ISSN: 2278-4632
Vol-13, Issue-08, August 2023

ADAPTIVE BLOCKCHAIN-BASED FRAMEWORK TRANSACTION VALIDATION FOR DATA ASSURANCE

[#1]AENUMALA YAMINI, *Department of MCA,*

[#2]Dr.V.BAPUJI, *Associate Professor& HOD,*

Department of Master of Computer Applications,

VAAGESWARI COLLEGE OF ENGINEERING, KARIMNAGAR, TELANGANA

ABSTRACT: The greatest barrier to universal data interchange is trust. Many data owners are unable to share their data because of a lack of infrastructure for establishing data trust, while data consumers are concerned about the quality of the shared data. Blockchain technology allows decentralized and transparent administration by allowing numerous parties to reach consensus on an immutable ledger. This program makes use of blockchain technology to create an end-to-end architecture for data trust, thereby improving data interchange integrity. In addition, depending on the projected trust value, we provide an adaptive method for calculating the number of transaction validators.

Keywords: Distributed, access control, data trust, and blockchain are all keywords.

1. INTRODUCTION

Data sharing has become a big concern regarding Data sharing has become a hot topic due to worries about privacy and confidentiality difficulties, data abuse, and moral and legal transgressions. Many data owners are unwilling to volunteer their data, which could be critical for a variety of research reasons, due to the lack of a clear and stable structure for data trust. The quality and dependability of the data supplied at the source is a major source of worry for both data users and data owners. As a result, both data owners and consumers are concerned about trust. Data trust is an innovative concept that promotes data sharing by requiring data consumers to be trustworthy while sharing and reusing their data. Data trust comprises legal, ethical, governance, and organizational structure considerations in addition to technical needs for supporting data exchange.

Blockchain technology can successfully supply the fundamental attributes for establishing a functional data trust framework without the need for intermediaries by altering current auditing methods and automatically enforcing the logic of smart contracts. Several more studies have looked into the possibilities of blockchain for data exchange, trust establishment, and access management. Blockchain can serve as a data trust link between data controllers and data users. Blockchain's distributed, secure, and dependable qualities can boost the credibility of the data trust system.

In this work, we describe a blockchain-based end-to-end architecture for data trust that ensures data consumers' integrity and quality at the source, as well as data owners' ethical and secure usage of data. We provide a trust model that assesses the reliability of input data sets using three criteria: data owner endorsement and reputation, data asset endorsement, and data owner confidence in the provided data set. These parameters, which are visible on the ledger, will be changed with each new transaction. We employ state-based endorsement with adaptive transaction validation depending on the dataset's trustworthiness in Hyperledger Fabric. We undertake a thorough performance analysis to show how well our system manages huge transactional databases and scales across numerous businesses. We declare that our system possesses all of the features required for data trust. It also benefits from smart

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

contract automation and the blockchain's transparency, immutability, and security.

2. LITERATURE REVIEW

Zavolokina et al. offered financial incentives for network involvement in addition to supplying high-quality data for automotive dossiers. The system believes that by punishing bad behavior, it may improve. Smart contracts are used to automatically calculate and implement incentives.

Shrestha et al. used blockchain technology and smart contracts to persuade data owners to donate their research data without giving up ownership or control. The system encourages people to join the network by giving them access to aggregated, anonymised data.

A subjective logic model was utilized to evaluate node reputation in order to ensure high-quality data transmission in the vehicular network. Dedeoglu et al. created a trust model to assess the integrity of data gathered by IoT sensor nodes. The model is based on the data source's authenticity and reputation, as well as observations from other adjacent sensor nodes. Furthermore, they use blockchain to monitor the quality of shared data by detecting incorrect or suspicious data collected by IoT devices or mobile crowd sensing.

Choudhury et al. ensured data quality while protecting data privacy. Regulatory bodies evaluate data precision as network participants. Data confidentiality is ensured by establishing activity-specific private channels. An et al. proposed delegated proof of reputation (DPoR), a lightweight consensus mechanism, to solve the difficult heavy computation problem applicable to crowd sensing node data quality control.

Huang et al. used smart contract verification processes to ensure the high quality of data collected from crowd sensing network sensor nodes. Su et al. devised a two-tiered incentive strategy based on reinforcement learning (RL) to encourage the sharing of high-quality data. Based on game theory, Casado et al. presented a cooperative edge computation layer method to support data quality and false data detection.

Shala et al. recently developed an incentive system to encourage peers in low-trust IoT networks to improve it. A critical component of the motivational system is control cycles with a goal trust score. Service providers with low trust ratings will be given a package of incentives, such as discounts on other services, to encourage them to provide better service in exchange for the promised benefits. The developers of devised an incentive-based method in order to encourage owners of high-quality (actual and useful) medical data to share their data and earn money, as well as miners who profit from participating in and validating transactions.

Wang et al. devised a strategy for an incentive that ensures anonymity so that crowd sensing contributors provide high-quality contributions. Participants are encouraged to exchange their superior sensing data for Bitcoin or Monero via the trust mechanism. Assuring data accuracy could also benefit data miners.

Dedeoglu et al. proposed a trust model for evaluating sensor node data precision in IoT networks. The model is based on evidence from nearby sensor node observations as well as the data source's dependability and reputation.

3. SYSTEM ARCHITECTURE

To that end, our proposed approach seeks to build a data trust architecture that is useful for both data creators and consumers. Two fundamental elements of our system design will be outlined in order to achieve this. For secure and verifiable access control management, a dependable system that can track and monitor access, as well as a trust model that can evaluate the quality and dependability of input data sets, are essential.

Our data trust framework is depicted in Figure 1. Using the available data, we create a trust model. Our methodology uses a blockchain-based application to establish the reliability of an initial dataset. Using this number, the system can ensure that only trustworthy data sets pass validation and that only trustworthy data assets are stored in the

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

ledger. Parameters and the mathematical language used to determine the trust value are discussed in Section V. If confidence is low, it means the data sets' veracity is suspect and more people will need to check them.

If the necessary details about a data set are recorded and kept as data assets on the ledger, then anyone who wants to get their hands on that data set can do so by submitting a formal request for access to that data set. The request will be forwarded directly to the data owner, who will determine the access terms and circumstances. Blockchain and smart contract systems provide decentralized, autonomous transaction enforcement, meaning no middlemen are required to complete the process. The owner of a resource can utilize a blockchain-based smart contract to set permissions for using the resource. In addition, owners of data can easily query the immutable and permanent storage of blockchain to see who has had access to their dataset in the past or who has access to it at the present time, independent of the response type shown.

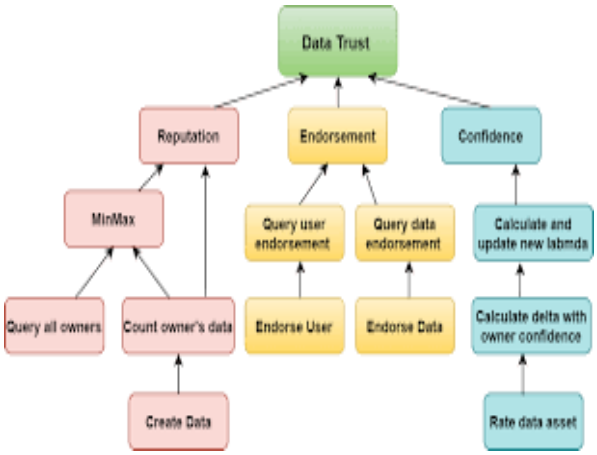

Fig 1: Presents our data trust framework architecture

TRUST MODEL

Here, we'll discuss the methodology behind deriving the reliability score from the provided data. In subsequent time periods, this value will be made available to individuals and institutions who have expressed an interest in acquiring the aforementioned data set. In addition, the system dynamically employs this value to settle on the optimal number of auditors for validating data collecting. The validation of the data set will be sped up and fewer verifiers will be needed as a result of the improved trust value.

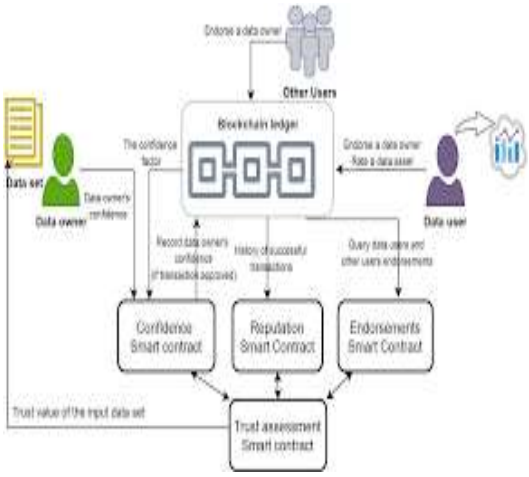

Fig: 2 Trust value for the input data sets

Terminology

In our distributed data trust architecture, data sets play an important role in both administration and security. As a result, these stores of information are commonly known as data assets. In the context of data management, the term "data owner" refers to the entity or person who has full authority over a certain data asset. A key or identifier is a special number used to access certain data. The data asset must be provided by the data owner. All decisions regarding who will have access to what data, for what purposes, and under what terms are ultimately made by the data owner. All permissions given and canceled by the data owner are recorded in an open and searchable database.

Reputation

In order to determine the data owner's reputation, min-max normalization is applied, which considers not only the data owner's own successful transaction history but also the minimum and maximum values of successful transactions across all users.

Endorsement

Data holders can receive one of two primary endorsements. A familiar system user may suggest a promising first-group participant, perhaps on the basis of previous successful collaborations with the data owner. Data users who have previously examined the dataset the current data owner is making available provide the second type of endorsement. The user makes a recommendation

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

to the data owner based on their assessment of the data's quality. The second type of endorsement is more weighted toward the data owner's final endorsement score for a given metric.

Confidence

The person in charge of entering the data will provide a confidence value between 0 and 1 (confidence [0, 1]) to show how much faith they have in the collected information. There needs to be a record so that the current data owner can factor in results from prior analyses when determining the data asset's reliability.

4. RELATED WORK

DATA TRUST

The idea of trust is multifaceted and is explored in fields as diverse as sociology, economics, psychology, computer science, and information technology. Different facets of trust and its application in various contexts are studied in these areas of study. Because it involves ethics, emotions, values, and knowledge from other fields, the concept of trust is intricate and difficult to grasp. Trust, at its essence, is a dynamic relationship between two parties—the trustor and the trustee—in which the trustor places confidence in the trustee according to established parameters. Comprehensive academic research on the topic of trust has led to the following definition: Trust is the trustor's inclination to embrace a subjective conviction that the trustee will demonstrate responsible conduct to optimize the trustor's interests in an uncertain situation. The trustor has formed this opinion after giving thought to how the trustee has treated them in the past.

A common definition of digital trust is the computational value established between a trustor and a fiduciary. It is measured in terms of trust qualities and evaluated according to a standard protocol. It is crucial to recognize the interdependence of many parts and individuals and to foster a climate of trust. O'Hara proposes a list of eight characteristics that are essential for the development of data trusts.

(1) Discovery
(2) Provenance
(3) Access Controls,
(4) Access
(5) Identity Management
(6) Auditing of Use
(7) Accountability
(8) Impact

- The individual suggested Web Observatory as a possible technology answer to the critical problem of data trust.
- Data discovery refers to the process through which users first learn about the nature and characteristics of the data they will be using.
- The term "provenance" is used to describe the data's connection to its background and the information that describes it.
- Access control is the process through which data's owners monitor and adjust who has access to their information.
- The term "access" refers to the operational procedure by which data consumers are granted permission to access and use data.
- through "identity management," we mean the process through which data owners can determine and confirm the identities of persons who request to use their data.
- In this context, "auditing of use" refers to the provision of a comprehensive log of data access.
- The installation of access control and use audits is crucial to ensuring accountability.
- In the context of data trust records, "impact" refers to the evaluation of data's worth, utility, and potential for profit.

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

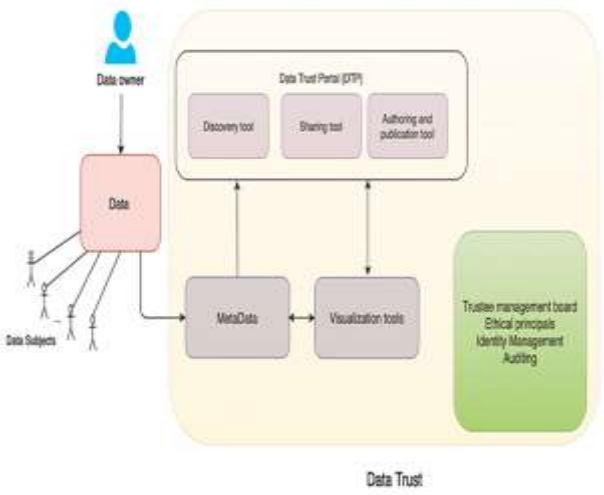

Fig: 3 Data Trust

DATA TRUST PORTAL ARCHITECTURE (DTP)

The picture shows how O'Hara incorporated ideas from web observatories into the design of his Data Trust Portal (DTP). The Data Transfer Protocol (DTP) itself does not involve any sort of data storage. Instead, it is up to the data owners to guarantee the safety of the data and to choose and implement a suitable interface method to enable access. With the DTP platform, you can easily establish a secure protocol for data discovery and sharing. Metadata about the characteristics and origins of data are used in this protocol.

Stalla-Bourdillon et al. present a standardized method for making efficient use of data that they address in their work. Within this paradigm, the authors stress the need for well-defined data governance responsibilities and processes. There are three main layers used to convey the idea of data trust in the visual representation

(1) The data layer
(2) The access layer
(3) The process layer

Fig: 4 Data Trust Portal Architecture

The picture shows how O'Hara incorporated ideas from web observatories into the design of his Data Trust Portal (DTP). The Data Transfer Protocol (DTP) itself does not involve any sort of data storage. Instead, it is up to the data owners to guarantee the safety of the data and to choose and implement a suitable interface method to enable access. With the DTP platform, you can easily establish a secure protocol for data discovery and sharing. Metadata about the characteristics and origins of data are used in this protocol.

Stalla-Bourdillon et al. present a standardized method for making efficient use of data that they address in their work. Within this paradigm, the authors stress the need for well-defined data governance responsibilities and processes. There are three main layers used to convey the idea of data trust in the visual representation

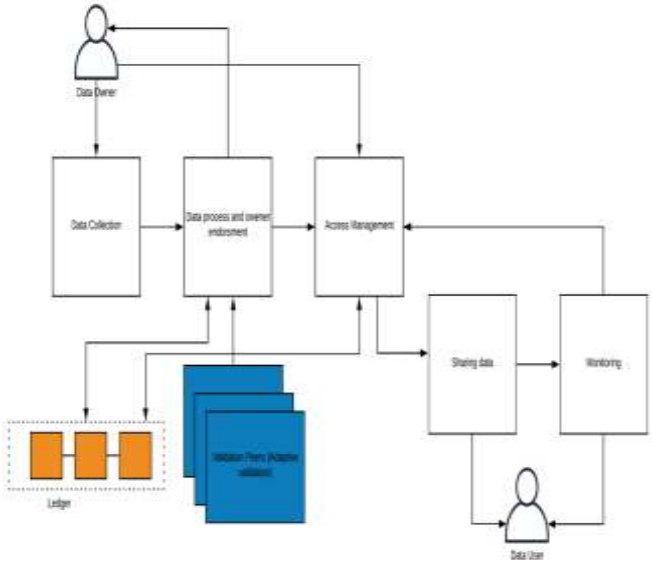

Fig: 5 End to end data trust architecture with adaptive validation

ACCESS MANAGEMENT AND SHARING DATA ASSETS

The picture shows how O'Hara incorporated ideas from web observatories into the design of his Data Trust Portal (DTP). The Data Transfer Protocol (DTP) itself does not involve any sort of data storage. Instead, it is up to the data owners to guarantee the safety of the data and to choose and implement a suitable interface method to enable access. With the DTP platform, you can easily establish a secure protocol for data discovery and sharing. Metadata about the characteristics and origins of data are used in this protocol.

Stalla-Bourdillon et al. present a standardized

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

method for making efficient use of data that they address in their work. Within this paradigm, the authors stress the need for well-defined data governance responsibilities and processes. There are three main layers used to convey the idea of data trust in the visual representation

5. SYSTEM ANALYSIS

Discovery

Using the system interface, authorized data users can learn about the available data assets, including the properties of data sets depicted as metadata. Our proposed method yields a trust value that informs data consumers about the quality of data assets.

Provenance

After adding their data sets as data assets to the system, data proprietors must provide metadata pertinent to data provenance, such as data origin, collection time, and collection method. Both transaction verifiers and data consumers can evaluate the data's quality with this information. In addition, when a data set is modified, a transaction is performed to update the data asset's properties on the ledger. By distinguishing the actual activities that have been performed on the data sets, it is possible to trace data evolution and query data provenance.

Access Control

Data assets are completely under the control of their proprietors. They have control over who can access their data, and smart contracts enable them to automatically enforce these regulations. In order to verify the legitimacy of submitted transactions, smart contracts also permit granular access checks.

Access

To avoid jeopardizing the interests of individuals by permitting access to their information, data sets containing personal information must be de-identified or anonymized prior to dissemination. In addition, Hyperledger Fabric provides private data and private communication, which may be of interest to data proprietors who do not wish for other system users to view information associated with their data. This feature allows them to share information about their data assets with interested parties. Smart contracts' adaptable policies can also be used to restrict access to data provenance. Data users, for instance, can request permission from data owners to examine the provenance of a data collection.

Identity Management

Before interacting with the Hyperledger Fabric permissioned blockchain, all actors and users must acquire a digital identity wrapped in an X.509 digital certificate. This identity is essential in a blockchain network for distinguishing the permissions individuals have over resources and information access. Additional identifiers may be added to a digital identity to identify the individual or organization possessing it. These capabilities allow data owners to identify individuals endeavoring to gain access to their data assets.

Auditing

The auditing of data is one of the primary reasons for implementing blockchain. Due to blockchain technology, we are able to audit every system connection and procedure. Blockchain provides an immutable audit trail of data revisions, access requests, access grants, and access revocations in the context of data sharing. Data owners can search for previous queries for access to their data assets and modifications to access rights. In addition to the revision history and source of the data assets, data consumers can observe the revision history of the data sets. Keeping a watch on the immutable log of data transactions to automatically build audit trails and record any data breach attempts makes it easier to identify potential threats.

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

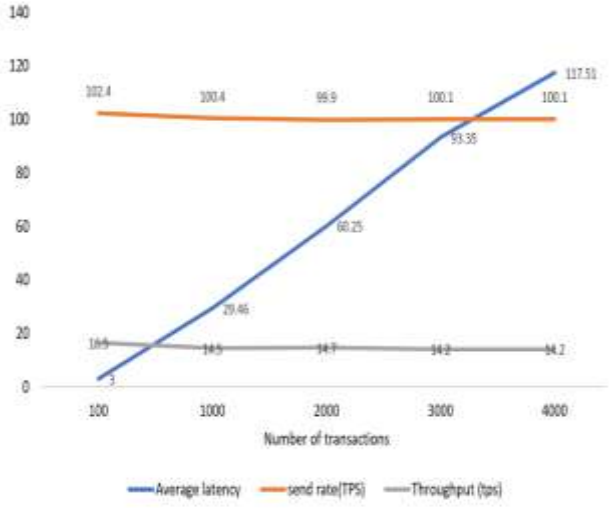

Fig: 6 Send rate, throughput and average latency for Create Data transaction

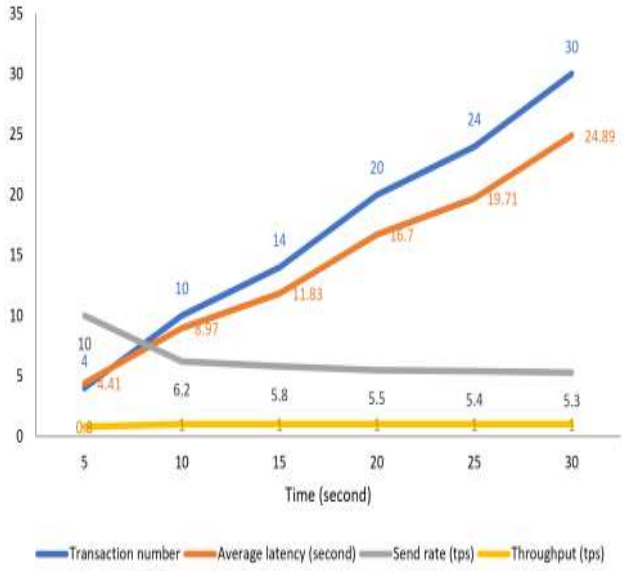

Fig: 7. Calculate and record the minimum and maximum number of data assets belong to a single user (MinMax)

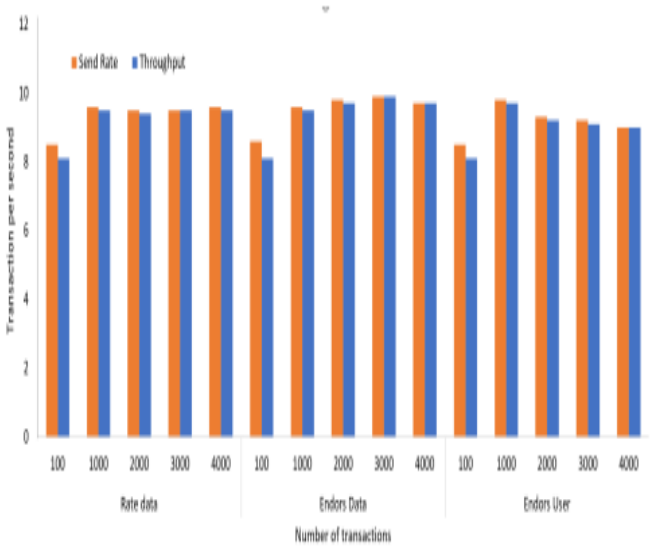

Fig: 8 Send rate and throughput for Rate Data, Endors Data, and Endors User transactions

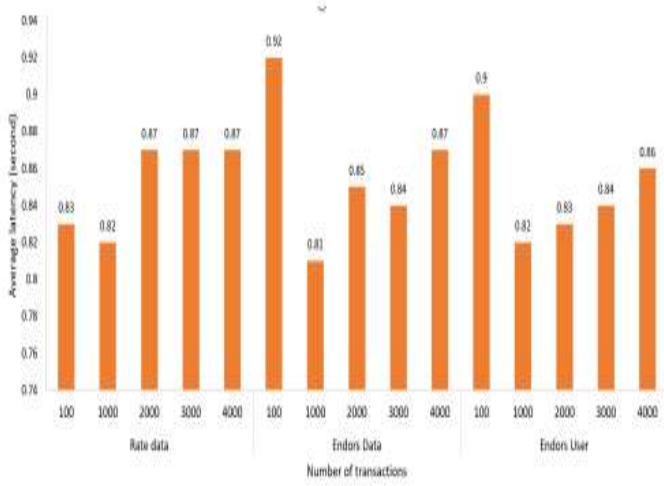

Fig: 9 Average latency for Rate Data, Endors Data, and Endors User transactions

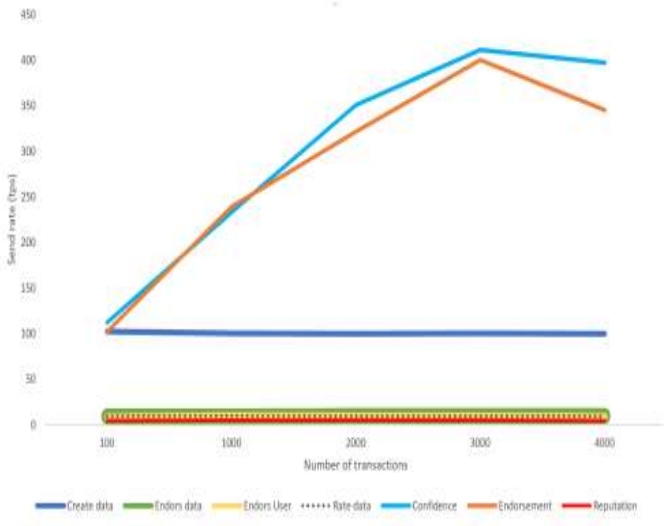

Fig: 10 Send rate for all transaction

ISSN: 2278-4632
Vol-13, Issue-08, August 2023

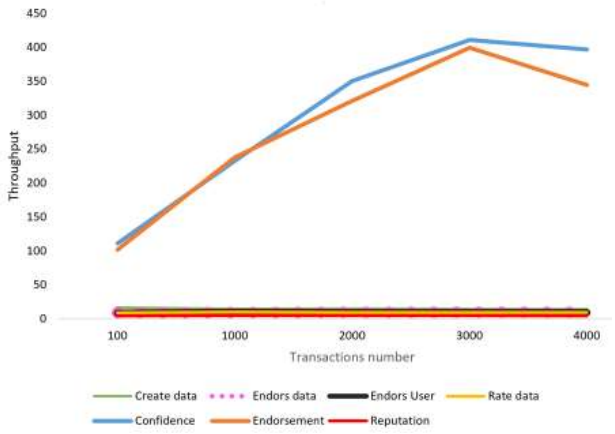

Fig: 11 Throughput for all transaction

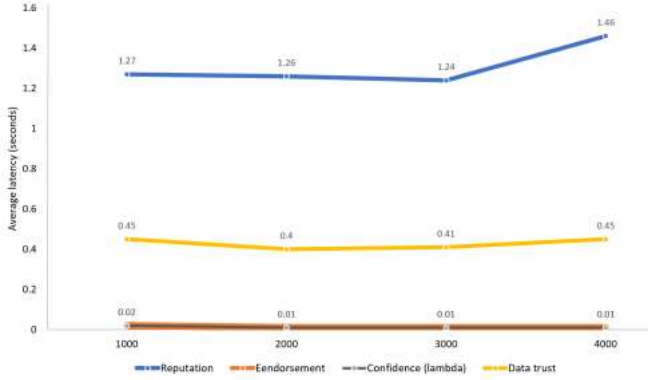

Fig: 12 Average latency for DataTrust, Reputation, Endorsement and Confidence transactions

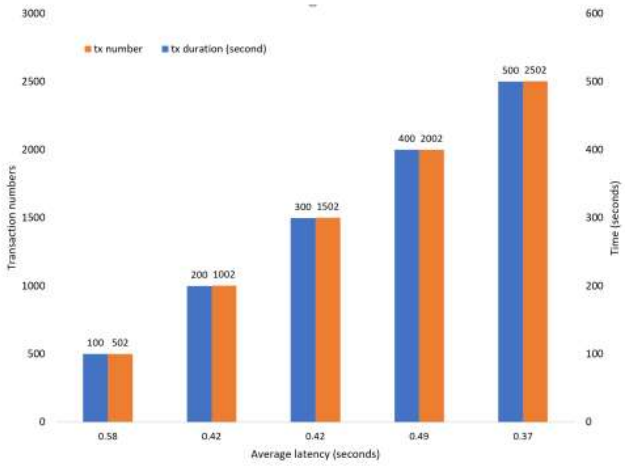

Fig: 13 Average latency for DataTrust based on time

6. CONCLUSION

Due to a dearth of mutual trust, current methods cannot provide a functional and transparent mechanism for data sharing. In this paper, we introduced a permissioned blockchain-based system for end-to-end data trust. Using a novel trust model that considers the data owner's reputation, recommendations, and confidence in the data being provided, our system evaluates the input data's quality. Consequently, data users ensure that the integrity of the public data set is continuously assessed and updated. Safe, open, and automatic access management based on smart contracts can also benefit data proprietors. They are the only actors in the system with complete authority over their data assets and the ability to independently manage access privileges. Using blockchain technology's provenance and auditability, data proprietors can track the history of access restrictions and modifications to their data assets. To identify potential threats, it is possible to extract from the ledger relevant records that provide a clear image of the system, detect suspicious requests, and reveal protocol violations. The evaluation results demonstrate the system's capacity to efficiently process numerous transactions involving the writing, modifying, and querying of trust parameter values.

REFERENCES

1. W. Yang, and M. Guizani, “An incentive mechanism for data sharing based on blockchain with smart contracts,” Computers & Electrical Engineering, vol. 83, p. 106587, 2020.
2. K. Shrestha and J. Vassileva, “User data sharing frameworks: A blockchain-based incentive solution,” in 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2019, pp. 0360–0366.
3. M. Shen, J. Duan, L. Zhu, J. Zhang, X. Du, and M. Guizani, “Blockchain-based incentives for secure and collaborative data sharing in multiple clouds,” IEEE Journal on Selected Areas in Communications, vol. 38, no. 6, pp. 1229–1241, 2020.
4. L. Yue, H. Junqin, Q. Shengzhi, and W. Ruijin, “Big data model of security sharing based on blockchain,” in 2017 3rd International Conference on Big Data

Computing and Communications (BIGCOM). IEEE, 2017, pp. 117–121.

5. Brandão, H. São Mamede, and R. Gonçalves, "Systematic review of the literature, research on blockchain technology as support to the trust model proposed applied to smart places," in World Conference on Information Systems and Technologies. Springer, 2018, pp. 1163–1174.
6. J. Kang, R. Yu, X. Huang, M. Wu, S. Maharjan, S. Xie, and Y. Zhang, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4660– 4670, 2018.
7. Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. Leung, "Blockchain-based decentralized trust management in vehicular networks," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1495–1505, 2018.
8. S. Stalla-Bourdillon, G. Thuermer, J. Walker, L. Carmichael, and E. Simperl, "Data protection by design: building the foundations of trustworthy data sharing," Data & Policy, vol. 2, 2020.
9. L. Hang, I. Ullah, and D.-H. Kim, "A secure fish farm platform based on blockchain for agriculture data integrity," Computers and Electronics in Agriculture, vol. 170, p. 105251, 2020.
10. A. Kushida, D. A. Nichols, R. Jadrnicek, R. Miller, J. K. Walsh, and K. Griffin, "Strategies for deidentification and anonymization of electronic health record data for use in multicenter research studies," Medical care, vol. 50, no. Suppl, p. S82, 2012.
11. S. Rouhani and R. Deters, "Blockchain based access control systems: State of the art and challenges," in IEEE/WIC/ACM International Conference on Web Intelligence, 2019, pp. 423–428.
12. E. Androulaki, A. Barger, V. Bortnikov, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Ferris, G. Laventman, Y. Manevich et al., "Hyperledger fabric: a distributed operating system for permissioned blockchains," in Proceedings of the Thirteenth EuroSys Conference, 2018, pp. 1–15.