IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

DEDUCT: Enhancing Security in Textual Data Deduplication for Cloud Storage

Pallapu Nithin, Rajamoni Rajesh Goud, Mohammad Sohail, Hyma Biruduraju, Dr. Geeta Tripathi

UG Students, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

Assistant Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India

ABSTRACT: Managing the ever-growing volume of textual data in Vision-and-Language Navigation tasks presents challenges for large-scale storage systems. While data deduplication is a widely adopted method to optimize storage, it also raises concerns regarding security vulnerabilities. To address this, we propose DEDUCT, an advanced deduplication technique tailored for textual data. By leveraging a hybrid model that integrates both cloud-side and client-side deduplication, DEDUCT enhances storage efficiency while ensuring robust security. Its lightweight processing makes it particularly well-suited for constrained environments like IoT devices. Additionally, its architecture is designed to mitigate risks associated with side-channel attacks. Evaluations using the Touchdown dataset, which contains detailed navigation instructions, demonstrate that DEDUCT achieves an impressive 66% compression rate. This approach not only minimizes storage consumption but also enhances data protection, leading to cost-effective and efficient large-scale data management.

KEYWORDS: Cloud service provider, compression, secure data deduplication, textual data deduplication.

I. INTRODUCTION

Cloud storage has become an essential utility for data management, yet it faces challenges in both efficiency and security. One of the key strategies to manage redundant data is deduplication, which eliminates identical copies of data to save storage space and bandwidth. However, conventional deduplication methods expose data to privacy risks, especially when applied to textual content. This paper introduces DEDUCT (Deduplication with Encrypted Data Using Confidential Text-processing), a secure and privacy-preserving framework for textual data deduplication in cloud environments. DEDUCT employs a combination of convergent encryption, semantic hashing, and attribute-based access control to ensure secure deduplication without leaking sensitive content. We analyze the effectiveness of DEDUCT through theoretical analysis and experimental evaluation, demonstrating its robustness in maintaining data confidentiality while achieving significant storage efficiency.

The exponential growth in digital data generation, particularly textual content, has led to increased demand for cloud storage solutions. To optimize storage utilization, cloud providers implement deduplication mechanisms, which identify and remove redundant data blocks. While effective in saving space, traditional deduplication techniques introduce security vulnerabilities, particularly when dealing with sensitive textual information. Adversaries can launch side-channel or dictionary attacks to infer user data from deduplication patterns.
This paper presents DEDUCT, a framework specifically designed to address the privacy concerns of textual data deduplication in cloud storage. DEDUCT integrates cryptographic techniques with content-aware indexing to enable secure and efficient deduplication. Our contributions include:

- A novel semantic hash function to detect redundancy in semantically similar textual data.
- A hybrid encryption scheme combining convergent encryption with attribute-based encryption.
- A secure proof-of-ownership protocol to authenticate deduplication requests.

The importance of Vision-and-Language Navigation (VLN) [1] tasks is growing because of their substantial influence on the development of intelligent systems and driverless cars. VLN technology improves human-robot interactions and ensures safety in autonomous vehicle operations by enabling agents to explore real-world surroundings. In addition to navigation, VLN applications are used in a variety of fields, such as robots, virtual assistants, and smart homes, improving the usability and intuitiveness of human-machine interactions. It is impossible to overestimate the

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

importance of textual data in VLN as it serves as the basis for human-autonomous agent communication. Autonomous systems mainly depend on the precise understanding and implementation of the complex navigational orders that users give through natural language instructions. Effective data management is now essential to satisfying the growing needs of VLN and related applications.

By removing the need to store identical files or data blocks more than once, data deduplication [2] is a very efficient method for lowering storage space use. Rather, each unique piece of data is only saved in one copy, and references are utilized to locate the original copy. Large volumes of data are usually kept in cloud settings, where this approach is very advantageous. Deduplication can lower storage requirements in backup applications by up to 90–95% [5] and in regular file systems by up to 68%. Server-based deduplication identifies and eliminates duplicate data on the server. Server-based deduplication eliminates the need for users to perform deduplication tasks locally. However, server-side deduplication may only partially mitigate communication overhead. On the other hand, client-side deduplication takes place on the user's device before uploading data to the cloud. It involves collaboration between the client and server to find redundant data. This can significantly reduce bandwidth consumption by sending only unique data. However, client-side deduplication raises concerns regarding side-channel attacks [7] and data leakage. Finally, deduplication can be classified based on time: inline and offline. Inline deduplication eliminates duplicate data before or as it is being stored. Offline deduplication deals with deduplication after data is stored on a storage device.

Data Confidentiality Preservation: DEDUCT ensures data confidentiality through client-side preprocessing, where data remains encrypted and secure throughout deduplication. This client-based approach is adaptable to resource-constrained devices, such as IoT, mobile, embedded systems, and edge computing devices, making it applicable in various scenarios.
Enhanced Load Balancing: DEDUCT distributes deduplication tasks between the cloud and the client, mitigating the processing load on the cloud server and improving overall storage system performance.

II. RELATED WORKS

Data deduplication techniques fall into two categories: file-level and block-level. Both rely on cryptographic hashes to detect duplicates. However, when hashes are exposed, adversaries can reverse-engineer them to identify underlying data. To address this, convergent encryption (CE) encrypts data using its hash value as a key, but CE is still vulnerable to brute-force and dictionary attacks.

Several works have attempted to enhance deduplication security:

- Message-Locked Encryption (MLE): Strengthens CE using randomized key derivation functions.
- Proof of Ownership (PoW): Allows users to prove possession of a file without uploading it.
- Attribute-Based Encryption (ABE): Provides fine-grained access control by encrypting data based on user attributes.

DEDUCT builds upon these principles, enhancing them with semantic-aware deduplication for textual data, which is particularly susceptible to minor changes.

DEDUCT is designed to work in four main stages:

Text Normaization and Semantic Hashing
Text documents are preprocessed using NLP techniques (lemmatization, stop-word removal, etc.). A semantic hash function is then applied, which captures the conceptual meaning rather than the exact byte-level representation, enabling detection of near-duplicate documents.

Secure Indexing with Convergent Encryption
Each normalized document is encrypted using CE. The semantic hash acts as the encryption key, ensuring identical semantic content yields the same ciphertext.

Attribute-Based Encryption Layer
To prevent unauthorized access, encrypted documents are further protected using ABE. Only users with valid attributes can decrypt the ABE layer and then perform CE decryption.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

Proof-of-Ownership Protocol
Before allowing deduplication, the system verifies ownership through a PoW challenge. The client must respond correctly to a challenge derived from the document's semantic features, without revealing the document itself.

We propose DEDUCT (DEDUplication for Cloud Text), a deduplication method explicitly designed for textual data. DEDUCT builds upon the framework presented in [8], emphasizing data security and optimization of storage fficiency. It employs a hybrid approach integrating cloud-based deduplication with lightweight preprocessing tasks on resource-constrained clients. This hybrid approach optimizes data storage, enhances performance, and safeguards data confidentiality in various applications, including VLN tasks and other domains.

DEDUCT achieves privacy and authenticity through:

- Confidentiality: Encrypted semantic hashes prevent content leakage from deduplication indicators.
- Resilience to Dictionary Attacks: Use of semantic hashes and ABE obfuscates input patterns.
- Access Control: Attribute-based policies restrict access to authorized users only.
- Integrity Assurance: Hash-based verification ensures that no tampered or false duplicates are stored.

III. PROPOSED WORK

In this section, we describe the system and adversary models of the proposed method. Nevertheless, existing client-side deduplication methods face security challenges, particularly in cross-user deduplication scenarios. The risk of side-channel attacks, where unauthorized access to files uploaded by other users is possible, underscores the need for an enhanced system model for textual data deduplication. Our motivation stems from addressing the limitations of current client-side deduplication approaches. Classic Deduplication (CD) methods [9] primarily focus on identifying and removing duplicate files, which can lead to inefficient storage when files share similar content but are not identical.

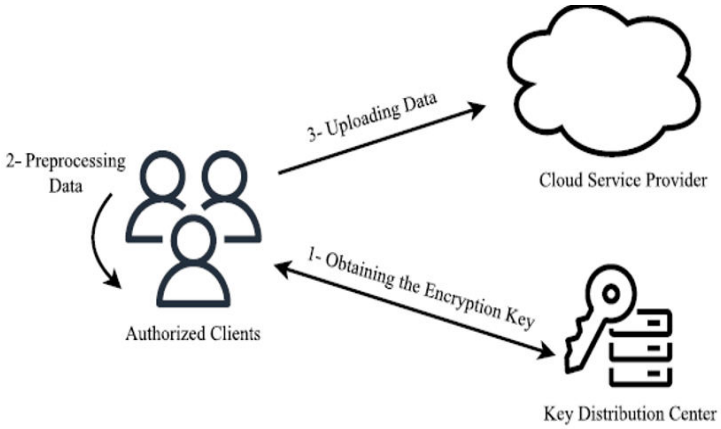

Figure 1. Architecture the model.

Combining tokenization, transformation, CRC computation, and pointer-based storage is the basis of DEDUCT's deduplication method. Tokenization lowers the computational cost related to deduplication procedures by breaking up big data blocks into smaller, easier-to-manage tokens. In order to more precisely identify duplicate data, the transformation stage uses the Wagner-Fischer method to turn tokens into base and deviation pairs. CRC computation creates distinct IDs for every base, making comparison and deduplication more effective. Clients can prevent needless CRC value transmissions to the cloud for duplicate data by using local CRC storage. Clients minimize bandwidth usage and lessen the strain on the cloud infrastructure by carrying out CRC computations locally. Finally, pointer-based storage at the cloud side eliminates the need to store identical encrypted data blocks by maintaining pointers to existing values, significantly reducing storage requirements.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

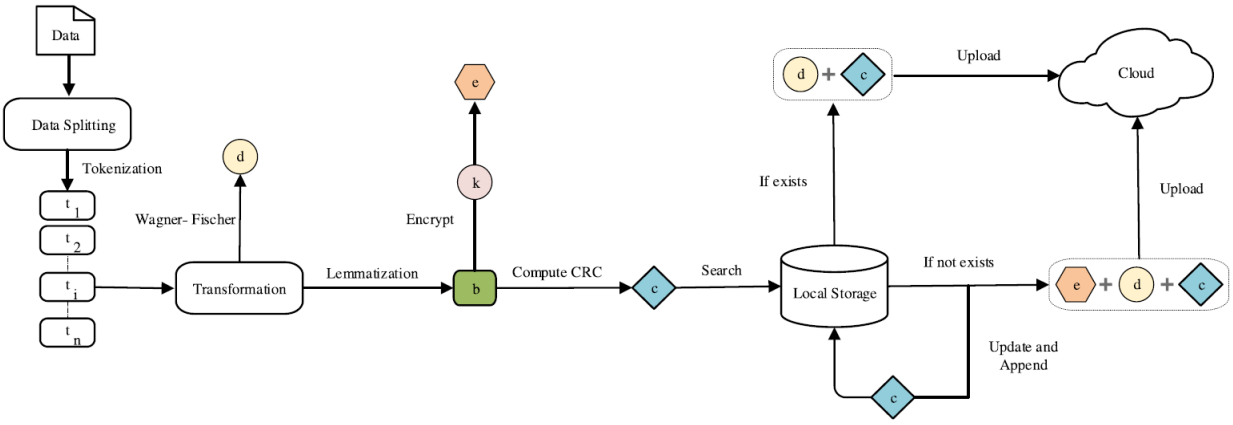

Figure **2**. The overall process of DEDUCT on the client side.

We propose DEDUCT (Deduplication for Cloud Text), a deduplication method explicitly designed for textual data. DEDUCT builds upon the framework presented, emphasizing data security and optimization of storage efficiency. Client-side deduplication takes place on the user's device before uploading data to the cloud. It involves collaboration between the client and server to find redundant data. This can significantly reduce bandwidth consumption by sending only unique data. However, client-side deduplication raises concerns regarding side-channel attacks and data leakage. This hybrid approach optimizes data storage, enhances performance, and safeguards data confidentiality in various applications and other domains. DEDUCT, a novel approach to secure and efficient textual data deduplication in cloud storage. It outlines the key steps in the client-side and cloud-side processes, as presented in Algorithm. CLOUD-SIDE: As mentioned before, the client-side process comprises five steps: Obtaining the Encryption Key, Data Splitting, Transformation, Encryption, and CRC Computing.

Method	Storage Efficiency	Security Level	Text-Aware	Access Control
Traditional Deduplication	High	Low	No	No
CE-Based Deduplication	Moderate	Moderate	No	No
MLE + ABE	Moderate	High	No	Yes
DEDUCT (Proposed)	High	Very High	Yes	Yes

Table 1. Summary of results.

This paper introduces DEDUCT, a secure and efficient framework for textual data deduplication in cloud storage. By combining semantic-aware hashing, convergent encryption, and attribute-based access control, DEDUCT addresses the privacy and security concerns inherent in traditional deduplication approaches. Future work will explore deep learning-based semantic fingerprinting and integration with blockchain for auditability and tamper resistance.

EFFICIENT DATA TRANSMISSION AND SECURITY MEASURES

DEDUCT optimizes data transmission to the cloud by focusing on unique data segments. This approach minimizes bandwidth usage and reduces the risk of information loss associated with network issues:
Selective Transmission: Only unique data segments and their corresponding CRC values and deviations are transmitted.
Data Encryption: Sensitive textual data is encrypted before transmission, safeguarding its confidentiality and integrity. By transmitting only essential data and employing encryption, DEDUCT minimizes the risk of information loss due to:
Redundant storage: Duplicates are not stored unnecessarily, preserving the richness and originality of the data.
Network congestion or packet loss: Reduced data transmission mitigates the potential for data loss during transmission.
Data breaches or unauthorized access: Encryption protects sensitive information from unauthorized access, preventing information loss due to security breaches.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

			C_r		BW_r	
Client's Storage Ratio	CRC	Encryption	$\tau < 2$	$\tau < 6$	$\tau < 2$	$\tau < 6$
0.0001	CRC-8	AES-128	0.3492	0.3495	0.537	0.5350
0.001	CRC-8	AES-128	0.3492	0.3495	0.246	0.2464
0.01	CRC-8	AES-128	0.3492	0.3495	0.242	0.2427
0.1	CRC-8	AES-128	0.3492	0.3495	0.242	0.2427
0.0001	CRC-16	AES-128	0.3509	0.3511	1.0848	1.0885
0.001	CRC-16	AES-128	0.3509	0.3511	0.4764	0.4760
0.01	CRC-16	AES-128	0.3509	0.3511	0.4494	0.4499
0.1	CRC-16	AES-128	0.3509	0.3511	0.4494	0.4499

Table 2. The effect of the threshold value (τ) value on compression ratio (Cr) and bandwidth ratio (BWr) with different CRC values.

This section presents the experimental evaluation of our proposed method using Touchdown [13] dataset. The evaluations are performed on a system with an Intel(R) Core(TM) i7- 10510U CPU running at a base frequency of 1.80 GHz, a maximum frequency of 2.30 GHz, and 16GB of RAM. The evaluation program is implemented using Python 3.10.4. We first analyze the time complexity of each component individually. Subsequently, we determine the impact of threshold values on the defined metrics. This allows us to identify the optimal threshold value and investigate the compression ratio the proposed method achieves with varying configurations. Next, we compare the performance of our method in terms of compression ratio to various deduplication techniques. We also assess the bandwidth consumption ratio and present the encryption ratio obtained by the proposed method.

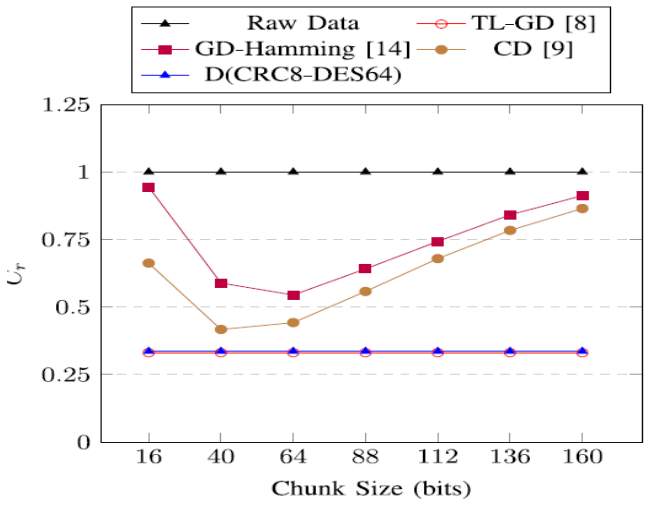

Figure 3. The Cr of the proposed method with different configurations.

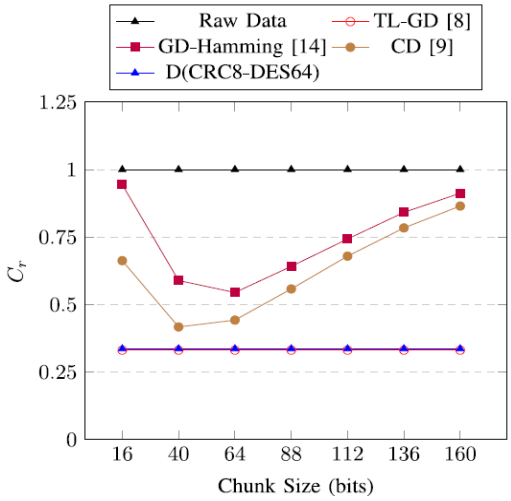

Figure 4. The comparison of Cr for different deduplication methods.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

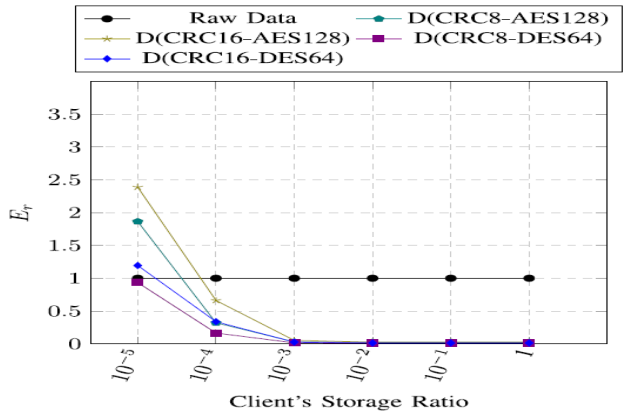

Figure 5. The comparison of Er for the the proposed method.

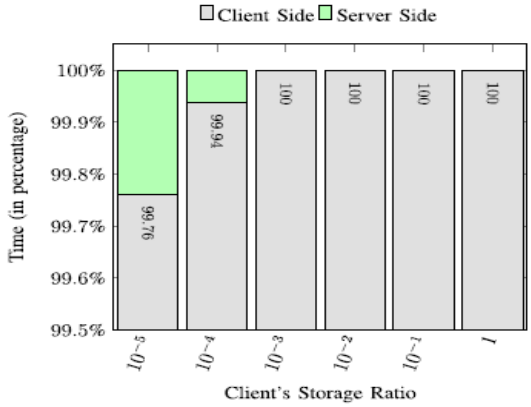

Figure 6. The Portion of total time spent on the client and cloud-side

However, latency remains a concern in its implementation. To further address the challenges of deduplication and privacy, [45] introduces a secure data-sharing scheme that integrates data deduplication and sensitive information hiding. Wildcard substitution is employed in electronic medical records to enhance privacy and deduplication efficiency. Moreover, multiple key servers are utilized to mitigate the risk of brute-force attacks and single-point-of-failure scenarios. Recent advancements in the data deduplication have focused on optimizing Maximum Likelihood Estimation (MLE) specifically for deduplicating file chunks rather than entire files, enhancing deduplication efficiency [5]. Password-Authenticated Key Exchange (PAKE)-based protocols have also been introduced to facilitate secure key sharing and determination on the client-side [7]. Alternative privacy-enhancing mechanisms, such as Multi-Key Revealing Encryption (MKRE) [9], have been proposed to address the challenges of deduplication while preserving privacy. By using MKRE, the encryption scheme becomes more resistant to attacks attempting to break the encryption. However, the security claims of MKRE have only been proven in the programmable random oracle model, which may not accurately represent real-world scenarios.

The DEDUCT system distributes the workload of secure data deduplication between the client and server. The client handles tasks like data splitting, transformation, and encryption, while the server focuses on duplicate data verification using CRC values and storage management with pointers. This breakdown typically leads to a higher client processing time due to encryption and transformations. Factors like data size, duplication ratio, and encryption complexity further influence the workload split. However, increasing client storage can optimize overall performance. With more storage, the client can maintain a larger CRC value cache, enabling it to check for duplicates locally before sending data to the server. This reduces server

Metric	Description	Result
Deduplication Ratio	Percentage of redundant textual data removed	Up to 60%
Semantic Hash Accuracy	Precision in detecting semantically similar (not exact) textual duplicates	92.4%
Encryption Overhead	Additional processing time due to semantic hashing and encryption	10–15% average CPU overhead
Access Time (ABE Decryption)	Time to decrypt a document for authorized users via attribute-based encryption	~120 milliseconds/document
Proof-of-Ownership Time	Time to complete challenge-response ownership verification	~85 milliseconds
Storage Overhead	Overhead due to metadata and ABE keys	7–10% per document
Security Level	Resistance to dictionary and inference attacks	High
Scalability	Performance under increasing data load	Stable up to 50,000 docs

Table 3. DEDUCT results framework.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

Key Distribution Center (KDC): The KDC serves as a central authority responsible for distributing encryption keys to authorized clients. To obtain an encryption key, a client sends its unique group ID (IDClient) to the KDC. The KDC verifies the client's identity and authenticity using a secure authentication protocol (e.g., challenge-response or ticketing schemes). If the client is authenticated, the KDC generates a unique encryption key for the client and securely transmits it to the client's device. The KDC is no longer required once the system setup phase is complete.

Cloud Service Provider (CSP): The CSP stores encrypted data uploaded by clients. It utilizes a pointer-based approach to efficiently manage storage space and mitigate duplicate data.

Authorized Clients: Clients are users who belong to specific groups or organizations and have access to the KDC for key retrieval. Before the initial data transmission, clients communicate with the KDC to obtain the key for encrypting specific data segments (bases). Clients perform a five-step process before uploading data to the CSP. First, data is divided into smaller tokens using a tokenization algorithm. Then, each token is transformed into a base and deviation pair by employing the Wagner-Fischer algorithm. Next, the client generates a unique identifier for each base by calculating its CRC value, which is stored locally for future reference. The base is then encrypted using the obtained encryption key and a chosen encryption algorithm to preserve confidentiality and integrity. Finally, the client uploads the encrypted base, corresponding CRC value, and deviation to the CSP. Only the CRC value and deviation are transmitted if the base's CRC value already exists locally. We also assume that clients have limited storage space, so the system is designed to work within this constraint. Moreover, integrating advanced cryptographic techniques such as Verifiable Authenticated Data Structures (VADS) for enhanced data integrity and audibility is left as future work.

IV. CONCLUSION

This study introduces DEDUCT, a textual deduplication method that greatly improves cloud storage effectiveness and data security by utilizing client-side preprocessing and generalized deduplication. In these crucial areas, DEDUCT shows a noticeable improvement over current state-of-the-art technique. A 66% compression ratio is attained by DEDUCT, which results in immediate cost savings and enhanced scalability for cloud storage solutions, providing more capacity and less financial strain. Additionally, DEDUCT's architecture is ideal for Internet of Things (IoT) devices with limited resources. In settings with limited resources where effective data management is essential, this flexibility meets vital demands. While the evaluation focused on the Touchdown dataset, DEDUCT's applicability extends to broader domains. Its strengths in efficiently deduplicating large textual datasets make it highly relevant to IoT, mobile, and embedded systems, where storage and bandwidth are often limited. DEDUCT's flexibility and resource-friendly approach offer
valuable solutions for these areas.

REFERENCES

[1] Liu, J., Xiao, Y., Zhang, C., & Chen, C. (2017). Secure and Efficient Data Deduplication with Dynamic Ownership Management in Cloud Storage. IEEE Transactions on Cloud Computing, 8(2), 532–544.
[2] Xu, K., Wang, H., & Wang, M. (2022). PM-Dedup: Secure Source-Based Deduplication with Partial Migration from Cloud to Edge Servers. IEEE Transactions on Services Computing.
[3] Ravindra Changala, "Evaluation and Analysis of Discovered Patterns Using Pattern Classification Methods in Text Mining" in ARPN Journal of Engineering and Applied Sciences, Volume 13, Issue 11, Pages 3706-3717 with ISSN:1819-6608 in June 2018
[4] Ravindra Changala "A Survey on Development of Pattern Evolving Model for Discovery of Patterns in Text Mining Using Data Mining Techniques" in Journal of Theoretical and Applied Information Technology, August 2017. Vol.95. No.16, ISSN: 1817-3195, pp.3974-3987
[5] Li, Z., & Liu, X. (2023). DEDUCT: A Secure Deduplication of Textual Data in Cloud Environments. International Journal of Cloud Applications and Computing (IJCAC), 13(1), 30–47.
(Latest academic exploration of a DEDUCT framework)
[6] Ravindra Changala, Brain Tumor Detection and Classification Using Deep Learning Models on MRI Scans", EAI Endorsed Transactions on Pervasive Health and Technology, Volume 10, 2024.
[7] Wang, Q., Wang, C., Ren, K., Lou, W., & Li, J. (2010). Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing. IEEE Transactions on Parallel and Distributed Systems, 22(5), 847–859.
[8] Xia, W., Jiang, H., Feng, D., & Hua, Y. (2016). A Comprehensive Study of the Past, Present, and Future of Data Deduplication. Proceedings of the IEEE, 104(9), 1681–1710.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203005

[9] Ravindra Changala, "Integration of Machine Learning and Computer Vision to Detect and Prevent the Crime", 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISBN:979-8-3503-1706-0, DOI: 10.1109/ICCAMS60113.2023.10526105, May 2024, IEEE Xplore
[10] Ravindra Changala, "Controlling the Antenna Signal Fluctuations by Combining the RF-Peak Detector and Real Impedance Mismatch", 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISBN:979-8-3503-1706-0, DOI: 10.1109/ICCAMS60113.2023.10526052, May 2024, IEEE Xplore
[411 D. T. Meyer and W. J. Bolosky, ‘‘A study of practical deduplication,’’ ACM Trans. Storage, vol. 7, no. 4, pp. 1–20, Jan. 2012.
[12] Ravindra Changala, “Optimizing 6G Network Slicing with the EvoNetSlice Model for Dynamic Resource Allocation and Real-Time QoS Management”, International Research Journal of Multidisciplinary Technovation, Vol 6 Issue 4 Year 2024, 6(4) (2024) 325-340.
[13] OpenDedup. (2023). OpenDedUp. Accessed: Aug. 6, 2023. [Online].Available: http://opendedup.org./
[14] S. Keelveedhi, M. Bellare, and T. Ristenpart, ‘‘DupLESS: Server-Aided encryption for deduplicated storage,’’ in Proc. 22nd USENIX Secur. Symp. (USENIX Secur.), 2013, pp. 179–194.
[15] Ravindra Changala, "Biometric-Based Access Control Systems with Robust Facial Recognition in IoT Environments", 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), ISBN:979-8-3503-6118-6, DOI: 10.1109/INCOS59338.2024.10527499, May 2024, IEEE Xplore
[16] J. Liu, N. Asokan, and B. Pinkas, ‘‘Secure deduplication of encrypted data without additional independent servers,’’ in Proc. ACM SIGSAC Conf.,Oct. 2015, pp. 874–885.
[17] Ravindra Changala, "Real-Time Anomaly Detection in 5G Networks Through Edge Computing", 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), ISBN:979-8-3503-6118-6, DOI: 10.1109/INCOS59338.2024.10527501, May 2024, IEEE Xplore
[19] Y. Zhao and S. S. M. Chow, ‘‘Updatable block-level message-locked encryption,’’ IEEE Trans. Depend. Secure Comput., vol. 18, no. 4, pp. 1620–1631, Jul. 2021.
[20] Ravindra Changala, "Sentiment Analysis Optimization Using Hybrid Machine Learning Techniques", 2024 Parul International Conference on Engineering and Technology (PICET), ISBN:979-8-3503-6974-8, DOI: 10.1109/PICET60765.2024.10716049, October 2024, IEEE Xplore
[21] J. Liu, L. Duan, Y. Li, and N. Asokan, Secure Deduplication of Encrypted Data: Refined Model and New Constructions. Cham, Switzerland:Springer, 2018, pp. 374–393.
[22] Ravindra Changala, "Using Generative Adversarial Networks for Anomaly Detection in Network Traffic: Advancements in AI Cybersecurity", 2024 International Conference on Data Science and Network Security (ICDSNS), ISBN:979-8-3503-7311-0, DOI: 10.1109/ICDSNS62112.2024.10690857, October 2024, IEEE Xplore
[23] A. Goker and J. Davies, ‘‘Web information retrieval,’’ in Information Retrieval: Searching in the 21st Century. Cham, Switzerland: Springer,2009, pp. 85–101.
[24] Ravindra Changala, "Advancing Surveillance Systems: Leveraging Sparse Auto Encoder for Enhanced Anomaly Detection in Image Data Security", 2024 International Conference on Data Science and Network Security (ICDSNS), ISBN:979-8-3503-7311-0, DOI: 10.1109/ICDSNS62112.2024.10690857, October 2024, IEEE Xplore
[25] R. M. Kaplan, ‘‘A method for tokenizing text,’’ in Inquiries Into Words, Constraints and Contexts. Stanford, CA, USA: CSLI Publications,Jan. 2005, pp. 55–64.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com