Blockchain for Identity Management: Review, Applications, Comparison, and Challenges

Priyanshi Agrawal
Viterbi School of Engineering
University of Southern California
Los Angeles, USA
pa92514@usc.edu

***Abstract*—Identity management systems require robust mechanisms that address security and privacy issues to efficiently work in the ever-increasing complex systems. The current methods like centralized identity management systems and large third-party identity providers, such as Google and Facebook, pose the problem of cross-domain authentication, interoperability difficulties, single point failure, data privacy, and scalability.**

Breakthrough in Blockchain, the decentralized ledger system, provides us with a solution, which single-handedly answers concerns like privacy and security, based on the immutability feature that has enabled us to address these problems.

Blockchain methodology enables building a new authentication paradigm called self-sovereign identity, allowing the user or entity to have ownership and control of their identification data without involving any other entities. The new approaches enhance de-centralization, transparency, and user control for managing identity information.

Although the advantages and application of blockchain technology for IoT devices, e-health, and banking sectors have enabled digital identity management in complex systems, it still cannot completely solve the above-mentioned problems.

This research aims to analyze various existing and proposed architectures for applications that use Blockchain technology and develop an understanding of how this technology can further improve security in systems.

Keywords— Blockchain, Identity Management, Decentralized System, Self-Sovereign Identity

I. INTRODUCTION

In today's world, digital Identity management technology serves as the basic security guarantee of computer and internet applications in all kinds of industries like banking and health sector. It is crucial that these systems provide secure and credible methods to carry out authentication tasks to avoid data breach, frauds and network attacks.

The conventional identity management systems which require user to enter user ID and password has posed problems of password management for different websites for the user and on other end the risks of system compromise which manages the passwords for all users. This centralized system creates risk of data breach of large databases and putting the user's personal information at risk. Even with the development of new technologies like Open Authorization (OAuth) which are widely used, there exists problems of central cross-domain authentication and involving third party in authentication [1].

Another problem that many citizens face is government-issued digital identities in different government data repositories is a way to spy on all citizens' daily activities. This leads to the actualization of concept of Self sovereign identity, which allows user to manage and maintain their digital identity.

To solve these problems, Blockchain technology has captured the attentions of academia and industry. Blockchain uses the distributed ledger technology which has the advantages like decentralization, peer-to-peer data management and security and hence is looked at as a way to solve problems traditional identity management systems.

Systems like uPort, Sovrin and ShoCard are developed. In this paper, these block chain-based models and proposed systems are analysed: what are they trying to solve and what are the limitations in terms of security.

The paper is organized as follows: Section II discusses the background about identity management,

authentication, block chain, decentralization, self-sovereign identity. Section III presents the blockchain based models review. Section IV presents the industry applications. Section V presents the analysis of the models. Section VI presents the future research challenges. Section VII presents the conclusions.

II. BACKGROUND

1. **Identity management**: Identity management refers to broad administrative area which supports entity authentication, authorization and de-provision of user account in timely manner. Sound identity management and governance are needed to manage identities for online services. In addition to excellent performance and scalability, access control and privacy protection are also necessary when designing identity authentication and management systems [15]. An identity management system typically performs the following tasks:
    a. Identity Registration: bind the specified identifier to the public key provided at registration.
    b. Identity Update: update the binding relationship between the identifier and public key.
    c. Identity Authentication: authenticate whether the given identifier and public key match each other.
    d. Identity Revocation: unbound the identifier from the public key [3].

2. **Independent identity management (Centralized Systems)**: The centralized identity management approaches do not give individuals complete ownership and control over their digital identities and the challenges that are faced are single point failure, huge data breach, reputation damage, identity fraud, and loss of privacy. Projects for decentralized identity management systems like Pretty Good Privacy / Web of Trust (PGP / WoT) and Blockstack have drawn a lot of interest from both academics and business.

3. **Federated identity management systems (Single sign on)**: It provides authentication and authorization capabilities across organizational and system boundaries. The user can sign up for service provider service and use the same identity to access any other services that are said to be available using the same identity. There is lower risk of replication of credentials as user account is managed independently by identity provider and no enterprise directory integration is required [8]. Examples of federated entities include Facebook and Google's single sign-on systems.

4. **Self-Sovereign Identity**: The idea of self-sovereign identification holds that people have complete ownership and control over their digital identities without the need for a middleman in the sense they can communicate with various service providers and control their identities on their own mobile devices or in the cloud [4]. This eliminates the risk of third parties that can result in data lost or misuse of sensitive information. Based on rules of need-to-know and need-to-retain, the owner of data can control the information [16]. This makes it possible for an ecosystem to develop that makes it easier for entities using these identities to propagate trust and collect and record attributes.

5. **Blockchain**: A blockchain is a shared distributed database or ledger between computer network nodes. Blockchain serves as an electronic database for storing data in digital form. It collects information in blocks, that have certain storage capacities and cryptographically linked. When new information is added a new block is created and added to the chain and becomes a part of irreversible timeline. The most well-known use of blockchain technology is for preserving a secure and decentralized record of transactions in cryptocurrency systems like Bitcoin.

    Block chain Terminology:
    1. Node and Block: Node is defined as a computer in a peer-to-peer network, which can be presented as the owner of transactions that are performed by a certain user.
       Block can be defined as an immutable page of distributed ledger in a blockchain. On getting consent of the transaction, a block is added in the blockchain.
    2. Consensus: The consensus mechanism is used to validate and process the transaction by approval of node decision.
       Widely used consensus algorithms include proof-of-work, proof-of-stake and Practical Byzantine Fault Tolerance.
    3. Scalability: In the current scenario, the node or performance scalability is provided depending upon access.
       Public blockchains such as Bitcoin and Ethereum provide node scalability whereas

Hyperledger, which is a private blockchain, offers performance scalability.

4. Smart Contract: The third-generation revolution of blockchain widened the application of blockchain in several different sectors, apart from asset management and cryptocurrency.

By defining arbitrary rules, the complex applications can be monitored by smart contract.

The functions in Ethereum smart steps and storage space. The gas cost is paid in cryptocurrency called ether.

5. Access: Based on consensus, there are three different types that a blockchain can be categorized into.

Public or permissionless blockchain provides the anonymity characteristic, but it lacks privacy, whereas the consortium and private blockchain is generally used at organization level.

The benefits of using Distributed Ledger Technology for Identity management include:

1. Decentralization: No single central authority owns or manages the ledger that contains the identity information [5].
2. Tamper-resistant: The DLT's historical transactions cannot be altered, and any updates to the data are made transparent.
3. Inclusive: New ideas for user identity bootstrapping can be developed that broaden the applicability of legal identities and lessen exclusion.
4. Cost-effectiveness: Sharing identification information has the potential to reduce the amount of duplicate personal information stored in databases and save money for relying parties.
5. User control: If users lose access to the services of a certain identity provider or broker, they cannot lose control of their digital identifiers.

The immutability of blockchain technology and consensus mechanisms are seen as the ideal solutions for distributed systems as they eliminate the need of centralized authorities to validate the transactions [17]. Modern blockchains offer smart contracts, which allow interacting parties to construct agreements based on preset rules and without the need for a trusted third party.[14]

New identity management strategies that attempt to challenge established methods for creating and using digital identities have emerged because of the blockchain data structure. In transactions involving identity information, these novel identity management (IdM) systems aim to increase decentralization, transparency, and user control [2].

III. Related Work And Proposed Methods

From the standpoint of digital identity management and the security and privacy of personal data, numerous systems and architectures have been created and proposed leveraging blockchains. Some of the examples of Self Sovereign identity include Sovrin, uPort, and OneName and examples of decentralized trusted identity include ShoCard, BitID, ID.me, and IDchainZ. In the following section we will discuss how proposed models try to use DLT for IDM and what issues are solved and what are the challenges.

There is no definitive evaluation scale available for the evaluation of proposed solutions and many of them have been evaluated and compared with other solutions based on a widely known framework law of identity. The laws involve: User control, minimal disclosure, justifiable parties, design for interworking technologies, consistent user experience, human integration, directed identity [2].

1. **uPort**: "Decentralized identification for all" is the goal of the open-source decentralized identity system known as uPort. Its uses Ethereum DLT for upcoming decentralized applications and for centralized services like email and banking.

   The users' and trustees' uPortIDs are linked in the underlying design, which creates the possibility for collaboration against a particular uPort user, and the trustees themselves could be one attack vector.

   It uses the JSON data structure and different attributes can be encrypted individually but it might reveal metadata about particular characteristics or connections to identity providers or reliant parties. And the privacy of user information in JSON data structure on message server can be compromised [2].

2. **Jolo**: Jolo is another self-sovereign identity management which is similar to uPort based on hierarchically deterministic keys (HD keys). It is also developed on top of Ethereum. The only difference between them representation of data in the systems and possess similar confidentiality issue [16].

3. **CertCoin**: CertCoin is a system that incorporates the best aspects of transparent Certificate Authorities and of the Web of Trust. This decentralized authentication system based on the NameCoin blockchain [8]. NameCoin is

a cryptocurrency designed to act as a decentralized DNS for .bit addresses.
CertCoin is a viable PKI, capable of replacing Certificate Authorities and PGP Webs of Trust. Our construction benefits from an entirely decentralized architecture offering inherent fault tolerance, redundancy, and transparency. Despite this, CertCoin supports the expected features of a full-fledged Certificate Authority including certificate creation, revocation, chaining, and recovery [7].

4. **Sovrin**: Sovrin is a permissioned ledger-based open-source identification network that stores identity records. It aims to equip users to fully control all aspects of their identity. Sovrin is public, but only stewards—trusted organizations that can include banks, colleges, governments, and other entities—can run nodes that participate in consensus processes, making the ledger permissioned [2].
   In this system, third party is not involved but users must rely on stewards which act on their behalf in the Sovrin Network. Every user can choose which attributes to disclose to the relying party for identity validation process to realise full control. But this possess risk as a lot of information could be with agency.

5. **iResponder**: iResponder is a biometric provider (based on Biometric Open Protocol Standard) that combines the biometric with blockchain technology [16]. Through its collaboration with the Sovrin Foundation, it provides users with a distributed identity. The iris scan data is kept on a private server, and a 12-digit random string serves as the private key for each distinct template.

6. **ShoCard**: ShoCard is a digital identity card on a mobile device. The user scans the document like passport or driver's license and sign it on the mobile app. The app generates a private and public key to seal the record. The cryptographic hashes are stored in Bitcoin transactions for later use. ShoCard's is used for verification of identity in face-to-face and online interactions [2].
   The delivery of encrypted certifications between ShoCard users and relying parties is managed by the ShoCard central server, which serves as an intermediary. By doing this, ShoCard poses less risk than if it had kept and disseminated identity data in plaintext.
   Storage for certificates using symmetric encryption is provided by the ShoCard server (known as envelopes). To allow the user to only share certifications with specific individuals, ShoCard never learns the encryption key. This information is encrypted in the card, this gives confidence that transaction information sharing is done with relying parties and is controlled by the end user.
   However, ShoCard's intermediary role does create uncertainty if company ceased to exist, users will not be able to use the system and in essence makes this system centralized.

7. **BIdM**: is a decentralized cross-domain identity management system based on blockchain. The aim is to solve the authentication center's single point of failure problem and improve the authentication performance for many cross-domain authentication scenarios [3].
   This model uses a decentralized identifier (DID) consists of a uniform resource identifier (URI) and a document. URI is used as an identifier, and the document stores the user's public key, identity information, and other data. When a user sends a message, the recipient can verify the user's public key information by looking up the DID to determine whether the message has been tampered with.
   In BIdM, the function of CA is function is provided by Identity provider which is a trusted authority such as the government, bank, individual, or private enterprise. DID is master identity which does not allow to trace information to IdP (shadow identity). This of separating user identifiers from specific user identity attributes is called a two-level identity architecture which helps in protecting user privacy.
   The performance testing of BIdM prototype is done on Ethereum. And the results show that the user's private key and accumulator credentials cannot be obtained by attackers, even if the negotiated key is exposed during an authentication session. As a result, they are unable to compromise the integrity of the signature while altering communication data. By using this functionality, BIdM can stop man-in-the-middle attacks. The simulated authentication gives 100% accuracy, but they are not enough to justify system's security for malicious attacks.

IV. APPLICATIONS

1. **Edge computing/IOT devices**:

   Internet of Things (IoT) provides ubiquitous connection between entities and realizes intelligent perception, recognition and

management of objects. IoT architecture usually relies on cloud computing platform to enhance data storage, transmission and analysis capabilities. Edge computing has benefits, but it also poses increased security and privacy risks. Messages passing on open edge networks are easily susceptible to eavesdropping and even tampering by malicious adversaries. Additionally, some IoT applications save a lot of personal data, such the name, ID number, bank card number, mobile phone number, etc. of the user.

Yifan et al. [6] present a privacy-preserving three-factor authentication scheme in the edge computing environment. Blockchain technologies are integrated to manage the public key of entities and track misbehave users. Smart contracts make it flexible to add or revoke users in the system. Security analysis and performance evaluation is presented to demonstrate that proposed protocol satisfies security requirements but in order to address the demands of authentication and privacy protection in many application situations future research will focus on using more sophisticated techniques such as group signature and zero knowledge proof, etc.

M. Mamdouh et al. [9] describes various available IoT device authentication mechanisms and identity management approaches including blockchain. The study showcases IoT devices may be maintained automatically, and the transmitted data can be synchronized over the blockchain. Also, blockchain-based solutions can defend IoT devices against a variety of threats, including data fishing and cache theft. It has several benefits outside of the Internet of Things, including stability, quick transactions, decentralization, time immutability, transparency, and high security for shared data.

2. **Open banking ecosystem**:
   The banking and financial industries are evolving toward "Open Banking," which can maximize client benefits through data exchange. A primary goal is to ensure consent when sharing customer financial data with different organizations. Blockchains can protect data integrity and customer privacy by controlling access, securing transactions, and making data tamper-proof.
   Blockchain-based identity management and access control (BIMAC) framework is based on a comprehensive review of open banking (OB) requirements. The framework was designed to meet the challenges associated with multiple digital identities and the need to monitor the use of personal information by third parties. The OB requirements are addressed to securely decentralize data and provide privacy-preserving application interfaces.
   The BIMAC framework offers complete access control while satisfying customers' "right to be forgotten". The decentralized identity integration module supports a change in identity management from a server-centric system to a decentralized self-sovereign system. A decentralized access control module supports the use of third-party service providers to provide novel integrated services while protecting privacy at a high level of security.
   Blockchain-based authentication and data sharing between banks and users help to improve KYC (know your customer) and anti-money laundering procedures, according to the World Bank. [10]

3. **E-health systems**:
   X. Xiang et al. [11] present a new Blockchain-based access control protocol for secure sharing in IoT-based medical applications in an e-health system. Users and service providers can generate pairwise secret keys in the early stages of the authentication and access control phase, allowing them to build session keys for secure communication. The efficiency and robustness of the protocol are analysed by BAN logic, demonstrates that the protocol has strong resilience to various adversarial attacks.

4. **Cloud-native applications**:
   Cloud-native uses MTLS (Mutual Transport Layer Security) based on PKI (Public Key Infrastructure) to provide an important guarantee for cloud-native network security. The heavy certificate management gradually makes the Public Key Infrastructure (PKI) a bottleneck for these systems. Although using certificateless public key techniques in cloud-native systems has many benefits, identity management is a key issue that cannot be ignored. Li et al. [13] proposes blockchain as a trust for service identity and public key distribution. Framework mostly concentrates on the management problem after the public key generation, ignoring the public and private key generation component of the certificateless encryption technique.

V. ANALYSIS

The various blockchain based identity solutions claims to fulfill the self-sovereignty and disrupt the traditional approach of identity management. Implementing proof of concept for solutions based on the self-sovereign identification principle that are currently being used as evaluation criteria is part of the ongoing research in this field. However, the importance of open standards, network scalability, and flexible identity, as well as the adoption of creative solutions and user-empowering identity objectives such as giving full control, has been emphasized.

P. Dunphy et al. [2] describe how uPOrt, Sovrin, and ShoCard adapt different approaches. As of now, none of the schemes that are reviewed are supported by a fresh, empirically based understanding of user interaction.

Sovrin uses blockchain technology in a manner that it relies on trusted stewards which is similar to depending on trusted third parties. If the ShoCard's intermediary role ceased to exists, the users will not be able to use the system as in case of centralized methods. The original challenge to provide usable end user key management is largely unaddressed. According to recent research, key management is still a major source of worry for Bitcoin users. Also, the promising idea of key recovery was put forth in uPort and Sovrin, digital identity strategies that do away with central authorities and rely on strong key management techniques from their users run the risk of alienating non-technical users who will be powerless to recover resources when things go wrong.

The classification and protection of data privacy, access based on customer consent, data for tracking TSP access, open API consensus mechanisms and administration, and decentralized self-sovereign identities are the main topics of current blockchain-based studies.

VI. CHALLENGES

Following are the challenges and trade-offs in building a feasible and effective Identity management that needs to be addressed:

1. Smart Contract Engineering: Externally imported data could be redundant, manipulated, bogus, and malicious, therefore, the smart contract should provide a mechanism to prevent the entry of unintended data on the blockchain.
2. Security and Privacy: The interaction between stakeholders on-the-chain and unknown stakeholders off-the-chain could potentially raise the privacy and security concerns. Therefore, new procedures are needed to solve both internal (such as DAO attacks, re-entry assaults, or DDoS attacks) and external (such as bringing fraudulent or corrupt data on the blockchain, reputation manipulation attacks, or identity theft attacks) security and privacy vulnerabilities.
3. Design Challenges: Blockchain oracles are always at risk of centralization, collusion, and Sybil attacks. A better approach is to enable multiple oracles reporting the same data based on on-chain consensus mechanism. The issue of scalability will also arise when blockchain needs to attest to the massive data streams from multiple oracle systems.
4. Scalability: Scalability of a blockchain is composed of two factors - node scalability and performance scalability. Node scalability refers to the extent to which a network can upload more nodes without a loss in performance. Performance scalability refers to the number of transactions processed per second. None of the currently used blockchains are scalable. Using proof-of-work (PoW) consensus processes, public blockchains like Bitcoin and Ethereum make this trade-off in favor of node scalability.
5. Flexibility: It is important to consider the technical and research work done on the portability of the digital identity. When existing platforms disappear for various causes, it must ensure a seamless transfer of identification with the least amount of identity data to new platforms.

CONCLUSION

Inclusion of blockchain distributed ledger technology into current software architecture brings both quality improvements and limitations. The immutability and transparency can ensure data integrity, while the underlying decentralization enhances the availability of whole system. In the above-described models and applications it can be observed that Distributed Ledger Technology is used for different techniques of decentralization, but the actual implementation does not provide complete and better authentication and Identity Management solutions as compared to traditional solutions.

REFERENCES

[1] Z. Song and Y. Yu, "The Digital Identity Management System Model Based on Blockchain," 2022 International Conference on Blockchain Technology and Information Security (ICBCTIS), 2022, pp. 131-137, doi: 10.1109/ICBCTIS55569.2022.00040.

[2] P. Dunphy and F. A. P. Petitcolas, "A First Look at Identity Management Schemes on the Blockchain," in IEEE Security & Privacy, vol. 16, no. 4, pp. 20-29, July/August 2018, doi: 10.1109/MSP.2018.3111247.

[3] R. Chen et al., "BIdM: A Blockchain-Enabled Cross-Domain Identity Management System," in Journal of Communications and Information Networks, vol. 6, no. 1, pp. 44-58, March 2021, doi: 10.23919/JCIN.2021.9387704.

[4] Y. Liu, Q. Lu, H.-Y. Paik, and X. Xu, “Design patterns for blockchain-based self-sovereign identity,” Proceedings of the European Conference on Pattern Languages of Programs 2020, 2020.

[5] R. Rana, R. N. Zaeem, and K. S. Barber, “An assessment of Blockchain Identity Solutions: Minimizing risk and liability of authentication,” IEEE/WIC/ACM International Conference on Web Intelligence, 2019.

[6] Y. Wang, X. Jia, Y. Xia, M. K. Khan, and D. He, “A blockchain-based conditional privacy-preserving authentication scheme for Edge Computing Services,” Journal of Information Security and Applications, vol. 70, p. 103334, 2022.

[7] H. Nusantoro, R. Supriati, N. Azizah, N. P. Lestari Santoso, and S. Maulana, “Blockchain based authentication for Identity Management,” 2021 9th International Conference on Cyber and IT Service Management (CITSM), 2021.

[8] S. Y. Lim, P. Tankam Fotsing, A. Almasri, O. Musa, M. L. Mat Kiah, T. F. Ang, and R. Ismail, “Blockchain technology the identity management and Authentication Service Disruptor: A survey,” International Journal on Advanced Science, Engineering and Information Technology, vol. 8, no. 4-2, p. 1735, 2018.

[9] M. Mamdouh, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, “Authentication and identity management of IOHT devices: Achievements, challenges, and Future Directions,” Computers & Security, vol. 111, p. 102491, 2021.

[10] C.-H. Liao, X.-Q. Guan, J.-H. Cheng, and S.-M. Yuan, “Blockchain-based identity management and Access Control Framework for open banking ecosystem,” Future Generation Computer Systems, vol. 135, pp. 450–466, 2022.

[11] X. Xiang, J. Cao, and W. Fan, “Decentralized authentication and access control protocol for blockchain-based e-health systems,” *Journal of Network and Computer Applications*, vol. 207, p. 103512, 2022.

[12] V. Dehalwar, M. L. Kolhe, S. Deoli, and M. K. Jhariya, “Blockchain-based trust management and authentication of devices in Smart Grid,” Cleaner Engineering and Technology, vol. 8, p. 100481, 2022.

[13] X. Li, J. Zhang, X. Niu, and J. Guan, “Blockchain-based Certificateless identity management mechanism in cloud-native environments,” Proceedings of the 2021 ACM International Conference on Intelligent Computing and its Emerging Applications, 2021.

[14] H. Al-Breiki, M. H. U. Rehman, K. Salah and D. Svetinovic, "Trustworthy Blockchain Oracles: Review, Comparison, and Open Research Challenges," in IEEE Access, vol. 8, pp. 85675-85685, 2020, doi: 10.1109/ACCESS.2020.2992698.

[15] D. Pöhn and W. Hommel, “An overview of limitations and approaches in identity management,” Proceedings of the 15th International Conference on Availability, Reliability and Security, 2020.

[16] K. Gilani, E. Bertin, J. Hatin and N. Crespi, "A Survey on Blockchain-based Identity Management and Decentralized Privacy for Personal Data," 2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS), 2020, pp. 97-101, doi: 10.1109/BRAINS49436.2020.9223312.