
- Removal of intermediaries: The peer-to-peer nature of the blockchain does away with the need of intermediaries.
- Speed: Transactions are much faster than a centrally controlled ledger.
- Security: Neither the node nor anyone else except the sender and the receiver can access the data sent across the blockchain.

1) The Merkle tree

The block is divided into two main categories which are the header and the body. The header has four components, a timestamp, a nonce, a hash reference to a previous block and a hashed list of all transactions that took place since the last created block. The blocks are stored in a multi-level data structure, a tree structure called the merkle tree. This structure is the key factor of the mining. The merkle tree or binary hash tree is a type of a binary tree, where the bottom of the tree contains the transactions (hashed), the intermediate tree nodes (leaves) contain the hash of the two nodes that made it, all the way till the top where it is a single hashed tree-node called the Merkle root (root hash).

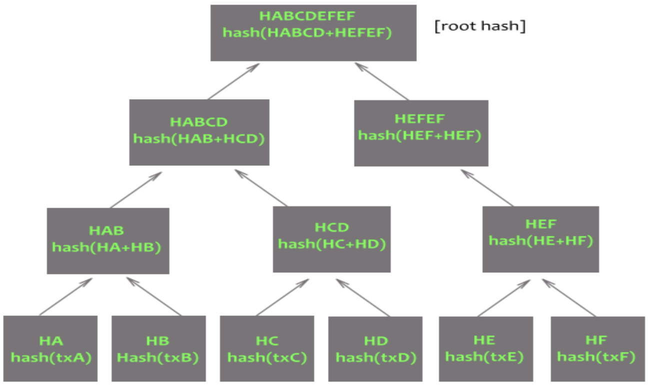

Fig. 2 Overview of a Merkle tree

With reference to Figure 2, there are six transactions (txA, txB, txC, txD, txE, txF,) with their hashes (HA, HB, HC, HD, HE, and HF) at the base/bottom of the tree. Concatenating any two hashes of the transactions together (HA + HB), (HC + HD) and (HE + HF) will give the first leaves of the tree. The next steps will be the hashing of the leaves H(HA + HB), H(HC + HD) and H(HE + HF) with the results HAB, HCD and HEF which will result in new leaves. The leaves will continue to concatenate. At the end of the process, hashes HABCD and HEFEF will be created. The process will result to a final and unique hash, the root hash which for this example is hash (HABCD + HEFEF) = HABCDEFEF called the Merkle root. Merkle root is placed in the block header mentioned as “hashed list of all transactions that took place since the last created block”. “Merkling” the hashes of child nodes in the tree help verify contents for parents and generally large data structures.

The advantage of utilizing merkle tree data structures is that any node in the network can check the historical backdrop of many transactions easily, and hence any individual is guaranteed that their duplicate of the blockchain is finished and alter-free. Confirmations are given as a feature of the centre blockchain code and guarantee that open private keys are substantial, transactions are being marked effectively and exchanges are legitimate the distance back to the root.

2) Consensus Algorithm

For a block to be accepted by the network peers, miners must complete the proof of work [30], which covers all the data in the block. The difficulty of this work is adjusted as to limit the rate of new block generation to one every 10 minutes (in Bitcoin blockchain) and can vary for other blockchain.

- Proof of Work

PoW is the calculation of hash functions to solve ‘mathematical puzzle’ in blockchain. Producing a proof of work is a random process and hence it requires a lot of trial and error. PoW algorithm is based on computation power. Miners are as powerful as the number and power of CPUs they own. This algorithm is the oldest and the most common one in the Blockchain technology, one of the problems that PoW has, is that it spends a great amount of electricity and bandwidth over the process of mining.

- Proof of stake

Driving the weaknesses of PoW algorithm, PoS algorithm was developed to make blockchain nodes as powerful as their stake. The earning reward for a miner is the function of the amount of stakes the miner holds. For example, if a node has 10% coins in account, it will earn 10% of any new coins created in the future because the probability of signing next block would be related to the amount of stake. In this case there is no need to solve a very hard mathematical challenges as in PoW, which prevents wasting resources like electricity. All seems good here but there is also a new issue that the owner of the oldest set of coins or the one has more coins get more rewards (rich get richer), the only thing that need to be done is to prove the ownership of its stake.

- Proof of Importance

Due to the problem with PoW and PoS, there was a new consensus algorithm call the Proof of Importance (PoI). The idea behind this algorithm is that the nodes are important as their activities on the network. Nodes that are active on the network will be rewarded. Each address is given a trust score, and activities on network gets higher, the more chance a node will be rewarded based on loyalty and effort.

- Practical Byzantine Fault Tolerance (PBFT)

This consensus mechanism is certainly one of many that can be utilized in permissioned blockchains, in which a new block is generated if more than 2/3 of all validating peers post the same reaction. Hyperledger fabric out of the box does not provide PBFT, however gives its users the feature of adding this consensus mechanism modularly.

3) Access to Data

Depending on the consensus, there are three types of blockchain which are: public, private and consortium Blockchain

- Public Blockchains

Public blockchains or permission-less blockchains are accessible for everyone and anyone can participate as a node in the decision-making process. Public blockchain achieve consensus without central authority and thus can be considered as decentralized. All users maintain a copy of the ledger on their local nodes and use a distributed consensus mechanism to reach decision or eventual state of the ledger. Bitcoin is the best example of a public blockchain -whenever a user does a transaction, it is reflected on every copy of the block.

- Private Blockchains

Private Blockchains are private and open only to a consortium or group of individuals or organizations that has decided to share the ledger among themselves. Only the owner of the Blockchain has the right to make any changes to it. For example, Blockstack [31] [32] [33] aims to provide the financial institutions with back office operations, including clearing and settlement on private Blockchain. However, the use cases of a private Blockchain are relatively small as compared to the public Blockchain. Some people may argue that private Blockchain is not of much used as the implementation concept does not differ much from that of the current systems. Nonetheless private blockchain can provide solutions to some of the problems which Bitcoin cannot, such as know-your-customer (KYC) or anti-money laundering (AML).

- Consortium Blockchains

This blockchain is basically a hybrid of public and private blockchains. The consensus process is controlled by a preselected set of nodes. Rather than allowing any node to participate in reviewing the transactional process, a consortium blockchain provides multiple defaults and distributed nodes for the process. A consortium platform offers many of the benefits associated with private blockchains, such as the efficiency and privacy of transactions. Besides, a consortium blockchain is generally faster, with higher scalability and provide more transaction privacy.

4) Permission Restrictions

Permission restrictions will determine which nodes are eligible to create blocks of records. A permissioned blockchains predefine the users to carry out transaction processing, as in Hyperledger fabric blockchain. Meanwhile for permission-less blockchains, there is no restriction on the identities of processors, therefore anyone can be a part of the network. This is the case in Bitcoin and Ethereum.

5) Scalability of blockchain structure

The scalability of blockchain structure is composed of the factor of node scalability and performance scalability. Node scalability in blockchain networks refers to the extent to which the network can upload more node without a loss in performance. Performance scalability on the other hand refers to the number of transactions processed per second. It is impacted by the latency among transactions and the block length.

A blockchain is considered scalable if it can add thousands of globally distributed nodes whilst still processing thousands of transactions per second. Currently, none of the prevailing blockchains are scalable. Public blockchains such as Bitcoin and Ethereum make this trade-off in favor of node scalability by using proof-of-work (PoW) consensus mechanisms. On the other hand, a Hyperledger fabric instance that modularly adds PBFT makes this trade-off in favor of performance scalability. For business structures of less than 20 nodes this might be a viable solution. However, if there are more nodes that takes place in PBFT, transaction throughput can be reduced significantly.

6) Governance

Governance [34] refers to the degree to which decision-making power is distributed within the blockchain network. It attempts to answer the question of who could make what decisions on a blockchain platform. Each blockchain platform needs to be developed and maintained. Usually, a core developer crew performs this task. As there are many stakeholders in a blockchain network, such as core developers, miners, currency-exchanges, decentralized applications (Dapps) developers, decisions making for new changes to the blockchain center protocol are very important and frequently controversial. This is a strong factor where blockchain systems differ from each other.

7) Anonymity on blockchain

Anonymity on the blockchain refers to whether the identity of a node is openly transparent. In public permission-less blockchains, such as Bitcoin and Ethereum, users are pseudonymous since they cover their identity behind a pseudonym, their public wallet address. In private permissioned blockchains, such as Hyperledger fabric, users usually know each other.

8) Native currency

Native currency refers to whether the blockchain has an inherent currency [3]. For example, Bitcoin uses its currency “Bitcoin” as a medium for exchange. Ethereum uses “Ether”. while Hyperledger fabric does not use an own currency.

9) Turing Completeness

Scripting refers to the degree to which a blockchain's programming features to support the development of Dapps. This function will allow the developer to check the Turing completeness of the blockchain. Turing completeness refers to any tool or device that in theory can calculate everything assuming sufficient resources (memories) is available. Ethereum and Hyperledger fabric are Turing complete so they provide developers with a Turing-complete scripting language (Solidity for Ethereum and Chaincode for Hyperledger Fabric), which allows developers to create smart-contracts that can interact with each other and form decentralized applications. While other blockchains, such as Bitcoin, only provide a very limited stack-based programming. This makes application development very tough and sometimes not possible.

10) Ethereum VS Hyperledger

- Ethereum

Ethereum [2], an open source project, provides a blockchain solution that allows distributed application to be deployed. Ethereum links smart contracts and blockchains. A smart contract is a credible contract that is completely controlled by computer program and does not depend on any agency. The contract is automatically executed once execution conditions are satisfied and no individual node can modify it. Ethereum can be seen as Bitcoin 2.0, a crypto currency with support for smart contracts. Ethereum is known for its high cost for performance scalability and privacy. Its built-in cryptocurrency is known as Ether (ETH).

- Hyperledger

Hyperledger [35] on the other hand, is a Linux Foundation banner project which covers frameworks like Hyperledger Fabric, Sawtooth, Iroha, Indy and Burrow. Hyperledger provides core modules and API to facilitate development and interoperability. Hyperledger satisfies major purposes of blockchain for business. In permissioned network, proof-of-work does not involve solving difficult cryptographic problems, also known as mining. Transaction can be confirmed within a short time which is a major business requirement. Besides, when mining is not required, there is no reliance on cryptocurrencies which is used to incentivise miner. It supports selective disclosure which gives businesses the flexibility to make transactions visible to selected parties. This can be achieved through the key management of encryption or signature operations. Above and beyond, Hyperledger leverages on smart contract to automate business processes, Byzantine fault tolerance (BFT) algorithm and fine-grained access control for its permissioned mode of operation.

There are fundamental differences among the blockchain technologies. For instance, participation of nodes in the decentralised network, consensus mechanism, scalability and native cryptocurrency. A detail comparison between Ethereum, Hyperledger and Bitcoin is presented in Table I.

TABLE I
COMPARISON OF ETHEREUM, HYPERLEDGER AND BITCOIN

Characteristics	Ethereum	Hyperledger	Bitcoin
Founded	July 2015	July 2017	January 2009
Permission restrictions	Permission-less	Permissioned	Permission-less
Access to data	Public or private	Private	Public
Consensus	PoW	PBFT	PoW
Scalability	High node-scalability, low performance-scalability	Low node-scalability, high performance-scalability	High node-scalability, low performance-scalability
Centralized regulation (governance)	Medium, core developer group, but EIP process	Low, open-governance model base on Linux model	Low, descentralized decision making by community
Anonymity	Pseudonmity, no encryption of transaction data	Pseudonymity, encryption of transaction data	Pseudonymity, no encryption of transaction data
Native currency	Yes- Ether (ETH)	No	Yes- Bitcoin(BTC)
Scripting	High possibility, Turing complete virtual machine, high-level language support (Solidity)	High possibility, Turing complete, scripting of chaincode, high-level Go-language	Limited possibility, stark-based scripting
Programing language	Golang, C++, Python	Golang, Java	C++

As a new technology with infinite opportunities, there are many individuals, companies as well as governments, which have started some researches and development on blockchain technology. While it may take years for blockchain technology to mature fully, many blockchain solutions and applications are already perfectly feasible in the near term, and new opportunities will continue to present themselves as the underlying technology evolves. In this section, blockchain solutions in the field of identity management and authentication from year 2014 to 2018 are presented. A summary of the related works is exhibited in Table II.

A. Sovrin

Sovrin [14] is a trust framework for decentralized, global public utility for self-sovereign identity. It is also the first global public utility exclusively for self-sovereign identity and verifiable claims. Self-sovereign aims to provide portable identity for any person or organization. Having a self-sovereign identity allows the holder to present verifiable credentials in a private way. These credentials can be gender, age, education background or employment information.

The Sovrin protocol is based entirely on open standards and open source Hyperledger Indy Project. All Sovrin identifiers and public keys are pseudonymous by default. The solution is pairwise-pseudonymous identifiers, a separate Distributed Identifier (DID) for every relationship.

As of time of writing over 20 stewards have signed on to operate under the Sovrin Trust Framework.

B. *MyData*

MyData [36] is a research commissioned by Finnish government for personal data management. This Nordic self-sovereign identity model is driven by the concept of human centric control, usability, accessibility and openness. MyData can be used to secure flow of data between sectors likes governments, healthcare and finances. The core of MyData authentication are user managed access, OpenID single sign-on and Oauth 2.0 which control access to Web APIs. Blockchain is used to distributed control of fraudulent activities to the entire network of stakeholders, as any attempt to tamper with the blockchain is easily detectable.

The research, which joint forces with Sovrin, aims at strengthening digital human rights while opening new opportunities for business to develop innovative personal data services. It is also aiming at addressing EU General Data Protection Regulation (GDPR) [37], new rules on controlling and processing personally information enforced since May 2018.

C. *Waypoint*

Waypoint [28] is a decentralized multi-factor authentication system that is deployed on the Ethereum Virtual Machine. This solution allows identity authentication to be performed on the Blockchain, with Web API based implementation.

With a mobile base apps and desktop version available, Waypoint allows application to secure multiple modules within one product by defining multiple functions. It provides feature to store user behaviour and perform analytics for real time behavioural based authentication. The commercial solution is currently at beta-stage.

D. *Bloom*

Bloom [38], a blockchain project for credit scoring and identity management that uses Ethereum and IPFS. it is an all-encompassing protocol it that it allows for each traditional and digital currency holders to serve as lenders to users who are unable to obtain a bank account or credit score. users will create an id contract (BloomID) to be attested by friends, family and corporation. The BloomIQ system then reports and tracks debt obligations, ensuing in a BloomScore as a metric of client's credit worthiness. The bloom protocol creates a globally portable and inclusive credit profile, reducing the need for classic banking infrastructure and opaque, proprietary credit scores.

E. *BlockStack*

Blockstack [31] [32] [33] provides decentralized services for naming (DNS), identity, authentication and storage. developers can use JavaScript libraries to build serverless apps and not worry about handling infrastructure. Blockstack will replace the contemporary client/server model; users control their information, apps run client-side, and the open Blockstack network replaces server-side functionality.

F. *ShoCard*

ShoCard [39] is a commercial mobile identity solutions that protects consumer privacy. It is basically a tiny file that only user can manipulate. When users create a ShoCard ID, through the App or via SDK, their identity document is scanned and signed. Then, the app will generate a private and public key to seal that record. The record is then encrypted, hashed and sent to the Blockchain where it cannot be tampered with or altered. Shocard Identity Platform is built on a public BlockCypher's blockchain infrastructure, data or keys that could be compromised are stored off-ledger.

G. *Uport*

Uport [40] is a secure system for self-sovereign identity. It aims to be an open identity system for a decentralized web. It operates on the Ethereum blockchain and enables users to send and request credentials, digitally sign transactions, as well as manage their keys and data in a secure manner. It allows the publication of identity data to other Blockchain such as Bitcoin and Ethereum.

Uport identities can be either individuals, devices, entities, or institutions. Examples of interactions powered by uPort include blockchain transactions such as buying shares on the Gnosis predication market, as well as making private statements to other uPort users or applications. uPort utilizes two protocols, namely the Identity and Claims Protocol. The Identity Protocol is an address on a decentralized network, controlled by a private signing key, and makes use of a decentralized public key infrastructure (PKI) that enables signature validation. On the other hand, the Claims Protocol refers to a standard message format that enables source attribution and facilitates interoperability between various blockchain and identity networks. The Claims Protocol supports the JSON Web Token (JWT) and Ethereum transactions. Among the products and tools offered by uPort is the self-sovereign wallet, where it allows its users to sign transactions and manage their keys and data in one simple, secure location. uPort also offers development tools to assist Simple Authentication and SSO for dapps or modern web applications. Although its seems very promising, Uport is still in the closed-beta stage.

H. *I/O Digital*

I/O Digital [41] provides an identity management that utilizes an improved blockchain called DIONS (Decentralized I/O Name Server) and secured using Proof of Stake (POS I/O). The DIONS blockchain enables storage of data, with capabilities of document and identity storage. DIONS also allows for message encryption using AES 256 block cipher and accompanied with a complete Alias system. The Alias system allows its users to store sensitive identity credentials and provides a way to manage reputation and control their data, as the user can choose to create a public (unencrypted) alias, private (encrypted) alias, or both. The aliases are easy to remember and fully transferable between users. The IOC data, messaging / alias system fees are redistributed to all active stakers in the network. This ensures further IOC distribution, and incentives users to stake while securing the network. Features such as alias creation and decryption, secure channel negotiation via a single Invite, secure file transfer, and secure instant message communication are available on a readily hard-coded into a HTML5 wallet system.

I. BlockAuth

BlockAuth [42] is franchised network of OpenID Connect providers that that enables user to own and operate its own identity registrar. User privacy falls within the control of the users by allowing them to choose what information they wish to make public. All user data will be encrypted. Information they wish to keep entirely private is encrypted with multi-part keys that require multiple parties to work in tandem to decrypt. Additionally, BlockAuth use their financial resources to help developers of open source projects by paying grants or bounties. BlockAuth is providing an easy-to-integrate authentication system through modern standards-compliant API. This framework is necessary to build an entire resilient decentralized ecosystem to perform the tasks of user authentication and verification.

J. UniquID

UniquID [43] is a decentralized identity and access management platform that provide digital keys. It aims to solve the increasing challenges attributed to the Internet of Things. This platform prioritizes identity before security. User's device would be saved inside their own private blockchain. This private blockchain would act as a digital vault to protect the user's digitally connected assets via secure authentications. UniquID also enables devices to be independent. This means that authentications are carried out device to device without the need of any third-party intermediaries. This concept is applied to deal with challenges related to cybersecurity and Internet of Things.

Besides that, UniquID's device centric solution does not require the usage of passwords, as it recognizes its users through personal connected objects, or integrated with fingerprint or other biometry on personal devices. Thus, this removes the risk associated with user generated passwords. It claims to be ready for deployment on custom hardware, servers, personal computers or smart phones and tablets. It is currently in a private beta stage.

K. Jolocom

Jolocom [44] aims to develop a solution to provide a decentralized identity based on hierarchically deterministic keys (HD keys). These keys are generated, provisioned, and controlled by the users themselves. This platform allows easy management of multiple personas and preservation of pairwise anonymity in context specific interactions. The derived key pairs can be recovered by using a simple seedphrase. Besides that, Jolocom also allows the modelling of IoT devices ownership for integrated human and machine identity. Jolocom is focused on providing a lightweight, global, and self-sovereign identity solution for decentralized systems that is easy to deploy for non-technical users. It also maintains an open source release to support the larger decentralized application community.

The Jolocom system architecture consists of the Jolocom Library, its user interface, a public distributed storage system, and a storage backend. The Jolocom library offers a comprehensive RESTful API for performing all available identity related functionalities: creating a new identity (Decentralized Identifiers, DID) and DID Document Object (DDO) which documents verifiable claims related to the identity. The Jolocom user interface is a fully decentralized mobile application to manage and use their decentralized digital identity. It currently allows for creation of new identities, creation and updating claims on identities, as well as verifying claims on other identities. Future development plans of Jolocom include interaction with Ethereum smart contract, integration with other blockchains, as well as management of tokens.

L. Cambridge Blockchain

Cambridge Blockchain [45] is founded with the mission of fostering Cambridge's blockchain ecosystem. It is working on an identity Blockchain for validating secure digital identity documents, processing electronic signatures, and recording transactions. Cambridge Blockchain's distributed architecture resolves the competing challenges of transparency and privacy, leading to stronger regulatory compliance, lower costs and a seamless customer experiences.

M. KYC.LEGAL

KYC.LEGAL [46] is an Ethereum based blockchain identity service that allows other services to verify users. It allows the identity of users to be established and documented, so that going forward online provider can register any services that require such verification by providing only that information which is required for each individual service. The product is made up of two parts: document verification through a mobile application, and verification of identity and documents with the help of a KYC.LEGAL agent.

N. CertCoin

CertCoin [47] is a decentralized authentication system based on the NameCoin [48] blockchain. This system carries the best aspects of transparent certificates authorities and web of trust. Certcoin is absolutely public and auditable. Certcoin helps the expected features of a full-fledged certificate authority such as certificate creation, revocation, chaining, and recovery. Domain purchases and transfers are executed with simple Bitcoin transactions to incentivize miners. The CertCoin layout additionally facilitates trusted key distribution that makes it more suitable for performance conscious applications. Besides that, it also addresses several issues inherent to current PKIs, such as the need for a trusted third party and limited accessibility.

O. Authenteq

Authenteq [49] uses a facial recognition algorithm to create a digital identity on a blockchain. Authenteq allows users to verify identity and create personal sovereign digital IDs which is stored in an encrypted blockchain. All personal data are owned and controlled by owner, and not accessible by any third party. Authenteq can be adopted by any type of online services. API and plugin are provided for business integration. This is one of the commercial blockchain IdM solutions which incorporates biometric features for authentication.

TABLE II
SUMMARY OF BLOCKCHAIN-BASED IDENTITY MANAGEMENT AND AUTHENTICATION SOLUTIONS

Solution	Description	Propose type	Blockchain	Network	ID Mgmt	Auth	Status
Sovrin [11]	Decentralized global public utility for self-sovereign identity	Non-profit foundation	Hyperledger Indy	Public Permissioned	Yes	No	Completed (September 2016)
MyData [36]	This Nordic initiative which joint forces with Sovrin to build self-sovereign identity and authentication mechanism	Government	Hyperledger Indy	Public Permissioned	Yes	Yes	On-going
Waypoint [28]	Decentralized multi-factor authentication system	Company	Ethereum	Private	No	Yes	Beta stage (October 2017)
Bloom [38]	Blockchain project for credit scoring and identity management	Open source	Hyperledger	Permissioned	Yes	No	Completed (January 2018)
BlockStack [31, 33]	Decentralized services for naming/DNS, identity, authentication and storage	Start-up	Ethereum	Private	Yes	Yes	Completed (October 2017)
ShoCard [39]	Identity platform to protect consumer privacy	Start-up	Ethereum	Public	Yes	No	Completed (December 2017)
Uport [40]	Identity management	Company	Ethereum	Public/Private	Yes	No	Completed (October 2016)
I/O Digital [41]	Identity management based on the Blockchain	Start-up	Ethereum	Private	Yes	No	Completed (January 2018)
BlockAuth [42]	Developing identity registrar base on the Blockchain	Start-up	Ethereum	Permission-less	Yes	No	Completed (July 2014)
UniquID [43]	Identity and access management of connected things	Open source	Ethereum	Permission-less	Yes	No	Beta Stage (June 2016)
Jolocom [44]	Applications for user to own their personal digital identity	Start-up	Ethereum	Public/Private	Yes	No	Development stage (February 2018)
Cambridge Blockchain [45]	Identity Blockchain	Start-up	Ethereum	Permission-less	Yes	No	Alpha Stage (June 2017)
KYC.LEGAL [46]	User identification and verification to prevent fraud	Company	Ethereum	Permission-less	Yes	No	Completed (February 2018)
CertCoin [47]	NameCoin based decentralized authentication system	Open source	Hyperledger	Permissioned	No	Yes	Completed (May 2014)
Authenteq [49]	Identity verification platform that uses a facial recognition algorithm to create a digital identity on a blockchain	Company	Ethereum	Permission-less	Yes	No	Completed (August 2014)

III. Result and Discussion

Even though there are many legislation issues surrounding the exchange of sensitive data attributes, personal privacy concerns are addressed inadequately or simply overlooked. Self-sovereign identity management, blockchain and Distributed Ledger Technology are going to patch the gap that current technology falls short of providing a secure and cost-efficient identity management framework. Blockchain authentication and self-sovereign identity management can be deployed by government agencies, financial institutions and enterprise business for providing a secure and reliable authentication and identity management solution.

The discovery of this new mechanism creates a secure platform for service providers to authenticate users with no single point of failure and prevent attacks and leakages of user data. This solution is a tamper-proof reference point to verify personal data without having to expose the actual data to a service provider.

Blockchain identity management and authentication solution by design is distributed, decentralized and fault-tolerant which decreases the deployment and maintenance cost. However, scalability seems to be the biggest challenge with public blockchain. Some argued that by centralizing some parts of the technology, blockchain identity management will be more cost effective and secure.

On the other hand, instead of on premise deployment of blockchain network, Blockchain-as-a-Service (BaaS) [50] allows customers to leverage cloud-based solutions to build, host and use their own applications and smart contracts on the blockchain. Cloud providers take over other necessary tasks to keep the infrastructure operational. Undeniably, BaaS is aiding the blockchain adoption across businesses. Companies such as IBM, Microsoft, or even google had started offering the cloud as a service business model based on blockchain technology.

Even though blockchain provides the technology to resolve identity management glitches, some parties argued that identity management has always been a business issue but not a technology problem [51]. Blockchain technology does not resolve access management issues such as key management problem that is inherent in server centric and federated identity environment. Another long-running problem with identity is around the verification of user identity, in which there is no one responsible and liable for vetting data, the same problem where federated identity projects have become stuck. The solution to this problem is probably to extend the notion of zero knowledge proof in self-sovereign identity management. This leads to a mechanism in which the prover demonstrates possession of knowledge without conveying any information apart from the fact that he or she possess the knowledge.

Besides, enhancement of Ethereum and Hyperledger blockchain is required which in turn could improve the performance of blockchain network. In real world implementations, it will require an overhaul or at least a focused effort to integrate this technology with exiting implementations of identity authentication to begin an initial acceptance of this technology in the market.

IV. Conclusions

Email and password credentials are notoriously easy to crack as can be witnessed in all the large-scale online account hacking. Current online services rely blindly on online providers to perform identity management and authentication. There should be an ideal form of identity management that only grants access to certain information and eliminates the need for each online service provider to store credentials for every client.

Blockchain can offer a solution by decentralizing the ownership of credentials and offering a universally available protocol for verifying one’s record in an immutable chain of data. Blockchain can create a secure platform for online service providers to authenticate users. Besides, this technology could also help to instill the trust back in users. Users should have full control over who has the right to use their data and what they can do with it once they gain access.

To facilitate this peer-to-peer exchange of data and consent, routing of requests, mechanisms for discovery and recording of events, a decentralized network that is publicly accessible, immutable and resistant to faults and tampering is needed. Distributed ledger technology and Blockchain is the revolution that makes this possible.

Acknowledgment

This research was supported by UNITAR International University Internal Research Grant.

References

[1] Nakamoto, S., Bitcoin: A Peer-to-Peer Electronic Cash System. 2008.

[2] Wood, G., Ethereum: A Secure Decentralised Generalised Transaction Ledger EIP-150 Revision. 2014.

[3] Swan, M., Blockchain: Blueprint for a new economy. 2015: O'Reilly Media, Inc.

[4] Alizadeh Mojtaba, A.S., Zamani Mazdak, Baharun Sabariah, Sakurai Kouichi, Authentication in mobile cloud computing: A survey. Journal of Network and Computer Applications, 2016. 61: p. 59-80.

[5] Shu Yun Lim, M.L.M.K., Tan Fong Ang, Security Issues and Future Challenges of Cloud Service Authentication. Acta Polytechnica Hungarica, 2017. 14(2): p. 69-89.

[6] TheStar, M’sia sees biggest mobile data breach, in TheStar. 2017.

[7] MalaysiaKini, After data leaks, Personal Data Protection Act needs review, in Malaysia Kini. 2018.

[8] Nagaraju, S. and L. Parthiban, SecAuthn: Provably Secure Multi-Factor Authentication for the Cloud Computing Systems. Indian Journal of Science and Technology, 2016. 9(9).

[9] Ghazizadeh E., M., J. L. A., Zamani, M., Pashang, A. A survey on security issues of federated identity in the cloud computing. in Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on. 2012.

[10] Meredith, S., Facebook-Cambridge Analytica: A timeline of the data hijacking scandal. 2018, CNBC.

[11] Andrew Tobin, D.R., The Inevitable Rise of Self-Sovereign Identity. 2017.

[12] Simon, H. SAML: The Secret to Centralized Identity Management. 2004.

[13] Zwattendorfer, B., et al., A Federated Cloud Identity Broker-Model for Enhanced Privacy via Proxy Re-Encryption, in Communications and Multimedia Security, B. De Decker and A. Zúquete, Editors. 2014, Springer Berlin Heidelberg. p. 92-103.

[14] Andrew Tobin, D.R., The Inevitable Rise of Self-Sovereign Identity (White paper). 2017: Sovrin Foundation.

[15] M I Awang, M.A.M., R R Mohamed, A Ahmad, N A Rawi, A Pattern-Based Password Authentication Scheme for Minimizing Shoulder Surfing Attack. International Journal on Advanced Science, Engineering and Information Technology, 2017. 7(3).

[16] Keszthelyi, A., About Passwords. Acta Polytechnica Hungarica, 2013. Vol. 10, No. 6.

[17] Recordon, D. and B. Fitzpatrick, OpenID Authentication 1.1. Finalized OpenID Specification, May, 2006.

[18] Celesti, A., et al. Three-Phase Cross-Cloud Federation Model: The Cloud SSO Authentication. in Advances in Future Internet (AFIN), 2010 Second International Conference on. 2010.

[19] Senk, C., Future of Cloud-Based Services for Multi-factor Authentication: Results of a Delphi Study, in Cloud Computing, M. Yousif and L. Schubert, Editors. 2013, Springer International Publishing. p. 134-144.

[20] Chaurasia, B., A. Shahi, and S. Verma, Authentication in Cloud Computing Environment Using Two Factor Authentication, in Proceedings of the Third International Conference on Soft Computing for Problem Solving, M. Pant, et al., Editors. 2014, Springer India. p. 779-785.

[21] Banyal, R.K., P. Jain, and V.K. Jain. Multi-factor Authentication Framework for Cloud Computing. in Computational Intelligence, Modelling and Simulation (CIMSim), 2013 Fifth International Conference on. 2013.

[22] Imran Naguru, N.K.R.B., Feature Matching in Iris Recognition System using MATLAB. International Journal on Advanced Science, Engineering and Information Technology, 2017. 7(5).

[23] Hahn, C. and J. Hur, Efficient and privacy-preserving biometric identification in cloud. ICT Express, 2016. 2(3): p. 135-139.

[24] Rathgeb, C. and A. Uhl, A survey on biometric cryptosystems and cancelable biometrics. EURASIP Journal on Information Security, 2011. 2011(1): p. 3.

[25] Markus Jakobsson, E.S., Philippe Golle, Richard Chow, Implicit authentication for mobile devices, in Proceedings of the 4th USENIX conference on Hot topics in security. 2009, USENIX Association: Montreal, Canada. p. 9-9.

[26] Jeong, H. and E. Choi, User Authentication using Profiling in Mobile Cloud Computing. Aasri Conference on Power and Energy Systems, 2012. 2: p. 262-267.

[27] Chow, R., et al., Authentication in the clouds: a framework and its application to mobile users, in Proceedings of the 2010 ACM workshop on Cloud computing security workshop. 2010, ACM: Chicago, Illinois, USA. p. 1-6.

[28] Ismail, R., Enhancement of Online Identity Authentication Though Blockchain Technology. 2017: Malaysia.

[29] Abdellaoui, A., Y.I. Khamlichi, and H. Chaoui, A Novel Strong Password Generator for Improving Cloud Authentication. Procedia Computer Science, 2016. 85: p. 293-300.

[30] Vukolić, M., The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication, in Open Problems in Network Security. iNetSec 2015. Lecture Notes in Computer Science. 2016, Springer.

[31] M. Ali, R.S., J. Nelson and M. J. Freedman, Blockstack: A New Internet for Decentralized Applications (Whitepaper). 2017.

[32] J. Nelson, M.A., R. Shea and M. J. Freedman, Extending Existing Blockchains with Virtualchain, in Workshop on Distributed Cryptocurrencies and Consensus Ledgers. 2016.

[33] M. Ali, J.N., R. Shea and M. J. Freedman. Blockstack: A Global Naming and Storage System Secured by Blockchains. in 2016 USENIX Annual Technical Conference. 2016.

[34] Atzori, M., Blockchain technology and decentralized governance: Is the state still necessary? 2015.

[35] Foundation, T.L., Hyperledger Overview. 2018.

[36] Panetta, R., & Cristofaro, Lorenzo, A closer look at the EU-funded My Health My Data project. Digital Health Legal, 2017. 10-11.

[37] Council of the European Union , E.P., Regulation (EU) 2016/679 of the European Parliament and of the Council Official Journal of the European Union, 2016.

[38] Jesse Leimgruber, A.M., John Backus, Bloom Protocol: Decentralized credit scoring powered by Ethereum and IPFS. 2018.

[39] Shocard, Identity Management Verified Using the Blockchain. 2017.

[40] Christian Lundkvist, R.H., Joel Torstensson, Zac Mitton, Michael Sena, UPORT: A Platform for Self-Sovereign Identity. 2016.

[41] Digital, I.O., I/O Digital Application Based Blockchain Whitepaper. 2016.

[42] BlockAuth, Powering a franchised network of OpenID Connect providers that verify user authentication and authenticity. 2014.

[43] Uniquid. Uniquid Blockchain Access Management. 2017; Available from: http://uniquid.com/.

[44] Charleen Fei, J.L., Eugeniu Rusu,Kasia Szawan, Kai Wagner, Natascha Wittenberg, Jolocom: Decentralization By Design. 2018.

[45] Blockchain, C. Identity compliance, simplified. 2018; Available from: https://www.cambridge-blockchain.com/.

[46] Legal, K. Blockchain identity verification. 2018; Available from: https://kyc.legal/en.

[47] Conner Fromknecht, D.V., Sophia Yakoubov CertCoin: A NameCoin Based Decentralized Authentication System. 2014.

[48] NameCoin. Namecoin. 2018; Available from: https://www.namecoin.org/.

[49] Authenteq. Identity Verification & KYC. 2018; Available from: https://authenteq.com/.

[50] Samaniego, M., & Deters, R. . Blockchain as a Service for IoT. in IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). 2016. IEEE.

[51] Kirk, J. Blockchain for Identity Management: It's Years Away. 2018 [cited 2018; Available from: https://www.bankinfosecurity.com/blockchain-for-identity-management-its-years-away-a-10598.