International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 May 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Improved Cauchy Reed-Solomon Codes for Cloud Data Retrieval and Secured Data Storage using Role-Based Cryptographic Access and forensic investigation

Mahesh B Gunjal*[1]
Department of Computer Science and Engineering,
Dr A. P. J Abdul Kalam University Indore(MP),452010
Email**:-** **maheshgunjal2010@gmail.com**

Dr. Vijay R Sonawane [2]
Department of Computer Science and Engineering,
Dr A. P. J Abdul Kalam University Indore(MP),452010
Email: vijaysonawane11@gmail.com

Abstract—Doling out client consent strategies to PC frameworks presents a huge test in guaranteeing legitimate approval, especially with the development of open frameworks and scattered stages like the cloud. RBAC has turned into a broadly involved strategy in cloud server applications because of its versatility. Granting access to cloud-stored data for investigating potential wrongdoings is crucial in computer forensic investigations. In cases where the cloud service provider's reliability is questionable, maintaining data confidentiality and establishing an efficient procedure for revoking access upon credential expiration is essential. As storage systems expand across vast networks, frequent component failures require stronger fault tolerance measures. Our work secure data-sharing system combines role (Authorized) based access control and AES encryption technology to provide safe key distribution and data sharing for dynamic groups. Data recovery entails protecting data dispersed over distributed systems by storing duplicate data and applying the erasure code technique. Erasure coding strategies, like Reed-Solomon codes, guarantee disc failure robustness while cutting down on data storage expenses dramatically. They do, however, also result in longer access times and more expensive repairs. Consequently, there has been a great deal of interest in academic and business circles for the investigation of novel coding strategies for cloud storage systems. The objective of this study is to present a novel coding method that utilizes the intricate Cauchy matrix in order to improve Reed-Solomon coding efficiency and strengthen fault tolerance.

Keywords-Role-based control of access, AES, multi-authority access control, cloud data security, Cauchy matrix, Reed-Solomon codes, forensic investigation, proxy key, data recovery, and revocable storage

I. INTRODUCTION

The main source of "big data" development in the modern day is social networking sites, and it produces huge amounts of data. It is huge in both size and complicated in structure. The use of mobile devices and the web is growing daily as a result of the current global issue. As a result, many companies produce large amounts of data at the Petabyte or Exabyte level. According to Wang et al. [1], traditional systems have a limited capacity to collect, store, and analyse such massive amounts of data.

Implementing access restrictions stands as a fundamental method to guarantee the privacy of data kept on cloud servers environments. Over time, multiple approaches to access control have been outlined in scholarly works. Authorization based on role (RBAC), a highly favored design in this field, aids in simplifying security management, particularly in expansive systems. RBAC uses roles to link users with resource privileges, bypassing direct assignment of permissions to individuals. Access to associated permissions and resources is granted solely to users assigned membership within a specific role. RBAC has seen widespread adoption across various systems since its formalization in the 1990s [1], with the RBAC standard proposed in 2000 [3], and the model revisited and expanded in 1996 [2]. Preserving data integrity serves as a primary objective in cloud storage. Our suggested effort attempts to protect data and make recovery easier when things go wrong. wherein a proxy server assumes responsibility for executing these operations. Users' information would be kept in cloud storage's public and private sections, with users limited to the entire cloud, thus upholding heightened security in the private cloud. In the event of unauthorized alterations, The original data kept in the private cloud would be recovered via an

intermediary(proxy). To maintain acceptable fault tolerance while achieving optimal speed, various redundancy configuration options are often available to cloud storage users.

Ensuring data accessibility in distributed storage systems is crucial, especially in scenarios involving frequent node failures. This study employs Authorization based on role a cloud forensic investigation method based on snapshots, and proposes an encryption technique called AES 128 for secure information storage and exchange. Additionally, through this effort, a server for backup mechanism has been built that allows all users to do ad hoc retrieval by using the machine as a proxy storage server. Before initiating work on the suggested system, various previously established methodologies with inherent knowledge gaps were examined, addressing concerns such as time consumption, space utilization, and the complexity associated with matrix construction commonly observed in current computer systems.

Advancements in processing and storage capabilities have opened up opportunities for managing, compiling, and analyzing extensive datasets. Big data analytics introduces innovative methods for scrutinizing large datasets across various sectors like healthcare, finance, law enforcement, instruction, and retail [4]. However, dealing with the surging volume of data poses numerous challenges, including safeguarding privacy, preserving data integrity, and implementing robust access controls. These challenges are crucial in defending data against threats like assaults that degrade data and man-in-the-middle attacks. The present research aims to develop models facilitating data restoration for dispersed datasets, utilizing a Cauchy matrix generation technique.

Yang et al. [5] propose a privacy-centric and cost-effective multi-cloud computing storage solution. Additionally, Chervyakov et al. [6] provide a distributed data storage framework for handling encrypting information and controlling calculations. Error detection and correction are addressed through the Double Atomic Numbering Scheme. Ongoing study focuses on a block chain security framework customized for dispersed cloud server storage. Li et al. [7] employ a formal algorithm for managing the backing up of file blocks amongst several clients and information centers. Because of the exponential expansion in data, Bhubaneshwar & Tharini [8] emphasize the necessity for flexible and dependable networked storage systems. In addition, studies on Low-Density Parity Check (LDPC) algorithms try to address the difficulties associated with large-scale data storage. Tang & Zhang [9] demonstrate simplifying encoder and decoder designs by concatenating a Vandermonde matrix with an identity matrix.

Extensive exploration of finite field theories has aimed to reduce singular matrices in the Cauchy matrix. Cauchy RS codes are used for their resilience against erasures. Make venko et al. in [10] illustrate heuristic modifications to enhance bit-matrices used for coding, particularly in reliable data storage. One essential requirement for cloud storage systems is the capacity to serve several clients at once. A new feature called the assistance rate is the most recent update. A Quality of Service (QoS) metric designed for programmable, networked contexts is presented by Kazemi et al. [11], which can handle multiple concurrent information access requests with ease. Chen & Ma [12] introduce a cost-effective and efficient disc recovery technique that minimizes data reads from the disc, evenly distributing the burden across surviving discs and requiring only minimal disc reads for failed disc restoration, although complete data recovery takes substantial time. Shen et al. [13] demonstrate a method for computing the ideal quantity of code word fragments for each XOR-based code, optimizing reconstruction matrices with minimal information required, notably improving the speed of input/output in cloud content systems with big chunk sizes, but requiring a substantial overhead in terms of material read in recoveries.

The Advanced Cauchy Reed Solomon (ACRS) approach complements the conventional "Reed Solomon" method in this study. Leveraging the XOR technique to bypass Erasure restoration is facilitated and the conversion process is streamlined by Galois Field Multiplication. This "ACRS" code uses a novel technique to manage "m" disc failures by converting "k" blocks of data into "m" coding blocks. Talks on cloud computing frequently center on spreading different assets, including processors, hardware, and software. The services are readily expandable and reasonably priced [14]. This appeal of cloud computing attracts both business owners and cybercriminals. Therefore, "computer forensic investigation" becomes critical in gathering evidence against these criminals. However, integrating modern technology and cloud computing practices presents challenges in the investigative process. The primary obstacles include navigating conflicting regulations applied to different data types stored across various locations, restricted access to evidence stored in the cloud, and the complexity of removing real proof to prove a point or demonstrate authenticity. As cloud computing continues to expand, a new form of data exchange emerges, promising benefits to all individuals or entities involved. Nevertheless, there's a risk that shared data might not securely access the resources within cloud computing. To address this concern, cryptography is employed to safeguard users' sensitive information and ensure complete data confidentiality. Furthermore, authorization to access protected data must be dynamic and revocable, especially when the authorization period expires [15].

Cloud environment always employ various security, deduplication, and data storage techniques. We developed a method cloud that will have the issues as a result of this gap analysis.

- Problem with unauthorized user access
- No balancing techniques have been used in earlier academic research on cloud storage.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 May 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

- In environments with many clouds, there are issues with data leakage and forensic investigation after unauthorized access.
- Revocable storage identity of users.
- An issue with distributed databases using proxy servers and cloud computing's irrational resource consumption and storage overhead.
- And main problem is managing duplicate data on the cloud to achieve fault tolerance data duplications.

In considering the problems mentioned above, a cloud storage system needs to find a reliable security strategy. Additionally, this method is wanted to be able to identify the authorized user and achieve fault tolerance and prevent data duplication.

II. RELATED WORK

This section reviews important studies on time, space, and complexity challenges in matrix generation for fault-tolerant systems in the data storage, and RBAC systems in the cloud. Additionally, forensic investigation tools and methods for grant revocation.

Although a number of trusted models have been developed and put forth in recent years for use in the context of the cloud, it is still difficult to present a model that is effective and can completely satisfy the desired security requirements. Tan et al. [16] suggested a dynamic RBAC model in 2011. environment for cloud computing. However, a comprehensive analysis of their work shows that it falls short of fully achieving the requisite security criteria. Barsoum and Hasan [17] suggested a cloud storage method with four crucial features in 2012. One of these The advantage lies in enabling trust between an owner and a cloud service provider. Li and Du in 2013 introduced the "Cloud-Trust," an adaptable trust model, facilitating users in selecting more dependable service providers. Lin et al. in 2014 introduced "MTBAC," a reciprocal dependability-based access control strategy that takes user behavior trust and cloud service node reliability into account. Recent approaches by Uikey and Bhilare [21] and Zhou et al. [20] aim to enhance RBAC in the cloud, focusing on trust. Zhou et al. [20] associated trust with RBAC, proposing "Owner-Role RBAC" and "Role-User RBAC" probabilistic trust models to augment cloud data storage security, allowing owners to determine role trustworthiness and roles to assess user trust. However, these models face security vulnerabilities like collusion or on/off attacks. Introducing "TrustRBAC," [21] proposes a novel trust-based RBAC paradigm, enhancing cloud-based systems' effectiveness and reliability. They advocate for an inclusive RBAC architecture built on trust and reputation, meeting stringent security standards, operating efficiently, and suitable for open environments such as the cloud [23].

To safeguard sensitive data from unauthorized manipulation, misuse, or destruction, corporations require a method to control access among their personnel. Authorization based on role (RBAC) offers a means to restrict data access based on an individual's role within an organization. Through RBAC, individuals only gain access to the necessary resources and information pertinent to their responsibilities. By aligning access privileges with designated roles, the risk of mishandling private information diminishes. RBAC proves especially beneficial in large organizations and businesses engaging independent contractors. As the number of approved suppliers fluctuates and the workforce expands, configuring unique credentials for each individual could become cumbersome. With a role-based access control system in place, administrators can categorize employees or contractors into predefined roles, granting them access to specific resource sets. Membership in these groups can be revoked upon completion of current assignments, thus restricting access appropriately. Moreover, administrators can adjust authorization levels for these groups, enhancing overall workforce management, productivity, and compliance. RBAC allows administrators to categorize users based on their multiple roles, enabling a single user to belong to several distinct groups simultaneously. Typically, employee access aligns with their active status, responsibilities, and pertinent security policies. The principle of granting minimal authorization necessary for task performance ensures data safety. Olanrewaju et al. proposed a combined approach using the AES and the Blowfish algorithm, which offers robust encryption for cloud data storage, ensuring reliability and improved performance compared to the AES algorithm. Their proposed code incorporates elements from both the AES and Blowfish algorithms. A different study by Hardwaj et al. covered the ideas behind symmetrical and asymmetrical techniques, block and stream cyphers, and their functions in security. It also focused on the RSA algorithm and the Diffie-Hellman key exchange mechanism. They assessed a number of symmetric and asymmetric techniques, emphasizing MD5's quicker encoding speed among them.

In order to provide fault tolerance, several storage systems use (RS) codes, which allow them to handle more faults than RAID [26], [27]. For many years, Reed-Solomon coding—which has been shown to be sound theoretically—has been an essential part of data storage. In data storage systems, it is frequently employed as an erasure code to provide robustness against multiple disc (m disc) failures. Data blocks can be converted into m coding blocks with k data segments and a positive number m by using RS codes. In a RS code, a binary word consists of w bits, where $2^w \geq k + m$. Particularly, these codes work with binary data words. Cauchy RS(CRS) codes significantly enhance the effectiveness of RS codes by employing advanced projection techniques to Galois Field multiplications can be transformed into XOR operations [28]. For storage devices, CRS codes are currently the most effective erasure codes [29]. Furthermore, CRS coding functions on entire stripes across different storage devices rather than on individual words. These stripes are split perfectly into w packets, each appreciably larger than the other. A matrix-vector

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 May 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

product called AXE = B, where A, X, and B are all made up of bits, is the fundamental building block of CRS codes. A is referred to as a Cauchy matrix, X is a set of data stripes, and B is a group of stripes with coding information. The number of XORs required for encoding or decoding operations directly affects the performance of CRS coding because it only uses XOR operations. While elements such as device latency and cache behavior also have an impact on performance, cutting down on XOR operations is a solid and practical way to improve a code's efficiency [30].

Identity-based encryption (IBE) was first proposed by Shamir [31] and quickly adopted by Boneh & Franklin [32], doing away with the need for a public key system (PKI). A process must be in place to remove clients from the framework when required, such as when their authority expires or their secret key is made public, regardless of whether the system is IBE or PKI. In the context of classical PKI, revocation has been thoroughly explored [33], [34], and [35], with commonly recognized solutions like certificates revocation lists and adding validity periods to certificates being suggested. Still, relatively few research works have examined revocation in the context of IBE.

An efficient RIBE approach immune to such dangers was recently presented by Seo et Emura [36] in response to the real threat posed by the exposure of decryption keys. Even in the event that the decryption keys for the present time period are compromised, this approach makes sure that the security of the keys for other time periods stays uncompromised. Inspired by earlier efforts [37], Liang et al. [38] presented a cloud-based revocable identity-based proxy re-encryption. To make the revocation process easier, they used a broadcast encryption technique [39] to encrypt the ciphertext of the update key without regard to users. As a result, the update key can only be decrypted by non-revoked users. Revocable storage identity-based encryption (RS-IBE) is used to provide an affordable data transmission system that satisfies three security goals. [40]. The use of TPM in the hypervisor, the implementation of multi-factor authentication, and the modification of cloud service provider policy to allow persistent storage devices are some suggested cures. By improving the cloud service's investigability and bringing it into line with modern digital forensic investigation procedures, these modifications hope to increase customer trust [41]. The size of typical digital forensic investigations is increasing at a rate of 36% per year, according to a recent FBI report [42], highlighting the rise in digital crimes. A further worry that comes up during the physical evidence seizure is the privacy of other clients' data because several cloud clients share every resource at the same time.

III. PROPOSED SYSTEM

To securely access and share data in vulnerable environments like the cloud while preventing data loss, our proposal advocates for Authorized Role-Based Access Control and the utilization of the Advanced Cauchy Reed-Solomon method. Users can safely get their master and secret keys from the gateway administrator within our technology. The Third-Party Authority (TPA) is responsible for managing communication among various parties. The system offers secure revocation, particularly beneficial for untrusted users. Furthermore, our research recommends implementing proxy keys. The system quickly invalidates active keys and creates fresh ones for shared users when a data holder withdraws access for a particular end user. By using these techniques, the system can simultaneously attain the highest levels of security and confidentiality. For multi-authority cloud storage platforms, it presents a decentralized approach to data access control that effectively handles attribute revocation. This decentralized and revocable system helps identify unauthorized users through detailed logs, alleviating the decryption burden on users in accordance with their attributes. The primary aim of deploying this robust attribute-based encryption technology is to ensure robust data security within cloud storage environments. This approach holds potential practical applications in online social networks and remote storage facilities, among other contexts.

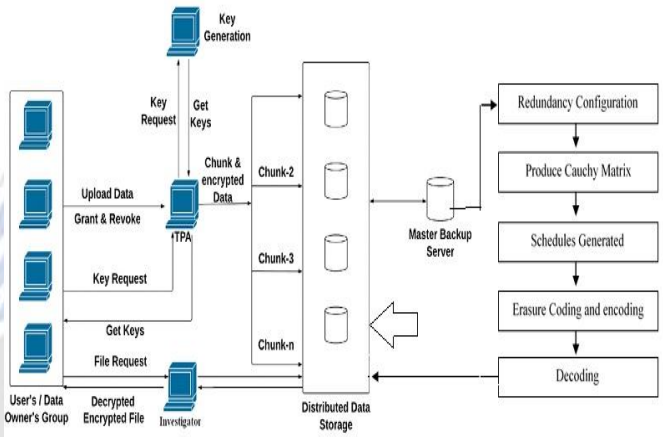

Figure 1: Proposed Systems Framework

The matrix architecture used to arrange and assess the cloud data hierarchy is shown in Figure 1. A node domain controller and a control broadcast channel make up this architecture. When tasks are scheduled or at predetermined intervals, data packets are transmitted from terminals to data sources. To simplify data decoding and address retrieval, the CaCo architecture incorporates additional self-address calibration units. Our proposed solution emphasizes reduced wait times and enhancements in system efficiency. Initially, the system computes values for (j, n, x), where J represents the total chunk count, N denotes the matrix nodes, and X signifies the sum of J and N. Once the vectors for j, n, and x are established, files can be uploaded. For instance, when j = 4, n= 2, and x = 6, the system generates 4 chunks and utilizes data nodes for storing the encrypted data. After data storage, it generates an [8 * 8] matrix for each algorithm. Determining conclusions becomes challenging as the "redundancy configuration" needs to adapt to user data. Additionally, individuals may choose to implement similar "matrix computation and redundancy configurations" based on research findings.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 May 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Module Description:

Registration and Authentication: The Certification Authority (CA) registers entities, including data owners, AAs (Authorized Agents), TTPs (Trusted Third Parties), and users, allowing them to build their profiles under CA.

Data Uploading: Once user identity verification is done, the data owner can upload the file. The Elgamal encryption method encrypts the data, and both keys are shared with TTPs and AAs during this process.

Data Sharing: Any file can be shared with members of the cloud group by the data owner, who has the responsibility for data sharing.

Control of Access & The revocation: Persons who hold the necessary authorization can view or utilise shared files under access control. The data owner has the ability to deny a user access to a particular file during revocation.

File retrieval & downloading: Users can request downloads from the cloud server, alongside verification from TTP (Trusted Third Party) and AA (Access Authorization).

File data restoration: The Advanced Cauchy Reed Solomon technique is used in case of data loss to achieve fault tolerance or restore the original data.

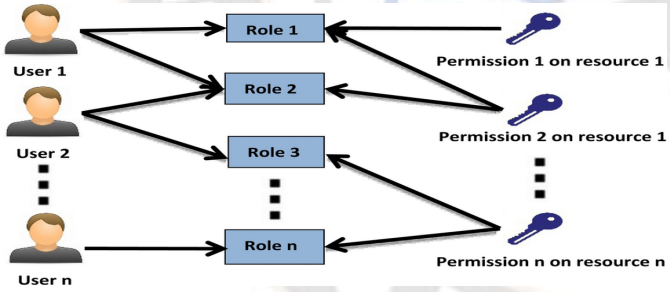

Figure 2: Role-Based Access Control

Take a look at Figure 2 illustrating the approach for controlling access based on roles. To safeguard sensitive data against unauthorized use, illegal alterations, or improper destruction, corporations require a means to limit access among their personnel. RBAC serves as a method for restricting data access according to an individual's role within an organization. By implementing RBAC, individuals are granted access solely to the resources and information essential for their job responsibilities. Their access privileges are determined by their roles, thereby reducing the risk of mishandling confidential data. RBAC methods prove highly advantageous, particularly in larger organizations and those employing independent contractors. As the number of approved suppliers fluctuates and employee numbers grow, it becomes challenging to create unique access configurations for each individual. With a role-based access control system, administrators can categorize employees or contractors into predefined groups (roles) and allocate them access to specific resources. Upon completion of their current assignments, employees' group memberships can be revoked, limiting ongoing access. Administrators can also adjust authorization levels for these groups, enabling them to efficiently manage staff on a broader scale, enhance productivity, and ensure compliance.

RBAC Algorithms

- **Input: Attribute Em-ID, F Data and F key-data**
- **Output: Policies or signatures serve as rules.**

1. Create the data string Sr-list [].
2. Initialize variables: a=0, k=0, User Email-ID.
3. Access Filedata and filekey.
4. Populate a with {fkey list [i to n]}.
5. Populate k with {Em-ID List [i to n]}.
6. For each item in a, read and assign to Sr-list.
7. If key-data matches a and User E-ID matches k,
8. Display the information of User File Share.
9. Otherwise, not show User File Share information.
10. End loop.
11. End Procedure.

Role BAC allows admin to categorize users into various groups based on their multiple roles, permitting a single user to belong to multiple groups simultaneously. Access for employees typically depends on their active status, responsibilities, and relevant security policies. The most effective approach involves granting users the least amount of authority necessary to carry out their duties, often termed as "minimum authorization." This principle, known as access privilege management, significantly contributes to safeguarding individual data.

Cauchy Reed-Solomon Coding matrix

This novel approach uses the Advanced CRS technique to solve fault tolerance. The application of this tactic entails show every possible configuration of the data-matrix format and multiple schedulers, and then selecting the most effective method from those that are offered. Since there isn't a specified rule or process for finding the ideal mix, this is the case. It is impossible to make any conclusions because the "redundancy configuration" needs to change depending on the user data. Additionally, a person may decide to implement the same type of "matrix computation as well as redundancy configuration" based on the research's findings.

Based on RS coding, the advanced Cauchy Coding (CC) technique is suggested. Figure 3 shows the CC model for data recovery. Here, several distinct entities are involved in the encoding and decoding process.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 May 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

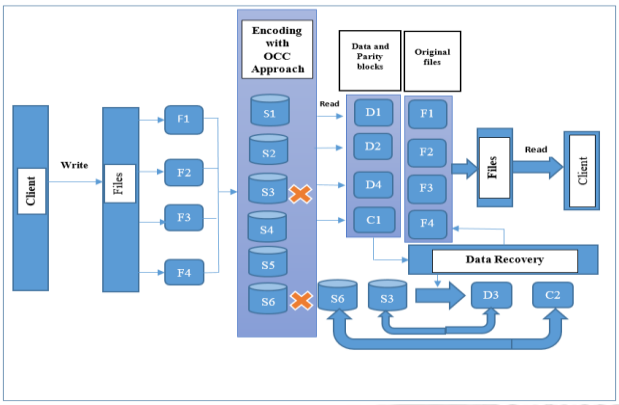

Figure 3: The encoding and decoding procedure

Based on RS coding, the advanced Cauchy Coding (CC) technique is suggested. Figure 3 shows the CC model for data recovery. Here, several distinct entities are involved in the encoding and decoding process.

When a client initially requests to store a data file, the request to write data undergoes an encoding phase, wherein the client-side files File1, File2, File3, and File4 can be stored across multiple server centers. To ensure data security, the original format of The original file is converted into a new format so that it can be stored on the server unit, rather than being stored in its initial form. The process of encoding entails encrypting the file that is received from the client in order to improve system durability against errors and data availability.

The CC technique assesses redundancy configurations (j, n, x) during the encoding process to construct a check matrix. In this setup, the quantity of coding units, parity blocks of data, and code data chunks bits per word are denoted as one, m, and n, respectively. This approach considers the balance between resource utilization and the time taken for matrix creation. The system is simulated under a dispersed environment, where individual Ca-Co coding matrices are generated for each segment through XOR operations. Cauchy matrices are constructed on GF, aiming to minimize both additions and multiplications by utilizing the Cauchy matrix alongside XOR operations. The efficiency of the encoding process in terms of time is influenced by the density of the Cauchy matrix.

Algorithm 1:

1. CC encoding algorithm
2. Input: {j,n,x} where j= block of data, ,n=parity blocksof data and x= code unit bits per word
3. Output: Code word check matrix

Algorithm 2:

Algorithm 1:
1. Encoding technique using CC (Cauchy Reed-Solomon) algorithm
2. Input: {j,n,x} - l denotes data blocks, m signifies parity blocks, and n represents coding unit bits per word
3. Output: Check matrix for the code word

Algorithm 2:
1. Recovery process for data
2. Input: Check matrix for the code word
3. Output: Original plain text data

The encoding process maintains data blocks and parity blocks for the input files among servers S1 through S6, once the check matrix is created in Algorithm 1. The check matrix created while encoding is used in the process of decoding, which is used when clients want to access the data. By reassembling the data, any missing data blocks can be retrieved.

A specific polynomial, we must first produce the components of GF(2^n) before producing its binary form matrix. In order to create a GF(2^n)using prime polynomials, we must first produce polynomials.

Create GF(2^3)for n=3 using prime polynomials. p(x)= x3+x+1

Generate a Cauchy coding matrix

Expand the Cauchy coding matrix of size $j \times m$ to the binary matrix of size $jm \times mn$. Consider the extended matrix as a matrix X.

Divide m data blocks into n parts similar $d_1, d_2, \ldots d_m$to into. $d_{1,1}, d_{1,2}, \ldots d_{m,n}$

$C = X \times C$

$C_{1,1} = (X_{1,1} \cdot d_{1,1}) \quad (X_{1,2} \cdot d_{1,2}) \quad (X_{1,mn} \cdot d_{m,n})$

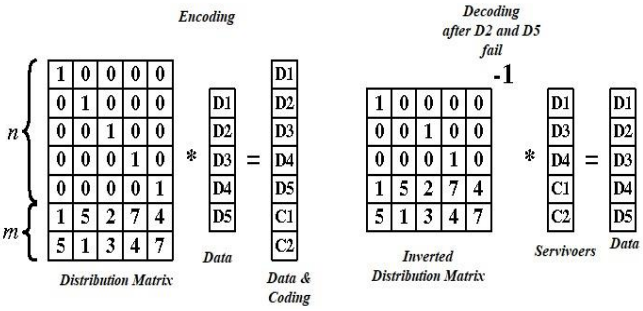

Figure 4: Encoding and decoding using Reed-Solomon codes with parameters n=5 and m=2 over the field(GF(2^3)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 May 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

To facilitate the computations, A column vector D = D1,..., Dn> contains n data words, each of which is depicted. D is multiplied by an identity matrix made up of the first n rows of an (n + m) × n distribution matrix. The result is a column vector D|C with n plus m elements, where C = C1,..., Cm > stands for the coding words. As a result, each row in the distribution matrix is a storage mechanism that houses either data or code words. To recalculate the lost data during decoding, the distribution matrix's entries associated with node failures are deleted, the resulting matrix is inverted, and then multiplied by the remaining words. Figure 4 shows the procedure.

Snapshot Based Forensic Approach

The suggested model revolves around the Cloud Service Provider (CSP) storing snapshots of a Virtual Machine (VM) flagged for malicious activities by an system for detecting assaults. Concurrently, the CSP is asked to provide log files from the suspected VM, which are then gathered and analyzed by the investigator to extract evidence. This research aims to implement this approach across multiple VMs and examine the ramifications of using cloud-based virtual machines to obtain evidence. Building a structure for the investigation of digital assets is the aim. within the Cloud Infrastructure as a Service (IaaS). For accurate and valid evidence collection, the suspected VM needs continuous monitoring even after the identification of malicious behavior. Extended monitoring ensures a higher level of certainty regarding the presence of malevolent actions. The suspicious virtual machine (VM) is moved to various nodes in order to protect the privacy, reliability, and validity of other VMs once the investigator has identified the origins of the proof. The VM evidence is protected from potential contamination and manipulation by transferring it or isolation.

The following algorithm outlines the proposed method for conducting forensic investigations utilizing VM snapshots as evidence.

Algorithm for snapshot based attacker detection

- **Input = {userRequest, Event, uploadDoc}**
- **Process = {Snapshotcreation, Storage, invistigate}**
- **Output ={detectTraitor, Blockuser, Unblockuser}**

1. These are the virtual machine group.
2. VM = {VM[i1], VM[i2]}… VM[in]} represents a collection of virtual machines.
3. Intermediate_Process comprises {cloud _service, Investigator, db}.
4. Client= {C1, C2…Cn} denotes the set of users.
5. Cloud _service receives inputs from ←{C[i]…C[n]}, where all users initially communicate with cloud servers.
6. Cloud _service_provider also establishes communication with database servers, denoted as cloud _service_provider ← db.
7. The request flow is as follows: {F_Investigator →cloud _service_provider → db}.
8. The response follows this path: {db → cloud _service_provider → F_Investigator}.
9. The Investigator acquires the complete set of specific users' evidence using the formula below.

$$Eve = \sum_{i=1}^{n} (\text{Snap[i]} \rightarrow \text{Malicios activity})$$

1. Investigator can take the action according the current evidence
2. Success and Failure conditions
3. If(Request !=Null)
4. Success Condition4
5. If(snap[i] set ==Null)
6. Failure Condition

A unique method has been introduced to facilitate digital forensics within the cloud setting, focusing on enhancing performance by utilizing VM snapshots as evidential data. This method integrates an intrusion detection system within both the Virtual Machine (VM) and Virtual Machine Monitor (VMM) to pinpoint malicious VMs, consequently refining cloud performance concerning storage capacity and time by archiving snapshots of identified malicious VMs. The suggested approach involves capturing snapshots of suspected VMs, securely storing them in persistent storage, thereby augmenting overall cloud performance.

III. EXPERIMENTAL FINDINGS AND DISCUSSION

As per the Role-BAC scheme, client roles are assigned based on the principle of granting minimum necessary privileges for accessing objects. Monitoring occurs for each entry, while any unauthorized access is logged. In summary, this approach is dynamic, scalable, and supports both active and inactive workflows within the system. Furthermore, it offers various advantages:

1. Defense against Distributed Denial-of-Service (DDoS) Attacks: By employing a limited access strategy, this method defends against DDoS attacks, preventing adversaries from blocking services due to restricted access.

2. Reduced Server Threats: Restricting user access mitigates the risk of hackers obtaining sensitive information. Moreover, adhering to the least privilege policy minimizes the potential misuse of information even by legitimate users.

3. Cost Reduction for Organizations: The model's scalability leads to adaptable administration, operational, and maintenance costs that align with the accessed services.

4. Enhanced Server Response Time and Operational Efficiency: Unlike traditional access control methods, limiting daily

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 May 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

accesses per user/role reduces server workload, resulting in higher operational efficiency and faster response times.

5. Role Segregation and Auditing: Dividing permissions into specific roles simplifies auditing and minimizes conflicts within the system.

6. Improved Security through Least Privilege: Enhancing security against internal and external threats, this approach reduces the chances of data leaks and hacker infiltrations, safeguarding valuable data assets.

The performance of a distributed system with respect to data encoding and decoding times is assessed in the next section. By dividing data blocks into n parity blocks, the encoder makes the storage device insensitive to numerous disc failures. If each word in binary data has n bits, then the advanced CC code can work as long as $2n \geq j + n$. Different combinations of data and parity blocks are used in redundancy configurations (j, m, x) for data encoding. The suggested system's encoding and decoding times are observed in the analysis for various redundancy settings. The term "encoding time" relates to how long it takes to convert a file's contents into a secure format for storage, and "decoding time" describes how long it takes to un-encrypt a sequence. Table II shows how long the coding takes for files ranging from 1MB to 10MB using the suggested CC coding method.

Table I. Encoding And Decoding Time According To Sizeof Input Files For Proposed CC (L=10, M=20)

File size	Encoding Time	Decoding Time
1 MB	0.13	0.146
5 MB	0.27	0.32
10 MB	0.55	1.4

Table II outlines various factors such as data availability, recoverability, efficiency, and storage overhead in case of disk failure. In the case where the sources are represented by j = 4 and n = 1, it exhibits exceptional performance, strong the ability to recover, and data stability. Additionally, the storage overhead is minimal under this configuration. However, as the number of parity blocks rises, the required storage space also increases.

TABLE II. PERFORMANCE ANALYSIS OF PROPOSED CC FOR VARIOUS INPUT PAIR VALUES (L, M)

(HI=HIGH,ME-=MEDIUM, LO=LOW)

Encoding	Availability	Recoverability	Efficiency	Storage Overhead
l=4, m=1	HI	HI	HI	LO
l=5, m=1	HI	HI	ME	LO
l=6, m=1	HI	HI	HI	LO
l=7, m=1	HI	HI	HI	ME
l=4, m=2	HI	HI	HI	LO
l=4, m=3	HI	ME	HI	LO
l=4, m=4	HI	HI	ME	LO
l=4, m=5	HI	ME	Low	HI

Group administrators have the option to store their files in cloud storage. These files remain in cloud storage. A group administrator creates a secret key and disperses it among all members of the group. The group owner will provide a new key that is utilized by every group member if this one isn't used. In the event that the owner of the data withdraws access for a certain end user, the system instantly invalidates all active keys and creates new ones for all users who have shared access. Both the private key and user information are stored in cloud storage. When a user accesses a file in the cloud, the file directly interacts with cloud storage, enabling the user to read and modify its contents. For a third-party authority (TPA) to access a specific file from cloud storage, the user must provide them with a secret key. The TPA is granted the ability to modify the specific file within a designated time frame. Any subsequent user's access to the same file within the given period will terminate once the allocated access time for that user expires.

IV. CONCLUSION AND FUTURE WORK

This paper delves into addressing trust issues within cryptographic Role-BAC (RBAC) systems to safeguard data storage in cloud platforms. It proposes trust models for owners and roles within cryptographic RBAC schemes, enabling flexible access policies. These schemes ensure compliance with these policies in cloud settings. Owners and roles within the RBAC system can leverage these trust models to evaluate the dependability of specific users and roles. Data duplication and recovery are significant concerns in cloud-based environments. Cloud storage systems commonly employ diverse redundancy techniques based on the intended ratio of fault tolerance to efficiency. This study enhances data retrieval by employing Advanced Cauchy Reed-Solomon coding, demonstrating improved performance. Furthermore, the paper illustrates the

application of trust models through a real-world example, highlighting how forensic investigation can mitigate risks and enhance the identification of unauthorized users. Cloud software offers substantial benefits, particularly in meeting the increasing demand for online data sharing. This paper introduces a secure data-sharing system tailored for cloud computing. It integrates identity revocation and ciphertext updates to prevent revoked users from accessing previously shared or future data. Looking ahead, the system aims to strike a balance between energy conservation and resource virtualization using Advanced Cauchy RS code. This optimization aims to efficiently manage network resources such as storage, power, and bandwidth. The future plan includes deploying this system within a hybrid cloud and Hadoop distributed file system environment.

References

1. J. Wang, Y. Yang, T. Wang, R. S. Sherratt, J. Zhang, "Big Data Service Architecture: A Survey," Journal of Internet Technology, vol. 21, no. 2, pp. 393-405, Mar. 2020.
2. R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," IEEE Comput., vol. 29, no. 2, pp. 38–47, Feb. 1996.
3. R. Sandhu, D. Ferraiolo, and D. Kuhn, "The NIST model for role-based access control: Towards a unified standard," in Proc. RBAC, 2000, pp. 47–63.
4. N. A. Ghani, S. Hamid, I. A. T. Hashem, E. Ahmed, "Social Media Big Data Analytics: A Survey", Computers in Human Behavior, Vol. 101, pp. 417-428, 2019, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2018.08.039.
5. J. Yang, H. Zhu, T. Liu, "Secure and Economical Multi-Cloud Storage Policy with NSGA-II-C", Applied Soft Computing, Vol. 83, 2019, 105649, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2019.105649.
6. N. Chervyakov, M. Babenko, A. Tchernykh, N. Kucherov, V. Miranda 1 Lopez, J. M. Cortes-Mendoza, "AR-RRNS: Configurable Reliable Distributed Data Storage Systems for Internet of Things to Ensure Security", Future Generation Computer Systems, Vol. 92, 2019, Pages 1080-1092, ISSN 0167-739X, https://doi.org/10.1016/j.future.2017.09.061.
7. J. Li, J. Wu, L. Chen, "Block-secure: Blockchain Based Scheme for Secure P2P Cloud Storage", Information Sciences, Vol. 465, pp. 219- 231, 2018, ISSN 0020-0255.
8.] P. V. Bhuvaneshwari & C. Tharini, "Review on LDPC Codes for Big Data Storage", Wireless Personal Communications, Vol. 117, Issue 2, pp. 1601-1625, 2021.
9. Y. J. Tang and X. Zhang, "Fast En/Decoding of Reed-Solomon Codes for Failure Recovery", IEEE Transactions on Computers, vol. 71, no. 3, pp. 724-735, 1 March 2022, https://doi: 10.1109/TC.2021.3060701.
10. M. Makovenko, M. Cheng and C. Tian, "Revisiting the Optimization of Cauchy Reed-Solomon Coding Matrix for Fault-Tolerant Data Storage," IEEE Transactions on Computers, https://doi: 10.1109/TC.2021.3110131.
11. F. Kazemi, S. Kurz, Soljanin and A. Sprintson, "Efficient Storage Schemes for Desired Service Rate Regions", 2020 IEEE Information Theory Workshop (ITW), pp. 1-21, 2020.
12. X. Chen and X. Ma, "Optimized Recovery Algorithms for RDP $(p, 3) $ Code," in IEEE Communications Letters, vol. 22, no. 12, pp. 2443-2446, Dec. 2018, https://doi: 10.1109/LCOMM.2018.2875468.
13.] Z. Shen, J. Shu and P. P. C. Lee, "Reconsidering Single Failure Recovery in Clustered File Systems," 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), 2016, pp. 323-334, doi: 10.1109/DSN.2016.37.
14. M. Damshenas, A. Dehghantanha, R. Mahmoud and S. bin Shamsuddin, "Forensics investigation challenges in cloud computing environments," Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), Kuala Lumpur, Malaysia, 2012, pp. 190-194, doi: 10.1109/CyberSec.2012.6246092.
15. K. Lee, "Comments on "Secure Data Sharing in Cloud Computing Using Revocable-Storage Identity-Based Encryption"," in IEEE Transactions on Cloud Computing, vol. 8, no. 4, pp. 1299-1300, 1 Oct.-Dec. 2020, doi: 10.1109/TCC.2020.2973623.
16. Z. Tan, Z. Tang, R. Li, A. Sallam, and L. Yang, "Research on trust-based access control model in cloud computing," in Proc. 6th IEEE Joint Int. Inf. Technol. Artif. Intell. Conf., 2011, pp. 339–344.
17.] A. Barsoum and A. Hasan, "Enabling dynamic data and indirect mutual trust for cloud computing storage systems," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 12, pp. 2375–2385, Dec. 2012.
18. X. Li and J. Du, "Adaptive and attribute-based trust model for service level agreement guarantee in cloud computing," IET Inf. Security, vol. 7, no. 1, pp. 39–50, 2013.
19. G. Lin, D. Wang, Y. Bie, and M. Lei, "MTBAC: A mutual trust-based access control model in cloud computing," China Commun., vol. 11, no. 4, pp. 154–162, 2014.
20. L. Zhou, V. Varadharajan, and M. Hitchens, "Trust enhanced cryptographic role-based access control for secure cloud data storage," IEEE Trans. Inf. Forensics Security, vol. 10, no. 11, pp. 2381–2395, Nov. 2015.
21. C. Uikey and D. S. Bhilare, "TrustRBAC: Trust role-based access control model in multi-domain cloud environments," in Proc. Int. Conf. Inf. Commun. Instrum. Control, 2017, pp. 1–7.
22. M. Ghafoorian, D. Abbasinezhad-Mood and H. Shakeri, "A thorough trust and reputation based RBAC model for secure data storage in the cloud," IEEE Tran. Paral. Distribute. Sys., vol. 30, pp.778-788, 2018.
23. Olanrewaju, R. F., Abdullah, K., & Darwis, H. (2018, November). Enhancing Cloud Data Security Using Hybrid of Advanced Encryption Standard and Blowfish Encryption Algorithms. In 2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT) (pp. 18-23). IEEE.
24. Bhardwaj, A., Subrahmanyam, G. V. B., Avasthi, V., & Sastry, H. (2016). Security algorithms for cloud computing. Procedia Computer Science, 85, 535-542.

25. J. S. Plank, “A tutorial on Reed-Solomon coding for fault1tolerance in raid-like systems,” Softw. Pract. Experience, vol. 27, pp. 995–1012, Sept. 1997.
I. S. Reed and G. Solomon, “Polynomial codes over certain finite fields,” J. Soc. Ind. Appl. Math., vol. 8, no. 2, pp. 300–304, 1960.
26. J. Blomer, M. Kalfane, M. Karpinski, R. Karp, M. Luby, and D. € Zuckerman, “A xor-based erasure-resilient coding scheme,” International Computer Science Institute, Berkeley, California, USA, 1995.
27. J. S. Plank, J. Luo, C. D. Schuman, L. Xu, and Z. Wilcox-O’Hearn, “A performance evaluation and examination of open-source erasure coding libraries for storage,” in Proc. 7th Conf. File Storage Technol., Berkeley, CA, USA, 2009, pp. 253–265.
28. G. Zhang, G. Wu, S. Wang, J. Shu, W. Zheng and K. Li, "CaCo: An Efficient Cauchy Coding Approach for Cloud Storage Systems," in IEEE Transactions on Computers, vol. 65, no. 2, pp. 435-447, 1 Feb. 2016, doi: 10.1109/TC.2015.2428701.
29. Shamir, “Identity-based cryptosystems and signature schemes,” in Proc. Adv. Cryptol., 1985, pp. 47–53.
30. D. Boneh and M. Franklin, “Identity-based encryption from the weil pairing,” SIAM J. Comput., vol. 32, no. 3, pp. 586–615, 2003.
31. S. Micali, “Efficient certificate revocation,” Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. TM-542b, 1996.
32. W. Aiello, S. Lodha, and R. Ostrovsky, “Fast digital identity revocation,” in Proc. 18th Annu. Int. Cryptol. Conf., 1998, pp. 137–152.
33. Goyal, “Certificate revocation using fine-grained certificate space partitioning,” in Proc. 11th Int. Conf. Financial Cryptography, 2007, pp. 247–259
34. J. Edmonds, “Paths, trees, and flowers,” Can. J. Math., vol. 17, no. 3, pp. 449–467, 1965.
35. J. S. Plank, C. D. Schuman, and B. D. Robison, “Heuristics for optimizing matrix-based erasure codes for fault-tolerant storage systems,” in Proc. 42Nd Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw., Washington, DC, USA, 2012, pp. 1–12.
36. Z. Zhang, A. Deshpande, X. Ma, E. Thereska, and D. Narayanan, “Does erasure coding have a role to play in my data centre?” Tech. Rep. MSR-TR-2010-52, Microsoft Research, Cambridge, England, May 2010.
37. J. S. Plank and Y. Ding, “Note: Correction to the 1997 tutorial on Reed-Solomon coding,” Softw. Pract. Exp., vol. 35, pp. 189–194, Feb. 2005.
38. J. Wei, W. Liu and X. Hu, "Secure Data Sharing in Cloud Computing Using Revocable-Storage Identity-Based Encryption," in IEEE Transactions on Cloud Computing, vol. 6, no. 4, pp. 1136-1148, 1 Oct.-Dec. 2018, doi: 10.1109/TCC.2016.2545668.
39. M. Damshenas, A. Dehghantanha, R. Mahmoud and S. bin Shamsuddin, "Forensics investigation challenges in cloud computing environments," Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), Kuala Lumpur, Malaysia, 2012, pp. 190-194, doi: 10.1109/CyberSec.2012.6246092.
40. Federal Bureau of Investigation (FBI), "Regional Computer Forensics Laboratory (RCFL)", Program Annual Report for Fiscal Year 2007, Washington, DC, 2008
41. Shaftab Ahmed, M. Yasin Akhtar Raja, "Tackling cloud security issues and forensics model", High-Capacity Optical Networks and Enabling Technologies (HONET), 2010, vol., no., pp.190-195, 19-21 Dec. 2010