Data Trust Framework Using Blockchain Technology and Adaptive Transaction Validation

ABSTRACT

Trust is the main barrier preventing widespread data sharing. The lack of transparent infrastructures for implementing data trust prevents many data owners from sharing their data and concerns data users regarding the quality of the shared data. Data trust is a paradigm that facilitates data sharing by forcing data users to be transparent about the process of sharing and reusing data. Blockchain technology proposes a distributed and transparent administration by employing multiple parties to maintain consensus on an immutable ledger. This paper presents an end-to-end framework for data trust to enhance trustworthy data sharing utilizing blockchain technology. The framework promotes data quality by assessing input data sets, effectively manages access control, and presents data provenance and activity monitoring. We introduce an assessment model that includes reputation, endorsement, and condence factors to evaluate data quality.

We also suggest an adaptive solution to determine the number of transaction validators based on the computed trust value. The proposed data trust framework addresses both data owners' and data users' concerns by ensuring the trustworthiness and quality of the data at origin and ethical and secure usage of the data at the

end. A comprehensive experimental study indicates the presented system effectively handles a large number of transactions with low latency.

**EXISTING SYSTEM**

Shala *et al.* [22] introduced an incentive mechanism to motivate low trusted peers in the IoT network to increase their trust score. The incentivization system uses control loops that contain a target trust score. For the service providers with low trust scores, a package of incentives, such as discounts for other services, will be sent to encourage them to offer a better service in exchange for the promised benefits. In [12], authors presented an incentive-based model to encourage medical data owners to share their high-quality data (real and practical) and earn revenues, as well as miners who get benefit by participation and validating transactions.

Wang *et al.* [19] introduced a privacy-preserving incentive mechanism to achieve high-quality contribution in crowdsensing. The trust model motivates participants to share their high-quality sensing data and profit in the form of Bitcoin or Monero cryptocurrencies. Miners verify the quality of data and earn revenue as well. Zavolokina *et al.* [20] provided a financial incentive for participating in the network and provides high-quality data for car dossiers. The system expects to \_x errors by punishing harmful behaviour. They employ smart contracts for automatically calculating and enforcing incentives. Shrestha and Vassileva [9] utilized blockchain and smart contracts to encourage data owners to share their research data without losing control and ownership of it.

In [25], a subjective logic model has been used to assess nodes' reputation to ensure high-quality data sharing in the vehicular network. Dedeoglu *et al.* [21] presented a trust model to assess the quality of data observed by sensor nodes in the IoT network. The model consists of three elements: evidence from other neighbour sensor nodes' observations, the data source's confidence, and its reputation. They also employ blockchain to control the quality of shared data by detecting inaccurate or suspicious data captured by IoT devices or mobile crowdsensing.

Choudhury *et al.* [30] ensured data quality while maintaining data privacy. Regulatory agencies assess the quality of data as network participants. Data privacy is ensured by creating activity-specific private channels. An *et al.* [18] presented a lightweight consensus mechanism called delegated proof of reputation (DPoR) to solve the heavy computation problem appropriate for data quality control in crowdsensing nodes. Huang *et al.* [17] ensured the quality of collected data from sensor nodes in the crowdsensing network through verification rules embedded in smart contracts. Su *et al.* [13] designed a two-tier reinforcement learning (RL)-based incentive algorithm to improve high-quality data sharing. Casado-Vara *et al.* [14] also presented a cooperative algorithm based on game theory in the edge computing layer to promote data quality and false data detection.

**Disadvantages**

* The system is less secured since blockchain techniques which are maintains trust between data are not implemented.
* Trust is not implemented in which a multidisciplinary and multifaceted concept that has been defined in various disciplines, such as sociology, economics, psychology, computation, information and computer science, to model different types of relationships.

Proposed System

In the proposed system, the system proposes an end-to-end framework for data trust based on blockchain, which ensures the trustworthiness and quality of the data at origin for data users and ethical and secure usage of data for data owners. First, we introduce a trust model to assess input data sets' trustworthiness using three parameters: data owner endorsement and reputation, data asset endorsement and data owner confidence level in the provided data set. All these parameters are recorded on the ledger, and they will be updated with every new transaction.

The system also applies adaptive transaction validation using Hyper ledger Fabric state-based endorsement based on datasets trust value. Finally, the system conducts a comprehensive performance analysis to demonstrate our system's efficacy in handling large sets of transactions and scaling across multiple organizations.

The system states that our system presents all the properties required for data trust. At the same time, it benefits from transparency, immutability, security offered by blockchain technology, and smart contracts' automation capabilities.

**Advantages**

The proposed implements the following data trust concepts

* Discovery refers to the process of discovering the quality and properties of data by data users in the first place.
* Provenance refers to the ability of data users to access the historical record and metadata about the data
* Access control refers to the ability of data owners to control and manage access permissions toward their data

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**Software Requirements:**

* Operating System - Windows XP
* Coding Language - Java/J2EE(JSP,Servlet)
* Front End - J2EE
* Back End - MySQL