"A PRE-VALIDATION ACCESS TO INTERMIDIARY INCORRUPTION IN BIG DATA AMBIENCE"

[1]DEEPALI NIHARE
Nagpur Institute of Technology, Nagpur, Department of Computer Science & Engineering
niharedeepali@gmail.com

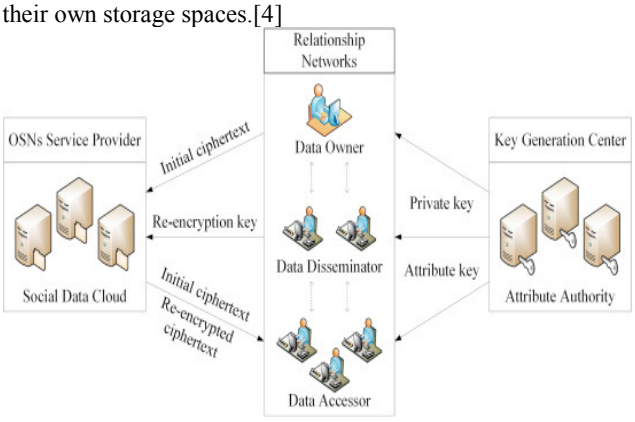
[2] PROF.JAGDISH PIMPLE
Nagpur Institute of Technology, Nagpur, HOD, Computer Science & Engineering
Jagdish.pimple@gmail.com

ABSTRACT: *The pre-validation mechanism combines the advantages of incorruption conditional re-intermediary multi-sharing mechanism with the attribute-based authentication technique, thus achieving attributes authentication before re-encryption, and ensuring the security of the attributes and data. the growing amount of data, the demand of big data storage significantly increases. However this paper proves that system is secure with encryption technique and the proposed pre-verification & validation mechanism could significantly increase & enhance the system & their technique with most security level in Big Data ambience.*

***Keywords*:** Re-encryption, pre-validation, privacy-preservation, BIG DATA

1. INTRODUCTION

Due to recent technological development, the amount of data generated by social networking sites, sensor networks, Internet, healthcare applications, and many other companies, is drastically increasing day by day. All the huge amount of data generated from different sources in multiple formats with very high speed is referred as big data. Big data has become a very active research area for last couple of years. The data generation rate is growing so rapidly that it is becoming extremely difficult to handle it using traditional methods or systems [1]. Meanwhile, big data could be structured, semi structured, or unstructured, which adds more challenges when performing data storage and processing tasks. Therefore, to this end, we need new ways to store and analyze data in real time. Big data, if captured and analyzed in a timely manner, can be converted into actionable insights which can be of significant value. It can help businesses and organizations to improve the internal decision making power and can create new opportunities through data analysis. It can also help to promote the scientific research and economy by transforming traditional business models and scientific values [2].Cloud computing provides seemingly unlimited resources as services to cloud users by rearranging various resources.[3] Cloud storage as one of the most popular cloud services enables cloud users to store tremendous amount of data in the cloud, which may exceed their own storage spaces.[4]



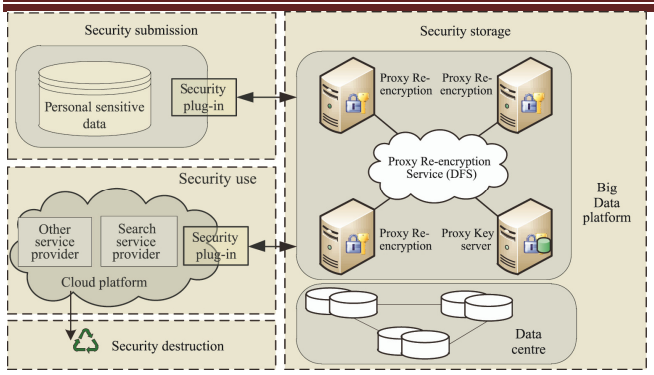
2. LITERATURE SURVEY

In this part, we review some related work on multi sharing Mechanism in big data. The comparison between our work and previous ones is shown afterwards.

Convergent Encryption Technique (CET) as the most prominent manifestation vulnerabilities of Message Locked Encryption was introducing [2]. In CTE technique, a user employs the hashing code of data as the key to encrypt the data, which results in any user with the same data can generate the same cipher text, thus realizing DE duplication. However, CE suffers from offline brute-force dictionary attacks. Moreover, it is hard to support user revocation. Bellare et al. proposed DupLESS to resist & assist the above-mentioned brute-force attack technique [3] by introducing a Key Server. But it still cannot control data access of other data users in a flexible way.[4] constructed a session-based-key & convergent key management scheme and a convergent key sharing scheme to recover the problem caused by recently changed ownership and data blocks. But this work requests all data owners communicate with each other to manage their session key. Liu et al. proposed a secure cross-user DE duplication scheme that supports client-side encryption without requiring any additional independent servers by applying a password authenticated key exchange protocol [5]. But this scheme requests that the data

3. SYSTEM ARCHITECTURE & DEFINITION

We take a fundamental of our system. We propose the system architecture mainly for many receivers to share a data our system can apply to the situation when the data is extremely large. When a user uploads his data to the cloud, he needs to encrypt the data to prevent that the data may be exposed to the cloud server. Then proxy, which is a semi-trust party, is used here for re-encrypting the cipher text. Every receiver who wants to share the data has his own key for decryption. Our system can let the receiver to obtain his desired data and ensure the rest of the data still been encrypted. Our system not only provides re-encryption operation, but can also ensure the authentication between data providers and receivers without leaking any privacy.



Different from schemes totally based on identity, we also adopt authentication that based on attribute which enhances the security to resist tracing and collusion attacks. With the two techniques, our system is able to meet the demand of flexibility and privacy preserving in big data storage.

3.1 PRIVACY IN DATA GENERATION PHASE

After cipher the Data, it is being stored in the HDFS. Then the sharing the data among the multiple receivers is done. So for a given cipher text, no one knows the identity information of the sender and receiver. Here in hiding the receiver information's anonymization is deployed of particular receiver. For Example, if the receiver enters the username that name will be anonym zed and after that it will send to the particular user

3.2 Algorithm & Equations

KeyGen (1κ**):** Let $\mathbb{S} = (p,\mathbb{G},\mathbb{G}T , e)$ be a bilinear map group system with randomly selected generators $g, h \in \mathbb{G}$, where $\mathbb{G},\mathbb{G}T$ are two bilinear groups of a large prime order p, $|p| = O(\kappa)$. Makes a hash function $Hk(\cdot)$ public. For a CSP, chooses a random number $s \in R\ \mathbb{Z}p$ and computes $S = gs \in \mathbb{G}$. Thus, $skp = s$ and $pkp = (g,S)$. For a user, chooses two random numbers α, $\beta \in R\ \mathbb{Z}p$ and sets $sku = (\alpha, \beta)$ and $pku = (g,h,H1 = h\alpha,H2 = h\beta)$.

TagGen(sk,F, $\mathcal{P}$**):** Splits F into $n \times s$ sectors $\{mi,j\}i\in[1,n],j\in[1,s] \in \mathbb{Z}n\times s$ p . Chooses s random $\tau 1, \cdots, \tau s \in \mathbb{Z}p$ as the secret of this file and computes $ui = g\tau i \in \mathbb{G}$ for $i \in [1, s]$. Constructs the index table $\chi = \{\chi i\}n$ i=1 and fills out the record χi a in χ for $i \in [1, n]$, then calculates the tag for each block mi as { $\xi(1) \leftarrow H\Sigma s$ i=1 τi (Fn), $\xi(2)$ $k \leftarrow H\xi(1)$ (Ck), $\xi(3)$ $i,k \leftarrow H$ ξ (2) k (χi), $\sigma i,k \leftarrow (\xi(3)$ i,k $)\alpha \cdot (\Pi s$ j=1 u mi,j j $)\beta$,

Where Fn is the file name and Ck is the CSP name of $Pk \in \mathcal{P}$. And then stores $\psi = (u, \xi(1), \chi)$ into TTP, and $\sigma k = \{\sigma i,j\}\forall j=k$ to $Pk \in \mathcal{P}$, where $u = (u1, \cdots, us)$. Finally, the data owner saves the secret $\zeta = (\tau 1, \cdots, \tau s)$.

4. PROPOSED SYSTEM

We proposed a new notation called pre-authentication mechanism in the model of IBNPRE. Different from the existing work, the proposed system can verify users'& validate it attributes before data sharing, thus satisfying the actual needs of users. The data of users can be shared with users having appointed attributes, and others have no access to the data. Our new proposed pre-authentication mechanism can provide multi-dimension privacy protection including data, user identities, and attributes. Only users whose attributes are authenticated could be qualified to share the data. This enhances the protection of user privacy preserving.

5. CONCLUSION

In this paper, we conclude that multi-sharing, anonymous and CCA-secure data sharing in big data context. Furthermore, we propose a new notion called pre-authentication in the proxy re-encryption system, which can ensure that only users whose attributes have been verified are permitted to obtain the data and provide well protection for the private attributes. The pre-authentication function greatly facilitates the needs of the users. Besides, we prove that users' data, identities and attributes are protected, and the pre-authentication process enhances the security of the system. To the best of our knowledge, we are the first to propose the concept of pre-authentication in this aspect.

6. REFERENCES

[1] Kun Wang, Jiahui Yu, Xiulong Liu, Song Guo, "A PreAuthentication Approach to Proxy Re-encryption in Big Data Context," *IEEE Transactions on Big Data*, 2332-7790 (c) 2016 IEEE

[2] Meyer, D.T., Bolosky, W.J.: A study of practical deduplication. ACM Trans. Storage 7(4),1–20 (2012)

[3] Ali, M., Dhamotharan, R., Khan, E., Khan, S.U., Vasilakos, A.V., Li, K., Zomaya, A.Y.:SeDaSC: secure data sharing in clouds. IEEE Syst. J. 99, 1–10 (2015)

[4] X.Y., Chen, J.J.: External integrity verification for outsourced big data in cloud and IoT: a big picture. Future Gener. Comput. Syst. 49, 58–67 (2015)

[5] Puzio, P., Molva, R., Onen, M., Loureiro, S.: ClouDedup: secure deduplication with encrypted data for cloud storage. In: Proceedings of IEEE 5th International Conference on Cloud Computing Technology and Science, pp. 363–370. IEEE (2013)

[6] Mulazzani, M., Schrittwieser, S., Leithner, M., Huber, M.: Dark clouds on the horizon: using cloud storage as attack vector and online slack space. In: Proceedings

[7] of USENIX Security Symposium, p. 5 (2011)X. Boyen and B. Waters, "Anonymous hierarchical identity-based encryption (without random oracles(lecture notes in computer science)," Advances in Cryptology, vol. 4117, pp. 290–307, Aug 2006.

[8] X. Liu, X. Xie, K. Li, B. Xiao, J. Wu, H. Qi, and D. Lu, "Fast tracking the population of key tags in large-scale anonymous rfid systems," IEEE/ACM Transactions on Networking, vol. 25, no. 1, pp. 278–291, 2017.

[9] K. R. M. Li, S. Yu and W. Lou, "Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings," Security and Privacy in Communication Networks -, Iternational ICST Conference, SECURECOMM, pp. 89–106, 2010.

[10] E. H. J. Benal, M. Chase and K. Lauter, "Patient controlled encryption: privacy of electronic medical records," ACM Cloud Computing Security Workshop, pp. 103–114, 2009.