Blockchain for Achieving Accountable Outsourcing Computations in Edge Computing

P. Karthi, Dr.S.R. Menaka, and Dr.G. Singaravel

Abstract--- *Since the emergence of cloud computing, data owners are being driven to move their intricate data management systems from on-site locations to private cloud providers in order to take advantage of increased flexibility and cost savings. However, sensitive data must be encrypted before being outsourced in order to safeguard data privacy. This renders outdated the conventional method of using data, which relies on plaintext keyword searches. It is therefore crucial to activate an encrypted cloud data search service. In order to meet the effective data retrieval, need, search services must support multi-keyword queries and offer result similarity rating, given the volume of data users and documents stored in cloud storage. Similar works on searchable encryption seldom distinguish between search results and concentrate on single keyword or Boolean keyword search. I define and solve the difficult problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE) for the first time in this work. I also outline a tight set of privacy requirements that must be met in order for such a safe cloud data utilization system to be implemented.*

Keywords--- *Fog-To-Cloud Computing, Secure Data Storage, Auditing Scheme, Edge Computing, Cloud Services.*

I. INTRODUCTION

Data security has become a top priority in the quickly changing world of Fog-to-Cloud computing, where processing and data storage are split between centralized cloud servers and edge devices (Fog). Strong auditing programs are becoming increasingly necessary to protect sensitive data as businesses use this hybrid architecture to take advantage of cloud and edge computing services. Creating an effective auditing system that not only takes into account the special features of Fog-to-Cloud environments but also maximizes resource consumption is one of the most important issues in this area. This introduction lays the groundwork for investigating a novel auditing scheme and investigating its potential to improve data storage security in Fog-to-Cloud computing while preserving scalability and efficiency. Our objective is to offer a comprehensive solution that conforms to the dynamic nature of contemporary computing paradigms and advances the continuous development of dependable and secure data management within the Fog-to-Cloud ecosystem by seamlessly integrating auditing methods.

1.1. Fog-to-cloud Computing

By fusing the advantages of centralized cloud services with edge computing, fog-to-cloud computing is a dynamic paradigm change in distributed computing. By taking advantage of the edge devices' close proximity to the data source, also referred to as the "Fog," this novel architecture maximizes processing power and minimizes latency. Fog-to-Cloud computing, which strikes a balance between the massive storage and computational resources provided by cloud servers and the real-time processing capabilities of edge devices, emerges as a key answer as data continues to rise exponentially. This combination has the potential to open up new opportunities in a number of industries, including autonomous systems and Internet of Things applications. Fog-to-Cloud computing opens the door for revolutionary developments in data processing, storage, and value-adding, ushering in a new age of efficiency and responsiveness in the digital world.

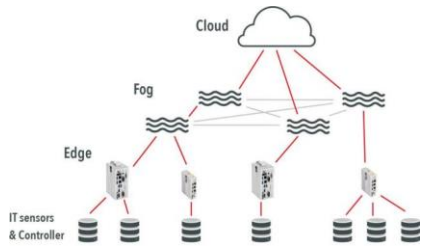


Figure 1: Fog-to-cloud Computing

1.2. Secure Data Storage

Figure 2: Secure Data Storage

Secure data storage is essential to maintaining information integrity and confidentiality in today's digital world. Our increasing dependence on data-driven technology necessitates strengthening the storage infrastructures that protect sensitive and important data. In addition to the physical storage, secure

P. Karthi, Second Year, M. Tech./IT, K.S.R. College of Engineering (Autonomous), Tamil Nadu. E-mail: karthimtech453@gmail.com
Dr.S.R. Menaka, Associate Professor/ IT, K.S.R. College of Engineering (Autonomous), Tamil Nadu. E-mail: menaka@ksrce.ac.in
Dr.G. Singaravel, Professor and Head/IT, K.S.R. College of Engineering (Autonomous), Tamil Nadu. E-mail: singaravelg@gmail.com

DOI: 10.9756/BIJSESC/V14I1/BIJ24002

data storage includes the complex web of protocols, encryption techniques, and access restrictions that protect data from loss, corruption, and unwanted access. Strong and flexible secure storage solutions are essential given the growing amounts of data and the increasing complexity of cyberattacks. This introduction lays the groundwork for a thorough examination of the many facets of safe data storage, covering the tactics, tools, and best practices that support the defense of our digital archives in a time when system resilience is critical.

1.3. Auditing Scheme

Creating and implementing a strong auditing program has become essential in the ever-expanding fields of information technology and data management. An auditing plan offers a methodical and all-encompassing way to track, monitor, and evaluate the different operations that take place within a system, acting as the watchful defender of data integrity. An effective auditing program does more than just make sure that compliance is followed; it also helps to proactively identify vulnerabilities, suspicious activity, and possible breaches, which strengthens a company's security posture. An advanced auditing framework is becoming more and more necessary as businesses struggle with the growing complexity of their digital ecosystems. This introduction lays the groundwork for a discussion of auditing schemes and their critical role in bolstering data security, improving regulatory compliance, and fostering openness in a setting where responsibility and trust are critical.

1.4. Edge Computing

The way we handle and manage data is being revolutionized by Edge Computing, which has emerged as a disruptive paradigm in the rapidly evolving field of information technology. By bringing computational resources closer to the data source than typical cloud computing models do, Edge Computing lowers latency and improves real-time processing capabilities. With the help of distributed computing, devices at the "edge" of the network—such as sensors and Internet of Things (IoT) devices—can process data locally, reducing the need for continuous connection with a centralized cloud server. Edge computing plays a key role in satisfying the demand for quicker response times and more effective resource usage as our digital environment grows more decentralized and data-intensive. This introduction provides an overview of Edge Computing and sets the stage for a deeper look at its uses, advantages, and revolutionary effects on a range of industries, including manufacturing and healthcare.

1.5. Cloud Services

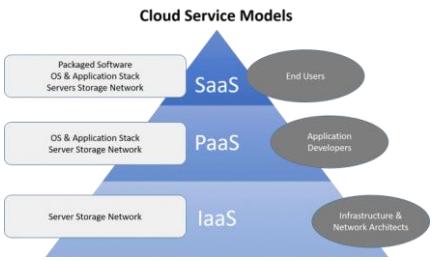


Figure 3: Cloud Services

Cloud services are becoming the cornerstone of flexible and scalable computing infrastructures in the modern era of digital transformation, completely changing the information technology environment. Fundamentally, cloud services involve the internet-based supply of computing resources, such as storage, processing power, and applications. Cloud-based models replace traditional on-premises solutions in this paradigm shift, offering unmatched accessibility, cost-effectiveness, and agility. Cloud services are being used by businesses all over the world to improve productivity, encourage creativity, and expand their IT capabilities on demand. In order to better understand how cloud services, which come in a variety of models—Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS)—catalyse digital advancements across industries, from startups to established enterprises, in the pursuit of a more connected and agile future, this introduction sets the stage for an exploration into the expansive domain of cloud services.

II. Literature Survey

2.1. Charm: A Rapidly Prototyping Framework for Cryptosystems

Aet Joseph. Charm, an adaptable framework for rapidly prototyping cryptographic systems, has been proposed and described in this study. Support for modular cryptographic building blocks, infrastructure for designing interactive protocols, and a large library of reusable code are just a few of Charm's features that explicitly make it possible to build new protocols. Our system likewise incorporates a bunch of explicit instruments that permit different cryptosystems to speak with each other. We utilized Appeal to build north of forty cryptographic calculations, including a few new ones that had never been created. This article describes our modular design, which includes a benchmarking module that can be used to compare the performance of Charm primitives to that of existing C implementations. We show that, in many cases, our methods reduce code size by an order of magnitude without affecting performance in an unacceptable way. Lastly, the research community can use the Charm framework for free, and we have a large and engaged user base to date.

2.2. Homomorphic MACS: MAC-based Network Coding Integrity

Shweta Agrawal and others suggested network coding in this study, which was found to improve the robustness and capacity of the network. Standard MACs and checksums, on the other hand, cannot be used to verify the integrity of data because intermediary nodes alter packets along the way. Pollution attacks, in which a single rogue node floods the network with incorrect packets, prevent the receiver from correctly decoding the packets, are another threat to network coded systems. Signature methods have been developed to stop such attacks, but they are typically too slow for online per-packet integrity. Additionally, they force the selection of network coding coefficients from a vast pool, which makes the network coding header larger. A homomorphic MAC for confirming the accuracy of network-coded data is presented in this paper. In network coding systems, our Homomorphic MAC is intended to substitute for standard MACs like HMAC.

2.3. Data Provable Possession in Untrusted Stores

We present a paradigm for proven data possession (PDP), which enables a client to determine without obtaining the original data that an untrusted server is the owner. The method generates probabilistic proofs of possession and substantially reduces I/O costs by selecting random blocks from the server. The client keeps a set amount of metadata to verify the evidence. By sending a small, consistent amount of data, the challenge/response protocol reduces network traffic. Subsequently, the PDP model for far off information confirmation can deal with enormous informational indexes in broadly scattered capacity frameworks. In comparison to systems with lower guarantees, our two provably safe PDP schemes perform better than previous ones. Specifically, the server cost is low (or maybe steady), rather than straight in information size. Explores different avenues regarding our execution affirm the attainability of PDP and show that its presentation is restricted by plate I/O instead of cryptographic handling. In this protocol, communication and client storage complexity are both important.

2.4. Storage Proofs from Homomorphic Identification Protocols

Giuseppe Ateniese et al. in this study offer Proofs of storage (PoS) are interactive protocols that enable a client to confirm the accuracy with which a server saves a file. Any homomorphic linear authenticator (HLA) can be used to generate storage proofs, as previous research has shown. The latter are signature/message authentication methods that allow "tags" from multiple messages to be homomorphically coupled to create a "tag" on any linear combination of these messages. A method for creating public-key HLAs from any homomorphic identifying protocol is presented by us. We next exhibit how to change over any open key HLA into a freely certain Po's with correspondence intricacy free of record length and an endless number of checks. By applying our transformations to a Showup variation of an identifying protocol, we produce the first unbounded-use PoS based on factoring (in the random oracle model), demonstrating their usefulness. A trend toward outsourcing data management to third-party service providers (also known as "servers") has been fueled by advancements in networking technology and the growing accumulation of information. Instead of spending a lot of money on hardware, software, and staff for "in-house" data maintenance, businesses can concentrate on their core operations.

2.5. Provable Dynamic Multicopy Data Possession in Cloud Computing Systems

Ayad F. et al. in this work suggest that businesses (CSPs) are increasingly outsourcing data to remote cloud service providers. For a monthly fee in gigabytes, customers can hire the CSP's storage infrastructure to store and retrieve nearly unlimited amounts of data. For improved scalability, availability, and durability, some customers may want their data replicated across multiple servers in various data centres. Clients are charged additional expenses when the CSP is expected to hold more duplicates. As a result, customers need solid assurance that the CSP is storing all copies of the data stipulated in the service contract and that these copies are compatible with the most recent customer updates.

III. Related Work

As a paradigm, edge computing offers services that let many end users outsource computations. A trust less environment is created when sampling-based replication calculation and block chain integration are utilized to confirm computation accuracy because edge nodes lack trust. However, due to its decentralized nature, the block chain is hampered by issues with excessive resource usage, making it impossible to directly implement computational overhead verification on the block chain. Thus, we suggest an off-chain block-based accountable verification system. The off-chain block satisfies a few Edge Computing needs, including lower service latency and diverse resource edge nodes. The off-chain block attempts to address the following two issues for trustworthy outsourcing computations: (i) how to accomplish responsible verification and (ii) how to construct the block securely and quickly. To be more specific, the block is built upon a Directed Acyclic Graph, where transactions pertaining to verification reports and computation results are updated in a fully decentralized manner. On the block chain, the block hash is kept track of. Furthermore, trustworthy verification is provided by the combination of off-chain verification and on-chain arbitration. Accountability for edge nodes is achieved using a trust evaluation paradigm. In addition, we performed the security analysis using a few performance characteristics. The scalability of our outsourced computations is demonstrated by simulating lightweight edge nodes using Raspberry Pis. To demonstrate the effectiveness of the suggested scheme's block chain updates, a consortium block chain with groups is also put into practice.

IV. Methodology

The representative privacy assurances in the related literature, such as searchable encryption, state that the server should only learn the search results. Using this general privacy definition, we investigate and develop a set of severe privacy criteria for the MRSE framework. Keyword privacy is very important because most people don't want other people, like the Edge-Fog-Cloud server, to see their searches. The most pressing concern is to conceal the keywords indicated by the appropriate trapdoor, which are what they are looking for. The possibility of the trapdoor generating function being randomised rather than deterministic is eliminated by trapdoor. The series of search results, where each search result is a collection of documents ranked in order, constitute the access pattern within the ranked search. To quickly achieve multi-keyword ranked search, we recommend utilizing "inner keyword similarity" to quantitatively evaluate the effective similarity measure "coordinate matching."

4.1. Edge-fog Cloud Setup Module

Instead of returning undifferentiated results, this module improves the schemes that enable multi-keyword queries and provide result similarity ranking for efficient data retrieval. Privacy-Preserving: to ensure privacy and stop the Edge-Fog cloud server from learning any more information from the index and dataset. Efficiency: The above objectives regarding

privacy and functionality ought to be met with minimal communication and computation overhead. A software or hardware module known as an edge-fog cloud setup module makes it possible to deploy and manage cloud, fog, and edge computing resources in a unified and integrated manner. The seamless processing and analysis of data is made possible by this module, which acts as a link between the cloud infrastructure, fog nodes, and edge devices. Data is processed at or near the edge of the network, closer to the source of data generation, in edge computing. By providing a more distributed architecture with intermediate nodes between the edge devices and the cloud, fog computing extends edge computing.

4.2. Chiper Text Coordinates Matching

An intermediate similarity measure known as "coordinate matching" measures the document's relevance to the query by counting the number of query keywords that appear in the document. Boolean queries perform well when the user specifies the precise subset of the dataset that needs to be recovered. Users are able to retrieve the most relevant documents in a ranked order and identify a list of keywords that indicate their concern with greater flexibility. By matching the cipher text's spatial or positional relationships to the plaintext, cipher text coordinates matching can be used to decrypt encrypted messages. A cipher text is an encrypted data or message that can't be read without the right decryption key or algorithm in cryptography. Sensitive data is shielded from unauthorized third-party access and interception by encryption. However, in order for the encrypted data to be useful to the intended recipient, it must be decrypted. The cipher text is transformed into the original plaintext during decryption, which can then be read and used. A technique known as cipher text coordinates matching makes use of the cipher text's spatial relationships to the plaintext that corresponds to it in a matrix or grid. The cipher text is derived from the plaintext using a mathematical algorithm, and the relationships between the two are preserved in the cipher text. This is the foundation of the strategy. Matching the letters or symbols in the cipher text to their positions in the matrix or grid is the method used to decrypt the text. The decrypted plaintext can be uncovered by making use of the spatial relationships that exist between the cipher text and the matrix or grid.

4.3. Data Privacy and Analysis

Before outsourcing, the data owner can use traditional symmetric key cryptography to encrypt the data and effectively prevent the cloud server from accessing the outsourced data. Index privacy, in the event that the cloud server concludes that encrypted documents and keywords are related to the index. As a result, the MRSE cloud server should be prevented from engaging in such an association attack by creating a searchable index. Analysis and data privacy are two essential aspects of data management that are crucial to preserving the data's integrity and security. The protection of private or sensitive data from unauthorized use, disclosure, or access is known as data privacy. On the other hand, data analysis is the process of taking valuable insights and information from data sets. Information security is fundamental to safeguard delicate data from falling into some unacceptable hands. Names, addresses, social security numbers, and financial data are examples of personally identifiable information. Since data breaches and cyber-attacks are becoming more common, data privacy is a major concern for individuals, governments, and businesses alike. On the other hand, data analysis entails gathering, analyzing, and interpreting data to discover patterns and insights. Marketing, healthcare, finance, and scientific research are just a few of the many areas where data analysis can be applied.

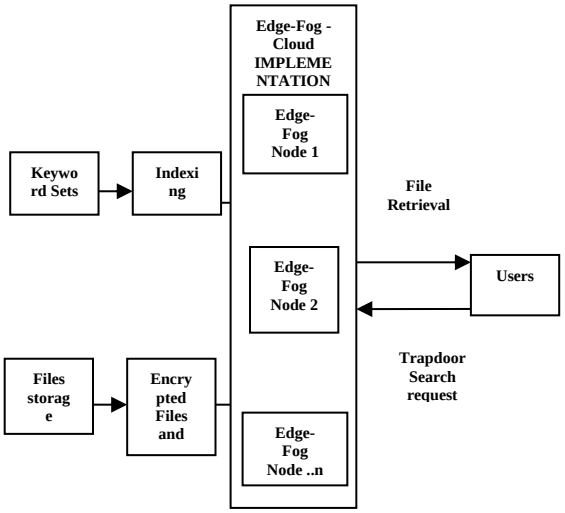
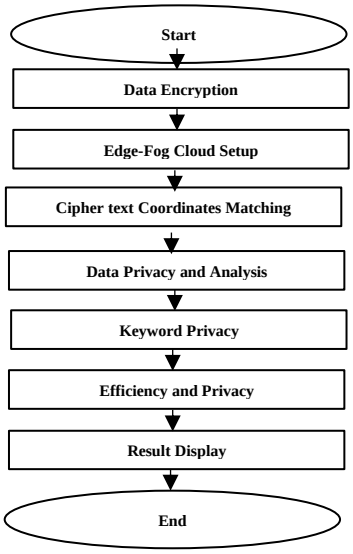


Figure 4: Block Diagram

4.4. Keyword Privacy

Users usually want their searches to be visible to others, like the cloud server, but the most important thing is to hide what they're looking for the keywords that the appropriate trapdoor specifies. To safeguard the query keywords, a cryptographic trapdoor can be created. The protection of search queries or keywords used by individuals when searching the internet is an essential component of data privacy. Keyword privacy refers to this aspect of the MRSE



4.5. Algorithm Definition

Data may be encrypted using a method called Multi-

Keyword Ranked Search Encryption (MRSE) and made searchable. Users may encrypt their data and search on it without having to first decode it thanks to this feature. Multiple keyword searching

4.6. Algorithm Implementation

1. Generate Secret and Public Keys the first step is to generate the secret and public keys using a secure random number generator. The secret key is used for encrypting the data, and the public key is used for verifying the encrypted data.
2. Encrypt Data Encrypt the data using the secret key. The encryption process should ensure that the data remains secure and confidential.
3. Generate Searchable Encrypted Index Generate a searchable encrypted index for the encrypted data. The index should be searchable using multiple keywords, and it should also maintain the ranking of the encrypted data based on the relevance to the search keywords.

V. RESULT ANALYSIS

Table 1: Comparison Table

Algorithms	Accuracy	Precision	Recall	F-Measure
MAC and HMAC	75%	78.20%	80.01%	79%
MRSE	81%	80%	91%	85%

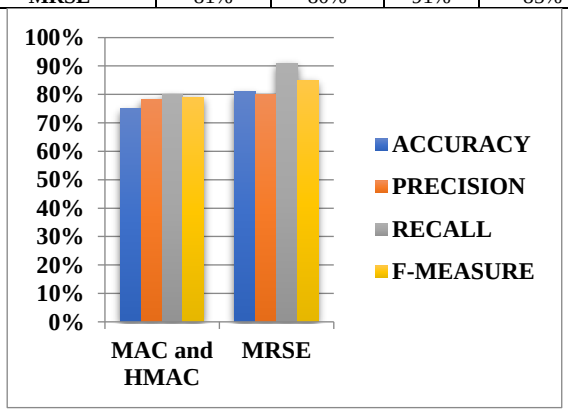


Figure 5: Comparison Graph

We examine and develop a set of stringent Privacy standards specifically for the MRSE architecture in this project using this broad definition of privacy. Due to the colossal number of information clients and archives in the Edge-Haze Cloud, various watchwords in the hunt demand should be permitted, and records should be returned in the order of their relevance to these keywords. In order to meet privacy requirements, the edge fog cloud setup module provides a means of preventing the Edge-Fog cloud server from learning additional information from the dataset and index.

VI. CONCLUSION

In conclusion, in terms of privacy protection, efficiency, scalability, and computational overhead, the suggested privacy criteria and solution for MRSE offer a considerable improvement over current MRSE frameworks. Through the resolution of the primary issues raised by MRSE users, the suggested solution enhances MRSE's feasibility and appeal for a range of application scenarios. The suggested method can be applied to a wide range of systems, including distributed file systems, local search engines on devices, and cloud-based search services.

VII. FUTURE WORK

In upcoming work, creating more advanced functions for creating trapdoors. A basic random trapdoor generation function is used in the current implementation. To better secure keyword privacy, more advanced trapdoor creation functions could be created. creating similarity metrics that are more effective. A basic similarity metric called coordinate matching is used in the current implementation. To enhance the system's performance, more effective similarity metrics, including inner keyword similarity, could be created. On developing real-time auditing schemes that can detect potential security threats and anomalies in real-time.

REFERENCES

[1] J.A. Akinyele, C. Garman, I. Miers, M.W. Pagano, M. Rushanan, M. Green and A.D. Rubin, "Charm: a framework for rapidly prototyping cryptosystems", *Journal of Cryptographic Engineering*, Vol. 3, Pp. 111-128, 2013.

[2] S. Agrawal and D. Boneh, "Homomorphic MACs: MAC-based integrity for network coding", *In Applied Cryptography and Network Security: 7th International Conference, ACNS, Paris-Rocquencourt. Proceedings*, Vol. 7, Pp. 292-305, 2009. Springer Berlin Heidelberg.

[3] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson and D. Song, "Provable data possession at untrusted stores", *In Proceedings of the 14th ACM conference on Computer and communications security*, Pp. 598-609, 2007.

[4] G. Ateniese, S. Kamara and J. Katz, "Proofs of storage from homomorphic identification protocols", *In Advances in Cryptology–ASIACRYPT: 15th International Conference on the Theory and Application of Cryptology and Information Security. Proceedings*, Vol. 15, Pp. 319-333. Springer Berlin Heidelberg.

[5] A.F. Barsoum and M.A. Hasan, "Provable multicopy dynamic data possession in cloud computing systems", *IEEE Transactions on Information Forensics and Security*, Vol. 10, No. 3, 485-497, 2014.

[6] F. Bonomi, R. Milito, J. Zhu and S. Addepalli, "Fog computing and its role in the internet of things", *In Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, Pp. 13-16, 2012.

[7] K.D. Bowers, A. Juels and A. Oprea, "Proofs of retrievability: Theory and implementation", *In Proceedings of the ACM workshop on Cloud computing security*, Pp. 43-54, 2009.

[8] J. Chang, B. Shao, Y. Ji, M. Xu and R. Xue, "Secure network coding from secure proof of retrievability", *Sci Chin Inf Sci*, Vol. 64, No. 12, Pp. 1-2, 2021.

[9] J. Chang, H. Wang, F. Wang, A. Zhang and Y. Ji, "RKA security for identity-based signature scheme", *IEEE Access*, Vol. 8, Pp.17833-17841, 2020.

[10] Y. Dodis, S. Vadhan and D. Wichs, "Proofs of retrievability via hardness amplification", *In Theory of Cryptography: 6th Theory of Cryptography Conference, TCC, Proceedings*, Vol. 6, pp. 109-127, 2009. Springer Berlin Heidelberg.