International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

Lightweight Centralised Digital Repository for Personal Data Management with Privacy Preservation Policy

Arun Kumar. S, Anbarasi M.S.

Abstract - The on-going increment in announced episodes of reconnaissance and security breaks bargaining clients' privacy raise doubt about the present model, in which third-parties gather and control huge measures of individual information. Bitcoin has shown in the budgetary space that trusted, auditable registering is conceivable utilizing a decentralized system of companions joined by an open record. In this paper, we portray centralized individual information the board framework that guarantees clients possess and control their information. We implement a centralized privacy preservation protocol that turns a block chain into an automated data verification manager that does not require trust in a third party for verification. Not at all like Bitcoin, exchanges in our framework are not carefully money related – they are utilized to convey directions, for example, storing , questioning and sharing information particularly verification . At last we examine use of centralized privacy preservation protocol for manage the personal data in centralized manner.

Index Terms: Bitcoin, Centralized ,Privacy,Personal data , Verification.

I. INTRODUCTION

Block chain Technology is becoming extremely popular among various industries due to its many applications such as payment services, smart contracts, digital identity, supply chain management and others [1]. To cope up with the increasing demand, Block chain must be able to scale and process transactions at a much faster rate than its current capabilities. Since most of the Block chain mainly focuses on security and decentralization, scalability is often sacrificed.

As a result, Block chains have incredibly slow processing speed. Over the years, many scalability solutions have been proposed. In this paper, we discuss about the scalability issues which are currently present in Block chain and further analyse various on-chain and off-chain scalability solutions of Block chain Technology.

As Block chain technology is becoming extremely popular among various industries, block chain must be able to scale and process transactions much faster than its current capabilities. It is hard for block chain to grow and support increasing numbers of transactions. Scalability can be achieved by Hard Forking, Private Channels, and Scalable Consensus mechanisms. Block chain innovation was initially gotten from the Bitcoin framework. Bitcoin is a noticeable decentralized advanced cash (or cryptographic money). As its significant advancement, block chain pulls in a ton of consideration after the extraordinary achievement of Bitcoin. Since the arrival of the first white paper, the market for digital money has flooded[12]. An ever-increasing number of ventures are beginning to acknowledge Bitcoin installments, including Dell, Reddit, Expedia, PayPal, and Microsoft. In spite of the fact that in the meantime, the Bitcoin innovation has seen numerous discussions, the hidden blockchain conventions and the disseminated processing engineering are generally acknowledged. Block chain is a decentralized and distributed ledger which records transactions securely without the involvement of any third party. These transactions are publicly recorded and are immutable.

Blockchains are inherently resistant to data modifications by design and purpose. It has the capacity of functioning as a distributed ledger and capable of effectively and verifiably and permanently recording transactions between two parties. Network participants can confirm transactions using block chain technology without a third-party access intermediary. It can be used in numerous technological applications such as payment, smart contracts, documentation, digital identity, supply management chains. Powerful applications include transfers of funds, voting, and many other uses.

While we as a whole receive the rewards of a data-driven society, there is a developing public worry about client privacy. Unified associations – both public and private, store up huge amounts of individual and touchy data. People have next to zero power over the data that is put away about them and how it is utilized. As of late, public media has over and again secured questionable occurrences identified with privacy. Among the better realized precedents is the tale about government reconnaissance [2], and Facebook's huge scale logical examination that was obviously directed without expressly educating members [8].

II. RELATED WORK

There have been various attempts to address these privacy issues, both from a legislative perspective ([3], [14]), as well as from a technological standpoint.

Manuscript published on 30 June 2019.
***** Correspondence Author (s)
Arunkumar. S*, Research Scholar, Department of Computer Science and Engineering, Pondicherry Engineering College, Pondicherry University, Puducherry, India 605014.
Anbarasi M.S, Assistant Professor, Department of Information Technology, Pondicherry Engineering College, Pondicherry University, Puducherry, India 605014.

OpenPDS

An as of late created system, shows a model for The initial two papers contributed similarly to this work. independent deployment of a PDS which incorporates a component for returning calculations on the data, in this way returning answers rather than the crude data itself [4]. Over the business, driving organizations executed their own restrictive validation programming dependent on the OAuth protocol [13], in which they fill in as concentrated confided in specialists. From a security point of view, specialists created different systems focusing on privacy concerns concentrated on close to home data.

Data anonymization

This strategies endeavor to secure by and by recognizable data. k-anonymity[10][11], a typical property of anonymized datasets necessitates that delicate data of each record is vague from in any event k−1 different records [15]. Other privacy-saving strategies incorporate differential privacy, a strategy that bothers data or adds commotion to the computational procedure preceding sharing the data [5], and encryption[9] conspires that permit running calculations and inquiries over scrambled data.

Fully Homomorphic encryption

In particular, fully homomorphic encryption (FHE) [7][18] plans enable any calculation to keep running over encoded data, however are at present too wasteful to even think about being generally utilized by and by.

Proxy signature

Proxy signature is a signature plot that an orig-inal signer delegates his/hersigning capacity to a proxy signer, and after that the proxy signer makes a signa-ture in the interest of the first signer.

Global signature

In this scheme , every group has common signature to sign the data. when ever the new the user added in group, global signature is provide to the new user.

Group signature

A group signature scheme is a method for allowing a member of a group to anonymously sign a message on behalf of the group.

RSA based Ring signature

A ring signature[16] is a kind of advanced signature that can be performed by any individual from a gathering of clients that each have keys. In this way, a message marked with a ring signature is supported by somebody in a specific gathering of individuals. One of the security properties of a ring signature is that it ought to be computationally infeasible to figure out which of the gathering individuals' keys was utilized to create the signature

As of late, another class of responsible frameworks developed. The principal such framework was Bitcoin, which enables clients to move cash (bitcoins) safely without a brought together controller, utilizing a publicly unquestionable open record (or blockchain). From that point forward, different activities (all in all alluded to as Bitcoin 2.0 [6]) showed how these blockchains can serve different capacities requiring confided in registering and auditability.

After conclude the disadvantages of the above models. We are need of the efficient centralized personal data management with privacy preservation policy.

Our Contribution:

1) We combine the concepts of homomorphism with ring signature for protect the privacy of data as well as data owner.

2) We propose the idea for create new centralised authenticated block chain private data management.

III. DATA AND USER PRIVACY ISSUE

All through this paper, we address the privacy concerns clients face when utilizing third-party administrations [17]. We centre explicitly around versatile stages, where administrations convey applications for clients to introduce. These applications always gather high resolution individual data of which the client has no particular knowledge or control. In our examination, we expect that the administrations are straightforward yet inquisitive (i.e., they pursue the protocol). Note that a similar framework could be utilized for other data privacy concerns, for example, patients sharing their therapeutic data for logical research, while having the way to screen how it is utilized and the capacity to immediately quit. In light of this, our framework ensures against the accompanying regular privacy issues:

Data Possession. Our framework centres around ensuring that clients claim and control their own data. All things considered, the framework perceives the clients as the proprietors of the data and the administrations as visitors with designated authorizations.

Data Transparency and Auditability. Every client has total transparency over what data is being gathered about her and how they are gotten to.

IV. THE PROPOSED SYSTEM

Our prosed system concentrates on two techniques such Hybrid hash Structure which provides robust data structure to store the personal data and Homomorpic Circle Signature which deals privacy protection mechanism against data verifier.

Hybrid hash Structure:

Data structure is a logical organization of data. It helps us to represent the data while storing and retrieving data to / from the data storage. The novel structure is called as Hybrid hash structure. It is the combination of blockchain, Merkle trees and versioning systems. Part 1 contains the hash value of previous node and hash value of current node. Part 2 is the root of all data block which contains hash information of its sub tree ie children. Part 3 consist of multi-level of hash records with versioning system. Part 4 is the last level of structure which contains all versions of data blocks with the help of versioning system. Versioning system is a scheme that archives every modification of a file or set of files over time so that it can be regenerate particular versions whenever we require for various purpose. The following figures illustrate the hybrid hash structure.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

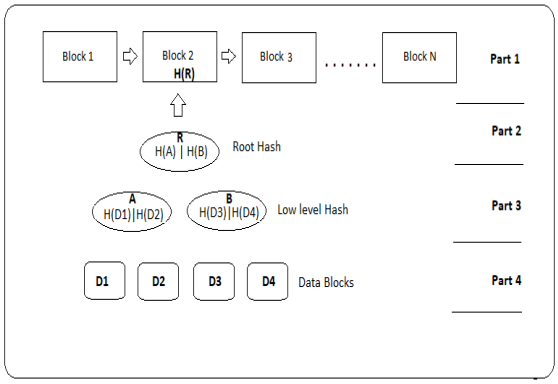

Fig no 1. Hybrid Hash Structure

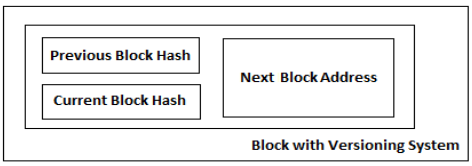

Fig no 2 : Internal structure of Block Node in Hybrid Hash Structure

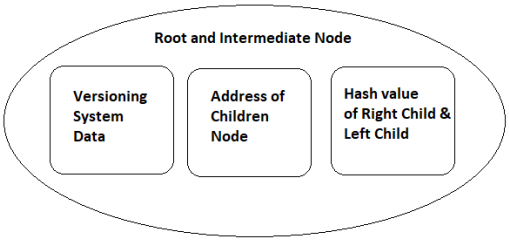

Fig no 3 : Internal structure of Root and intermediate Node in Hybrid Hash Structure

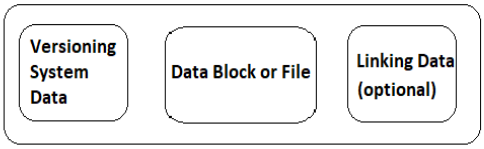

Fig no 4 : Internal structure of Root and intermediate Node in Hybrid Hash Structure

Homomorpic Circle Signature:

Whenever file is created or modified by any one of the user, file or data block is signed by same user. In future user wants to check the correctness of data. He/she will calculate the hash value of data. then he/she will match the same value with signed hash value. he / she will decided correctness of data based on comparison result.

If it is muti_owner data or group data, the above mentioned signature technique is not supported and. Because data may be modified by several users. Another challenge is that verifier needs the public key of particular user for audit the user's data this process reveals the identity privacy of user. To protect privacy from verifier modified signature system is required. some of signature system with privacy preservation are proxy signature and global signature etc. those signature technique are creating signature like as instead of user, someone will sign the data by using his / her private key and common private key is created and distributed to all users.The above signature techniques are protecting the privacy. but first one is single point failure and global private key will created and revoked whenever new user will participate the group and any existing user exits group. It is overburden of group manager who is controlling the group. The following algorithm expresses the Homomorpic Circle Signature creation and verification

Homomorpic CircleSign Algorithm

Given n group member's public keys (pk1,..,pkn) = (w1,..,wn), an audit block m ∈ Zp , the identifier of that block id .Circle signature is generated as follows:

1.User ut (called as signer) where 1<= t <= n , picks random no ri from Zp for all remaining users (ie $i \neq t$), where value of i is from 1 to n

2. Consider $\sigma_i = g_1^{r_i} \in G_1$

3. Signer computes

$$\phi = H(id, pk_1, pk_2, ...pk_n) g_1^m \in G1 .....(1)$$

4. Then fixes,

$$\sigma_t = \left(\frac{\phi}{\tilde{I}\left(\prod_{i=1}^{n,(i \neq t)} \omega_i^{r_i} \right)} \right)^{1/r_t} \in G1 .....(2)$$

DataAudit Algorithm

The Circle signature of an audit block m is

σ = (σ1, σ2, .. σn) Ɛ G1

In Circle siganture σ, σt is one of elements. So TPA can verify the signataure is signed by one of the authorised group user, but cannot learn the identity of the perticular user.

Consider two Blocks m1 and m2, Circle siganture σ1= (σ1.1,σ1.2 … σ1.n) for m1 and σ2= (σ2.1,σ2.2 … σ2.n)for m2, Identifiers of m1,m2 are respectivly id1,id2 and two random values v1,v2 Ɛ Zp.

Super block m' is combined by cloud using following formula

$$m' = v_1 m_1 + v_2 m_2 \in Z_P$$

TPA can verify m' without downloading the m1 amd m2 by verifying

$$\mu(\phi', g_2) \stackrel{?}{=} \prod_{i=1}^{n} \mu(\sigma_i', \omega_i) .....(4)$$

Where

$$\phi' = H(id_1, pk_1, ...pk_n)^{v_1} .. H(id_2, pk_1, ...pk_n)^{v_2} . g_1^{m'}$$

$$\sigma_i' = \left(\sigma_{1.i}^{v_1} . \sigma_{2.i}^{v_2} \right)$$

If super block m' is correct, then TPA trusts that m1 and m2 are correct. Figure.4 shows super block creating process.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

This process is not only protect the privacy also reduces the communication cost while auditing processIn the above way without revealing any single useful detail to the TPA the integrity of the data is verified correctly.Here novel signature will overcome of challenges discussed.it is called as Homomorpic circle signature. In group, any user can calculate the circle signature using the public key information of other user. After modification, data block will be signed by created circle signature. So verifier will understand that one of the user is created the circle signature, but not who is user. So the probability that verifier succeed to discover the identity of signer is 1/n. It also support the Homomorpic authentication ie block less verification.

V. PERFORMANCE EVALUATION

In this section we evaluate the performance of verification process (auditing) in aspects of auditing efficiency

For experimental analysis the proposed system was designed in Visual Studio 2010 as a ASP.NET web application with C# as the programming language. The system used Windows 7 operating system with configuration of Intel Core I5, 2.5 GHz Processor and 4 GB RAM. The data used varied in size from 10KB to 2 MB. A total of 2 GB data were used to perform the analysis in various dimensions. The ASP.NET C# namespace System.Security.Cryptography was utilized to write the encryption, decryption and Hash generation. The classes under the namespace System. Diagnostics has been used to record and analyze the performance of the system. The performance of auditing data is evaluated by carrying out a mock audit with various sizes of data. From the analysis performed, it is observed that the time taken to generate the audit data and to perform the audit differs with the size of the data. The fig.no5 depicts how the audit performance differs with the size of the audit data. One could see that with the increase in size of data, the time for performing the audit also increases linearly. This is because the time for generating hash and, signing and encrypting the super block all takes some time. While performing audit all these functions are run again to determine the existing integrity of a given data and the value is compared to the value stored at the time of uploading or modification of the data.

Table no 1: Time taken to Audit and generate audit data

File Size(In KB)	Time taken for Audit data(in milliseconds)	Time taken for generating Audit data(in milliseconds)
22	10	8
128	18	16
291	22	21
976	31	28

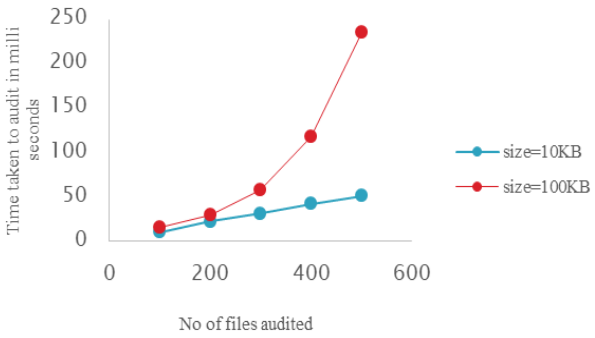

Fig no 5: Time taken to audit w.r.t number of files audited

Apart from the size of the file with the increase in size of the group, which apparently leads to more number of shared data, might increase the time to complete the audit. However, the time taken to complete the audit still depends on the size of the file as we have all other inputs ready at the time of audit. Also, the data transfer rate plays a key role in performance, which is common to any network based communication system. The experiment was carried out under a broad band network with a bandwidth of 2MB per second.

VI. CONCLUSION AND FUTURE WORK

Personal data and touchy data when all is said in done, ought not be confided in the hands of outsiders, where they are powerless to assaults and abuse. Rather, clients should possess and control their data without bargaining security or restricting organizations' and experts' capacity to give personalized administrations. Our technique allows this by merging a centralized block chain, with privacy preservation storage solution. Users are not required to trust any third-party and are always aware of the data that is being collected about them and how it is used. In addition, the user can allow any third party or verifier to verify the user data without worry about your privacy of data concerned. Future, We extent the technique in decentralized manner for some other application such as e-voting, e-auction etc.

.

REFERENCES

1. A. Chauhan, O. P. Malviya, M. Verma, and T. S. Mor, “Blockchain and Scalability,” 2018 IEEE International Conference on Software Quality, Reliability and Security Companion, 2018
2. James Ball. Nsa’s prism surveillance program: how it works and what it can do. The Guardian, 2013.
3. EUROPEAN COMMISSION. Commission proposes a comprehensive reform of data protection rules to increase users’ control of their data and to cut costs for businesses. 2012.
4. Yves-Alexandre de Montjoye, Erez Shmueli, Samuel S Wang, and Alex Sandy Pentland. openpds: Protecting the privacy of metadata through safeanswers. PloS one, 9(7):e98790, 2014.
5. Cynthia Dwork. Differential privacy. In Automata, languages and programming, pages 1–12. Springer, 2006.
6. Jon Evans. Bitcoin 2.0: Sidechains and ethereum and zerocash, oh my!, 2014.
7. Craig Gentry. Fully homomorphic encryption using ideal lattices. In STOC, volume 9, pages 169–178, 2009.

8. Vindu Goel. Facebook tinkers with users' emotions in news feed experiment, stirring outcry. The New York Times, 2014.
9. Federal Information and Processing Standards. FIPS PUB 180-4 Secure Hash Standard (SHS). (March), 2012. [12] Don Johnson, Alfred Menezes, and Scott Vanstone. The elliptic curve digital signature algorithm (ecdsa). International Journal of Information Security, 1(1):36–63, 2001.
10. Michael Lesk. How much information is there in the world? [14] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In ICDE, volume 7, pages 106–115, 2007.
11. Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond kanonymity. ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1):3, 2007.
12. Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Consulted, 1(2012):28, 2008.
13. Juan Perez. Facebook, google launch data portability programs to all, 2008.
14. Rt.com. Obama announces legislation protecting personal data, student digital privacy, 2015.
15. Latanya Sweeney. k-anonymity: A model for protecting privacy. International Journal of Uncertai
16. R. L. Rivest, A. Shamir, and Y. Tauman, "How to Leak a Secret," in Proc. ASIACRYPT. Springer-Verlag, 2001, pp. 552–565
17. K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69-73, 2012.
18. Wang, B., Li, B., Li, H.: Knox: Privacy-Preserving Auditing for Shared Data with Large Groups in the Cloud. Tech. rep., University of Toronto (2012).

AUTHORS PROFILE

Arun Kumar S received the Master of Technology in Computer Science and Engineering from SRM University, Chennai, India. He had 10+ years of teaching experience in engineering education .He currently working as Assistant Professor in Department of CSE in SRM University, Chennai, India. He has been the author and co-author of more than 10 articles published in referred journals. His Research interests include cloud computing and information security

ANBARASI.M.S received the Doctorate from ANNA University, Chennai, India. She has 15+ years of teaching experience in engineering education. She is currently working as Assistant Professor in Department of IT in Pondicherry engineering college, India. She has been the author and co-author of more than 20 articles published in referred journals Her Research interests include Data Mining, Software Engineering, and Cloud Computing.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)