FORTIFYING CLOUD STORAGE: SAFEGUARDING SENSITIVE INFORMATION IN OUTSOURCING

Priya M.[1], Shylaja V.[2]
[1] Teaching Associate, Dept. of Artificial Intelligence and Machine Learning
[2] Teaching Associate, Dept. of Artificial Intelligence and Data Science
[1,2] Ramaiah Institute of Technology, Bengaluru, India

Journal	Samvakti Journal of Research in Information Technology ISSN (Online) : 2583-3979 https://www.sjrit.samvaktijournals.com Volume 6 Issue 1 (2025) Page No : 55 – 65.
Discipline	Cloud Computing
Conference	**Global Synergies: Innovations in Business, Technology and Education - INNOBTE 25**
Conference Dates	Start Date: March 21, 2025 End Date : March 22, 2025
Institute Name	Bangalore Integrated Management Academy

Date Received : February 22, 2025
ID : sjrit.2025.12
Dol No. : 10.46402/sjrit.2025.12
Publication Date: April 28, 2025
Paper Type : Conference Paper
Dol Url : https://dx.doi.org/10.46402/sjrit.2025.12

ABSTRACT

In response to the rapid growth of cloud-based services and the increasing storage of sensitive data, a robust biometric authentication system is urgently needed. This project addresses this demand by ensuring secure user identity in cloud computing environments. Our approach follows a tripartite framework involving data owners, users, and the cloud, ensuring biometric data confidentiality during transmission and storage. Encrypted biometric information is securely transferred from the data owner to the cloud when a user requests authentication. Upon receiving the encrypted query, the cloud server processes it and returns an index to the database owner, enabling biometric attribute similarity computation. The proposed framework integrates advanced encryption and privacy-preserving mechanisms to protect users' biometric data. Additionally, it mitigates collusion attacks between users and the cloud, strengthening system security. This research aims to establish a benchmark for secure, privacy-enhancing biometric identification in cloud environments, advancing cloud-based services with enhanced usability and stringent data security.

Keywords: Cloud Computing Security, Biometric Authentication Systems, Cryptographic Techniques.

INTRODUCTION

The landscape of user identification is shifting, with biometric authentication emerging as a promising alternative to traditional password and ID card-based methods. Biometric identification offers high reliability and convenience, utilizing distinct characteristics such as fingerprints, iris patterns, and facial features, easily captured by various sensors. Due to its ubiquity, it is widely used across multiple sectors. However, as biometric systems gain traction, organizations managing vast biometric databases, such as the FBI's national fingerprint database, face significant challenges. The sheer volume of data demands robust mechanisms for effective management and protection. To reduce high computing and storage costs, database owners often outsource data to cloud servers provided by industry leaders like Amazon. However, preserving biometric data privacy remains a critical concern, requiring encryption before outsourcing to safeguard each individual's unique biometric signature.

In practice, this encryption process presents challenges. For instance, if an FBI partner, such as a local police department, needs to verify someone's identity, they contact the FBI and generate an encrypted biometric query, which is sent to the cloud to find a matching profile. The challenge lies in designing a biometric identification framework in cloud computing that balances privacy protection and efficiency. The growing volume of sensitive data stored in cloud-based services underscores the urgency of this issue.

R.S. Sharma, chairman of the Telecom Regulatory Authority of India (TRAI), highlighted this concern. After issuing a public challenge on social media, ethical hackers accessed and exposed his private data within hours, revealing vulnerabilities in current systems. This incident underscores the need for robust security safeguards in handling biometric data, especially in cloud computing environments.

Several approaches have been proposed to ensure biometric identification with strong privacy protection. However, achieving a balance between privacy and efficiency remains a challenge. This study aims to bridge this gap by proposing a novel biometric identification system that ensures data privacy while enhancing efficiency and security. Through rigorous analysis and empirical validation, this research seeks to advance biometric identity systems, paving the way for cloud-based services that offer both convenience and robust data protection.

LITERATURE REVIEW

The use of biometric identification has significantly increased in recent years, prompting many database owners to consider outsourcing to the cloud to reduce costs. However, this shift has raised significant privacy concerns. A technique was introduced that involved encrypting biometric data before sending it to the cloud for identity verification. To enhance security, the query data was also encrypted before processing. A thorough security analysis confirmed the approach's resistance to forged requests, while experimental data demonstrated superior performance in both preparation and identification compared to earlier protocols Rao et al., 2023[1] .

Despite its potential, cloud computing has long faced persistent security concerns. Cloud storage schemes play a crucial role in protecting data, particularly for securely sharing personal health records (PHR). Fuzzy techniques were used to encrypt data into ciphertext, enhancing confidentiality. However, cloud efficiency in data sharing remained inadequate. To address this, fuzzy identity biometric encryption (FIBE) was introduced, merging fuzzy techniques with biometric authentication to improve PHR system security and user experience. This system eliminated the need for complex passwords while providing secure and efficient PHR data sharing in the cloud (Venkatachalam et al., 2023)[7]

While biometric identification is reliable and convenient, privacy concerns persist. Existing protocols often sacrifice either computational efficiency or security. To overcome these drawbacks, two privacy-preserving biometric identification outsourcing protocols were proposed. One utilized an invertible linear transformation and random split approach for protection against known-plaintext attacks, while the other employed Householder transformation and permutation. Theoretical analysis confirmed their robustness, and evaluations showed efficiency improvements over previous methods. These advancements enhance both security and efficiency, supporting broader deployment while safeguarding biometric data (Yang et al., 2021)[3].

Cloud data outsourcing for cost-effective IT services has raised privacy concerns, particularly for resource-constrained devices like mobile phones. To mitigate risks, a method for secure cloud data outsourcing was introduced, using ranked keyword searches with probabilistic public key encryption. This approach reduced computational and communication overhead by transmitting only relevant files. Security and performance assessments validated its effectiveness, addressing privacy concerns while ensuring usability for mobile devices (Xu et al., 2021)[9]

A method prioritizing data privacy while maintaining computational and communication efficiency was introduced. Before transmission, both biometric and query data were

encrypted. The cloud server processed the encrypted data and returned the index of final matches, allowing the system to verify biometric legitimacy. Security analysis confirmed its resilience against strong attacks, and testing results showed improved computational and communication efficiency over existing biometric identification methods (Shahrukh et al., 2019)[9]

Although facial recognition has enhanced privacy protection in cloud computing, it faces challenges such as noise, variability, processing delays, and susceptibility to spoofing attacks. To address these limitations, a biometric cloud identification method using multispectral palmprints was introduced. Encrypted multispectral palmprint features ensured privacy and anonymity during both online identification and offline registration. Experimental results verified its accuracy and effectiveness in protecting cloud data (Gumaei et al., 2019)[7] .

Advancements in cloud computing have increased the adoption of biometric identification, allowing database owners to reduce storage and processing costs by outsourcing tasks and datasets. However, this transition raises significant privacy concerns. A secure and efficient biometric identification outsourcing method was proposed, where query data was encrypted before cloud processing. The cloud detected the encrypted dataset and provided results. Security analysis confirmed resistance to collusion and attacks, while experimental results demonstrated superior performance over previous protocols in both preparation and identification phases (Zhu et al., 2018)[2]

The rapid development of biometric identification highlights the sensitivity of biometric data and the need for privacy-enhancing measures. Matrix-transformation-based techniques have been criticized for security flaws, while homomorphic encryption-based systems offer lower computational efficiency. A recently proposed matrix-transformation method by Zhu et al. was found vulnerable to chosen-plaintext and known-plaintext attacks, underscoring the need for improved security. The study emphasized the necessity of scalable schemes that enhance biometric data privacy and security (Liu et al., 2019)[3]

A survey of cloud computing and biometric authentication concepts highlighted the shift toward cloud service providers for security. It examined the limitations of traditional authentication methods and emphasized the potential of biometric-based solutions for secure access control. Privacy concerns were a key focus, particularly the risk of biometric template misuse. The survey also reviewed recent proposals aimed at improving security and user privacy in biometric authentication for cloud computing (Al-Assam et al., 2019)[5]

Proposed system

To recover data from cloud storage using a biometric method that maintains privacy and withstands security attacks, we proposed an efficient and secure biometric identification scheme. The system architecture of the proposed work is depicted in *Figure 1*: System Model Suggestion below.

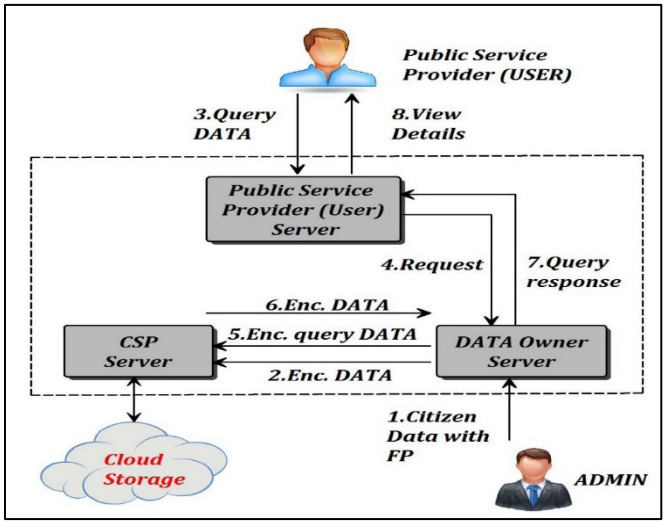

Figure 1: System Model Suggestion

Our system consists of three main components: the Cloud Service Provider (CSP), users (Public Service Providers), and the administrator (Database Owner). The administrator uploads biometric and secure data to the Data Owner Server, which encrypts the data and generates a hash code using the encrypted data and user ID. This information is then sent to the CSP web server, where it is stored in cloud storage, with Drive HQ used as the storage system.

To initiate the process, a Public Service Provider (User) submits an application package to the Data Owner Server, including the user ID and fingerprint (FP) for client verification. The Data Owner Server converts the user ID into a hash tag and sends it to the CSP server. The CSP server retrieves the relevant data based on the received hash tag and returns it to the Data Owner Server.

Upon decrypting the received data and extracting the FP, the Data Owner Server compares both sets of data. Based on the comparison results, it determines the authenticity of the query data. Finally, the Public Service Provider server receives the verification result and takes appropriate action.

Systems for Recognizing Fingerprints

This system employs minutiae-based fingerprint recognition, the most widely used biometric method for identification. Fingerprint identification establishes a person's identity, while fingerprint verification confirms it. Over the past 30 years, automated

fingerprint authentication has become more prevalent than alternatives like facial recognition and signature authentication. Once primarily used for criminal identification, it is now increasingly applied in civilian areas such as access control and financial security.

Fingerprint identification typically relies on minutiae—distinct features in ridge patterns. The two most common minutiae types are ridge endings and bifurcations *Figure 2*, which can be combined to form more complex fingerprint features. Minutiae matching involves using geometric transformations to align a subset of minutiae from the input fingerprint with those in the database template.

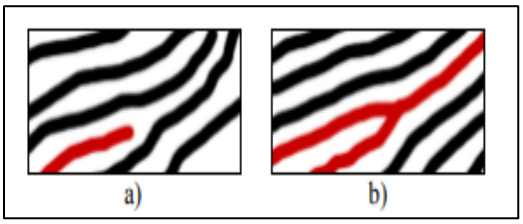

Figure 2: Example of a) ridge ending and b) bifurcation.

Feature Deletion

Several methods exist for extracting minutiae, with the most effective and widely used approach based on ridge thinning and binarization.

Ridge Thinning Method

The Crossing Number (CN) concept is used to process the binary ridge minutiae image before extracting finer details. First, the ridges are binarized and thinned until they are only one pixel wide *Figure 3.* Since skeletonization is crucial in many recognition systems, various approaches are available in the literature. Using a 3x3 window, each pixel's surrounding area in the ridge-thinned image is scanned to identify minutiae points *Figure 4.* Subsequently, the CN value is computed as half the sum of the differences between adjacent pixel pairs P_i and P_{i+1}. The ridge pixel is then categorized as a ridge termination, bifurcation, or non-minutiae point using the CN features listed in *Table 1.* This extraction method is the most commonly used.

Figure 3:Fingerprint image (a) binarization and (b) skeletonization.

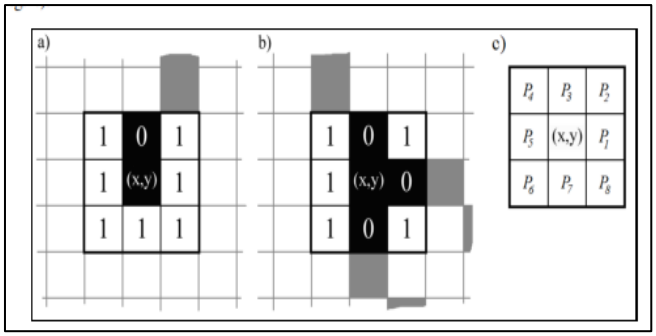

Figure 4:(a) Ridge ending and (b) bifurcation in (c) 3×3 window

CN	Property
0	Isolated point
1	**Ridge ending**
2	Continuing ridge
3	**Bifurcation**
4	Crossing

Table 1:Properties of the Crossing Number

Score

A high-quality rolled fingerprint image typically contains around 70–80 minutiae points, whereas a latent fingerprint has only 20–30. A minutiae-based fingerprint matching system identifies the number of matched details between the query and reference fingerprints to generate a similarity score. According to forensic standards, two fingerprints are considered to belong to the same individual if at least 12 minutiae points match.

The matching algorithm in *Equation 1* compares two minutiae sets: the similarity score S(T,I) is derived by taking the template T={m_1,m_2,…,m_j} from the reference fingerprint and matching it with I={m_1,m_2,…,m_i} from the query. Minutiae pairs m_i and m_a are considered matched only if the differences in their positions and orientations are within acceptable tolerance limits.

$$sd(mi, mj) = 1 \leftrightarrow \sqrt{(xi - xj)2 + (yi - yj)2} \quad \leq r0$$

Equation 1:Matching Algorithm

Alphabet for Fingerprints:

Step 1: Step 1: The colors initial.
Step 2: Launch the buffered picture creation process.
Step 3: Construct the binary image
Step 4: Produce a Greymap
Step 5: Local Binary Outcome
Step 6: Remove Noise (Use the mean technique to remove noise from the binary image)

Step 7: Skeletonization (We continue to skeletonize until two iterations do not differ from one another)
Step 8: Direction (Transform the direction matrix into a picture with a direction buffer)
Step 9: Minutiae (intersections: Take out the binary image's intersection points, which is a type of minutiae)
Step 10: Minutiae (endpoints - A type of minutiae is to extract end points from the binary image)

To enhance data security in our outsourcing project, we employ cryptography techniques using both symmetric and asymmetric algorithms. We generate hash functions using **SHA-1** and encrypt data with the **AES algorithm**, a widely adopted symmetric encryption standard known for its efficiency and security. AES replaced the **Data Encryption Standard (DES)** due to its superior speed and effectiveness. It operates as a block cipher with three variants, ensuring robust encryption. By integrating AES and SHA-1, our system provides a secure framework for protecting outsourced data. The AES algorithm ensures secure encryption by expanding the encryption key into a set of round keys. It begins with the Sub Bytes Transformation, a non-linear substitution process using a substitution table to replace each data byte. The Shift Rows Transformation shifts data rows by varying offsets, while the Mix Columns Transformation mixes data columns to enhance diffusion. The Add Round Key Transformation applies a bitwise XOR operation to combine each byte of data with a round key. AES operates through multiple rounds depending on key length—10 rounds for a 128-bit key, 12 for a 192-bit key, and 14 for a 256-bit key—ensuring secure encryption and decryption.

Hashing is a cryptographic method that converts input data into a fixed-length character string, typically represented as a hexadecimal integer. This process, performed by a hash function, is one-way and irreversible, meaning the original data cannot be easily retrieved from the hash. Hashing plays a crucial role in ensuring data integrity, secure password storage, and various cryptographic applications. Key points: (1) A wide range of data types, such as passwords, files, and messages, can be hashed to ensure security and data integrity. (2) The hash function generates a fixed-length output, usually a string of characters. (3) The same input always produces the same hash. (4) Hashing is one-way, meaning reversing the process is computationally infeasible. (5) Strong hash functions prevent two different inputs from producing the same hash (collision resistance). (6) Hashes verify data integrity by ensuring it has not been altered.

Message Digest 5 (MD5) is a well-known cryptographic hash technique that converts input data into a fixed-length 128-bit hash, typically displayed as a 32-character

hexadecimal string. Key aspects of MD5: (1) It is a one-way hashing function used to generate checksums and verify data integrity. (2) It produces a 128-bit fixed-length output. (3) MD5 is fast and efficient, making it suitable for quick hashing. (4) Despite being designed as a one-way function, it is vulnerable to preimage attacks, where attackers attempt to reconstruct the original data from the hash. (5) MD5 is also susceptible to collision attacks, where two different inputs produce the same hash. (6) Due to these vulnerabilities, MD5 is not recommended for security-critical applications. More secure alternatives, such as SHA-256 and SHA-3, are preferred.

Evaluation

This system meets the required specifications and has been tested and deployed in a hybrid cloud structure. Both user features functioned effectively. *Figure 5* shows a comparison of block upload and download times, while *Figure 6* presents the block verification time requirements.

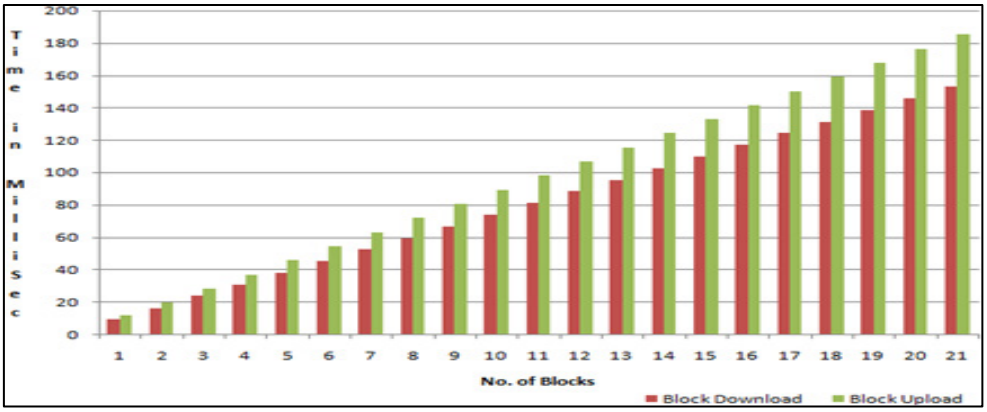

Figure 5:Block Uploading & Downloading

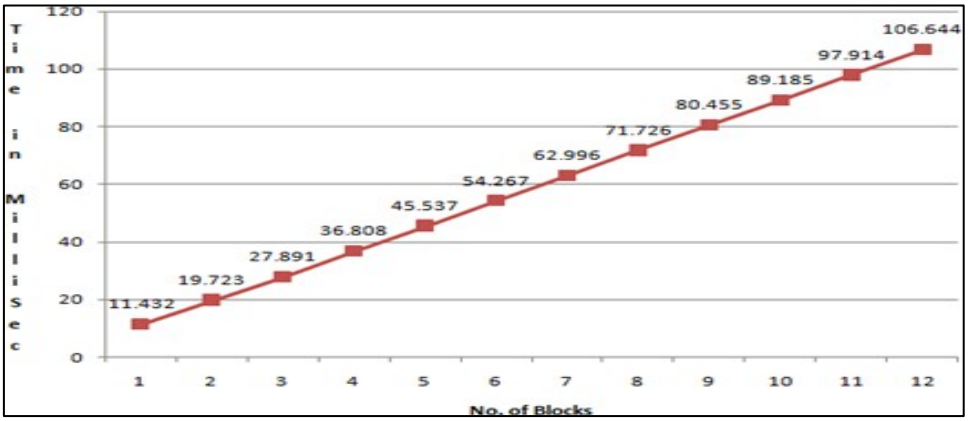

Figure 6:Time Requirements

CONCLUSION

Fortifying Cloud Storage: Safeguarding Sensitive Information in Outsourcing" concludes by presenting a reliable biometric authentication method for cloud computing. Our methodology enhances security in response to the growing reliance

on cloud services. Advanced encryption effectively secures sensitive data through collaboration between users, data owners, and CSPs. By implementing privacy-protecting features, our technology sets a benchmark for secure cloud storage. The results promise improved user convenience and enhanced data security, contributing to a safer digital ecosystem. Ongoing research and development in privacy-preserving and cloud security technologies will further strengthen consumer confidence in cloud-based services.

REFERENCES

[1] Al-Assam, Hisham, Waleed Hassan, and SheraliZeadally. "Automated biometric authentication with cloud computing." Biometric-based physical and cybersecurity systems (2019): 455-47.

[2] Gumaei, Abdu, et al. "Anti-spoofing cloud-based multi-spectral biometric identification system for enterprise security and privacy-preservation." Journal of Parallel and Distributed Computing 124 (2019): 27-40.

[3] L. Yang, C. Tian, G. Zhang, L. Li and H. Wang, "Efficient Biometric Identification on the Cloud With Privacy Preservation Guarantee," in IEEE Access, vol. 10, pp. 115520-115531, 2022, doi: 10.1109/ACCESS.2022.3218703.

[4] L. Zhu, C. Zhang, C. Xu, X. Liu and C. Huang, "An Efficient and Privacy-Preserving Biometric Identification Scheme in Cloud Computing," in IEEE Access, vol. 6, pp. 19025-19033, 2018, doi: 10.1109/ACCESS.2018.2819166.

[5] Liu, Chun, et al. "An efficient biometric identification in cloud computing with enhanced privacy security." IEEE Access 7 (2019): 105363-105375.9

[6] Rao, C. H. V., and M. Bhavani. "BIOMETRIC IDENTIFICATION." Journal of Engineering Sciences 14.04 (2023).

[7] Shahrukh, Mohammed, Rafi Hussain, and Abdul Rais Shaikh Farhan. "An efficient and privacy preserving biometrics identification scheme in cloud computing." J Eng Sci 10.1 (2019): 13-15.

[8] Venkatachalam, Chandrasekar, K. Manivannan, and Shanmugavalli Venkatachalam. "Securing Data in the Cloud: The Application of Fuzzy Identity Biometric Encryption for Enhanced Privacy and Authentication." International Conference on Image Processing and Capsule Networks. Singapore: Springer Nature Singapore, 2023.

[9] Xu, L. Zhang, L. Zhu, C. Zhang and K. Sharif, "Achieving Efficient and Privacy-preserving Biometric Identification in Cloud Computing," 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Shenyang, China, 2021, pp. 363-370, doi: 10.1109/TrustCom53373.2021.00063.

[10] Yang, X., Zhu, H., Wang, F. et al. MASK: Efficient and privacy-preserving m-tree based biometric identification over cloud. Peer-to-Peer Netw. Appl. 14, 2171–2186 (2021). https://doi.org/10.1007/s12083-021-01120-7

End