International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072

Privacy Protection of Big Data Using RingSignature

Anushree M.R[1], Renuka Patil[2]

[1]PG Student, Department of CSE, Global Academy of Technology, Bengaluru, Karnataka, India
[2]Associate Professor, Department of CSE, Global Academy of Technology, Bengaluru, Karnataka, India

---

Abstract - *In recent years, big data became a hot research topic. The increasing quantity of huge information also will increase the prospect of breaching the privacy of people. Since huge information need high process power and enormous storage, distributed systems are used. As multiple parties are concerned in these systems, the danger of privacy violation is augmented. There are varieties of privacy-preserving mechanisms developed for privacy protection at completely different stages (e.g., information generation, information storage, and information processing) of a big data life cycle. The goal of this project is to produce a comprehensive summary of the privacy preservation mechanisms in big data and present the challenges for existing mechanisms. Above all, during this project, we are proposing new mechanism for the protection of big data privacy called as ring signature.*

Key Words: **Big Data, Attribute Based Encryption, Identity Based Encryption, Homomorphic Encryption, and Ring Signature**

1. INTRODUCTION

All the huge quantity information generated from different sources in multiple formats with terribly high speed is referred as big data. Big data has become a really active analysis space for last few years. The information generation rate is growing rapidly so it is difficult to handle using traditional strategies or systems. Meanwhile, huge knowledge may be structured, semi-structured, or unstructured, that adds additional challenges once playing data storage and process tasks [1]. Therefore, to this end, we'd like new ways that to store and analyze data in real time.

Users' privacy may be breached under the following circumstances[2-3] :

- User personal information is combined with external datasets may lead to the inference of new facts about the users. Those facts may be secretive and not supposed to be revealed to any others.
- Sometimes personal information is collected and it will give value to business. For example, individual's shopping habits may reveal a lot of personal information about the user who will shop in online.
- Sometimes sensitive information are stored and processed in a location that is not secured properly. In such cases information leakage may occur during storage and processing phases

Protecting privacy in big data may be a quick growing analysis space. Whereas this paper introduced the essential idea of privacy protection in big data.

2. LITERATURE SURVEY

As we have said already the size and the variety of the data is growing rapidly, the tools and techniques that are used to handle such data should also be upgraded. Here there are some existing privacy preserving techniques.

V. Goyal, O. Pandey, A. Sahai, and B.Waters [5] worked on paper that proposes attribute based encryption scheme. It is one of the encryption techniques which ensure end to end big data privacy in cloud storage system.

In attribute based encryption, access polices are defined by the data owner and information are encrypted under those policies. The information can only be decrypted by the users whose attributes satisfy the access policies defined by the data owner. When dealing with big data one may often need to change data access policies as the data owner may have to share it with different organizations. The current attribute based encryption does not support policy updating. The policy updating is a challenging task in this type of encryption scheme. The reason for this is once the data are outsourced to the cloud storage, the data owner would not keep the local copy in their system. If the data owner wants to update the policy, he has to transfer the data back to the local system, he has to re-encrypt the data under new policy and store it back on the cloud server. This leads to very high communication overhead and high computational cost [6].

X. Boyen and B. Waters [8] proposed identity based encryption in their paper. Identity based encryption is an alternative to Public key encryption which is proposed to simplify key management system in a certificate-based public key infrastructure by using human identities like email address or IP address as public keys. The identity based encryption scheme was proposed to preserve the privacy of sender and receiver. Identity based encryption does not support the update of cipher text receiver.

C. Gentry made study on homomorphic encryption [10]. Public cloud is more vulnerable to privacy violation because of multi-tenancy and virtualization. The different cloud users may share the same physical space. In such a scenario the chances of data leakage also high. One way to protect the data on cloud is to first encrypt the data and store them on cloud to provide privacy. Then allow the cloud to perform

computations over encrypted data. Fully homomorphic encryption is the type of encryption technique which allows functions to be computed on encrypted data. Given only the encryption of a message, one can obtain an encryption of a function of that message by computing directly on the encryption. Homomorphic encryption provides privacy but the drawback is computational complexity and sometimes it is very hard to implement with existing technologies.

3. PROPOSED SYSTEM

In the proposed system, new privacy protection scheme is introducing, that is the Ring signature. Ring signature is the closed group file authentication system, where the data owner in the group can able to securely share his files within the group without outsider's inference. And simultaneously integrity of the file is maintained. For each and every file the Ring Signature will be created. With the help of Ring Signature the group members can able to successfully decrypt and download the file from the Hadoop.

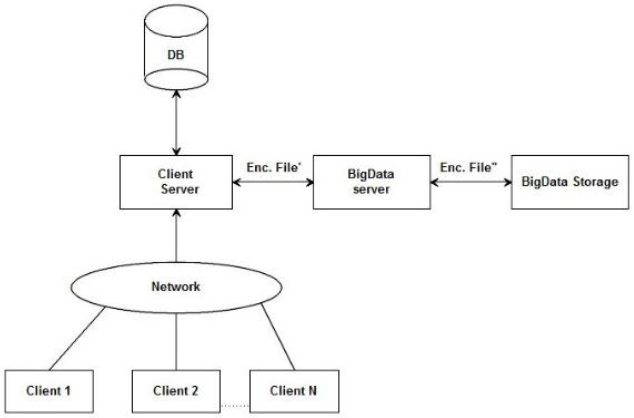

Fig -1: Proposed System Architecture

The above architecture explained about how we are using ring signature in order to provide security in big data. Admin will create a group by collecting their id. Then he will provide security keys to all users in the group. If client1 send the file, it will be stored in client server. In client server one database is maintained. After that file will be encrypted with the ring signature and it is send to big data server. File will be second time encrypted before going to the big data storage by using AES key. If anybody wants to download the file he has to upload the ring signature. Otherwise he cannot download the file from bigdata storage

3.1 Ring Signature Creation

User will get the public keys of all the members present in the user's group. Then he will generate the Hashcode for the uploading file. Using this, do XOR operation of the Hashcode with the public keys of all the members present in the user's group. The final result of XOR operation is the SecureMD (Secure Message Digest). Then Write the obtained SecureMD in to the file by that Secured file will be created. Then Encrypt the Secured file with the private key of uploading user, the Encrypted file is called Ring Signature of the uploading file. After creating Ring Signature user sends the Ring Signature file to the selected members of the group through the email.

3.2 File Encryption Process

- Fetch the AES Key
- Encrypt the uploaded file using AES Encryption Technique.

3.3 File Decryption Process

- Fetch the AES Key
- Decrypt the downloaded file using AES Decryption Technique

3.4 Ring Signature Verification Process

The destination user will get the public keys of all the members present in the user's group. Generate the Hashcode for the downloaded file. After that, do XOR operation of the Hashcode with the public keys of all the members present in the user's group. The final result of XOR operation is the SecureMD (Secure Message Digest), let us call it as SecureMD1. Decrypt the Ring Signature File with the public key of the uploaded user and get the SecureMD, let us call it as SecureMD2. Compare the SecureMD1 and SecureMD2. If matches then Ring Signature verification process successful and file will be successfully downloads to the Client system. Else Ring Signature verification process failure and it will give the error message that Ring Signature mismatching.

4. CONCLUSIONS

The amount of data is growing every day and it is impossible to imagine the next generation applications without producing and executing data driven algorithms. We have investigated privacy challenges in each phase of big data life cycle and discussed some advantages and disadvantages of existing privacy preserving technologies in the context of big data applications. A lot of works have been done to preserve the privacy of users from data generation to data processing phase, but still exist several open issues and challenges.
Sometimes the data owned by an organization does not have sufficient information to discover useful knowledge in that domain, and acquiring that data may be costly or difficult due to legal constraints and fear of privacy violation. To solve such problems, we need to design privacy preserving distributed analytic systems which are able to process different datasets from different organizations while preserving the privacy of each dataset. Secure multiparty computation techniques such as homomorphic encryption can be deployed to solve such issues. The main challenge in

deploying homomorphic encryption in the context of big data analytics is to keep the computational complexity as low as possible.

REFERENCES

[1] J. Manyika et al., Big data: The Next Frontier for Innovation, Competition, and Productivity. Zürich, Switzerland: McKinsey Global Inst., Jun. 2011, pp. 1_137.
[2] A. Katal, M. Wazid, and R. H. Goudar, ``Big data: Issues, challenges, tools and good practices,'' in Proc. IEEE Int. Conf. Contemp. Comput., Aug. 2013, pp. 404_409.
[3] Z. Xiao and Y. Xiao, ``Security and privacy in cloud computing,'' IEEE Commun. Surveys Tuts., vol. 15, no. 2, pp. 843_859, May 2013.
[4] S. Singla and J. Singh, ``Cloud data security using authentication and encryption technique,'' Global J. Comput. Sci. Technol., vol. 13, no. 3,pp. 2232_2235, Jul. 2013.
[5] V. Goyal, O. Pandey, A. Sahai, and B.Waters, ``Attribute-based encryption for _ne-grained access control of encrypted data,'' in Proc. ACM Conf. Comput. Commun. Secur., Oct. 2006, pp. 89_98.
[6] J. Bethencourt, A. Sahai, and B. Waters, ``Ciphertext-policy attributebased encryption,'' in Proc. IEEE Int. Conf. Secur. Privacy, May 2007, pp. 321_334.
[7] K. Yang, X. Jia, and K. Ren, ``Secure and verifiable policy update outsourcing for big data access control in the cloud,'' IEEE Trans. Parallel Distrib. Syst., vol. 26, no. 12, pp. 3461_3470, Dec. 2015.
[8] X. Boyen and B.Waters, ``Anonymous hierarchical identity-based encryption (without random oracles),'' in Proc. Adv. Cryptol. (ASIACRYPT), vol. 4117. Aug. 2006, pp. 290_307.
[9] M. Green and G. Ateniese, ``Identity-based proxy re-encryption,'' in Proc. Int. Conf. Appl. Cryptogr. Netw. Secur., 2007, vol. 4521. pp. 288_306.
[10] C. Gentry, ``A fully homomorphic encryption scheme,'' Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, 2009.