International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 08 | Aug - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

A Hybrid Cloud Approach for Secure Authorized Deduplication System Using Facial Authentication

V.KEERTHANA CSE& SVIST
J.Mangaiyarkarasi CSE&SVIST
A.N.Arun CSE&SVIST

Abstract - Over the use of peer-to-peer encryption, businesses are able to establish private communication channels between devices or parts of devices, ensuring that no intermediary devices are able to decipher important data as it travels over the network. To safeguard sensitive information like credit card details, PIN numbers, and passwords, privacy is essential. Rather of being handed down orally, this kind of knowledge may be exchanged digitally using messaging programs such as WhatsApp or Facebook Messenger.Most of these apps use what is known as peer-to-peer encryption, a kind of two-way communication in which no one other than the two parties involved can decipher the communications transmitted and received. It is necessary to remember the master password in order to access the password safe, which stores all the passphrases in one account. Other malevolent bots and hackers are prompted to plot an assault and attempt to breach the system by this strategy.Therefore, we provide an encryption technique in our suggested work that relies on the authentic user's facial landmarks. Since it is very difficult, if not impossible, to replicate all of the face characteristics in order to fool the authentication system, this method is extremely strong. To determine the legitimate user's identity, we use the Localised Binary Pattern method in conjunction with the Gabor Wavelet on the locality preserving projection model. Both the legitimacy on the peer-to-peer network and the encryption of data communicated across the network channel are supported by the face structure that has been extracted. As a result, this forms a very strong component for the authentication and encryption of sent communications.

***Key Words*:** peer-to-peer, system, encryption, authentication, breach, transmitted

I. .INTRODUCTION

Authentication of data has an inverse relationship with data redundancy. The likelihood of authenticating specific duplicate data decreases as redundancy grows. Data authentication is more likely to occur if data redundancy is minimised. Despite the meteoric rise of cloud computing and large data, deduplication has emerged as a popular issue in the last several years. Using the deduplication procedure, which prevents storing the same data multiple times in real time, greatly reduces the cost of cloud storage. Encrypting client data on the server enables secured deduplication. Clients must have faith in the service provider for it to be effective. Deduplication with security is often not supported by existing approaches. To guarantee deduplication while maintaining data security, we used biometric approaches in this study [1].

Same as In order to avoid having the same file again, data may be uploaded at several levels. One level is the file level, and another is the block level. Data deduplication is incompatible with many outdated encryption techniques, even when they provide data discretion. In particular, traditional encryption requires individuals with different encryption keys to encrypt their data. Therefore, it is not feasible to deduplicate distinct cypher messages by comparing data copies of odd users. The goal of convergent encryption is to make deduplication more likely while simultaneously enforcing data discretion. A Convergent key, obtained by calculating the cryptographic hash value satisfied with the copy, is used for encryption and decryption. Key maintenance and cloud-based cypher text transmission follow key manufacturing and data encryption. Once the encryption process is set up and derived from the data content, any two copies of the data that are identical will produce the same convergent key and, by extension, the same cypher text [2].

While data deduplication has many advantages, it also poses privacy and security risks as users' sensitive information is accessible to both insiders and outsiders of the cloud infrastructure system. Because several users would produce different ciphertext for the same data using these approaches, deduplication cannot take place, typical data encryption methods will not be able to handle this circumstance. These issues are being addressed by using convergent encryption algorithms [3].

One authorised de-duplication approach is differential authorisation duplicate check, which involves assigning a set of rights to each user when the system is initialised. This permission set defines which types of users may access files and run duplication checks [4].

An individual's biological traits are what make them special. Impersonation attacks, in which a malicious actor pretends to be a legitimate user in order to get access to protected resources, are a common problem with authentication systems. This might be one way to address this issue. A large number of studies have addressed authentication-related topics. To the best of our knowledge and the literature, this study is the first to attempt to highlight the fact that users' biometric uniqueness can be utilised as an authentication factor with no additional requirements other than the factory-fitted System Accessories [5].

High-availability storage and cheap, highly-parallel computing resources are both offered by modern cloud service providers. More and more people are storing data on the cloud and sharing it with other users who have certain rights that determine who may access it. This trend is driven by the widespread use of cloud computing.

The fundamental idea behind this project is that we need privacy in order to safeguard sensitive information like credit card details, PINs, and passwords. Rather of being handed down orally, this kind of knowledge may be exchanged digitally using messaging programs such as WhatsApp or Facebook Messenger.

II. .LITERATURE SURVEY

To safeguard data security, Jin Li et al. [6] suggested authorised data deduplication, which incorporates user permission differentials into the duplicate check. In addition, we introduced a number of novel deduplication structures that bolster authorised duplicate check in hybrid cloud architecture, whereby a private cloud server uses private keys to produce file duplicate-check tokens. In regards to the insider and outsider assaults outlined in the suggested security model, security analysis proves that these techniques are safe.

Within the context of a two-level multi-domain architecture, Xue Yang et al. [7] offered a huge data deduplication technique that was both efficient and respectful of users' privacy. During the process of multi-domain deduplication, this technology not only safeguards the confidentiality of data but also withstands brute-force attacks. This is because it creates a random tag and a predetermined number of random ciphertexts for each and every file. By restricting the capacity to conduct intra-deduplication to the agent and inter-deduplication to the cloud service provider, our method may be able to protect the message equality information from being disclosed or disclosed to a certain level.

An effective approach of reducing data transferred over the network and stored in data systems was developed by Laura Conde-Canencia et al. [8] using data deduplication algorithms and models. Focussing on the scenario where duplicate files are created due to editing mistakes, we use a theoretical approach to investigating data files. Data storage types such as main, backup, and archiving may all benefit from our research. Our novel variable-length block-level deduplication approach decreases computational cost by concentrating on pivots and outperforms previous work.

In order to meet the requirements of shared-nothing storage systems, AWAIS KHAN et al. [9] suggested GRATE, a high-performance inline cluster-wide data deduplication. Specifically, GRATE ensures excellent storage space economy without compromising speed by removing duplicate copies across the cluster. We use a distributed deduplication metadata shard that effectively handles duplicate fingerprint lookup I/Os and high-performance deduplication metadata without adding any single point of failure. Using the content fingerprint of chunks, cluster-wide data and deduplication metadata placement is done.

Xue Yang and colleagues [10] proposed the idea of a user-defined access control–supporting, energy-efficient, and secure deduplication system. By allowing only the cloud service provider to approve data access on behalf of data owners, our method has the potential to reduce duplication to the greatest extent possible without compromising the safety and privacy of cloud users. Furthermore, according to the findings of our security study, our approved safe deduplication approach reliably preserves data confidentiality and guarantees tag consistency, all while being resistant to brute-force attacks. Also, according to exhaustive simulations, our solution excels in the areas of computational overheads, communication overheads, storage overheads, and deduplication efficiency. All of these areas are important.

III.PROPOSED METHOD

The two most important things that cloud computing and the Internet of Things (IoT) do are safeguard user privacy and promote open sharing. When people normally share data with one another, it's so they can upload it to a cloud server and keep it there, which solves their internal storage problems. Both data owners and data consumers have complicated concerns about the security and privacy of stored data in the typical data exchange procedure. Therefore, the suggested method of data integrity verification based on blockchain is created to counteract the downsides. System and user authentication processes are made more secure by the suggested technique. To improve system security, the present research makes use of blockchain technology.

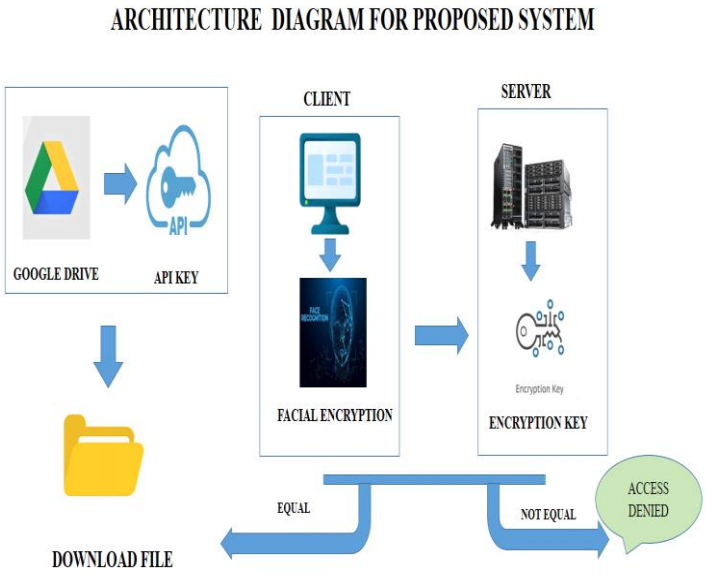

Fig 4.1 Architecture Explanation.

As part of our planned research, we provide an encryption method that uses the unique facial features of a genuine user. Since it is very difficult, if not impossible, to replicate all of the face characteristics in order to fool the authentication system, this method is extremely strong.To determine the legitimate user's identity, we use the Localised Binary Pattern method in conjunction with the Gabor Wavelet on the locality preserving projection model.

Both the legitimacy on the peer-to-peer network and the encryption of data communicated across the network channel are supported by the face structure that has been extracted. As a result, this forms a very strong component for the authentication and encryption of sent communications.

As part of our planned research, we provide an encryption method that uses the unique facial features of a genuine user. Since it is very difficult, if not impossible, to replicate all of the face characteristics in order to fool the authentication system, this method is extremely strong. To determine the legitimate user's identity, we use the Localised Binary Pattern method in conjunction with the Gabor Wavelet on the locality preserving projection model.

Both the legitimacy on the peer-to-peer network and the encryption of data communicated across the network channel are supported by the face structure that has been extracted. As a result, this forms a very strong component for the authentication and encryption of sent communications. Biometric authentication : We employ advanced COTS matchers to compare fingerprint and facial images. Due to licencing restrictions, we cannot publish the vendors' names, however these matchers score in the top three among NIST fingerprint and face evaluations.The Levenshtein distance between two strings is the shortest one-character action required to modify them. The Damerau-Levenshtein distance could help transpose neighbouring letters. Editor distance is like Levenshtein distance, however inserting and deleting are separate. Before normalising in the [0, 1] range, the edit distance is split by the highest conceivable edit distance between two strings of the same lengths as the given pair [11].

IV.RESULT AND DISCUSSION

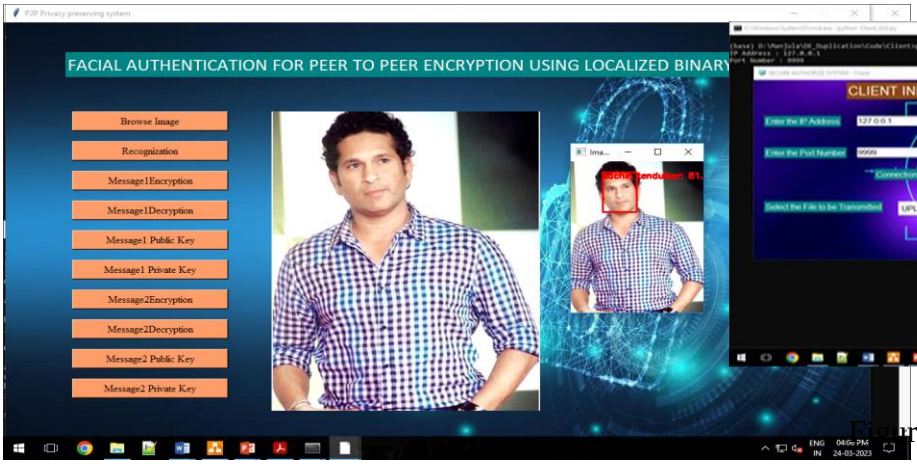

Figure 2. Facial recognition system for Data authentication

The figure 2 shows a GUI with LBP for recognition that is intended for face authentication in peer-to-peer encryption. A menu on the left side for secure communication, including "Browse Image," "Recognition," "Message/Encryption," "Message/Decryption," and public key management. A smaller pane on the right verifies a match with the caption "Matched Face: 97%," showing a high degree of confidence in the identification process. The primary display has a centre picture of a human, is used for facial recognition. Using LBP as a method for analysing face characteristics, the program purports to authenticate users using facial recognition to provide private and secure peer-to-peer communication.

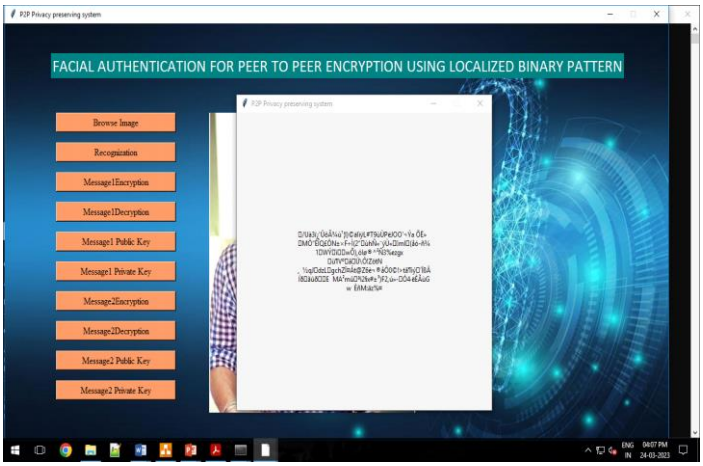

Figure 3. Graphical User Interface of the proposed system

An example of a P2P privacy-preserving system is shown in the figure 3. This system displays encrypted text or a ciphertext message. A communication has been encrypted by the system, and the user's face recognition has been used as a method of authentication in order to get access to the encryption or decryption operations. Not only does the program verify users via the use of face recognition, but it also makes it possible to encrypt and decode secure messages, so guaranteeing that communication is conducted in a manner that is both private and confidential.

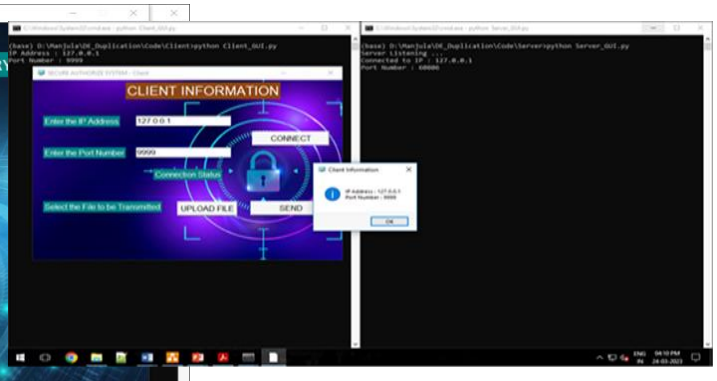

Figure 4. Client Information page

The above figure shows client information page of the above application. Here the details to be uploaded are IP address, Port number, and connect to the server. Once connection processes transfer file to be send.

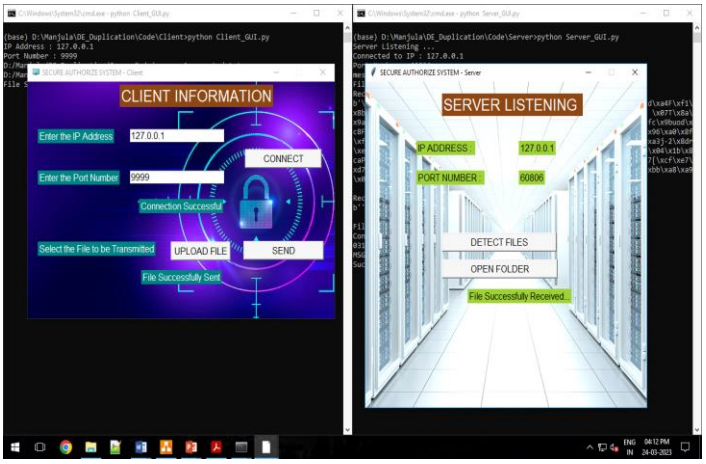

Figure 5. Client server connection

The figure 5 shows file transfer between client and server. Once connection is processed the server starts listening for any file from client. Detected files are encrypted and saved to the folder in an encrypted format.

V.CONCLUSION

To safeguard data security, this project presented authorised data deduplication, which incorporates user permission differentials into the duplicate check.Using private keys on a private cloud server, this project implements a number of novel deduplication constructs that enable authorised duplicate check in a hybrid cloud architecture.

Within this project, we execute testbed tests on a prototype of our proposed authorised duplication check mechanism to demonstrate its feasibility. Our authorised duplication check approach has negligible overhead when compared to convergent encryption and network transmission, as shown in this project.

Looking ahead: It does not take into account potential security issues with implementing the current approach in practice. The nation's safety is also enhanced. By removing unnecessary data duplication, it frees up RAM and gives us more than enough. It safeguards critical information while granting permission to private companies.

REFERENCE

[1]. Venkatachalam, K., Prabu, P., Almutairi, A., & Abouhawwash, M. (2021). Secure biometric authentication with de-duplication on distributed cloud storage. *PeerJ Computer Science*, *7*, e569.

[2]. Devi, B. R., Prakash, C. B., & Narayana, G. L. (2019). An Enhanced Approach for Securing Authorized Deduplication in Hybrid Clouds.

[3]. Meshram, T. B., & Deshmukh, S. Towards Security and Authorization Based Data Deduplication Using Hybrid Cloud.

[4]. Patil, M. A. V., & Kale, M. N. D. (2014). Survey on secure authorized de-duplication in hybrid cloud. *International Journal on Recent and Innovation Trends in Computing and Communication*, *2*(11), 3574-3577.

[5]. Otta, S. P., Panda, S., Gupta, M., & Hota, C. (2023). A systematic survey of multi-factor authentication for cloud infrastructure. *Future Internet*, *15*(4), 146.

[6]. Li, J., Li, Y. K., Chen, X., Lee, P. P., & Lou, W. (2014). A hybrid cloud approach for secure authorized deduplication. *IEEE transactions on parallel and distributed systems*, *26*(5), 1206-1216.

[7]. Yang, X., Lu, R., Shao, J., Tang, X., & Ghorbani, A. A. (2018). Achieving efficient and privacy-preserving multi-domain big data deduplication in cloud. *IEEE Transactions on Services Computing*, *14*(5), 1292-1305.

[8]. Conde-Canencia, L., & Hamoum, B. (2020, July). Deduplication algorithms and models for efficient data storage. In *2020 24th International Conference on Circuits, Systems, Communications and Computers (CSCC)* (pp. 23-28). IEEE.

[9]. Khan, A., Hamandawana, P., & Kim, Y. (2020). A content fingerprint-based cluster-wide inline deduplication for shared-nothing storage systems. *IEEE Access*, *8*, 209163-209180.

[10]. Yang, X., Lu, R., Shao, J., Tang, X., & Ghorbani, A. A. (2020). Achieving efficient secure deduplication with user-defined access control in cloud. *IEEE Transactions on Dependable and Secure Computing*, *19*(1), 591-606.

[11]. Sudhish, P. S., Jain, A. K., & Cao, K. (2016). Adaptive fusion of biometric and biographic information for identity de-duplication. *Pattern Recognition Letters*, *84*, 199-207.