IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Challenges & Solutions For Reliable And Secure Data Exchange In Data Markets

[1]Mr. Kewalkumar K. Mohod, [2]Dr. Avinash P. Jadhao, [3]Prof. Devendra G. Ingale

[1]M.E Student, [2]Associate Professor & HOD, [3]Assistant Professor

M.E Student, Dr. Rajendra Gode Institute of Technology & Research, Amravati, India

Guide, Dr. Rajendra Gode Institute of Technology & Research, Amravati, India

M.E Incharge, Dr. Rajendra Gode Institute of Technology & Research, Amravati, India

**ABSTRACT :**

In the era of block chain, ensuring data truthfulness and protecting the privates of data generators are both important to the long term healthy improvement of data market . On one hand , the main aim is to maximize the profit of the customer. Yet, to reduce operation cost, A strategic service provider may provide data services based on the whole raw data set, or even return a fake result without processing the data from designated data sources. The guarantee data confidentiality content of raw data should not be disclosed to data consumers if the real identity of data generator is hidden. Reliable and secure data exchange is developed internally in associate degree encryption and the signature using , partially encryption and by using identity-based signature. Also we have to maintain data processing, batch verification and outcome verification to maintain security and confidentiality of data. By using profile matching service with addition to reliable and secure data exchange we can improve the performance of information distribution service.

**KEYWORDS** – DATA MARKETS, DATA TRUTHFULNESS, DATA CONFIDENTIALITY

## I. INTRODUCTION

In the digital world large amount of data storage is available online to fulfill the requirement of the society to get the particular information where the specialist will collect the data from different sources and information patrons. There after this information is added to information administrators so that data will be purchase by different customer. [1] In information exchanging layer, the information purchasers will have to identify is real user or fake user and the information collected is correct. So we have to maintain confidentiality of the user who is a buyer who may be the individual user.

We have propose reliable and secure data exchange, which will be used in which productively coordinates Truthfulness of data and Privacy protection of data is maintain in Data Markets. reliable and secure data exchange will be is organized inside in an Encryption-then-Signature model, which will be utilizing to some degree homomorphic encryption method and personality based mark of data. It at the it will maintain time encourages cluster check, handling of information also result confirmation while

keeping up character security and information classification.[2]In additional we instantiate reliable and secure data exchange with a maintaining of profile-coordinating administration. Our practical result of examination in reliable and secure data exchange have some attractive properties which include the overhead corresponding in expansive low calculation to scale the information showcase.

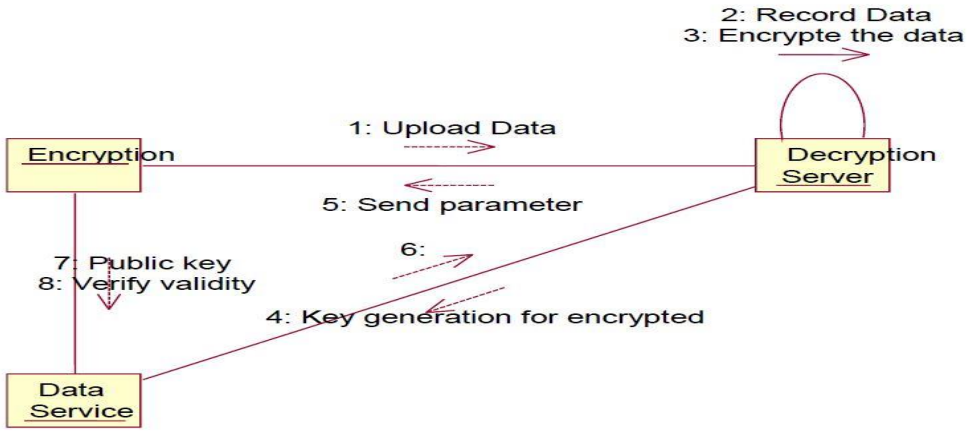



*Fig – 1 : Sequence diagram of reliable and secure data exchange*

## II. LITERATURE REVIEW

[1] A novel privacy preserving authentication and access control scheme for pervasive computing environments: Privacy and security are two important but seemingly contradictory objectives in a pervasive computing environment (PCE). On one hand, service providers want to authenticate legitimate users and make sure they are accessing their authorized services in a legal way. On the other hand, users want to maintain the necessary privacy without being tracked down for wherever they are and whatever they are doing.

[2] Automatic Enforcement of Data Use Policies with Data Lawyer : Data has value and is increasingly being exchanged for commercial and research purposes. Data, however, is typically accompanied by terms of use, which limit how it can be used. To date, there are only a few, ad-hoc methods to enforce these terms. We propose Data Lawyer, a new system to formally specify usage policies and check them automatically at query runtime in a relational database management system (DBMS). We develop a new model to specify policies compactly and precisely. We introduce novel algorithms to efficiently evaluate policies that can cut policy-checking overheads to only a few percent of the total query runtime. We implement Data Lawyer and evaluate it on a real database from the health-care domain.

[3] Account Trade: Account table protocols for big data trading against dishonest consumers : We propose Account Trade, a set of accountable protocols, for big data trading among dishonest consumers. To secure the big data trading environment, our protocols achieve book-keeping ability and accountability against dishonest consumers who may misbehave throughout the dataset transactions. Specifically, we

study the responsibilities of the consumers in the dataset trading and design Account Trade to achieve accountability against the dishonest consumers who may try to deviate from their responsibilities. Specifically, we propose uniqueness index, a new rigorous measurement of the data uniqueness, as well as several accountable trading protocols to enable data brokers to blame the dishonest consumer when miss behaviour is detected. We formally define, prove, and evaluate the accountability of our protocols by an automatic verification tool as well as extensive evaluation in real-world datasets. Our evaluation shows that Account Trade incurs negligible constant storage overhead per file (<;10KB), and it is able to handle 8-1000 concurrent data uploading per server depending on the data types.

[4] Anonymous Publication of Sensitive Transactional Data : Existing research on privacy-preserving data publishing focuses on relational data: in this context, the objective is to enforce privacy-preserving paradigms, such as k-anonymity and ℓ-diversity, while minimizing the information loss incurred in the anonymizing process (i.e., maximize data utility). Existing techniques work well for fixed-schema data, with low dimensionality. Nevertheless, certain applications require privacy-preserving publishing of transactional data (or basket data), which involve hundreds or even thousands of dimensions, rendering existing methods unusable. We propose two categories of novel anonymization methods for sparse high dimensional data. The first category is based on approximate nearest-neighbour (NN) search in high dimensional spaces, which is efficiently performed through locality-sensitive hashing (LSH). In the second category, we propose two data transformations that capture the correlation in the underlying data: 1. reduction to a band matrix and 2. Gray encoding-based sorting. These representations facilitate the formation of anonymized groups with low information loss, through an efficient linear-time heuristic. We show experimentally, using real-life data sets, that all our methods clearly outperform existing state of the art. Among the proposed techniques, NN-search yields superior data utility compared to the band matrix transformation, but incurs higher computational overhead. The data transformation based on Gray code sorting performs best in terms of both data utility and execution time.

## III. PROPOSED WORK

For the purpose of data market security we have considered double layer system model. It consist of data trading layer and data acquisition layer .In this layer there are four different kind of items or entities which include data contributor who will generate the data, a service provider who will provide the service ,the consumer who will consume the data and the registration center for the purpose of registration.. In the data acquisition layer, the service provider will produce a large amount of massive raw data from the data contributors who will contribute the data, from different social network users, different mobile and smart devices, smart meters, and so on. In order to have insensitive more data contributors to be actively submit high- quality standard data, then the service provider who is providing the service needs to reward those valid ones to compensate their data collection costs to find the valid data from large amount of data. For

the purpose of security, from each registered data contributor is equipped with a tamper-proof device so that data is secured which can be implemented with the help specific hardware or software. The data security is provided with the help of cryptography keys, code and data. Then the service provider which is cloud based will accumulate the collected data with different resources with the help of network bandwidths and storage space. Then we will identify the rich and value added data services which will be provided by to consumer instead of giving sensitive data directly, *e.g.*, social network analyses, probability distributions, personalized recommendations, and aggregate statistics. Then the registration center maintains an the online database of registrations done, and assigns each registered data to the contributor an identity and a password for purpose of security to activate the tamper-proof device. Besides, this also maintains an official website, called certificated bulletin board [4], on which the legitimate system that will show the participants who can publish essential information, *e.g.*, white lists, blacklists, resubmit-lists, and reward-lists of data contributors. Also the main another duty of the registration center is to set up the parameters for security by using a signature scheme and a cryptosystem.

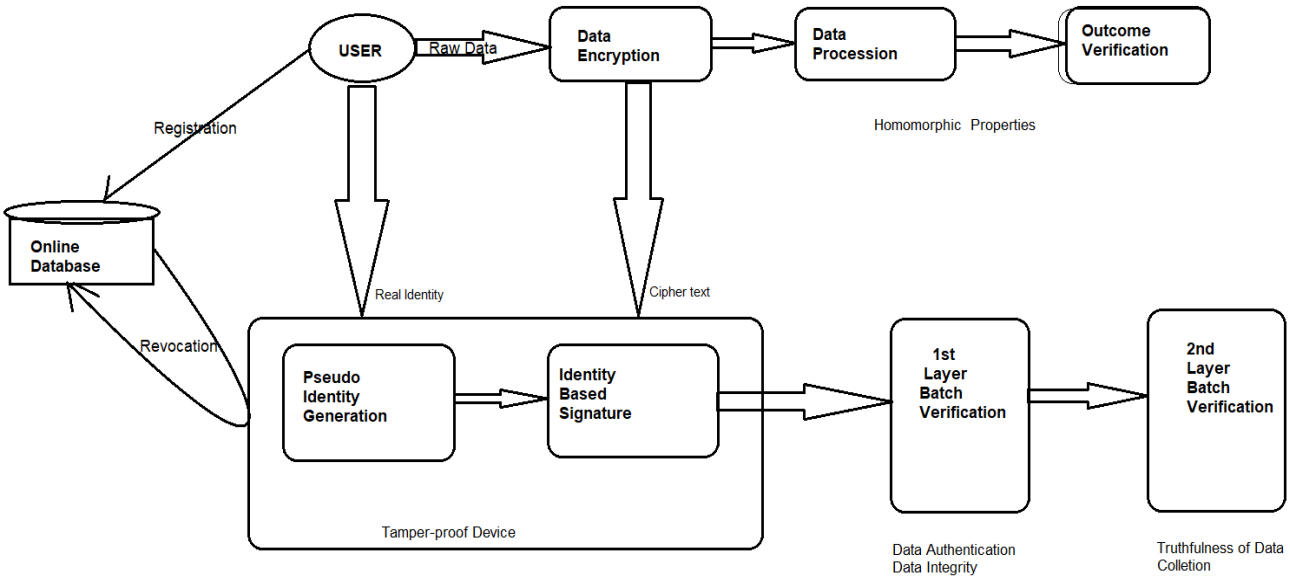



*Fig-2 : System architecture of reliable and secure data exchange.*

## IV. MODULE DESCRIPTION

In the system we have maintain four module each having specific function as follows

### A. User Interface

The user interface module is develop for intermediate user who will use it for updating the information of different merchant who is producing the data which is to identified and analysis which is then use by public users. The data submitted is identified, analysis and submitted.

### B. Server Module

In this module we have maintain the centralized server which will deployed for monitoring all the activities of public who will provide the reaction regarding handing and using of particular data product with feedback. All activities of user are maintain in the server. All Question and answer are stored and maintain by the server.

### C. Review Questionnaire And Analysis Module

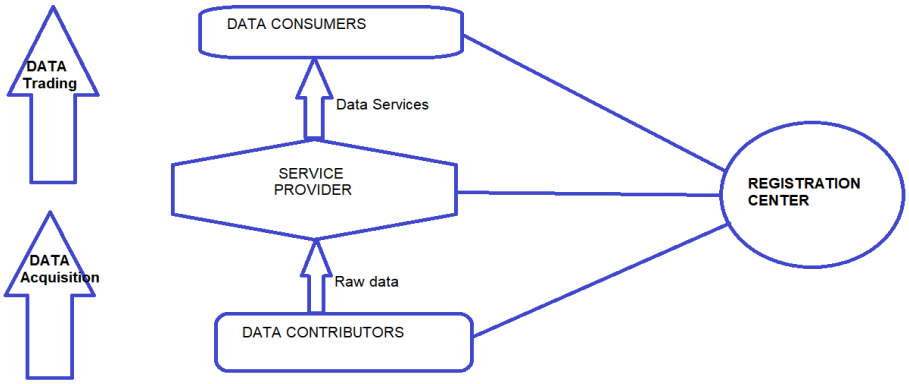
In this module the best answer for the given question will be provided so that the consumer will get the information or result which is best suited for that question. The product information will be given in the form of answer. After that and analysis module will provide the satisfaction of the customer regarding the product. The opinion given by customer can be used for further improvement and give the best data to the customer.

### D. Block Chain Formation

The block chain formation module issue for the purpose of security in which secure data analysis data is stored in main server using the cryptography algorithms so the data will be highly secure. Block chain formation is done by using Ethereum Software and based cryptographic algorithm.

## V. METHODOLOGY

The two-layer batch verifications only hold when all the signatures are valid, and fail even when there is a single invalid signature. In practice, a signature batch may contain invalid one(s) caused by accidental data corruption or possibly malicious activities launched by an external attacker. Traditional batch verifier would reject the entire batch, even if there is a single invalid signature, and thus waste the other valid data items. Therefore, tracing and/or recollecting invalid data items and their corresponding signatures are important in practice. If the second-layer batch verification fails, the data consumer can require the service provider to find out the invalid signature(s).




we consider a two-layer system model for data markets. The model has a data acquisition layer and a data trading layer. There are four major kinds of entities, including data contributors, a service provider, data consumers, and a registration center. In the data acquisition layer, the service provider pro cures massive

raw data from the data contributors, such as social network users, mobile smart devices, smart meters, and so on. In order to insensitive more data contributors to actively submit high-quality data, the service provider needs to reward those valid ones to compensate their data collection costs. For the sake of security, each registered data contributor is equipped with a tamper-proof device. The tamper-proof device can be implemented in the form of either specific hardware or software . It prevents any adversary from extracting the information stored in the device, including cryptographic keys, codes, and data. We consider that the service provider is cloud based, and has abundant computing resources, network bandwidths, and storage space. Besides, she tends to offer semantically rich and value-added data services to data consumers rather than directly revealing sensitive raw data, e.g., social network analyses, probability distributions, personalized recommendations, and aggregate statistics. The registration center maintains an online database of registrations, and assigns each registered data contributor an identity and a password to activate the tamper-proof device. Besides, she maintains an official website, called certificated bulletin board , on which the legitimate system participants can publish essential information, e.g., white lists, blacklists, resubmit-lists, and reward-lists of data contributors. Yet, another duty of the registration center is to set up the parameters for a signature scheme and a cryptosystem. To avoid being a single point of failure or bottleneck, redundant registration centers, which have identical functionalists and databases, can be installed.

## VI. RESULT AND DISCUSSION

The main aim of preserving the data privacy is replace the traditional data mining technique with high secure security double layer security model the main goal of this project is to save the dataset in high secure model when it is collected and analysis. The data prevention and consistency is maintain so that no authorization user can modify the data which is placed in server. The customer data is maintain highly secure by using Cryptography tools to build data mining models sources.

## VII. CONCLUSION

In this paper we have proposed the highly secure model which will provide the efficient for data truth-fullness and security of information. The data contributors will have confidence that whatever data is submitted is highly secure, Along with this customer will get the consistent data which will be highly secured. In real world it will provide high security to the user data. Besides, the service provider is enforced to truthfully collect and process data. Furthermore, both the personally identifiable information and the sensitive raw data of data contributor are well protected. In addition, we have instantiated reliable and secure data exchange with two different data services, and extensively evaluated their performances on two real-world datasets. Evaluation result shave demonstrated the scalability of reliable and secure data exchange in the context of large user base, particularly from computation and communication overheads. At last, we have shown the feasibility of introducing the semi-honest registration center with detailed theoretical analysis and substantial evaluations.

## VIII. REFERENCES


[1] K. Ren, W. Lou, K. Kim, and R. Deng, "A novel privacy preserving authentication and access control scheme for pervasive computing environments, "IEEE Trans. Veh. Technol., vol. 55, no. 4, pp. 1373– 1384, Jul.2006.

[2] M. Balazinska, B. Howe, and D. Suciu, "Data markets in the cloud: An opportunity for the database community,"Proc. VLDB Endowment, vol. 4, no. 12, pp. 1482–1485,2011

[3] P. Upadhyaya, M. Balazinska, and D. Suciu, "Automatic enforcement of data use policies with data lawyer," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2015, pp.213–225.

[4] T. Jung, X.-Y. Li, W. Huang, J. Qian, L. Chen, J. Han, J. Hou, and C. Su,"Account Trade: Accountable protocols for big data trading against dishonest consumers,"Proc. IEEE Conf. Comput. Commun., 2017,

a. pp. 1–9.

[5] G. Ghinita, P. Kalnis, and Y. Tao, "Anonymous publication of sensitive transactional data,"IEEE Trans. Knowl. Data Eng., vol. 23, no. 2, pp. 161–174, Feb.2011.

[6] B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preservingdata publishing: Asurveyofrecent developments,"ACM Comput. Surveys, vol. 42, no. 4, pp. 1–53, Jun.2010.

I. Bilogrevic, M. Jadliwala, V. Joneja, K. Kalkan, J. P. Hubaux, and I. Aad,"Privacy-preserving optimal meeting location determination on mobile devices,"IEEE Trans. Inf. Forensics Security, vol. 9, no. 7, pp. 1141–1156, Jul.2014.

[7] R. Zhang, Y. Zhang, J. Sun, and G. Yan, "Fine-grained private matching for proximity-based mobile social networking," Proc. IEEE INFOCOM, 2012, pp. 1969–1977.