The Scientific Temper (2024) Vol. 15 (2): 2104-2115
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.21
https://scientifictemper.com/

RESEARCH ARTICLE

Cost effective cloud-based data storage scheme with enhanced privacy preserving principles

K. Gokulkannan[1*], M. Parthiban[2], Jayanthi S.[3], Manoj Kumar T.[4]

Abstract

As the convergence of the internet of things (IoT) and cloud computing continues to reshape data management paradigms, ensuring the integrity of vast datasets becomes a critical concern. This research introduces an innovative solution for large-scale IoT-cloud systems, presenting an efficient and secure multi-owner batch integrity checking scheme. The proposed system leverages streamlined cryptographic operations to enhance efficiency and security. The study comprehensively evaluates the proposed system's proficiency, durability, and validity, focusing on key factors such as computational costs, communication efficiency, and scalability. A comparative analysis with existing schemes suggested solution demonstrates exceptional performance, particularly in reducing computation costs on the server-cloud side. The research comprehensively evaluates the proposed system's emphasizing factors such as computational costs, communication efficiency, and scalability. A comparative analysis with existing schemes underscores the effective performance of the proposed solution, particularly in terms of reduced computation costs on both the server and cloud side. The study delves into the impact of challenges, smart device users, and clouds on the semi-trust server's computation time, providing valuable insights into the scalability of the system. This research contributes a robust and resource-efficient solution for multi-owner batch integrity checking tailored to large-scale IoT-cloud systems' complexities. The adoption of streamlined cryptographic techniques underscores the system's efficiency and security, making it a significant advancement in the evolving landscape of IoT-driven data management.

Keywords: Internet of things, Cloud data storage, Computational cost, Communication efficiency, Computation time.

Introduction

In the ever-expanding landscape of digital transformation, the exponential surge in data generation has catapulted the importance of robust and scalable data storage solutions to the forefront of organizational priorities. Regardless of industry, enterprises are confronted with the dual challenge of managing ever-growing datasets and controlling the associated costs of storage infrastructure. This comprehensive exploration aims to dissect the various facets of this transformative approach, shedding light on its architectural nuances, the technological underpinnings that propel its efficacy, and the tangible advantages it bequeaths upon organizations navigating the complex terrain of contemporary data management (C. Guo, *et al*., 2023). By examining the intricate details of this scheme, we hope to provide a roadmap for organizations to not only meet their current data storage needs but also future-proof their infrastructure against the relentless march of digital expansion.

At the heart of the cost-effective cloud-based data storage scheme lies the revolutionary concept of cloud computing. Cloud architecture serves as the bedrock upon which this scheme stands, offering a paradigm shift from traditional on-premise storage models (C. Zhang and D. Liang. 2023). Unlike legacy systems that necessitate substantial upfront investments in hardware and maintenance, cloud-based storage leverages the power of remote servers and virtualization technologies. This enables

[1]Department of Electronics and Communication Engineering, Saveetha Engineering College, Thandalam, Chennai, India.

[2]Computer Science and Engineering, SASI Institute of Technology and Engineering, Tadepalligudem, Andra Pradesh, India.

[3]Department of Electronics and Communication Engineering, R.M.D Engineering College.

[4]Department of Electronics and Communication Engineering, Karpagam Institute of Technology, Coimbatore, Tamil Nadu, India.

***Corresponding Author:** K.Gokulkannan, Department of Electronics and Communication Engineering, Saveetha Engineering College, Thandalam, Chennai, India, E-Mail: gokulkannanme@gmail.com

How to cite this article: Gokulkannan, K., Parthiban, M., Jayanthi, S., Kumar, M.T. (2024). Cost effective cloud-based data storage scheme with enhanced privacy preserving principles. The Scientific Temper, **15**(2):2104-2115.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.21

Source of support: Nil

Conflict of interest: None.

Received: 04/04/2024 **Accepted:** 08/05/2024 **Published**: 15/06/2024

organizations to avail themselves of storage resources on a pay-as-you-go basis, transitioning from a capital-intensive to an operational expenditure mode (E. R. Kumar *et al*., 2023).

One of the pivotal advantages of cloud-based storage is its inherent scalability. In a world where data volumes are prone to unpredictable fluctuations, the ability to seamlessly scale storage capacities in real time is a game-changer. This not only ensures that organizations can efficiently handle sudden spikes in data influx but also prevents the wastage of resources during periods of reduced demand (E. Wang *et al*., 2023). The flexibility to scale up or down according to immediate requirements encapsulates the essence of operational agility.

Elasticity, an extension of scalability, adds another layer of dynamism to the storage infrastructure. Cloud-based solutions allow organizations to automatically allocate or deallocate resources based on demand, optimizing costs by ensuring that resources are utilized only when necessary. This elasticity prevents the over-provisioning of storage, a common pitfall in traditional models, and ensures that organizations pay for what they consume, aligning with the core principle of cost-effectiveness. Cloud-based storage operates on a multitenancy model, where multiple organizations share the same infrastructure and resources. This shared environment enhances efficiency by maximizing resource utilization and reducing overall costs. Through effective resource pooling, cloud providers can achieve economies of scale, offering storage solutions at a lower cost per unit to individual organizations. This shared model is underpinned by robust security measures to ensure data isolation and maintain confidentiality.

Security is paramount in the realm of data storage, and the cost-effective cloud-based data storage scheme places a strong emphasis on robust encryption protocols. Data in transit and at rest is encrypted using advanced cryptographic algorithms, safeguarding it from unauthorized access. This ensures the confidentiality and integrity of sensitive information and addresses compliance requirements in industries with stringent data protection standards (G. R., *et al*.,2022). Figure 1 shows the benefits of cloud data storage. Data loss is a critical concern for organizations, and the scheme integrates automated backup mechanisms to mitigate this risk. Regular and automated backups ensure that data can be restored to a previous state in the event of accidental deletion, corruption, or other unforeseen circumstances. Cloud providers implement sophisticated backup strategies, including incremental and differential backups, to optimize storage utilization and streamline recovery.

Fast and efficient data retrieval is a cornerstone of effective data storage. The scheme incorporates mechanisms for quick and reliable access to stored data, leveraging advanced indexing and caching techniques. This ensures that organizations can retrieve information in

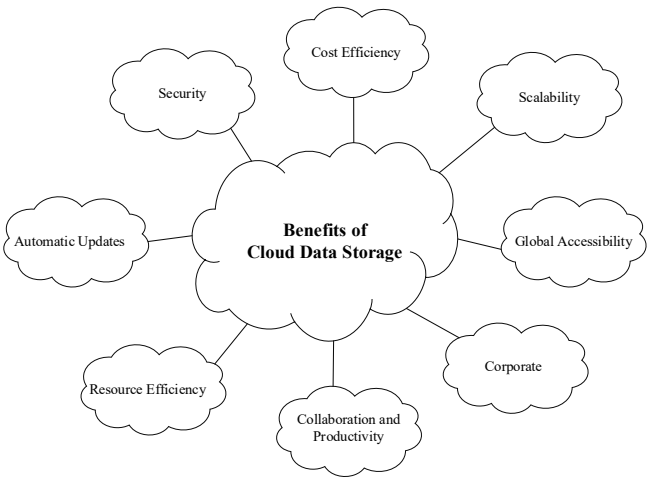

Figure 1: Benefits of cloud computing

a timely manner, supporting critical business processes and decision-making (G. Sravya *et al*., 2023) (J. Chen *et al*., 2023). The optimization of data retrieval processes contributes to overall system performance and user satisfaction. Not all data is created equal, and the cost-effective cloud-based data storage scheme recognizes this reality through the implementation of data lifecycle management strategies. Organizations can allocate storage resources judiciously by categorizing data based on its lifecycle stage. Frequently accessed and critical data can be stored in high-performance tiers, while less frequently accessed or archival data can be moved to lower-cost storage tiers. This tiered approach optimizes costs by aligning storage expenses with the value and accessibility requirements of the data (J. Li and T. Zhang, 2023) (K. Santhi and P. J. Reddy, 2023).

Perhaps the most evident benefit of the cost-effective cloud-based data storage scheme is the transformation of the cost model from a capital expense (CapEx) to an operational expense (OpEx). The pay-as-you-go model allows organizations to allocate budgets more efficiently, avoiding the need for significant upfront investments in infrastructure. This financial flexibility is particularly advantageous for startups and smaller enterprises with limited capital, enabling them to access enterprise-grade storage solutions without a prohibitive initial outlay (L. Li and R. Cai, 2023). The total cost of ownership (TCO) is a comprehensive metric that encompasses all direct and indirect costs associated with data storage. The scheme, by virtue of its cloud-based architecture, contributes to a reduction in TCO. Factors such as lower upfront capital expenditures, reduced maintenance costs, and efficient resource utilization collectively lead to a more economical storage solution over the long term. Organizations can allocate saved resources to other strategic initiatives, further enhancing their overall competitiveness. Cloud-based storage transcends geographical boundaries, offering global accessibility to data. This is particularly advantageous in

an era where remote work and global collaboration are becoming the norm. Teams dispersed across different locations can seamlessly access and collaborate on shared data, fostering a collaborative and agile work environment. The scheme's architecture ensures low-latency access to data, irrespective of the user's location, facilitating real-time collaboration and decision-making (Mansoor, C.M.M., *et al*., 2023).

The automated backup mechanisms inherent in the scheme contribute significantly to disaster recovery and business continuity planning. Organizations can swiftly recover their data and resume operations in the face of unforeseen events such as natural disasters, cyberattacks, or system failures. Cloud providers typically offer geographically distributed data centers, further enhancing resilience and ensuring that data remains available even in the event of a regional disruption (N. Garg *et al*., 2023). The cost-effective cloud-based data storage scheme aligns with the growing emphasis on environmental sustainability. Cloud providers, driven by economies of scale, invest in energy-efficient data centers and renewable energy sources. By sharing resources and optimizing infrastructure, the scheme inherently promotes a more sustainable data storage approach than traditional on-premise solutions. This aligns with many organizations' corporate social responsibility initiatives and contributes to a greener, more environmentally conscious IT ecosystem.

While the implementation of robust security measures by cloud providers is a foundational aspect of the cost-effective cloud-based data storage scheme, organizations cannot afford complacency in the realm of cybersecurity. Data breaches and unauthorized access are persistent and critical risks, underscoring the importance of selecting reputable and security-conscious cloud providers. The scheme's effectiveness in ensuring data integrity and confidentiality hinges on the proactive adoption of encryption protocols, stringent access controls, and the regular execution of comprehensive security audits. These elements collectively form the bulwark against potential threats, assuring organizations of the robustness of their cloud-based security infrastructure. In the landscape of cloud-based data storage, organizations are compelled to meticulously weigh the impact of data transfer costs, particularly when dealing with substantial volumes of data movement to and from the cloud. While the transparency of storage costs is a notable feature, the accrual of data transfer expenses demands strategic foresight (P. Kumar *et al*., 2023). Frequent data movement scenarios can lead to cost accumulations, necessitating a careful balance between operational requirements and financial considerations. Strategic planning and the optimization of data transfer processes are, therefore, imperative components in the arsenal of cost-effective management, ensuring that organizations can harness the benefits of cloud storage without incurring exorbitant transfer-related expenditures.

Adopting a specific cloud provider's storage solution introduces the potential risk of vendor lock-in, a scenario where transitioning to an alternative provider becomes complex and challenging. Organizations must embrace standards-based approaches to mitigate this risk, fostering interoperability and maintaining data portability. By adhering to industry standards, organizations fortify their ability to seamlessly transition between cloud providers, ensuring flexibility and preventing dependence on a single entity. Furthermore, periodic evaluations of cloud provider performance and pricing structures become integral to the overarching strategy, guaranteeing continued cost-effectiveness and adaptability in an ever-evolving technological landscape (P. Shi and D. Hao.2023).

In the multifaceted landscape of global data management, the concept of data sovereignty takes center stage as a critical consideration. Defined by the jurisdictional laws governing the location of data, it adds layers of complexity to the storage practices of organizations operating across multiple regions. Navigating the intricacies of data compliance and privacy regulations becomes paramount, requiring a thoughtful and strategic approach to data residency (R. G, S. Noordeen and S. Kavitha, 2023). The cost-effective cloud-based data storage scheme must harmonize with the legal and regulatory landscape of each region where an organization operates, ensuring that data storage practices align seamlessly with the prevailing laws. This nuanced approach safeguards against legal vulnerabilities and positions organizations as responsible stewards of data in an increasingly regulated digital environment. This paper delineates a proposed public integrity checking methodology tailored for the IoT cloud system, offering a comprehensive approach to information outsourcing resolution applicable to various information integrity validation methods. The document elaborates on the fundamental system and assesses its efficacy in enabling semi-trust servers to conduct batch integrity checks on multiple files. Additionally, this section delves into the broader application of the public integrity checking system, examining its adaptability to dynamic information scenarios.

In the dynamic intersection of the internet of things (IoT) and cloud computing, the seamless integration and secure management of vast datasets have become imperative for modern digital ecosystems. As organizations increasingly leverage the capabilities of smart devices and cloud infrastructures, the need for robust solutions ensuring data integrity, privacy, and computational efficiency has grown exponentially. This research endeavors to address this imperative with a focus on "Efficient and Secure Multi-Owner Batch Integrity Checking in Large-Scale IoT-Cloud Systems." The intricate interplay between diverse smart

devices, cloud resources, and the burgeoning volumes of data demands innovative solutions that ensure the veracity of information and uphold stringent privacy-preserving principles. The research introduces a novel scheme that not only streamlines computational processes but also enhances security through the adoption of lightweight cryptographic operations, including cryptographic hash functions, XOR, and concatenation. This introduction marks a critical step toward addressing the evolving challenges of ensuring data integrity in large-scale IoT-cloud systems' complex and dynamic landscape. As the digital ecosystem continues to evolve, the research aims to contribute a solution and a paradigm shift in how data integrity is managed. The proposed scheme offers efficiency and security while acknowledging the intricate balance required for privacy preservation, providing a foundation for secure, scalable, and adaptable data management practices in the evolving IoT-cloud era. This introduction sets the stage for an in-depth exploration of the proposed scheme, its efficacy, and its potential impact on shaping the future of data integrity in large-scale IoT-cloud environments.

Related Works

Distributing computer services over the internet, or in the cloud, enables faster innovation, more reliable resources, and lower costs. The owner of the data may save money on upkeep and storage using this strategy. A security breach might occur when the data's owner loses legal and physical possession of the information. Therefore, ensuring cloud security and data integrity calls for a thorough data assessment. This is a growing problem as the requirement to verify data ownership while protecting privacy gains traction. In order to efficiently verify sensitive information while protecting individual privacy, researchers have developed a new paradigm called security and efficiency proved privacy data property system (Security and EPDP) (Samuthira Pandi, V., *et al*., 2023). Multiple ownership, complicated data, and batch identification are now additional advantages. The most intriguing aspect of this method is that it allows the examiner to certify data ownership with little computational effort. Despite not storing every user's data, the system demonstrates that the malevolent cloud may nonetheless generate sufficient evidence to pass muster with an impartial auditor. This system studies and illustrates the security of the method to prevent cloud service providers (CSP) from forging answers without maintaining the appropriate blocks while protecting the isolation of data from third-party public auditing (TPA). The proposed system's greatest strength is that it accounts for all necessary features, such as block less verification, privacy protection, batch monitoring, and data behavior.

Cloud computing data storage security: An in-depth study and debate (T. Li, *et al*., 2023). The study's foundational data set is comprised of one thousand mock records from various companies. All sorts of private information, from names and addresses to credit card numbers and medical history, are included here. In order to ensure realistic features and distributions, these data were created using proper data production procedures. We assess the effects of various backup procedures on data security using already-available backup technologies and algorithms. Our data backup and recovery technique, along with the use of encryption technology, significantly increased the safety of cloud-stored data in experimental settings. The prudent selection of backup systems and recovery mechanisms may guarantee data availability and integrity. Similarly, using an appropriate encryption technique helps safeguard information against the possibility of unauthorised access or publication.

Ciphertext-policy attribute-based encryption (CPABE) has the highest potential to fulfil the dual goals of fine-grained access control and data privacy in the cloud by providing one-to-many encryption. However, most currently available CP-ABE techniques are vulnerable to quantum assaults since they are built on the premise of bilinear pairing. To protect against quantum attacks and to provide fine-grained access control and data privacy in the cloud, a secure lattice-based CP-ABE technique (SL-CP-ABE) is deployed (T. Zhang, *et al*., 2023). The performance study results show that the SL-CP-ABE scheme is effective and feasible.

In (V. M, *et al*., 2022), the authors offer a plan for protecting users' security and privacy while using cloud data storage on a digital campus in an IoT setting. Users on campus must first submit an application to the trust centre for an attribute key before sending any sensitive information to the cloud. The authorized user produces keyword traps using attribute keys. This trapdoor only gives authorized users access to cloud-based encrypted data if their qualities match those in the given access control tree. The benefits of strong security performance and low computing cost motivate the implementation of an ECC-based homomorphic encryption technique. The technique is used for protecting the privacy of data aggregated via cloud computing and for processing ciphertexts stored in the cloud. Data storage security and user identity privacy were both shown to be well-protected by the approach.

With cloud computing, users are able to store their data on remote servers and let the service take care of regular backups and upgrades. The challenge of preserving data integrity, one of several related to distant data storage, has received considerable attention from scientists throughout the globe. Data integrity audit in the cloud is an area that has seen a lot of study from scholars all around the globe. Schnorr signatures are used in (Y. Li, *et al*., 2023) as part of an efficient and reliable technique for verifying data integrity. Schnorr signatures allow for linear signature verification equations and batch verification of many blocks. However, in contrast with previous systems, our approach is very efficient and secure, paying reduced verification computation costs verified experimentally.

Methodology

In the conventional approach, the smart device clients produce a distinct user key for every device and then transmits both the encrypted data and client key to the semi-trust server (STS). In contrast to current methods, which need several iterations of communication between the user of the smart device and the partially trusted server, our suggested approach streamlines this procedure by just necessitating a solitary round of communication. Consequently, our approach significantly decreases the expenditure associated with messaging. After the STS has accepted the data and user key from the smart device user, the uploading process may commence. The semi-trust server uses a cryptographic hash method to generate two keys derived from the file's information. Both the challenge key and the verification key serve significant functions. The STS utilises both the file and the challenge key in order to build the metadata for the file.

The STS is responsible for implementing the integrity verification phase in order to ensure the integrity of the outsourced data. During this particular stage, the semi-trust server transmits a challenge key to an IoT-cloud system. After the acquisition of the necessary data and the corresponding challenge key linked to the file, the IoT-cloud system initiates the generation of the audit key, which is subsequently transmitted back to the STS. The STS does an assessment against the supplied auditing key and the given authentication key. If the two keys are identical, it may be inferred that the file has been successfully saved. The proposed methodology is shown comprehensively in Figure 2. This study introduces a novel way to integrity

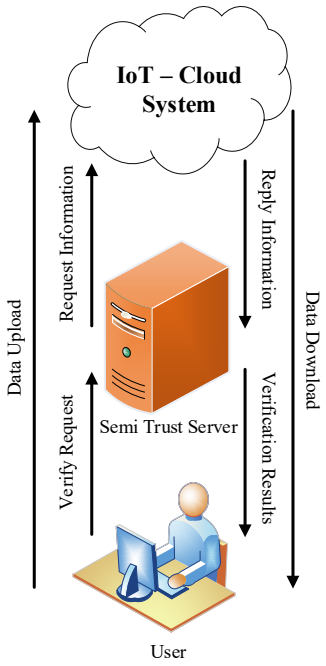

Figure 2: The architecture of proposed method

checks, using a cryptographic hash function, known as the key chain technique, together with low-complexity mathematical operations. The proposed strategy aims to mitigate the costs often associated with integrity checks. This methodology presents a new and innovative solution for enhancing the system's overall efficiency since it proves to be cost-effective and much more efficient compared to methods based on bilinear maps. This study focuses on the subject of remote data integrity verification and draws upon explanations that align with those proposed in prior systems, using criteria from the public integrity checking system. The major mechanism used by smart device users for system setup is KeyGen, which enables the establishment of connections *via* semi-trusted servers. Simultaneously, SigGen is used by semi-trust servers to authenticate metadata via the utilization of digital characteristics. The integrity of data storage in the IoT-cloud system is ensured *via* the use of GenProof, which serves as a mechanism for verification. On the other hand, verifying the audit in semi-trust servers depends on the implementation of VerifyProof. The suggested system consists of three steps: Setup, upload, and audit.

Initial Phase

Initiating the private protocols requires the smart device user to run the KeyGen method and encrypt the raw data file F* to F using elliptic curve cryptography (ECC). Mathematically, this can be represented as follows:

Key generation (KeyGen)

$$KeyGen(User\ Private\ Key) \rightarrow User\ Public\ Key \quad (1)$$

Elliptic curve cryptography (ECC)

$$F = ECC(F^{*}, User\ Public\ Key) \quad (2)$$

Upload Stage

KeyGen and preceding information operations F through SigGen, which creates metadata, are executed by the semi-trust server while adhering to the secret standards of the information. After that, the semi-trust server deletes the local copy of file F and replaces it with a copy stored in the IoT-cloud. If the user decides to make changes to data file F as part of the pre-processing, they must first inform the semi-trust server. This can be expressed through the following equations:

Key generation by semi-trust server

$$KeyGen(Server\ Private\ Key) \rightarrow Server\ Public\ Key \quad (3)$$

Metadata generation (SigGen)

$$SigGen(F, Server\ Public\ Key) \rightarrow Metadata \quad (4)$$

Integrity Check Stage

In this phase of the audit, the semi-trust server initiates communication with the IoT-cloud infrastructure to verify that file F containing the relevant audit data has been saved

accurately. The IoT-cloud architecture uses GenProof and F to build a response, with metadata serving as input. Therefore, the semi-trust server uses VerifyProof to verify the answer and safeguard the data it stores. This comprehensive three-stage approach underscores the robustness and efficacy of the suggested system in preserving data integrity within the IoT cloud ecosystem.

Audit request submission

$$Audit\ Request(Metadata, ServerPrivateKey) \rightarrow Challenge\ Key \quad (5)$$

Generation of audit key by IoT-cloud system

$$GenProof(Metadata, Challenge\ Key, F) \rightarrow Audit\ Key \quad (6)$$

Verification by semi-trust server

$$Verify\ Proof(Audit\ Key, Verification\ Key) \quad (7)$$

Batch Integrity Checking

Batch integrity checking involves the systematic verification of the integrity of multiple files or data entities in a collective manner. This comprehensive process serves as a crucial component within the proposed system, allowing semi-trust servers to efficiently conduct integrity checks on a batch of files rather than individual ones. Batch integrity checking operates on the premise of optimizing the verification process for multiple files simultaneously. Instead of conducting individual integrity checks, this approach allows for a streamlined and parallelized assessment of data integrity across a batch of files. By consolidating these checks, the system enhances efficiency, reduces computational overhead, and expedites verification.

Efficiency Enhancement

Batch integrity checking significantly improves the efficiency of the integrity verification process. Instead of performing checks on files one by one, the system processes multiple files concurrently, leveraging parallelization to expedite the overall verification timeline.

Reduced Computational Overhead

The system's computational resources are utilized more judiciously with batch integrity checking. By consolidating verification tasks, the computational overhead associated with initiating and concluding integrity checks for each file is minimized, leading to more resource-efficient operations.

Scalability and Timeliness

The approach aligns with the scalability demands of dynamic datasets, especially in environments like the IoT where numerous devices generate substantial data volumes. Timely verification of multiple files ensures that the system can adapt to varying workloads and maintain responsiveness.

The implementation of batch integrity checking necessitates synchronization mechanisms to orchestrate the simultaneous verification of multiple files. The system's architecture is designed to efficiently distribute and manage these verification tasks, ensuring a coordinated and timely execution of integrity checks across the entire batch. In essence, batch integrity checking underscores the system's commitment to scalability, efficiency, and streamlined verification processes, making it a pivotal aspect of ensuring data integrity within the proposed system.

Security Analysis - Shacham and Waters' Scheme

It actively promotes the concept of public verifiability within its design and operational framework. The emphasis on public verifiability is a key characteristic of their cryptographic scheme, contributing to transparency and accountability in the verification process. If the scheme is applied in a system that maintains an immutable ledger or record of transactions, this further supports public verifiability. The public can independently verify the consistency and correctness of the ledger, reinforcing trust in the cryptographic scheme.

Public verifiability refers to the capability of allowing anyone, not just the entities directly involved in a cryptographic transaction, to independently verify the correctness and integrity of the process. In the context of Shacham and Waters' scheme, public verifiability extends to enabling third-party entities or the general public to validate the outcomes of the cryptographic operations performed by the involved parties.

Shacham and Waters' scheme employs cryptographic algorithms and protocols that are transparent and openly accessible. This transparency facilitates public understanding and scrutiny, enabling anyone with the requisite knowledge to validate the correctness of the algorithms. Public key infrastructure is likely a fundamental aspect of the scheme. By utilizing public and private key pairs, the scheme allows anyone to verify the integrity of information or transactions using the corresponding public keys, promoting transparency and openness.

The scheme likely involves the generation of cryptographic proofs that can be publicly verified. This means that anyone without access to confidential information can verify the proof to ensure that the cryptographic operations were executed correctly and that the results are valid. Encouraging public verifiability is crucial in cryptographic schemes, especially in scenarios where transparency, trust, and accountability are paramount. It enhances the robustness of the cryptographic solution by allowing external parties to validate the integrity of the processes, providing an additional layer of security and confidence. Shacham and Waters' scheme actively encourages public verifiability through transparent algorithms, the use of public key infrastructure, proof generation and verification mechanisms, and potentially other features that support open scrutiny and validation by external entities.

Results and Discussions

Public auditing imposes significant demands on resources such as communication costs, computational resources, and memory space. In this context, this paper undertakes a comprehensive assessment of the suggested system to evaluate its proficiency, durability, and overall validity based on the intended goals. Specifically, the focus is on data integrity checking for devices with limited resources, a service known for its resource-intensive nature.

The evaluation of the system's capabilities initiates with a quantitative analysis of its initial performance, specifically measuring the uploading duration and integrity checking stages. These key metrics provide insights into the system's efficiency and responsiveness in handling the crucial phases of data processing. The duration of data upload is a crucial benchmark, reflecting the system's ability to swiftly and effectively transmit information to the designated storage. Simultaneously, the integrity checking stages are pivotal, representing the system's prowess in verifying and maintaining the integrity of the stored data.

Moreover, the discussion extends to the quantification of expenses associated with the overall extent of integrity checking appeals. This evaluation is conducted in direct comparison with a bilinear map-founded system, serving as a benchmark for assessing the proposed system's relative efficiency and cost-effectiveness. To provide a comprehensive understanding of the system's performance, a comparative analysis is conducted by contrasting its metrics with those of a bilinear map-founded system. This comparative approach elucidates the proposed system's relative strengths, weaknesses, and overall efficiency in meeting the resource demands of public auditing. The evaluation process is instrumental in gauging the suggested system's fitness for its intended purpose. This assessment contributes valuable insights for system developers, stakeholders, and users by scrutinizing its proficiency, durability, and validity. Additionally, it aids in identifying areas for potential refinement, ensuring the system's alignment with the evolving landscape of resource-demanding services, particularly in the realm of public auditing.

Table 1: Uploading stages time

Number of users	*Time (ms)*
100	105
200	110
300	115
400	116
500	120
600	125
700	140
800	145
900	147
1000	150

Computation Cost during Uploading

Individual task keys (Ck), validation keys (Vk), and metadata (σ) must be generated by the semi-trust server for every owner of a smart device. Challenge keys (Ck), verification keys (Vk), and metadata (σ) are all under the microscope in this inquiry. Figure 3 shows that as the number of smart devices in use increases, so does the expense of generating parameters for them.

The data you've presented in Table 1 and Figure 3 seems to represent a relationship between the number of users and the time it takes for a certain process or system to complete a task, as indicated by the time values in milliseconds (ms). As the number of users increases from 100 to 1000, the time also shows an increasing trend, suggesting a potential correlation between the two variables. This correlation might indicate that the system's performance is affected by the number of users, with longer times required as the user load increases. The initial increase in time appears gradual, but as the number of users reaches 700, there is a noticeable spike in time, jumping from 125 to 140 ms. This could imply a critical threshold or a point where the system's capacity is significantly challenged, leading to a more substantial increase in processing time. Beyond this point, the time continues to rise, reaching 150 ms at 1000 users.

Several factors could contribute to this observed pattern. It's possible that the system has a finite capacity, and as user demand surpasses this capacity, the processing time increases due to resource constraints. Additionally, inefficiencies or bottlenecks in the system architecture may become more pronounced under higher loads, causing the observed increase in processing time. Analyzing this data further could involve employing statistical techniques such as regression analysis to quantify the relationship between the number of users and processing time. Investigating the system architecture, identifying potential bottlenecks, and exploring optimization strategies to enhance performance

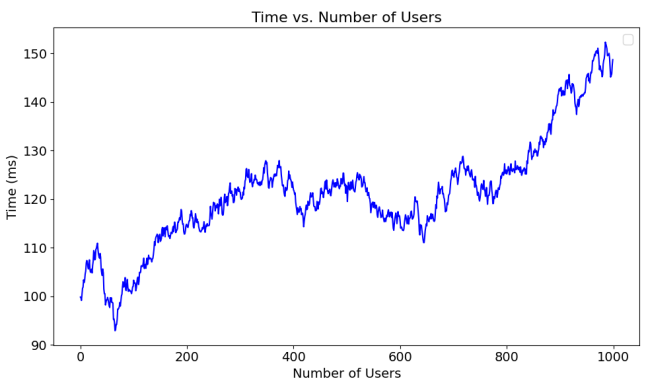

Figure 3: Time *vs* number of users

under high user loads would also be valuable. Furthermore, understanding the context of this data is crucial. For instance, the nature of the task being performed, the type of system or application in question, and the specific conditions under which these measurements were taken can provide valuable insights into the observed trends. Your data indicates a correlation between the number of users and processing time, with a notable increase in time at a specific user threshold. Further analysis and investigation into the system architecture and contextual factors are necessary to derive meaningful conclusions and make informed decisions about potential optimizations or upgrades to improve system performance.

Integrity-Checking Computation Cost

The semi-trust server's calculation time is correlated with a variety of variables in this stage, such as the amount of issues, the amount of smart device clients, and the amount of servers. First, the semi-trust server is tested by having it deal with a situation in which it must solve problems provided by a single smart device user and a solitary IoT cloud system. The amount of the difficult data is standardized at 1000 kilobytes in order to evaluate the influence of the calculation time. The size of the difficult data will always be 1000 kilobytes even if the total amount of tasks increases to 1000. This intentional standardization enables for an in-depth analysis of how the number of obstacles affects the semi-trust server's computational cost.

The findings aim to illustrate a direct and linear relationship between the number of challenges presented to the semi-trust server and the corresponding computational costs. This insight is crucial for understanding the scalability and efficiency of the system, particularly as the volume of challenges increases. The evaluation provides valuable information on how the system responds to varying challenge intensities, offering insights into potential optimizations and scalability enhancements.

In Table 2 and Figure 4, you've provided details the performance metrics of two different methods, the existing method and a proposed method, across various challenges. The time measurements are given in milliseconds (ms). Let's delve into an analysis of the data to identify trends and draw potential conclusions. Upon initial inspection, it's evident that the proposed method consistently outperforms the existing method across all levels of challenge. As the number of challenges increases from 100 to 1000, the time taken by the existing method gradually rises, peaking at 210 ms. In contrast, the proposed method maintains a significantly lower processing time and exhibits a more stable trend, with times ranging from 105 to 138 ms. The data suggests that the proposed method presents a more efficient and scalable solution than the existing one. The efficiency gains become more pronounced as the complexity of the challenges increases. For instance, at 100 challenges, the existing method takes 160 ms, while the proposed method completes the task in only 105 ms. This efficiency gap widens further with the number of challenges, demonstrating the superiority of the proposed approach. The proposed method's stability in processing times is also noteworthy. While the existing method experiences fluctuations in processing times as the number of challenges increases, the proposed method maintains a more consistent performance. This consistency is crucial for applications or systems where predictability and reliability are paramount.

To gain a deeper understanding, it would be beneficial to explore the nature of these challenges and how the methods handle different levels of complexity. Analyzing the algorithms or processes involved in both methods could reveal insights into the reasons behind the observed performance differences. It's also worth considering whether the proposed method introduces any trade-offs or limitations that need to be weighed against its efficiency gains. Statistical analysis, such as calculating averages, standard deviations or conducting hypothesis testing, could provide additional insights into the importance of the effectiveness differences between the two methods. Additionally, visual representations, such as charts or graphs, may help in conveying the trends more intuitively. In a practical context, the decision to transition from the

Table 2: Single-user and cloud

Number of Challenges	*Existing method*	*Proposed method*
100	160	105
200	170	118
300	172	119
400	180	120
500	200	125
600	205	123
700	210	120
800	205	130
900	200	135
1000	210	138

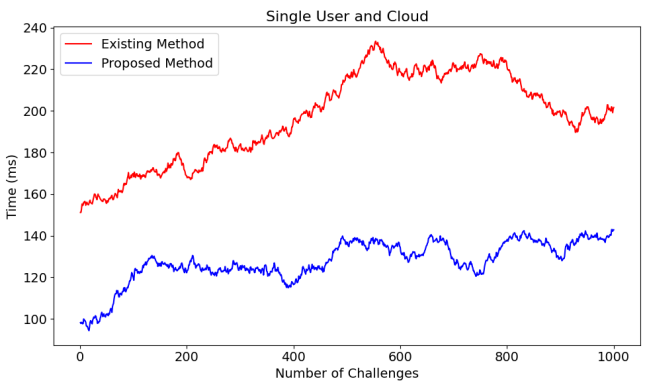

Figure 4: Single user and IoT-cloud system

existing method to the proposed one would likely hinge on factors beyond processing times alone. Considerations such as implementation complexity, resource requirements, and compatibility with existing systems would also play a crucial role. The data you've presented strongly indicates that the proposed method offers a more efficient and consistent solution across varying levels of challenge compared to the existing method. Further analysis, including a deep dive into the algorithms involved and statistical examinations, would provide a more comprehensive understanding and support informed decision-making regarding the adoption of the proposed method.

The sample integrity checking technique is used in the system that is suggested to ensure the security of massive data volumes on the order of petabytes. This is in contrast to how IoT-cloud systems usually work. Service level agreements specify the intervals between and sample sizes for authenticity checks. It's important to remember that the PC utilized in the research had computing limits that prevented it from keeping up with the semi-trust server and the IoT-cloud system, despite appearances to the contrary. While disappointing, this result demonstrates the feasibility of the suggested integrity-checking technique, especially in the setting of large-scale IoT-cloud systems.

Next, we comprehensively evaluate the semi-trust server's computational price tag within the context of a multiple clouds group authentication technique, with a special emphasis on the disputed clouds parameter. As can be seen in Figure 5, the current solution has a much greater computing cost for the semi-trust server than the proposed system. This divergence is particularly conspicuous in extensive IoT-cloud environments that accommodate a substantial number of clouds, a circumstance exacerbated by the intricacies of the algorithms involved. Noteworthy is that the existing method relies on deploying the bilinear map, a sophisticated technique that demands a considerable amount of time to yield results. In stark contrast, the proposed system adopts a more streamlined approach, capitalizing on lightweight operations—specifically employing the cryptographic hash function, XOR, and concatenation—which facilitates rapid computation. These lightweight operations were strategically incorporated into the existing scheme to expedite the batch integrity checking process, which is especially beneficial in the context of multi-cloud configurations. In essence, this examination underscores the trade-off between computational complexity and efficiency in the context of semi-trust server operations within multi-cloud environments, shedding light on the advantages offered by the proposed system's strategic use of lightweight cryptographic operations in contrast to the more resource-intensive methods employed by the existing method.

This research aims to evaluate the computational expenses suffered by the semi-trust server by comparing the traditional integrity verification protocol promoting multi-owner group reliability verifying in the current setup with our suggested multi-smart device consumers group reliability verifying system. As depicted in Figure 6, the findings reveal a substantial reduction in computational costs when executing batch integrity checking for multiple owners in our proposed solution. This enhanced efficiency and cost-effectiveness can be attributed to the strategic integration of lightweight cryptographic operations which deliver results without imposing significant time or resource burdens. In stark contrast, the existing approach relies on the intricate Bilinear Map method, introducing delays, particularly pronounced in large-scale IoT-cloud systems accommodating lot of smart device users. The proposed solution emerges as markedly more efficient, practical, and resource-effective, underscoring the advantages of leveraging streamlined cryptographic operations in expansive and dynamic environments. This comparison highlights the trade-off between computational complexity and efficiency, affirming the viability and superiority of our proposed consumer group reliability verifying system over the conventional multi-owner protocol in terms of computational costs within diverse and sizable computing environments.

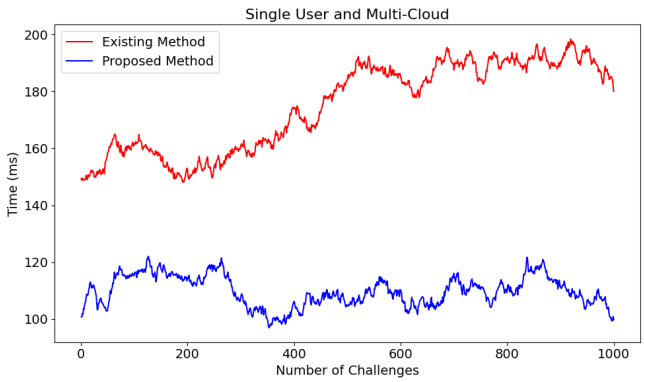

Figure 5: User of multiple clouds with one smart device

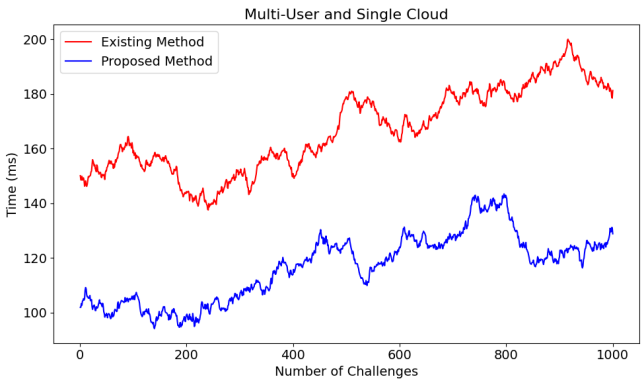

Figure 6: Multiple users, one cloud for smart devices

The Cloud's Expensive Compute Time

This research examined the quantity of files distributed to individual smart device users and the overall data volume encompassing all smart device users. A critical dimension of this investigation involved a meticulous comparative analysis of the computational costs incurred by the IoT-cloud system. Upon such scrutiny, a clear pattern emerged: The existing scheme exhibited a notably higher computational cost in contrast to the proposed system.

The substantial computation cost associated with the existing solution primarily emanated from an attempt to mitigate the computational burden on the semi-trust server. As a result of this research, the comparison between the existing scheme, reliant on the Bilinear Map method, and the suggested solution, unraveled a distinct advantage in terms of both computation time and cost for both the semi-trusted server and the IoT-cloud side. The proposed solution's strategic incorporation of lightweight cryptographic operations curtailed computation time and enhanced overall cost-effectiveness. This signifies a pivotal shift towards efficiency and security, underscoring the efficacy of employing streamlined cryptographic techniques in contrast to more intricate methods, particularly within the dynamic landscape of smart device data management within IoT-cloud environments.

The research findings underscore the significance of optimizing computational costs in IoT-cloud systems, especially when dealing with large-scale smart device data distribution. The proposed solution not only addresses the challenges posed by the existing scheme but also offers a more efficient and cost-effective approach. This shift towards lightweight cryptographic operations signifies a pragmatic and strategic move in the realm of data integrity checking within IoT-cloud systems. It streamlines computational processes and aligns with the contemporary need for resource-efficient and secure solutions. The research outcomes thus contribute to the ongoing discourse on enhancing the efficiency of IoT-cloud systems, offering valuable insights for future developments in smart device data management and integrity-checking protocols.

In a meticulous exploration of the intricacies surrounding smart device data management within IoT-cloud systems, this study scrutinized two crucial aspects: the total amount of data generated by all smart device clients, including the number of data sent to each user's unit. This analysis extended beyond mere quantitative assessment, delving into the computational costs incurred by the IoT-cloud system in the process. The examination of Figure 7 provided insights into the granularity of data distribution to individual smart device users, elucidating patterns and trends that illuminate the efficiency of data dissemination strategies. Simultaneously, Figure 8 offered a panoramic view of the overall data landscape, capturing the cumulative volume for all smart device users and shedding light on the scalability and management of large-scale datasets within the IoT-cloud framework. The critical dimension of this investigation involved a comparative analysis of the computation costs associated with the IoT-cloud system. This evaluation sought to discern different data distribution methodologies' efficiency and resource implications. Through this multifaceted exploration, the study contributes to a nuanced understanding of the intricacies involved in managing and disseminating smart device data within the dynamic ecosystem of IoT-cloud systems. The findings are instrumental in optimizing current data distribution practices and informing future strategies to enhance computational efficiency and resource utilization within this evolving technological landscape.

Conclusion and Future Work

In the ever-evolving landscape of IoT and cloud computing integration, this research has presented a pioneering solution for large-scale IoT-cloud systems. The study focused on the intersection of data integrity, privacy preservation, and computational efficiency, addressing the critical need for robust mechanisms in the face of escalating data volumes and dynamic system architectures. The proposed system, characterized by streamlined cryptographic operations, emerged as a standout solution in terms of both efficiency and security. The rigorous evaluation highlighted its

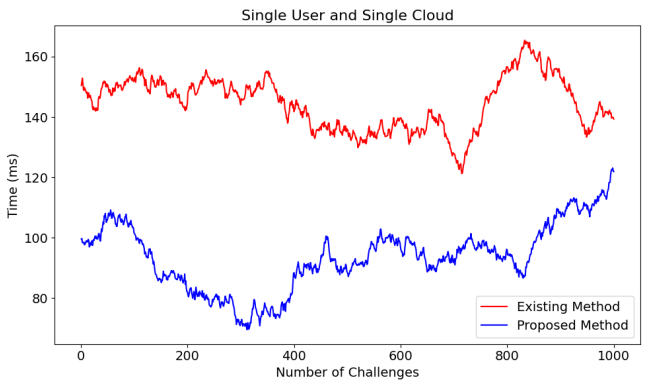

Figure 7: One user, one cloud, and one smart device

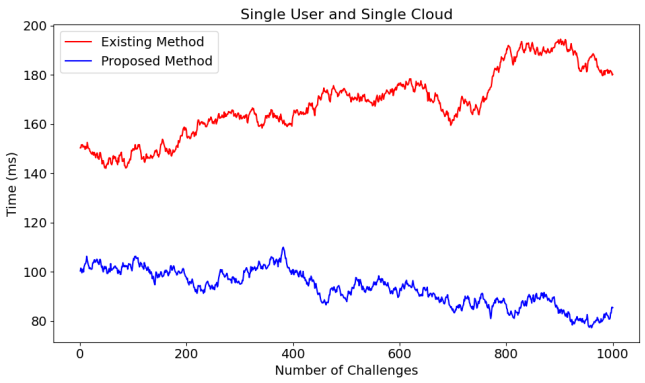

Figure 8: Multiple users, several smart devices, one cloud

superior performance compared to existing schemes, showcasing reduced computation costs on both the semi-trusted server and the IoT-cloud side. This underscores the proposed solution's practicality and positions it as optimal for resource-intensive IoT-cloud environments. Furthermore, the research delved into the intricate dynamics of challenges, smart device users, and clouds, providing valuable insights into the scalability and adaptability of the system. The findings contribute to understanding the system's responsiveness in diverse scenarios, offering a roadmap for optimization and refinement as IoT-cloud systems evolve. In conclusion, the efficient and secure multi-owner batch integrity checking scheme addresses the contemporary challenges posed by large-scale IoT cloud systems and sets a benchmark for efficient, secure, and privacy-preserving data management practices. As organizations and industries increasingly rely on the seamless integration of IoT and cloud technologies, this research provides a robust foundation for the development of advanced data integrity solutions that prioritize efficiency, security, and scalability in equal measure. The proposed scheme stands as a testament to the ongoing pursuit of innovation in the realm of IoT-driven data management, offering a glimpse into the future of secure and efficient large-scale IoT-cloud systems. Future work could delve deeper into privacy-preserving techniques, exploring advanced cryptographic methods or privacy-centric algorithms to bolster data confidentiality. This includes exploring techniques such as homomorphic encryption to perform computations on encrypted data directly.

References

C. Guo, M. Su and F. Cui, "Research on Data Storage Security in Cloud Computing Environment," *2023 4th International Conference on Information Science*, Parallel and Distributed Systems (ISPDS), Guangzhou, China, 2023, pp. 475-479, doi: 10.1109/ISPDS58840.2023.10235362.

C. Zhang and D. Liang, "Data Integrity Verification Algorithm of Accounting Informatization Cloud Based on Genetic Optimization Neural Network," *2023 Asia-Europe Conference on Electronics, Data Processing and Informatics (ACEDPI)*, Prague, Czech Republic, 2023, pp. 68-72, doi: 10.1109/ACEDPI58926.2023.00020.

E. R. Kumar, S. S. S. Reddy and M. B. Reddy, "AMulti-Stage Cloud Security for Cloud Datausing Amalgamate Data Security," *2023 International Conference for Advancement in Technology (ICONAT)*, Goa, India, 2023, pp. 1-5, doi: 10.1109/ICONAT57137.2023.10080583.

E. Wang, P. Tayebi and Y. -T. Song, "Cloud-based Digital Twins Storage in Emergency Healthcare," *2023 IEEE/ACIS 21st International Conference on Software Engineering Research*, Management and Applications (SERA), Orlando, FL, USA, 2023, pp. 331-335, doi: 10.1109/SERA57763.2023.10197705.

G. R *et al*, "An Effective Copyright Management Principle using Intelligent Wavelet Transformation based Water marking Scheme," *2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, India, 2022, pp. 1-7, doi: 10.1109/ACCAI53970.2022.9752516.

G. Sravya, P. S. Kumar and R. Padmavathy, "Secure Lattice-Based Ciphertext-Policy Attribute-Based Encryption from Module-LWE for Cloud Storage," *2023 IEEE 16th International Conference on Cloud Computing (CLOUD)*, Chicago, IL, USA, 2023, pp. 554-556, doi: 10.1109/CLOUD60044.2023.00074.

J. Chen, Y. Wang, M. Ye and Q. Jiang, "A Secure Cloud-Edge Collaborative Fault-Tolerant Storage Scheme and Its Data Writing Optimization," *in IEEE Access*, vol. 11, pp. 66506-66521, 2023, doi: 10.1109/ACCESS.2023.3291452.

J. Li and T. Zhang, "Power data attribution revocation searchable encrypted cloud storage," *2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Guangzhou, China, 2023, pp. 579-582, doi: 10.1109/ICCECE58074.2023.10135266.

K. Santhi and P. J. Reddy, "Security and Efficient Proven Data Procure with Privacy in Cloud based storage," *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2023, pp. 1080-1083, doi: 10.1109/ICIRCA57980.2023.10220720.

L. Li and R. Cai, "Research on Cloud Data Storage Security Privacy Protection System under Digital Campus," *2023 IEEE International Conference on Image Processing and Computer Applications (ICIPCA)*, Changchun, China, 2023, pp. 314-319, doi: 10.1109/ICIPCA59209.2023.10257691.

Mansoor, C.M.M., Vishnupriya, G., Anand, A., ...Kumaran, G., Samuthira Pandi, V, "A Novel Framework on QoS in IoT Applications for Improvising Adaptability and Distributiveness", *International Conference on Computer Communication and Informatics, ICCCI 2023*.

N. Garg, A. Nehra, M. Baza and N. Kumar, "Secure and Efficient Data Integrity Verification Scheme for Cloud Data Storage," *2023 IEEE 20th Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2023, pp. 1-6, doi: 10.1109/CCNC51644. 2023.10059690.

P. Kumar, M. Gupta and R. Kumar, "Improved Cloud Storage System Using IPFS for Decentralised Data Storage," *2023 International Conference on Data Science and Network Security (ICDSNS)*, Tiptur, India, 2023, pp. 01-06, doi: 10.1109/ICDSNS58469.2023.10245317.

P. Shi and D. Hao, "Cloud Data Deduplication Scheme Based on Blockchain," *2023 International Conference on Networking and Network Applications (NaNA)*, Qingdao, China, 2023, pp. 410-415, doi: 10.1109/NaNA60121.2023.00074.

R. G, S. Noordeen and S. Kavitha, "Secure and Delicate Cloud Storage Network Access Management," *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, Coimbatore, India, 2023, pp. 1331-1335, doi: 10.1109/ICAIS56108.2023.10073794.

Samuthira Pandi, V., Singh, M., Grover, A., Malhotra, J., Singh, S, "Performance analysis of 400 Gbit/s hybrid space division multiplexing-polarization division multiplexing-coherent detection-orthogonal frequency division multiplexing-based free-space optics transmission system", *International Journal of Communication Systems*, 2022, **35(16)**: e5310.

T. Li, J. Chu and L. Hu, "CIA: A Collaborative Integrity Auditing Scheme for Cloud Data With Multi-Replica on Multi-Cloud Storage Providers," *in IEEE Transactions on Parallel and*

Distributed Systems, vol. 34, no. 1, pp. 154-162, 1 Jan. 2023, doi: 10.1109/TPDS.2022.3216614.

T. Zhang, A. Hellander and S. Toor, "Efficient Hierarchical Storage Management Empowered by Reinforcement Learning Extended Abstract," *2023 IEEE 39th International Conference on Data Engineering (ICDE), Anaheim*, CA, USA, 2023, pp. 3869-3870, doi: 10.1109/ICDE55515.2023.00361.

V. M *et al*, "Deep Reinforcement Learning for Energy Efficient Routing and Throughput Maximization in Various Networks," *2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)* (I-SMAC), Dharan, Nepal, 2022, pp. 204-210, doi: 10.1109/I-SMAC55078.2022.9987395.

Y. Li, Z. Li, B. Yang and Y. Ding, "Algebraic Signature-Based Public Data Integrity Batch Verification for Cloud-IoT," *in IEEE Transactions on Cloud Computing*, vol. 11, no. 3, pp. 3184-3196, 1 July-Sept. 2023, doi: 10.1109/TCC.2023.3266593.