Design and Analysis of Self Sovereign Zero Knowledge Identities on blockchain

Mohsin Hafeez
Department of Computer Science
LUMS
Lahore , Pakistan
24020477@lums.edu.pk

Zartash Afzal Uzmi
Department of Computer Science
LUMS
Lahore, Pakistan
zartash@lums.edu.pk

Naveed Ul Hassan
Department of Computer Sciences
LUMS
Lahore, Pakistan
naveed.hassan@lums.edu.pk

***Abstract*—With the rise of the popularity in blockchain, there has been a major shift in its adoption towards real-world applications. In this paper, we explore the already available identity models and propose new methods to increase their scalability. It also talks about the "Laws of Identity" and how Self-Sovereign Identities abide by these laws. We explored different blockchains for identity models and proposed a framework to analyze their architecture and constraints regarding the Identity models. In the end we also discussed some real world applications and the companies currently working on decentralized identities.**

***Index Terms*—Blockchain, Self-Sovereign Identity, Zero-Knowledge Proofs, Distributed Ledger Technology**

I. Introduction

One of the biggest innovations of 21st century, the emergence of Blockchain technology, has the potential to convert the already present centralized systems into decentralized and secured protocols. This enables users to control their data in a more secure and transparent way. Every action that happens on the blockchain is stored in the form of a transaction that could be verified later on. The use cases of blockchain are not only limited to the cryptocurrency rather they are being used in many other useful applications as well. Some examples of these include use cases in DAOs, DeFi and Financial Derivatives and as well as in notary and land record management systems. One of the potential use cases in digital services is the use in a digital identity system by making it more secure and user governed. With the increase in online presence and digital products, it's really important to protect one's identity.

Digital identity refers to the online representation of an individual's identity, including personal information such as name, age, and address. In today's digital age, the use of digital identity is becoming increasingly prevalent as more and more transactions and interactions are conducted online. However, the reliance on digital identity also presents new challenges, particularly in terms of security and privacy.The only method available as of now is to authenticate the identity using the username and password. And for each app or service that we use there is a different Username and Password. Moreover, it has become really hard for organizations to verify an individual's identity.

Blockchain technology has emerged as a potential solution to these challenges, offering a decentralized and secure way to store and manage digital identity. By leveraging the decentralized nature of blockchain, individuals can have greater control over their personal information and who has access to it. In this research paper, we will explore the concept of digital identity and its importance in the digital age. Later on we'll drive a framework like Chianmaster to classify the different existing Blockchain identity solutions according to the requirements and the functionalities. We will also delve into the role of blockchain in enabling secure and efficient verification of digital identity, as well as the potential benefits and challenges of using blockchain for this purpose. After the advent of the internet, it was really important to make a set of attributes for the identity of a person. Without knowing someone's identity on the internet exposes us to different serious exploits and growing dangers.

II. Laws of Identity Kim Cameron

The "Laws of Identity" are a set of principles for identity and identity management systems, proposed by Kim Cameron, then Chief Architect of Identity at Microsoft. The laws are intended to provide a framework for understanding and designing identity systems that are secure, privacy-preserving, and user-centric.

A. User control and consent

Technical identity systems should only disclose information that identifies a user if the user has given their consent for the information to be revealed in order to protect the user's personal information and ensure that it is not shared without their permission

B. Minimal Disclousre for Constrained Use

The best long-term solution is the one that reveals the least amount of identifying information and limits its use to the greatest extent.

C. Justifiable Parties

Digital identity systems should be designed only to allow the disclosure of identifying information to parties who have a legitimate and justified role in a specific identity relationship

D. Directed Identity

A universal identity system should have both "omni-directional" identifiers for public entities and "unidirectional" identifiers for private entities to allow for discovery while protecting against unnecessary disclosure of correlation handles. In other words, the system should have identifiers that can be used by anyone to find and identify public entities, and identifiers that can only be used in one direction (from the private entity to the public entity) to protect the privacy of the private entity

E. Pluralism of operator and operating systems

A universal identity system should be able to support the integration and collaboration of different identity technologies used by multiple identity providers

F. Human integration

The universal identity metasystem is designed to safeguard against identity threats by establishing humans as part of a decentralized system that communicates with machines in a clear and precise manner

G. Consistent experience across contexts

These laws are intended to provide guidance for designing identity systems that are secure, private, and user-centric. They are not intended to be a strict set of rules that must be followed, but rather a set of principles that should be considered when designing and implementing identity system.

III. Identity Management System (IMS)

An identity management system (IMS) is a software application or set of applications that enable an organization to manage and control access to network resources by enabling the creation, maintenance, and use of user identities. An IMS typically includes a database of user information and a set of tools for managing user access to network resources. There are different types of Identity Management Systems that have evolved over the year. In this paper, we will discuss the implications of the Self Sovereign Identities and how they can be used to integrate with the existing blockchains with improved security, efficiency and interoperability.

IV. Digital Identity

Digital identity is a set of credentials that are used by individuals to give the same level of trust as in a face-to-face interaction. This level of trust is obtained by a set of attributes that are associated to a person and are stored in some databases. These systems also have the ability to differentiate different identities based on the credentials provided. There are certain characteristics that are fixed and cannot be changed, such as ethnicity, gender, and date of birth. In addition to these static attributes, there are also characteristics that can change over time, such as age, job position, and address. Some attributes are specific to an individual and help to uniquely identify them. Attributes that uniquely identify an individual, such as email address and passport number, are called identifiers. Within the identity ecosystem, a single identity can have multiple attributes, some of which may be shared by multiple users. Email is the most commonly used identifier for online services among ordinary users

V. Terminologies

A. Abbreviations and Acronyms

In this Paper we are using different Abbreviations and terminologies such as DID (Decentralized Identifier), IdP (Identity Provider) , SP(Service Provider), zk-proofs (Zero Knowledge Proofs)

B. Terminologies

- **Entity :** Objects that have a distinct existence, such as individuals and organizations, are referred to as entities. These entities can be either physical or logical. They are divided into three categories based on their function: identity providers, service providers, and users. Identity providers are responsible for verifying the identity of users, service providers offer services to users, and users consume these services
- **Attribute :** An attribute defines the traits of an entity that must have two basic properties of Possession and Endorsement.
- **Identifier** An identifier is a characteristic that can be used to identify something within a specific context uniquely. In the case of blockchain-based identity solutions, the blockchain address often serves as the identifier and is controlled by the corresponding private key.
- **Claim**Claim is a digital signature by some identity providers stored in a smart contract and works as the endorsement for the attributes of the user.
- **Hashchain** A hashclaim is a hash value that combines a user's public key, an identifier for a specific attribute, and a random value called a salt. This hash value acts as proof that the attribute has been given to the owner of the public key that is used in the hashclaim. The hashclaim allows for verification of the transfer of the attribute.

VI. Conventional Identity Models

Multiple identity models have been introduced in the past where different identities and attributes were used to connect people online. These are as following :

A. Centralized Model

Most internet identities are controlled by a single entity, which means that the user's credentials are owned and managed by that entity. However, this model has some drawbacks. Users do not have control over their own identity records, and they can be taken away or misused by the identity provider.

B. Federated Identity Model

Federated identity management systems allow for authentication and authorization across different organizations and systems by establishing agreements that recognize an individual's identity at one provider and establishing agreements on data ownership between providers

C. User-Centric Model

Self Sovereign Identity Model) (The type of identity model where identity and the data is user-controlled and owned.

In this research paper we only focus on the Self Sovereign Identity model. The following section covers the properties for a self-sovereign identity.

VII. Self-Sovereign identity

Self-Soverign Identity (In Search of Self-Sovereign Identity Leveraging Blockchain Technology) . Self-sovereign identity (SSI) is a concept in the field of digital identity that refers to individuals or organizations having full control over their own digital identity and personal data. In an SSI system, the individual or organization is the sole owner and custodian of their own identity, and has the ability to disclose personal information to other parties as needed selectively.

One of the main principles of SSI is decentralization, meaning that there is no central authority that controls or manages identity information. Instead, identity information is stored and managed on a distributed ledger or another decentralized platform, and is controlled directly by the individual or organization. This allows for greater privacy and security, as the individual has full control over who has access to their personal data.

A. SSI management on the Blockchain :

SSI has the potential to revolutionize the way that individuals and organizations interact online by giving them greater control over their personal information and enabling more secure and trusted online transactions. It has applications in a wide range of areas, including finance, healthcare, education, and government. In this paper we'll see how SSI are generated using different blockchains and how they can be integrated in different real life purposes. There are some characteristics of the SSI which also align with the laws of the identity. These are as follows :

- The existence (User must have the autonomy to manage identity)
- Ownership (User should be the owner of his own identity)
- Access (User should have unrestricted access to their identity)
- Transparency (Systems that verify the identity should be open source)
- Persistence (identities should exist as long as the user wants)
- Portability (Data about identity should be easily portable)
- Interoperability (Identities should be used anywhere cross borders)
- Consent (Users should have the authority to allow or disallow use of their identity and data)
- Minimization (discourse of data should be minimized as much as possible)
- Protection (Right of entities must be protected in any case)

In this paper, we will be only discussing the already available models for the SSI identities and the possible solution for better security and efficiency.

VIII. Blockchain-Based Identity Models

A. Identity Model on Ethereum

In the preliminary model (Self-Sovereign Identity using Smart Contracts on the Ethereum Blockchain) the user deploys two smart contracts on the blockchain. The contract after deployment returns a DID which is also a universally unique identifier not the same for two persons. This is done by using an identity contract, and it stores users three main things :

- Owner key
- Recovery contract
- IPFS hash

The owner key refers to the public portion of a key pair stored on a user's device. The recovery contract is the address of a contract that is used to recover user data. Any requests to modify user data will only be accepted if they are sent from the specified public key or recovery contract. The steps to create identity are as follows :

- **Identity Creation**The process starts with generating a public-private key pair on their device that is being used as an identifier (pubic key) and for signing the transactions (private key). The identity smart contract also publishes a recovery contract that contains logic for recovering an identity in case of a mishap
- **User Attributes** User can add different attributes related to their identity such as name . address, email, and mobile phone and sign it with their private keys. The user data is then stored in form of a JSON on IPFS. The uploaded content on IPFS generates a hash value which is signed by the user to update their identity.
- **Third-Party Attributes:**The attributes issued by the third party are also integrated in this identity model. These attributes are verified b a trusted third party. For getting these attributes, it's really important that user requests third party to sign the attributes and then save it in their wallet along with the signature.

This is a very simple SSI model on the blockchain which has a limited output due to it's implementation on public blockchain Ethereum. Moreover, there is also an opportunity to improve the security and the privacy of the user by using zero-knowledge proofs. In the below-improved identity model we'll cater to all of these requirements.

IX. Improved Zero-Knowledge Identity Model:

In the improved identity model when the identity provider (IdP) creates a name-value pair. The additional step here that improves the security of the protocol is using the ZKP (zk-SNARK) for better privacy. The hashclaim issued by the IdP will transfer the possession of the attribute without showing this to any other user on the system. So for example if Alice

Characteristics	Hyperledger	Ethereum	Polygon
Permission Restrictions	Permissioned	Permissionless	Permissionless
Access to Data	Private	Could be Private + Public	Public
Consensus	PBFT	PoS	PoS
Scalability	High Performance, low node scalability	High Node but Less Performance	High Node and Scalable
Governance	Open Governance Model	Core Group+Voting	Open Governance
Anonymity	Encryption of Data (Pseudonym)	Pseudonmity	Pseudonmity + Encryption of Data
Native Currency	No	Eth	Matic
Scripting	Turing Complete + Highly compatible	Scripting Compatible + Turing Complete	Smart Contracts + EVM compatible
Compatible Programming Language	Java , GoLand and Rust	C++ , Solidity, Rust , Golang	Substrate, Rust, Solidity , C++

reqiuests for a loan from bank the bank wil respond to Alice's request without showing this to any other user on the system and this happens using the hashclaim generated by the ZKP (zk-SNARK). The possession of an attribute is masked in a hashcalim and its endorsement comes from the IsPs action using ZKP.

This improved identity model could be further improved for throughput using the L2 scaling solution but in that case there is a trade off between security and the scalability that needs to be considered before taking any decision. The mathematical explanation of this is given in this paper. In the below section we'll look into the newly introduced Zero-Knowledege Idenity on Polygon which has a much better throughput and has a tendency to scale it on a larger number of users.

X. Polygon L2 zero knowledge identity:

Polygon claims to give the Ethereum Layer 1 security using different cryptographic tools and now with the integration of zk-Prooofs they have introduced a new scalable identity solution that could be integrated with the financial institutions, DeFi, Smart cities and with the health care management system. It uses Iden3 (idenity protocol) and for zk-proofs it uses the circom 2.0 toolkit to reduce the complexity and the efficiency of the system. There are few characteristics of the Polygon ID. These are as following :

- Decentralized and Self Govern
- Private by default
- d-Web reputation system integrated
- Open and permissionless
- Verifiable credentials instead of NFTs

XI. Improved Security and Efficiency by using ZoKrates

Integration of ZoKrates as discussed in the above models, for offloading some on-chain data and further using it in the zk-Proofs has improved the overall efficiency and the privacy of the system. This step will gain the trust of users more and allow them to adapt these identity models. ZoKrates reduces the complexity of the Zero-knowledge proofs and give a higher-level programmable view of what needs to be done. It also reduces the typical running time required and provides better integration with the other Dapps and DeFi protocols or any other service providers.

XII. Hyperledger Indy Model:

Permissioned blockchain hyperledger indy also provides an open-sourced, permissioned Self Sovereign Identity infrastructure on the distributed ledger. It has multiple frameworks that are being extensively used for the adoption of decentralized identities using Sovrin. An implementation of the Indy is done by the KivaProtocol in Sierra Leone in 2019, where they helped register the 5 Million people living below the poverty line for providing them with microloans. The whole process that used to take 4+ hours was reduced to a 5 minutes process with the Hyperledger Indy. They selected Indy because it was fast, cheap and secured blockchain. With all of this though, they were able to implement the Digital Identity system on Blockchain but they had a central governance body involved with no Self Sovereign Identity system. Technically they used blockchain as the fastest way to register and then authenticate the users.

XIII. Preliminaries:

A. Zero-Knowledge Proofs

Zero-knowledge proofs allow a party (the prover) to demonstrate to another party (the verifier) that they are aware of a certain secret value, without disclosing the value itself. The purpose of these proofs is to enable the prover to show that they have certain knowledge without providing any additional details about the knowledge in question. The one that are used in both of our above protocols is ZK-SNARK

XIV. Use Cases

A. DeFi

The identity management system can be integrated with the different DeFi protocols which allow certain activities based on the attributes present in the user's Identity. Moreover, to make it more practical we need to work on the Interoperability of the cross-chain data.

B. Integartion with IoT

With the advent of the internet more and more devices are connected on the internet. IoT devices are resource constrained so we need a protocol that can verify the IDs and send relevant data with secured encryption. The models discussed above can help us achieve this scalability without revealing user data. An IoT device node is a device that is used to transmit data, such as sensor readings or other types of information, as blockchain transactions to an IoT base station. These device nodes do not typically perform calculations or participate in proof of

work mechanisms, but they do typically encrypt the data they transmit in order to protect it from unauthorized access or tampering. The base station is responsible for recording the data on the blockchain and performing any necessary calculations

C. Land Registry

Self Sovereign Identities can be used in the Land registry management system where privacy and security remain intact and that too with the transparent possession of the attributes

D. Smart Cities

In the modern digital world, the concept of smart cities has gained a lot of traction but for it to properly work, it's really important to identify someone's identity by the Government. SSI can be integrated here using the blockchain.

XV. Related Work

Multiple companies are working on these digital identities on blockchain for efficient and secured identification and onboarding. Some of these are given below :

a) Evernym: The Evernym team has developed a publicly distributed ledger called Sovrin that is specifically designed for self-sovereign identity and privacy-preserving cryptography. This ledger is permissioned, meaning that access to it is restricted to certain individuals or groups

b) uPort ID: uPort is a decentralized identity system that allows individuals to create a secure digital identity and manage their personal information on a blockchain. uPort IDs are unique digital identities that are stored on the Ethereum blockchain and can be used to identify and authenticate users online. With a uPort ID, users can securely and privately share their personal information, such as their name, email address, and location, with other parties in a way that is verifiable and cannot be forged. uPort IDs can be used for a variety of purposes, including logging into websites, verifying the identity of individuals online, and conducting transactions on the blockchain

c) Trusted Key Identity: Trusted Key Identity is a company that helps manage and verify online identities through the use of Ethereum technology. They provide a mobile app and web-based services to confirm the identity of users, allow them to log in without a password, and prevent identity fraud through secure transactions and document signing. These services are used for onboarding, customer due diligence (KYC), and verifying identity and documents

XVI. Conclusion

Blockchain technology enables processes and transactions to be more transparent, decentralized, democratic, and secure by eliminating the need for a third-party organization to be involved. It is likely to play a crucial role in the near future as a method of digital identity, allowing individuals to authenticate and prove their identity for various digital services in today's interconnected world. In this paper we discussed the important underlying concepts in identity management and blockchain technology. We discussed the important protocols on Ethereum, Hyperldeger and the Polygon with an analysis of their functionalities. Each approach to identity management has its own advantages and disadvantages. For example, self-sovereign identity systems offer more control over one's identity, decentralized identity systems using blockchain technology offer decentralization, and systems that allow for easy verification by multiple entities. This paper presents a way to use zk-SNARK, a tool for proving knowledge of secret information without revealing the information itself, in an existing model for managing identities in a way that preserves privacy. The method involves the use of privacy attribute tokens to represent certain attributes that a user wants to keep private and the secret transferring ownership of these tokens between users. Two particular computations are employed to facilitate this transfer of ownership and to confirm that a user is the owner of the privacy attributes represented by the tokens. This allows individuals to prove they have certain attributes without revealing their identity or the values of those attributes. This paper gives opportunity to explore more into use of Zero Knowledge Proofs and layer 2 solutions for identity management systems.

References

[1] Dunphy, Paul, and Fabien AP Petitcolas. "A first look at identity management schemes on the blockchain." IEEE security privacy 16.4 (2018): 20-29.

[2] Rivera, Rogelio, et al. "How digital identity on blockchain can contribute in a smart city environment." 2017 International smart cities conference (ISC2). IEEE, 2017.

[3] Bhattacharya, Manas Pratim, Pavol Zavarsky, and Sergey Butakov. "Enhancing the security and privacy of self-sovereign identities on hyperledger indy blockchain." 2020 International Symposium on Networks, Computers and Communications (ISNCC). IEEE, 2020.

[4] ZoKrates - Scalable Privacy-Preserving Off-Chain Computations

[5] Analysis of Identity Management Systems Using Blockchain Technology

[6] Liu, Yue, et al. "Design pattern as a service for blockchain-based self-sovereign identity." IEEE Software 37.5 (2020): 30-36.

[7] Team, Polygon. "Introducing Polygon ID, Zero-Knowledge Identity for Web3 - Polygon: Blog." Polygon, Polygon, 4 Aug. 2022, https://polygon.technology/blog/introducing-polygon-id-zero-knowledge-own-your-identity-for-web3.

[8] Sora Identity: Secure, Digital Identity on the Blockchain

[9] Pöhn, Daniela, and Wolfgang Hommel. "An overview of limitations and approaches in identity management." Proceedings of the 15th International Conference on Availability, Reliability and Security. 2020.

[10] Evernym. "The Self-Sovereign Identity Company." Evernym, 23 Nov. 2022, https://www.evernym.com/.

[11] Design Pattern as a Service for Blockchain-Based Self-Sovereign Identity

[12] "Decentralized Identifiers (Dids) v1.0." W3C, https://www.w3.org/TR/2021/WD-did-core-20210103/.

[13] A zero-knowledge-proof-based digital identity management scheme in blockchain

[14] Cameron, Author Kim. "The Laws of Identity on the Blockchain." Kim Cameron's Identity Weblog, 27 May 2018, https://www.identityblog.com/?p=1658.

[15] Zhu, Xiaoyang, and Youakim Badr. "Identity management systems for the internet of things: a survey towards blockchain solutions." Sensors 18.12 (2018): 4215.

[16] Schäffer, Markus, Monika di Angelo, and Gernot Salzer. "Performance and scalability of private Ethereum blockchains." International Conference on Business Process Management. Springer, Cham, 2019.

[17] Ferdous, Md Sadek, Gethin Norman, and Ron Poet. "Mathematical modelling of identity, identity management and other related topics." Proceedings of the 7th International Conference on Security of Information and Networks. 2014.
[18] Lin, Chao, et al. "A new transitively closed undirected graph authentication scheme for blockchain-based identity management systems." IEEE Access 6 (2018): 28203-28212.
[19] Mingxiao, Du, et al. "A review on consensus algorithm of blockchain." 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, 2017.
[20] Schäffer, Markus, Monika di Angelo, and Gernot Salzer. "Performance and scalability of private Ethereum blockchains." International Conference on Business Process Management. Springer, Cham, 2019.
[21] Stokkink, Quinten, and Johan Pouwelse. "Deployment of a blockchain-based self-sovereign identity." 2018 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE, 2018.
[22] Zhu, Xiaoyang, and Youakim Badr. "Identity management systems for the internet of things: a survey towards blockchain solutions." Sensors 18.12 (2018): 4215.