International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

SECURE CLOUD DATA DE DUPLICATION WITH EFFICIENT RE-ENCRYPTION

Putta Srivani[1],G. Bhoomika[2],J. Srivigna[3],M. Parimala Sai[4]

[1]Associate professor,Department of IT ,Malla Reddy Engineering College For Women(Autonomous Institution), Maisammaguda, Dhulapally,Secunderabad,Telangana-500100

[234]UG Scholar, Department of CS,Malla Reddy Engineering College for Women, (Autonomous Institution), Maisammaguda,Dhulapally,Secunderabad,Telangana-500100
Email Id: Pulla.srivani@gmail.com

ABSTRACT

Missing data is often an issue in the credit scoring industry, especially in cases where data is MNAR, which means the missing information relates to variables that are not observable. To solve the problem of MNAR data for credit scoring tasks, the proposed method here is an innovative approach known as RMT-NET (Rejection-based Multi-task Network). The three major tasks which RMT-NET simultaneously learns utilizing a multi-task learning architecture are credit score prediction, imputation of missing data, and rejection of untrustworthy data points. The model successfully identifies and removes biased or uncertain data by combining a rejection mechanism, thereby ensuring high accuracy and fairness of credit score forecasts. Employing one common feature representation for each of the tasks through separate branches for every task,. According to experimental results, RMT-NET performs better than traditional techniques in handling missing data and generates credit ratings that are more robust and reliable, especially for datasets containing a significant proportion of MNAR data. The proposed approach is a potential solution for real-world credit scoring applications with imperfect or faulty data because it not only improves the accuracy of predictions but also reduces bias.

Keywords-Cloud Computing,Data Deduplication,Secure Data Storage,Re-encryption,Data Privacy,Convergent Encryption,Proof of Ownership (PoW),Proxy Re-encryption.

I. INTRODUCTION

Cloud capacity administrations have turned into the core activities that handle huge volumes of data; they are flexible and offer cost-effectiveness. From the ways of optimizing capacities, information deduplication has emerged to become a fundamental solution, hence deleting duplications on one's stored information. However, it poses significant security issues though effective. Conventional strategies often presuppose access to plaintext data or rely on weak encryption, which leaves sensitive data vulnerable to unauthorized access and malicious attacks. Even more so, in case of key repudiation or access upgrades, reencrypting the entire dataset can incur significant computational overhead, rendering those strategies impractical for the large-scale frameworks. Against these challenges, this extend proposed a secure and efficient data deduplication framework.Levers on the one-way property of hash capacities, the framework protects the information and prevents stub-reserved attacks. The area determination strategy based on the Sprout filter makes efficient identification of information while the Focalized All-Or-Nothing Change (CAONT) tool allows partial re-encryption of some fragments or perhaps the whole dataset than encrypting the entire thing. This approach does not only improve security but also fundamentally reduces the computational costs during the process of re-encryption. This proposed system gives complete security proofs and test validations making it a commonsensical and versatile arrangement for ensuring secure cloud capacity especially delicate spaces like healthcare, back, and venture information administration.

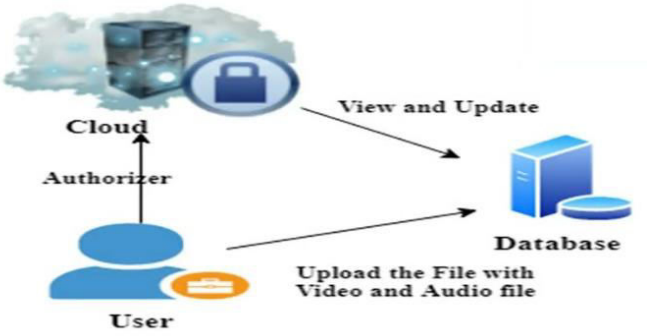

Fig 1: System Architecture

II. RELATED WORK

"Irrefutable Computation over Huge Databases with Incremental Updates"

Authors: X. Chen, J. Li, J. Weng, J. Ma, and W. Lou(2016)

This paper focuses on irreducible computation over large databases, dealing with the issues that come with judging computations while handling energetic databases that go through incremental updates. The authors suggest ways of affirming computations without necessarily asking the verifier to perform them, hence making it efficient for cloud-based computations. The method uses cryptographic techniques to ensure the correctness of operations on the large datasets, which is fundamental to secure outsourcing of database management.

"Pics-on-wheels: Photo reconnaissance in the vehicular cloud"

Authors: M. Gerla, J. Weng, and G. Pau(2013)

This paper introduces the concept of vehicular cloud computing, where vehicles collaborate to give administrations such as photo observation. The creators propose an design called Pics-on-Wheels, which empowers vehicles to capture and store pictures, uploading them to a cloud stage for advance investigation. This framework encourages urban checking and upgrades the security of transportation frameworks, focusing on the part of vehicle-based information in the developing Web of Things (IoT) landscape.

"Advances in modern calculations of secure outsourcing of secluded exponentiations"

Authors: X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou(2014)

In this work, authors present calculations for safely outsourcing secluded exponentiation errands, which are commonly utilized in public-key cryptosystems. The consider focuses on the issue of appropriately outsourcing cryptographic computations to the cloud whereas keeping security against malicious cloud suppliers. It is proposed that such computations be designed to allow secure verification of outsourced operations without revelation of private keys, and the significance of such calculations for cloud-based encryption services is obvious.

"DedupDUM: Secure and adaptive information deduplication with energetic client management"

Authors: H. Yuan, X. Chen, T. Jiang, X. Zhang, Z. Yan, and Y. Xiang(2018)

This paper introduces DedupDUM, a system that focuses on secure and adaptive information deduplication with energetic client administration. Deduplication is a technique for eliminating duplicate copies of data, which is particularly crucial in cloud storage. The designers handle the problem of ensuring information security and security in the deduplication prepare, presenting a plan that enables to perform effective deduplication but maintaining the ability to control active client access, even as clients are added or deleted.

"Bitcoin-based fair remunerations for outsourcing computations of fog devices"

Authors: H. Huang, X. Chen, Q. Wu, X. Huang, and J. Shen(2018)

This paper examines the use of Bitcoin in outsourcing computational tasks in a way that ensures fair payments in the context of mist computing devices. The authors describe a system

through which computation services can be paid using Bitcoin, thus ensuring fair pay for both providers and customers. The system integrates approaches for securing fairness in payments and preventing frauds to overcome the problems in the dispersed computing environment where trust cannot be established.

"The advanced universe of openings: Wealthy information and the expanding esteem of the Web of Things"

Authors: EMC Corporation(2014)

This report by EMC Enterprise talks about the developing esteem of the Web of Things (IoT) and the gigantic sums of information it creates, frequently alluded to as the advanced universe. The paper emphasizes the windows of opportunities and challenges portrayed by the growing volume of information, including how such information can be loaded with commerce and mechanical advancement. In addition, it underscores the role of cloud computing in managing and processing the information generated by IoT devices.

"Single instance store in Windows 2000"

Authors: W. J. Bolosky, S. Corbin, D. Goebel, and J. R. Douceur(2000)

This paper introduces the concept of single occurrence capacity (SIS) in the Windows 2000 working framework. Optimizing capacity by ensuring, as it were, just one duplicate of indistinguishable information is put away, even though it might occur in different places. Paper details the usage of SIS in the record framework of Windows 2000 and its benefits, which would include reduced capacity requirements and sophisticated efficiency in situations with copy records.

III. IMPLEMENTATION

The implementation of the secure cloud data deduplication framework with successful re-encryption encompasses the merging of key cryptographic and data organization techniques. Data deduplication is satisfied using blended encryption, where data is mixed with a key determined from its hash regard, ensuring undefined data pieces produce the same ciphertext for profitable deduplication while keeping up mystery. A Bloom filter-based zone choice methodology is utilized to rapidly and precisely plan hash qualities to capacity territories, updating the deduplication prepare. For re-encryption, the framework makes utilization of the Focalized All-Or-Nothing Alter (CAONT) to incorporate particular re-encryption of particular information bundles or rather than the aggregate dataset, in its entirely diminishing computational overhead. Re-encryption assignments are named to a mediator server that updates ciphertext underneath a unused key without revealing plaintext data. The system solidifies overwhelming key organization through a Trusted Master (TA) and executes fine-grained get to control utilizing attribute-based encryption (ABE). Security is development updated through the utilize of one-way hash capacities, which ensure resistance to stub-reserved attacks and unauthorized data get to. Reenactment tests in a cloud environment outline the framework's ampleness, showing up basic diminishments in re-encryption time and computational costs while fulfilling tall capacity efficiency and strong data security.

IV. ALGORITHM

The process of secure cloud data deduplication and able re-encryption starts with the hashing of the data piece through one-way hash work that results in a unique identifier. A Grow channel checks for duplicates in the cloud capacity. In case the hash exists, then the piece is checked as a duplicate, and metadata is overhauled without uploading the square once more. Something else, the data fragment is blended utilizing a focalized encryption strategy with the owner's key and put away in the cloud. For re-encryption, enacted by key denial or procedure upgrades, the Trusted Pro (TA) delivers a mediator re-encryption key. Utilizing the Consolidated All-Or-Nothing Change (CAONT), as it were the significant parcels of the data piece are re-encrypted, reducing computational overhead. Get to control is accomplished using attribute-based encryption

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

(ABE), such that only authorized clients can decrypt and access data. Proof of Ownership (PoW) is used for authenticating legitimate blue get access requests. The algorithm is designed to withstand stub-reserved attacks, preserve data privacy, and also reduce computational overheads so that it is both secure and capable for cloud storage scenarios.

RESULT

Fig 1: User Login

Fig 2: User Details

Fig 3: External Attacker

Fig 4:Accuracy Bar Chart

CONCLUSION

In conclusion, the ask around and progressions shown in the overviewed composing highlight essential movements in the regions of secure cloud computing, data deduplication, and the compelling outsourcing of computational errands. The works emphasize the noteworthiness of ensuring data security, security, and efficiency, particularly in circumstances where large-scale data planning and capacity are required, such as in cloud computing and the Web of Things (IoT). Approaches such as focalized encryption,secure outsourcing calculations, and enthusiastic client organization for deduplication donate reasonable courses of action to ensure fragile data while optimizing capacity efficiency. Besides, the utilize of creative techniques like Grow channels, Bitcoin-based installments, and single event capacity has opened unused streets for updating system execution and faithful quality. These courses of action play a pressing portion in tending to the challenges related with computational offloading, key organization, and keeping up respectability in spread systems. By and huge, the integration of these progresses in cloud circumstances and passed on systems talks to a step forward in making cloud capacity and computation more secure, flexible, and down to soil for both clients and advantage providers.

REFERENCES

[1] X. Chen, J. Li, J. Weng, J. Ma and W. Lou, "Verifiable computation over large database with

incremental updates", IEEE Trans. Comput., vol. 65, no. 10, pp. 3184-3195, Oct. 2016.
M. Gerla, J. Weng and G. Pau, "Pics-on-wheels: Photo surveillance in the vehicular cloud", Proc. Int. Conf. Comput. Netw. Commun., pp. 1123-1127, 2013.1.
[3]X. Chen, J. Li, J. Ma, Q. Tang and W. Lou, "New algorithms for secure outsourcing of modular exponentiations", IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 9, pp. 2386-2396, Sep. 2014.
[4]H. Yuan, X. Chen, T. Jiang, X. Zhang, Z. Yan and Y. Xiang, "DedupDUM: Secure and scalable data deduplication with dynamic user management", Inf. Sci., vol. 456, pp. 159-173, 2018.
[5] H. Huang, X. Chen, Q. Wu, X. Huang and J. Shen, "Bitcoin-based fair payments for outsourcing computations of fog devices", Future Gener. Comput. Syst., vol. 78, pp. 850-858, 2018.
[6]"The digital universe of opportunities: Rich data and the increasing value of the Internet of Things", 2014,[online]Available:https://www.emc.com/leadership/digital-universe/2014iview/index.htm..
[7]W. J. Bolosky, S. Corbin, D. Goebel and J. R. Douceur, "Single instance storage in windows 2000", Proc. Conf. Usenix Windows Syst. Symp., pp. 2-2, 2000.
[8]"Netapp deduplication helps duke institute for genome sciences and policy reduce storage requirements for genomic information by 83 percent", 2008, [online] Available: http://www.netapp.com.
[9]M. Dutch, "Understanding data deduplication ratios", SNIA Data Manage. Forum, pp. 1-13, June 2008.
[10] T. Jiang, X. Chen, J. Li, D. S. Wong, J. Ma and J. K. Liu, "TIMER: Secure and reliable cloud storage against data re-outsourcing", Proc. 10th Int. Conf. Inf. Security Practice Experience, pp. 346-358, 2014.
[11] X. Chen, B. Lee and K. Kim, "Receipt-free electronic auction schemes using homomorphic encryption", Proc. 6th Int. Conf. Inf. Security Cryptology, pp. 259-273, 2003.
[12]X. Zhang, X. Chen, J. Wang, Z. Zhan and J. Li, "Verifiable privacy-preserving single-layer perceptron training scheme in cloud computing", Soft Comput., vol. 22, no. 23, pp. 7719-7732, 2018.
[13]J. R. Douceur, A. Adya, W. J. Bolosky, P. Simon and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system", Proc. 22nd Int. Conf. Distrib. Comput. Syst., pp. 617-624, 2002.
[14]L. D. Stein, "The case for cloud computing in genome informatics", Genome Biol., vol. 11, no. 5, pp. 207-207, 2010.
[15]J. Hur, D. Koo, Y. Shin and K. Kang, "Secure data deduplication with dynamic ownership management in cloud storage", IEEE Trans. Knowl. Data Eng., vol. 28, no. 11, pp. 3113-3125, Nov. 2016.
[16]J. Li et al., "Secure distributed deduplication systems with improved reliability", IEEE Trans. Comput., vol. 64, no. 12, pp. 3569-3579, Dec. 2015.
[17]T. Jiang, X. Chen, Q. Wu, J. Ma, W. Susilo and W. Lou, "Secure and efficient cloud data deduplication with randomized tag", IEEE Trans. Inf. Forensics Security, vol. 12, no. 3, pp. 532-543, Mar. 2017.
[18]W. C. G. III, A. Shull, S. Myers and A. J. Lee, "On the practicality of cryptographically enforcing dynamic access control policies in the cloud", Proc. IEEE Symp. Security Privacy, pp. 819-838, 2016.
[19]J. Li, C. Qin, P. P. C. Lee and J. Li, "Rekeying for encrypted deduplication storage", Proc. 46th Annu. IEEE/IFIP Int. Conf. Depend. Syst. Netw., pp. 618-629, 2016.
[20]C. Qin, J. Li and P. P. C. Lee, "The design and implementation of a rekeying-aware encrypted deduplication storage system", ACM Trans. Storage, vol. 13, no. 1, pp. 9:1-9:30, 2017.
[21] R. L. Rivest, "All-or-nothing encryption and the package transform", Proc. 4th Int. Workshop Fast Softw. Encryption, pp. 210-218, 1997.