Private Message Sharing System On Distributed Servers With Erasure Coding Method

Sakshi R. Awadhiya
PG Scholar
Department of Computer Science and Engineering
Government College of Engineering
Amravati-444 604,Maharashtra,India

Prof. R.V. Mante
Assistant Professor
Department of Computer Science and Engineering
Government College of Engineering
Amravati-444 604,Maharashtra,India

***Abstract*:- Recently due to large number of confidential information, a secure storage distributed system is generated using private information retrieval(PIR) process. Such system are designed such that it provides security to both user privacy as well as data stored. To share the information on distributed system secret sharing is used and for backup purpose in case of node failure secure erasure coding scheme is used.**

Keywords—Private Information Retrieval, secret sharing scheme, data security, Erasure coding method.

I. INTRODUCTION

A Secure Distributed Servers are distributed systems which store data, shared between different systems, securely or in this case taking investigation process as an example where information related cases which are shared between officials are confidential. Private Information Retrieval(PIR) problem is used for the situation when the information is being shared in the presence of an eavesdropper must be secured. Our work is divided into two cases for security: i) user privacy (concealing the index of the desired message) from each of the databases and the other is data security from an eavesdropper who can access to one of the databases or link between the database and the user to obtain information about messages. A new secure distributed storage system and its corresponding PIR scheme is studied that simultaneously protect user privacy from the non-colluding databases and data security from an eavesdropper. A secret sharing scheme is made for distributed databases to preserve data security, and our PIR scheme relies on an existing PIR scheme to keep user privacy in the secure distributed databases. In contrast to existing secure distributed storage systems, the redundant secret shares which are exploited as a side information to increase the rate of the PIR are possessed by databases. To different cases are considered according to the availability of the data about which secret shares are stored in other databases. The first case is that there should be nothing about the secret shares that are stored in any databases. The second case is that each database is aware of the stored secret shares in other databases which can invade user privacy during a PIR procedure. In the design of secure distributed databases and PIR process, we prevent an eavesdropper or third party from obtaining information about the messages about different cases by preventing the individual database from obtaining any information about the messages from the stored data. In our proposed system, we will discuss about security of data on distributed systems where multiple servers will run at a time. This new technique is used to develop a secure communication system for highly secured cases generally investigated by Intelligence Bureau. Here the message will be partitioned into n shares which will be distributed over multiple servers. Due to which the availability of the message is increased in case of any attack made on particular database. To achieve data availability in any case of node failure, it became important to store backup of all the shares. But in other case, the server space is remained consumed by the same shares of messages on multiple servers. Therefore to reduced server space required to store backup of data we proposed secure erasure coding method, in which the shares will be stored in polynomial format on backup servers. Two level encryption technique can be used which can improve the security of the existing system.

Here we proposed secure message sharing over distributed network with integrity checking and data availability. In the design of secure distributed databases and PIR process, we prevent an external party from obtaining information about the messages by preventing the individual database from obtaining any information about the messages from the stored data. We provide a secure scheme against not only an external eavesdropper who can access to the stored data in the database but also a possible internal eavesdropper (eavesdropping database) which intends to leak confidential information about messages and user privacy to the adversaries. We will also focus on the lower-bound on the capacity of PIR that can be further improved by modifying our scheme especially when the number of databases is less. To retrieve shares privately, the user must generates queries and send it to databases. The queries are generated without information about the messages at the user, thus they are independent of the messages. For data security from the eavesdropper, databases store securely encoded data of the messages, and thus the eavesdropper obtains no information on the messages from the stored data and the link between databases and the user. We consider two possible scenarios whether or not each database knows which secret shares are stored in other databases by coordination between databases:

- Databases without coordination (Scenario 1): The databases do not coordinate each other, thus they do not know the set of the indices of the secret shares stored in other databases, i.e., DB 1 has no information about its shares. This scenario can be realized when the main server operates the storage system does not share this information or it outsources the external databases to use their storage capabilities.

- Databases with coordination (Scenario 2): The databases have the knowledge of the set of the indices of the stored secret shares in other databases by coordination.

II. LITERATURE SURVEY

To guarantee the essential attributes of storage systems such as reliability, security, and so on, the use of regenerating codes facilitates storage systems to efficiently cope with node failures where the message shares are stored in the

form of codes and transferred to multiple servers. The message shares can be recovered using the codes hence there is no need to store replica of each share[1]. But lack of security and integrity check provide us a new way to do more research on the topic. In Generic repair schemes, a linear secret sharing schemes can be securely repaired. The author proposed another scheme where codes are stored in the form of polynomial. In case of any failure the shares can be recovered by solving the polynomials[2]. To study the secure repair bandwidth under the general repair model when the secret sharing scheme being repaired is one of the open problem. To study secure repair where active adversarial nodes are present that may deviate from the prescribed repair protocol is one of the interesting problem. Also in this scheme complexity level is too high to go through this.

In [3], the multi-round private information retrieval over distributed network is being studied where it proves that the capacity of multi-round PIR is the same as the capacity of PIR of single round. The result includes T-privacy constraints. There is a drawback of storage overhead and no advantage in terms of capacity from multi-round over single-round schemes, nonlinear over linear schemes.

In [4], a recursive techniques is proposed to hide extra information in between the parts of Shamir's secret sharing schemes. This hidden information may be used for validation of shares at the time of secret reconstruction.

Simultaneous node failures can be revealed by DSS which is needed to be recovered with local connections. The design of coding schemes for DSS satisfy these properties. No major encryption technique is justified[5].

The data stored should be right even when some servers failed are considered in Secure storage and retrieval of information [SSRI] [6]. SSRI extend a property where an adversary can corrupt servers totally but some during given time interval. It is assumed that faults can occur at reconstruction time which is major shortcomings. The capacity of PIR is especially significant because of the central role played by PIR across a diverse array of problems that include locally decodable and batch codes, secure multiparty computation, instance hiding, secret sharing, and oblivious transfer[7].

- Codes for Distributed Storage

The regenerating codes model introduced in [8] considers optimizing two important resources: the storage capacity required by each node, and the repair-bandwidth. There exists a substitute between two resources, and lower bounds on their requirements were derived. Subsequent to their work, several explicit codes were constructed for the MSR and the MBR regimes of regenerating codes, many of which meet these bounds.

- Shamir's secret sharing :

A possible method to ensure information-theoretic security from passive eavesdroppers is to employ Shamir's secret sharing scheme[14], where the data is encoded and stored in a set of n nodes such that the entire data can be retrieved from any k nodes, while access to data in any (k-1) or fewer nodes provides zero information about the data. During repair of a failed node, this scheme requires a download of the entire data to a central location, following which the replaced node's data is re-encoded. Thus, the repair operations are inefficient in classical erasure codes, mandating significant network resources.

- Secure Network Coding :

The literature on secure network coding [15] primarily considers a multicast setting where a single source of data and every destination is interested in obtaining all the data sent by the source. Furthermore, with respect to security from passive eavesdroppers in the multicast setting, only the scenarios where the eavesdropper can access subsets of links is well understood in the literature.

III. PROPOSED METHODOLOGY

In our proposed system, we proposed secure message sharing over distributed network with integrity checking and data availability. In our system, the message will be encrypted on client side before transfer to the server. Then the message will be divided into n shares on server, shuffle their indices, perform second level encryption and store on multiple servers randomly. At the time of decryption, user need to specify the key, the message will be downloaded from different servers and on client side the message will be rejoined in the sequence to get combined message. After that the re-combined message will be decrypted on client side by the system automatically and deliver it to user. The keys required for client side encryption, will be maintained on the basis of upload date, time and some user defined algorithms.

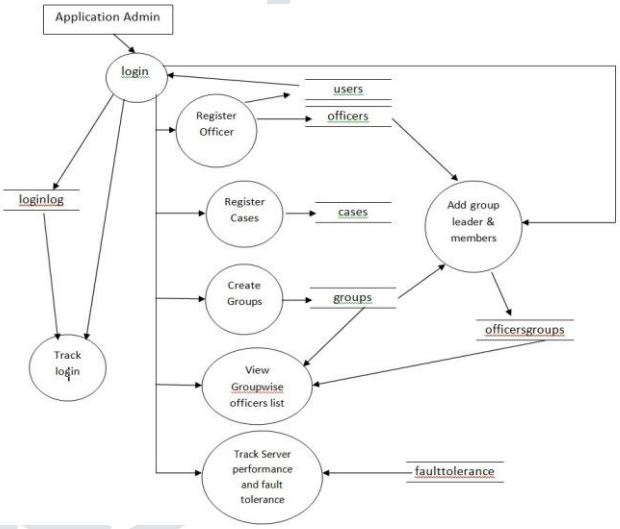

Fig a. Application Administrator

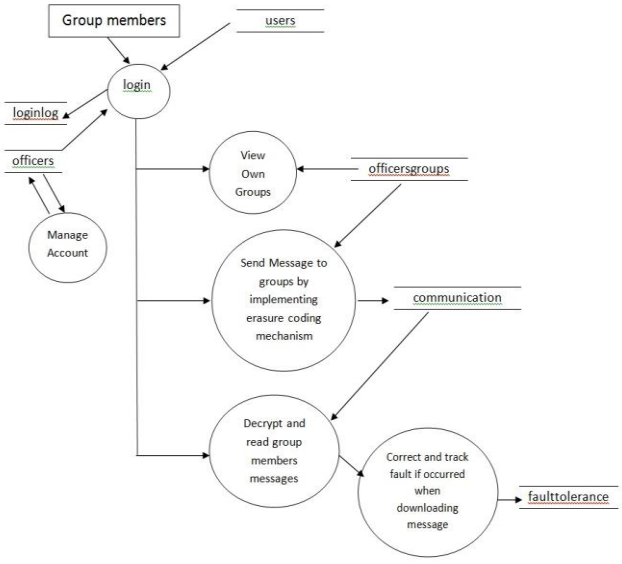

Fig b. Officers

Taking Intelligence Bureau into account there are various sections in which process is divided. Firstly administration

panel page will be shown where various options will be given as shown in fig a. where admin will login. As admin is generally the head or controller of all the process happening, there are various work under him such as registering the officers and also registering the cases. He will also create groups where officers will be divided accordingly and can assign the cases to different groups. Group wise list is also shown in fig a. which comes in handy when new case will come. Officer group must have leader and members where leader can assemble its group members according to the work to be done on the case and also give direction to them according to the progress on investigation. In this page admin can track the server performance as well as fault tolerance where in case of any server failure, system will automatically recover the missing share using erasure codes stored on backup server.

Officer can also login using id and password which will be mailed to them as soon as admin will register them. As shown in fig b. officer can later change their password of their interest. Officer has contact with its group members and can send message about the assigned case to every member of their group members securely. The messages shared between group members will be encrypted firstly on client side by making key using the date and time of said message. The algorithm used for encryption is user defined and use time and date as input to secure the key as well as message. Server side encryption can also take place for additional security to message as well as all information backup in case of node failures.

IV.CONCLUSIONS

A PIR problem is taken into account for distributed databases in the presence of an eavesdropper. Depending on whether the data indices in other databases are known at a database, we proposed two PIR schemes to ensure user privacy from each database and also security of data (in case there is an eavesdropper) at the same time. We overcome many of the existing drawbacks by managing space as well as security. Hence our system is more efficient.

References

[1] A. Fazeli, A. Vardy, and E. Yaakobi, “Codes for distributed PIR with low storage overhead,” in Proc. IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, pp. 2852-2856, Jun. 2015.

[2] W. Huang and J. Bruck, “Generic secure repair for distributed storage,” arxiv preprint arXiv:1706.00500, Jun. 2017.

[3] H. Sun and S. A. Jafar, “Multiround private information retrieval: capacity and storage overhead,” arXiv preprint arXiv:1611.02257, Nov. 2016.

[4] A. Parakh and S. Kak, “Recursive Secret Sharing for Distributed Storage and Information Hiding,” Information Sciences, vol. 181, no. 2, pp. 335-341, Dec2009.

[5] O. O. Koyluoglu, A. S. Rawat, and S. Vishwanath, “Secure cooperative regenerating codes for distributed storage systems,” IEEE Transactions on Infomation Theory, vol. 60, no. 9, pp. 5228-5244, Sep. 2014.

[6] J. A. Garay, R. Gennaro, C. Jutla, and T. Rabin, “Secure distributed storage and retrieval,” Theoretical Computer Science, vol. 243, no. 1-2, pp. 363-389, Jul. 2000.

[7] H. Sun and S. A. Jafar, “The capacity of private information retrieval with colluding databases,” in Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, Dec. 2016.

[8] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, “Network coding for distributed storage systems,”IEEE Transactions on Information Theory, vol. 56, no. 9, pp. 4539–4551, 2010.

[9]N. B. Shah, K. V. Rashmi, P. V. Kumar, and K. Ramchandran, “Distributed storage codes with repair-by-transfer and nonachievability of interior points on the storage-bandwidth tradeoff,” IEEE Transactions on Information Theory, vol.58, no.3, pp.1837–1852, Mar. 2012.

[10] C. Tian, “Rate region of the (4, 3, 3) exact-repair regenerating codes,” in IEEE International Symposium on Information Theory, Istanbul, Jul. 2013.

[11] B. Sasidharan, K. Senthoor, and P. V. Kumar, “An improved outer bound on the storage-repair-bandwidth tradeoff of exact-repair regenerating codes,” arXiv preprint arXiv:1312.6079, 2013.

[12] Y. Han, R. Zheng, and W. Mow, “Exact regenerating codes for Byzantine fault tolerance in distributed storage,” in Proc. IEEE International Conference on Computer Communications (INFOCOM), Florida, USA, March 2012.

[13] F. Oggier and A. Datta, “Byzantine fault tolerance of regenerating codes,” in IEEE International Conference on Peer-to-Peer Computing, 2011, pp. 112–121.

[14] A. Shamir, “How to share a secret,” Communications of the ACM, vol. 22, no. 11, pp. 612–613, 1979.

[15] J. Feldman, T. Malkin, C. Stein, and R. Servedio, “On the capacity of secure network coding,” in Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing, 2004.

[16] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, “Private information retrieval,” J. ACM, 45, 1998. Earlier version in FOCS 95.

[17]Hua Sun and Syed A. Jafar “Blind Interference Alignment for Private Information Retrieval”,2016.