IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 16-20

Design of an Efficient and Robust Secret Sharing Method for Big Data Analysis

Miss. Purva D. Thakare[1], Dr.Mrs.S.N.Kale[2], Dr.V.M.Thakare[3]
Student in Department of Computer Science SGBAU, Amravati , Professor in Electronics department SGBAU, Amravati, Head of Department of Computer Science ,SGBAU, Amravati

Abstract: *Big data is a phrase used to mean data sets that are too large or complex. It contains both structured and unstructured data. It becomes difficult to process data using traditional database and software techniques. Big- data can obtain large amount of individual users' sensitive data by sorting, analyzing, and mining the data. Secure sharing is a major issue. Big data demands a strong infrastructure for secure sharing. This paper is focused on five different techniques such as secure social multimedia big data sharing using a scalable JFE in TSHWT domain, Secret(N,N) threshold QR sharing approach, Verifiable secret sharing scheme using AMBTC, aggregation scheme based on secret sharing with fault tolerance and Fast secure computation based on XOR scheme. The method proposed is" secret sharing method for big data analysis using a third party auditor on hadoop storage."*
Keywords: *Privacy-preserving, data-confidentiality, security, encryption, cipher-text, decryption.*

I. Introduction

In the era of big-data large amount of data is being shared daily. As the world is entering in the age of modern information -technology, billions of people with mobile and devices such as sensors, actuators, robots are generating tremendous amount of data. Data sets whose size is beyond the ability of commonly used software systems to store, manage processes is called big data. However secure sharing is problematic. Secret sharing schemes are needed as it is necessary to keep the information private shared by the user. This paper, discusses five different secret sharing schemes such as secure sharing on big data such as secure social multimedia big data sharing using a scalable JFE in TSHWT domain, Secret (N,N)threshold QR sharing approach, Verifiable secret sharing scheme using AMBTC, aggregation scheme based on secret sharing with fault tolerance and Fast secure computation based on XOR scheme. These methods are secure but they have some drawbacks. To overcome these drawbacks, this paper proposed secret sharing method for big data analysis using a third party auditor on hadoop storage which makes sensitive information more secure so that only authenticated user can access the data and it also maintains the integrity.

II. Background

As per studies on sharing in big data many sharing schemes and algorithms for safe transfer of data have been develop to make the sharing secure and efficient. The sharing scheme in recent past years are:

Secure social multimedia big data sharing using a scalable JFE in TSHWT domain which maps hierarchical community structure of social networks into a tree structure of Haar wavelet transform for fingerprinting and encryption.It provides double layer of protection. [1].Secret(N,N)threshold QR sharing approach prevents dishonest participants from obtaining the data and can verify cheaters before revealing the shared secret [2].Verifiable secret sharing scheme using AMBTC it can specifically designed for images which offers honest participants a reconstructed secret image verification mechanism by using embedded watermark without having to send additional information through another channel [3].Aggregation scheme based onsecret sharing with fault tolerance hides user's identity through anonymity and provides fault tolerance with substitution strategy.It uses Paillier encryption [4].Fast secure computation based on XOR scheme which solves the processing problem using TUS2 method which uses protocol based on Shamir's secret sharing scheme[5].This paper introduces five sharing schemei.e. secure social multimedia bigdata sharing using a scalable JFE in TSHWT domain, Secret(N,N)threshold QR sharing approach, Verifiale secret sharing scheme using AMBTC, aggregation scheme based on secret sharing with fault tolerance, Fast secure computation based on XOR scheme.The paper is organized as follows: **Section I** Introduction. **Section II** discusses Background. **Section III** discusses previous work. **Section IV** discusses existing methodologies. **Section V** discusses attributes and parameters and how these are affected on mobility models.

Section VI proposed method and **Section VII** outcome result possible. Finally **section IX** Conclude this review paper.

III. Previous Work Done

In research literature, many sharing schemes in big data have been studied which influences privacy during sharing of data. Some sharing schemes provide secrecy so that the information should not get leaked with low computational complexity.

Conghuan ye et al. [1] has proposed Secure Social Multimedia Big Data Sharing Using a Scalable JFE in the TSHWT Domain basic idea is to map hierarchical community structure of social networks into a tree structure of Haar wavelet transform. It provides double layer of protection for social multimedia sharing in social network environment but also avoids big data superposition effect.

Pei-Yu Lin et al. [2] has proposed Distributed Secret Sharing Approach with Cheater Prevention based on QR Code basic idea is to design efficient and feasible secret QR sharing system to protect private QR data in real world applications which can verify cheaters before revealing the shared secret.

Kon-You Cai et al. [3] has proposed A Verifiable Secret Sharing Scheme based on AMBTC in which novel (t,n)-threshold verifiable scheme is proposed that can be applied to both grayscale and color images by combining AMBTC and Shamir's secret sharing scheme .It does not increase computational complexity.

Zhitao Guan et al.[4] proposed Protecting User Privacy Based on Secret Sharing with fault tolerance for Big Data in Smart Grid which proposes aggregation scheme based on secret sharing with fault tolerance in smart grid, ensures that control center gets the integrated data without revealing user's privacy. It also provides fault tolerance ability during the data aggregation.

KyoheiTokitaet. al[5] Fast Secure Computation Based on a Secret Sharing Scheme for $n < 2k - 1$ it improves the TUS2 method to allow high-speed processing. The TUS2 method requires polynomial processing because it is based on S-SSS. It proposes secure computation based on the XOR schemethe computation cost of this scheme is much lower than those of conventional methods.

IV. Existing Methodologies

Many sharing schemes have been implemented over the last several decades. There are different methodologies that are implemented for different sharing schemes i.e.secure social multimedia big data sharing using scalable JFE in TSHWT domain, Secret(N,N)threshold QR sharing approach , Verifiale secret sharing scheme using AMBTC, aggregation scheme based on secret sharing with fault tolerance, Fast secure computation based on XOR scheme.

SecureSocial Multimedia Big Data Sharing using a Scalable JFE in TSHWT domain: It is proposed to deal with social multimedia sharing According to the research , there has been no other method yet on the implantation of secure multimedia big data sharing in the TSHWT domain with a fingerprinting and encryption scheme in social network environments. First, it describes a method for fingerprint code produced by the dendrogram of hierarchical structure of social networks and conduct and experiment to get a discontinuity point vector (DPV) for TSH wavelet decomposition. Finally it proposes a fingerprinting and encryption method in the TSHWT domain. And contents protected are distributed via hybrid multicast-unicast using SNA [1].

Secret(N,N) threshold QR code sharing approach : It modifies the QR modules directly and can satisfy the essentials of steganography ,readability, robustness, security, low computational complexity and feasibility for distributed QR application .It also prevents dishonest participants from obtaining the data, and can verify cheaters before revealing the shared secret.It provides more robustness as compared to other methods [2].

Verifiable secret sharing scheme: It is combination of absolute moment block truncation coding (AMBTC) and Shamir's secret sharing scheme. The shadow size in this scheme is reduced to the one third of secret image. It offers honest participants a reconstructed secret image verification mechanism that uses an embedded watermark without having to send additional information through another channel in that manner it provides authenticity. It can increase each participant's trust more effectively [3].

Data Aggregation scheme based on secret sharing scheme and paillier encryption: It proposes a data aggregation scheme based on secret sharing with fault tolerance in smart grid, which ensures that control center gets the integrated data without revealing user's privacy. Meanwhile, it also considers fault tolerance during the data aggregation. It is secure than other conventional methods [4].

Fast Secure Computation Based on a Secret Sharing Scheme for $n < 2k - 1$: It improves the TUS2 method to allow high-speed processing. The TUS2 method requires polynomial processing because it is based on S-

SSS. The SSSS is affected by the computational cost. So that ,it proposes secure computation based on the XOR scheme.The computation cost of this scheme is much lower than those of conventional methods [5].

V. Analysis And Discussion

A joint fingerprinting and encryption (JFE) scheme based on tree-structured Haar wavelet transform (TSHWT) is proposed with the purpose of protecting media distribution in social network environments such as facebook, you-tube [1].Secret (N,N) threshold QR code sharing approach which splits the secret and conveyed with QR tags in the distribution application, and the system can retrieve the lossless secret when authorized participants cooperate [2]. Verifiable secret sharing scheme in which, the shadow size is about one third of the original secret image. It also offers the cheating prevention function [3]. Data Aggregation scheme based on secret sharing scheme and paillier encryption which ensures that control center gets the integrated data without revealing user's privacy. It also provides fault tolerance capability [4]. It uses Shamis's secret sharing scheme and XOR scheme for fast computation [5].

Table-1: Comparisons between different secret-sharing schemes

Sharing Scheme	Advantages	Disadvantages
Secure Social Multimedia Big Data Sharing using Scalable JFE in TSHWT domain	The proposed scheme can apply to social multimedia sharing, as it is scalable. Time complexity is low.	The limitation of the proposed method is the dynamical property of social networks
Secret(N,N) threshold QR code sharing approach	It can store a larger data payload and possesses the capability of correcting errors.	It lacks of accuracy.
Verifiable secret sharing scheme	It provides cheating prevention function.	It is restricted to images only.
Data-Aggregation scheme based on secret sharing scheme and paillier encryption	This scheme is scalable. It is resistant to differential attacks. It provides anonymity and fault tolerance.	The limitation of the proposed method is that it uses substituting strategy which is complex.
Fast Secure Computation Based on a Secret Sharing Scheme for n < 2k – 1	It is efficient than other methods. It has high processing speed.	It requires high storage capacity.

VI. Proposed Methodology

Secure Sharing Schemes are essential in the age of big-data. To make the sharing secure and so that the privacy should be maintained, this paper explains five different methods. This method enhances security, privacy. But it also has some drawbacks" secret sharing method for big data analysis using a third party auditor on hadoop storage" is proposed to overcome this drawbacks. The proposed method is the hybrid method i.e. it is the combination of above methods.

For Sharing:- In the proposed method the user first uploads the file on account and then shares it with the receiver which user wants to. Then the log entry of data is created every-time and data is replicated when data is shared. The key generation takes place public key is stored in the database and private key is sent with the data .Then recipient receives the data and access it by decrypting the private key. The decrypted key and the public key is then matched by the third party auditor .If it does not match then it generates error message otherwise data is accessed by authenticated user.

For data Integrity:- For data integrity purpose, if the intruder wants to modify or delete the data .The original file is already stored in the mirror directory and modified file in main style. When modification occurs it generate a message and send it to user for approval. If user return yes then modification is done by authenticated user otherwise the message is corrupted.

Basic steps of algorithm:

Step1: First the sensitive data is fetched and uploaded on hadoop and the data is shared.

Step2: Log entry of the data is generated.

Step3: With the help of proper algorithm the key is generated.
Step4: Recipient decrypts the key and access the data.
Step5: Both the public and private keys are matched with the help of third party auditor if it does not matched then error message is generated otherwise user can access the data.
Step 6: If intruder tries to modify the file then modified file is compared with original file and message is generated and the send to the user for approval of modification, if the user sends approved message the changes are made otherwise not.
Diagrammatic representation of proposed method is shown as follows:

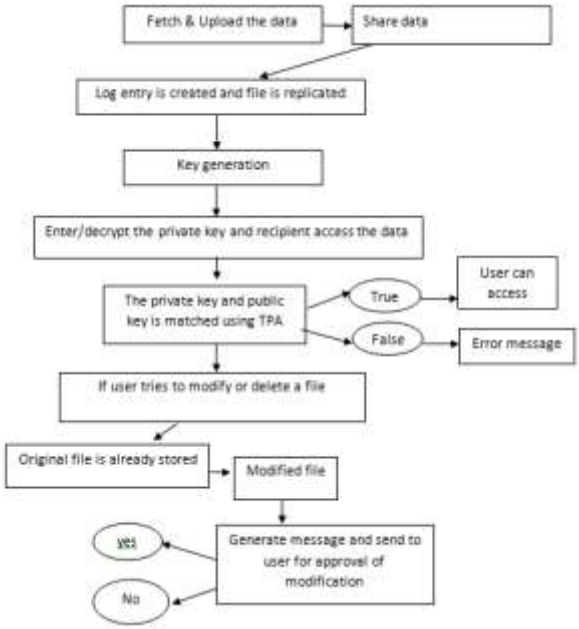

Figure-1: *Data flow diagram of an efficient and robust secret sharing scheme for big data*

VII. Outcome And Possible Result

In this way the proposed method by performing the encryption make sensitive information more secure so that only authenticated user can access the data and it also maintains the integrity of data so that the intruder cannot modify the data and user can retain the original data.

VIII. Conclusion

This paper is focused on the study of various sharing scheme i.e. secure social multimedia big data sharing using a scalable JFE in TSHWT domain, Secret(N,N)threshold QR sharing approach ,Verifiale secret sharing scheme using AMBTC, aggregation scheme based on secret sharing with fault tolerance, Fast secure computation based on XOR scheme. But some drawbacks are tracked in this existing scheme so to overcome this drawback, this paper proposed "secret sharing method for big data analysis using a third party auditor on hadoop storage" . Due to this method, the information or data shared is made secure by using proper encryption algorithm and avoids unwanted user to access the data. It also helps to maintain data integrity.

IX. Future Scope

From observations of the proposed method the future work will include to design a method with less computational overhead.

References

[1]. Conghuan Ye, Hefei ling, ZenggangXiong, Fuhao Zou Cong Liu, Fang Xu "Secure Social Multimedia Big Data Sharing Using Scalable JFE in the TSHWT Domain",*ACM Trans. Multimedia Computing Communication*, Vol. 12, No. 4s,61.1-61.23, Vol. 12, No. 4s,61.1-61.23.

[2]. Pei-Yu Lin" Distributed Secret Sharing Approach with Cheater Prevention based on QR Code", *IEEE Transactions on Industrial Informatics* , VOL. 129, 955-966, February 2011.

[3]. Kun-You Cai, Shin-Shian Wang, Pei-Feng Shiu, Chia-Chen Lin," A Verifiable Secret Sharing Scheme based on AMBTC", *ACM Access* , VOL. 129, 955-966 , February 2011

[4]. Zhitao Guan, Guanlin Si, Xiaojiang Du, Peng Liu, Zijian Zhang, Zhenyu Zhou"Protecting User Privacy Based on Secret Sharing with Fault Tolerance for Big Data in Smart Grid",*IEEE ICC 2017 SAC Symposium Big Data Networking Track* , Vol. 12, No. 4s,61.1-61.23, 2017.

[5]. KyoheiTokita, Keiichi Iwamura "Fast Secure Computation Based on a Secret Sharing Scheme for $n < 2k–1$",*ACM Trans. Multimedia Computing Communication* , Vol. 12, No. 4s,61.1-61.23, June 2018.