Secure biometric authentication with de-duplication on distributed cloud storage

M Vinoth Kumar [1] , **K Venkatachalam** [2] , **P Prabu** [3] , **Abdulwahab Almutairi** [Corresp., 4] , **Mohamed Abouhawwash** [5]

[1] Department of Computer Science and Engineering, Anna University,University College of Engineering Dindigul, Dindigul, Tamilnadu, India

[2] School of Computer Sceince and Engineering, VIT Bhopal University,Bhopal, Bhopal, Madhya pradesh, India

[3] Department of Computer Science, CHRIST (Deemed to be University), Bangalore, Bangalore, Karnataka, India

[4] School of Mathematics, Unaizah College of Sciences and Arts, Qassim University, Saudi Arabia, Saudi Arabia

[5] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt. Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA., Mansoura University, Mansoura, Mansoura, Egypt

Corresponding Author: Abdulwahab Almutairi
Email address: almutairi2017@qu.edu.sa

Cloud Computing is one of the evolving fields of technology, which allows storage, access of data, programs, and their execution over the internet with offering a variety of information related services. With cloud information services, it is essential for information to be saved securely and to be distributed safely across numerous users. Cloud information storage has suffered from issues related to information integrity, data security, and information access by unauthenticated users. The distribution and storage of data among several users are highly scalable and cost-efficient but results in data redundancy and security issues. In this article, a biometric authentication scheme is proposed for the requested users to give access permission in a cloud-distributed environment and, at the same time, alleviate data redundancy. To achieve this, a cryptographic technique is used by service providers to generate the bio-key for authentication, which will be accessible only to authenticated users. A Gabor filter with distributed security and encryption using XOR operations is used to generate the proposed bio-key (biometric generated key) and avoid data deduplication in the cloud, ensuring avoidance of data redundancy and security. The proposed method is compared with existing algorithms, such as convergent encryption (CE), leakage resilient (LR), randomized convergent encryption (RCE), secure de-duplication scheme (SDS), to evaluate the de-duplication performance. Our comparative analysis shows that our proposed scheme results in smaller computation and communication costs than existing schemes.

SECURE BIOMETRIC AUTHENTICATION WITH DE-DUPLICATION ON DISTRIBUTED CLOUD STORAGE

M Vinoth Kumar[1]**, K Venkatachalam**[2]**, P. Prabu**[3]**, Abdulwahab Almutairi**[4]**, and Mohamed Abouhawwash**[5]

[1]**Assistant Professor, Department of Computer Science and Engineering,Anna University, University College of Engineering Dindigul, Dindigul,Tamilnadu,India**

[2]**Assistant Professor(Senior),School of Computer science and Engineering,VIT Bhopal,Bhopal.**

[3]**Department of Computer Science, CHRIST (Deemed to be University), Bangalore, India**

[4]**School of Mathematics, Unaizah College of Sciences and Arts, Qassim University, Saudi Arabia**

[5]**Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.**

[5]**Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824 USA.**

abouhaww@msu.edu

Corresponding author:
Abdulwahab Almutairi [5]

Email address: almutairi2017@qu.edu.sa

ABSTRACT

Cloud Computing is one of the evolving fields of technology, which allows storage, access of data, programs, and their execution over the internet with offering a variety of information related services. With cloud information services, it is essential for information to be saved securely and to be distributed safely across numerous users. Cloud information storage has suffered from issues related to information integrity, data security, and information access by unauthenticated users. The distribution and storage of data among several users are highly scalable and cost-efficient but results in data redundancy and security issues. In this article, a bio-metric authentication scheme is proposed for the requested users to give access permission in a cloud-distributed environment and, at the same time, alleviate data redundancy. To achieve this, a cryptography technique is used by service providers to generate the bio-key for authentication, which will be accessible only to authenticated users. A Gabor filter with distributed security and encryption using XOR operations is used to generate the proposed bio-key (biometric generated key) and avoid data duplication in the cloud, ensuring avoidance of data redundancy and security. The proposed method is compared with existing algorithms, such as convergent encryption (CE), leakage resilient (LR), randomized convergent encryption (RCE), secure de-duplication scheme (SDS), to evaluate the de-duplication performance. Our comparative analysis shows that our proposed scheme results in smaller computation and communication costs than existing schemes.

1 INTRODUCTION

Data redundancy is indirectly proportional to the data authentication. When redundancy increases, then possibility of authentication of certain redundant data is very less. If data redundancy is reduced, possibility of data authentication is high. Deduplication is hot topic in recent years, inspite of rapid growth in cloud computing and big data. Cost of cloud storage is highly reduced by using deduplication process which avoids storing of same data at multiple times in real time. Secured deduplication is provided by encrypting client data in server. It must bring confident among client to believe service provider. Usually traditional techniques does not support deduplication with security. In this article, we implemented biometric techniques to ensure deduplication with data security.

1.1 Cloud authentication issues:

The advancement in data sharing and processing over the cloud makes consequence of innovative technologies like smart mobile devices, mobile applications with sensors, Internet spread, and usage, data availability in social media. These technologies make wider influences on big data in our day-to-day activity. Many organizations like Amazon, flip kart, Netflix performs data collection, mining, and analysis from various sources. Sharing of large volumes of data over the network has been made easy for accessible through cloud storage. The increasing need for storage disks over the network looks for authentication of stored data has resulted in security concerns in the cloud and distributed storage. Large amounts of storage in cloud infrastructure are occupied by duplicate data records. Aiming to address these technical concerns, researchers have focused on techniques for data de-duplication using biometric de-duplication with user authentication. The link between intrinsic individual characters with their behavior, physical, physiological is used to authenticate an individual with biometric data recognition. While comparing with knowledge-based authentication, biometric can provide stronger security guarantees. Building a biometric-enabled technology on the cloud is important for safety as well as security enhancement. The security is provided to areas such as forensics, surveillance, defense, banking, and personal authentication. Further biometric-based authentication process has been proven to provide stronger security guarantees and robustness in contrast to traditional methods for sensitive applicationsAh Kioon et al. (2013)

1.2 Cloud data redundancy issues:

Data de-duplication is the process of avoiding redundant data copies reducing the overhead by eliminating duplicate stored data. Today's world is hyper-connected with online communication, payments, ticket services, using managed or unmanaged networks, and devices scaling across several endpoints. These devices are currently being protected with security technologies and enabled with cryptographic single factor, two factors, multi-factor authentication methods Alomar et al. (2017). Human in every communication uses multi-factor methods for fast, reliable, and user-friendly authentication while accessing online services. The digital age of information allows the distribution and replication of data across the network. During this process, the same data may be shared, circulated, stored multiple times. This highlights the need for smart technologies to tackle the challenge of data deduplication along with authentication methods for users to access the data.

1.3 Cloud data authentication for deduplication:

This smart connected world makes the data secure with users through authentication processes Alomar et al. (2017). A user needs to identify himself in the system by sending authentication messages. When message 'A' is sent by the user, the system computes F(A) randomly and checks with stored data 'B'. A single authentication password alone cannot ensure the authentication of the user Benarous et al. (2017)Mohsin et al. (2017). Accessing sensitive data offline or online requires a fundamental security system for authentication Mohsin et al. (2017),Balloon (2001),Ometov et al. (2018) (Figure 1). Traditional authenticated transactions like applying seals, wax seals are physical security systems Konoth et al. (2016). Sender-based information validation alone cannot provide standard authentication Ibrokhimov et al. (2019). Figure 1 shows how the technology is evaluated from single security techniques to multiple security techniques.

Initially, a single data factor was used for authentication, which was eventually compromised by the research community Kim and Hong (2011),Dasgupta et al. (2016). Examples of single-factor security are user ids and passwords. Password authentication is considered the weakest level of security Bonneau et al. (2015),Wang and Wang (2015). Sharing of security information like a password can easily lead to
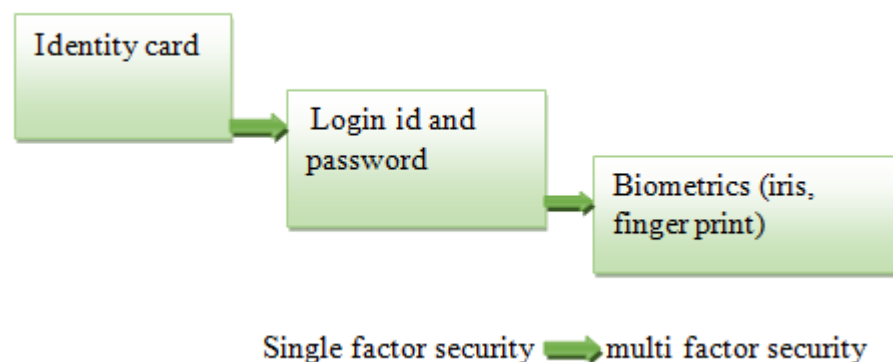
Figure 1. Security development stages

compromised security. Several attacks have occurred by unauthorized entities like dictionary hacking Ah Kioon et al. (2013), rainbow attacks Heartfield and Loukas (2015), and other social engineering attacks Grassi et al. (2016). When users choose password-based authentication, the complexity of the authentication must be ensured Gunson et al. (2011). High protection of the accounts cannot be ensured by using a single authentication factor Sun et al. (2014). The next level involves a two-factor authentication, which is achieved by an identity or security question Bruun et al. (2014),Harini and Padmanabhan (2013). At present, three categories of groups are available for connecting individuals with security credentials Scheidt and Domangue (2006):

- Ownership authentication – requires ID cards, smartphones.

- Knowledge authentication –requires passwords, secret keys.

- Biometric authentication – biometric data (fingerprints, iris scans, face scans)

Multi-factor authentication provides and ensures higher safety levels, with two or more credentials Bhargav-Spantzel et al. (2007),Banyal et al. (2014),Council and Committee (2010). Biometrics is widely used in the multi-factor authentication process based on individual biological and behavioral characteristics Huang et al. (2014). The higher level of security is offered explicitly by biometrics recognition using more security factors Tahir and Tahir (2008) and the evolutionary history of the authentication is described in Figure 2.

High-security infrastructures utilize multiple authentication factors for protecting the information. For example, ATM cash withdrawal has a combination of ownership patterns (card) which is accessed by using knowledge factors, such as PIN, to transfer money and manage accounts Coventry et al. (2003),Ometov et al. (2018). To make this system even more robust, a card with a PIN is further authenticated using a one-time password, while accessing the sensitive data Aloul et al. (2009). Facial recognition methods are also suggested for user authentication purposes Ometov et al. (2018). In a recent survey, it is also pointed out that, most business enterprises choose multi-factor-based security and authentication for the secure processing of transactions. At present, most enterprises use biometric systems for employee verification, bank management, lockers, vehicles, etc., making the multi authentication factor stronger and interesting, which paved the way to introduce innovative biometrics, among the research community.

Most of the electronic devices use multiple authentication techniques for security-based access and allow only authenticated owners to use the devices Symeonidis et al. (2016). One such potential application could be the usage of biometrics in a vehicle to authenticate its owner. With respect to market-based applications, authentication techniques can be broadly categorized as commercial, governmental, and forensic applications. Commercial purposes may include account access and ATM banking. Government needs may include document identification, government ID process, social security applications, military, border security controls. Lastly, forensic applications may include investigations, evidence identification, criminal identification Liu et al. (2018),Nor et al. (2015),Grigoras (2009).
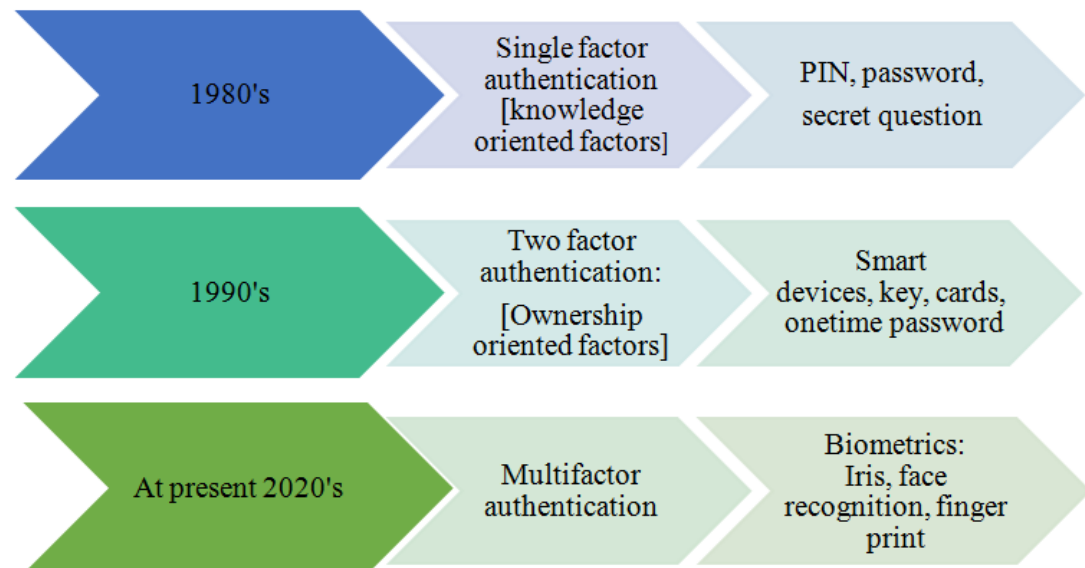
Figure 2. Evolution of various authentication levels

One of the challenges in multi-factor authentication techniques is the association between users and the sensors being used [51]. Regarding security, this relationship must be established so that only authorized users, authenticated in advance, should receive access rights. In this context, the use of biometrics (e.g., fingerprints, face recognition) is considered a user-friendly technique.

1.4 Contribution

In this research work, a multi-factor authentication technique with biometrics is proposed for the verification of users in cloud environments. The bio key of the data owner is first generated for the data stored in the cloud. Contributions on security and data redundancy are addressed in this paper through the following components:

1. A multi-factor authentication technique is used for bio key generation. Finger print of owner is processed for selecting appropriate features using edge detection with hashing function. Bio key generated based on extracted feature set.

2. We newly introduce bio key from data owner's fingerprint. It is chosen for the generation of bio keys which helps data to be shared with owners knowledge.

3. The encrypted data is stored in the cloud so that the data can be accessed by the user only when the bio key shared for the authentication process is satisfied.

4. The redundancy of stored data is eliminated while keeping data security in mind.

This article is organized into 5 sections. Section 2 presents a literature survey and section 3 presents the proposed methodology and its design. Section 4 describes the evaluation results of the proposed method and finally, section 5 concludes our paper and discusses future work.

2 LITERATURE REVIEW

[Kathrineet al, 2017] proposed a secure biometric authentication scheme for user identification and mutual authentication using elliptic curve cryptography for key generation and key exchange with optimal communication cost. [Farhana et al,m 2017] proposed a cloud-based mobile biometric authentication framework (BAM Cloud) using dynamic signatures and user authentication. The data is captured with a handheld mobile device and subsequently, storage, preprocessing, and training are done in a distributed manner on the cloud. The proposed method was implemented using Map Reduce on the Hadoop platform

and for training a Levenberg-Marquardt backpropagation neural network model was used, achieving a speedup of 8.5x and an accuracy of 96.23

Al-Assam et al. (2019) surveyed various biometric-based authentication methods in cloud environments. The traditional password-based authentication lacks security when it comes to cloud data. To this end, multifactor authentication is suggested which allows two or more authentication parameters along with password-based security. This review focuses on the various available biometric authentication models and their advantages and disadvantages.

Wong and Kim (2012) provided the concepts used in biometric-based authentication in cloud computing. They discussed the challenges and limitations of traditional methods, along with attack scenarios, such as misuse of biometric data, to track individuals and leak confidential information related to health, gender, ethnicity, etc. At the same, they also argued that the privacy of cloud-based biometric authentication is not going to resolve the authentication issues technically, rather new legislation to enforce privacy-aware measures on cloud service providers related to the biometric collection, data processing, and template storage is opined, leading secure authentication in cloud environments.

Dulari and Bhushan (2019) proposed a authentication method for user in cloud computing based distributed environment. In this process users biometric information are stored as template in cloud server. Further user verification is done with several participants. Here users feature vector query is compared with template saved in cloud server. In this method, homomorphic based encryption is used for matching the protocol. This matching protocol helps to compare the queried vector with template available as encrypted file . The metrics used for measuring output are Square of Euclidean distance, sensitive preservation of information and processing of authentication with high security .

Dulari and Bhushan (2019) proposed a security system called TORDES for cloud storage. This work uses legion containers in cloud storage. The data is stacked in TORDES for authentication with crypto-biometric systems to avoid unauthorized access.Indu et al. (2018). surveyed the different biometrics applied on cloud security issues to identify malware. Ziyad and Kannammal (2014) analyzed cloud security and threat possibilities. They analyzed several security mechanisms and suggested the robust ones for both academic and industry environments. Ziyad and Kannammal (2014) proposed a multifactor biometric authentication system for cloud computing. The biometric features considered in this work were palm vein and fingerprints, where palm vein biometric data was stored in multicomponent smart cards and fingerprint data in the central database of a cloud server.

Zahrouni et al. (2017) developed an application that serves as an extra layer of security on top of a pre-existing banking application by using a Biometric lock application, which allows a user to add a layer of security to their credit and debit cards, at the expense of minimal time overhead. Malathi and Raj.R (2016) proposed a biometrics-based user identification scheme using the features of a palm print, fingerprint, and iris for providing accurate personal identifications.

Wang et al. (2018) carried over Amin et al. (2018) protocol as a case study to provide ideas for designing secure protocols for cloud environments in order to overcome existing security weaknesses in the protocol. They further improved the protocol using BAN logic and also used heuristic analysis to prove the security of the protocol. Obergrusberger et al. (2012) proposed biometric observer techniques and provides a idea on fundamental trust rules in web as a prototype for implementation . The enrollment of users are supervised by an observer. further observers are those who enhance to provide authentication to biometric template. It provides more trustworthiness model for biometric identities. strong trust is build between observers and other due to best relation between both observers and individuals observed in the Database of system . Chandramohan et al. (2017) proposed a privacy-preserving model to prevent digital data loss in the cloud, helping the cloud requester/users) to trust their proprietary information and data stored in the cloud. Table 1 below represents the recent deduplication research problems and their methodology to overcome the identified problem. It is noticed that biometric based deduplication and authentication of data is not much used recently. Our proposed work shows advanced deduplication handling process with high security.

Shabbir et al. (2021) noticed profits of Mobile Cloud Computing (MCC) in medical healthcare. It faces more challenges in security and privacy of customer data. Here they implement layered security modeling using Modular Encryption Standard (MES) to increase the security of MCC. The performance is better than other encryption techniques. Rehman et al. (2021) addressed invehicle communication problems using controller area network and electronic control units. Security during communicating inside the vehicle is tackled using novel approach CANintelliIDS. It detects the vehicle intrusion attack

Table 1. Literature Survey of Existing Techniques

Authors	Problem	Methodology	Advantages	integrity	confidentiality
Youngjoo Shin et al [62] (2020)	Data deduplication and security in MEC	server less efficient encrypted deduplication (SEED) +Lazy encryption	It takes 1000ms for processing128MB	No	yes
Shangping wang et al [63] (2020)	Double payment avoidance and data deduplication	Block chain technology	Avoids third party	yes	yes
Jiaojiao Wu et al [64] (2020)	File deduplication and integrity	Confidentiality preserving deduplication using public auditing (CPDA)	Computation of authentication tag is done by CSP	No	Yes
Wenting shen et al [65] (2020)	Deduplication and brute force dictionary attacks	Light weight cloud storage auditing for deduplication	light-weight computation on the user side	yes	No
Xueyan liu et al [66] (2020)	Data deduplication in files	Verifiable ABKS over encrypted cloud data is proposed	realize indistinguishable keywords, unforgetability of signature and confidentiality of cipher texts	No	Yes
Jianli Bai et al [67] (2020)	cloud storage auditing and deduplication literatures fail to support the modifications of ownership	re-encryption algorithm and the secure identity-based broadcast encryption technology	ownership modification and integrity is maintained	Yes	No
Shynu P. G. et al [68] (2020)	Data deduplication	Modified Elliptic Curve Cryptography (MECC) algorithms	Recognize data redundancy at the block level	No	Yes

across controller area network. at result it gains 10.79Naeem et al. (2021) discusses energy efficient WSN for high performance network lifetime. In this article hybrid technique called Distance aware residual energy-efficient stable election protocol is used with energy efficient election protocol for optimal transmission routes. In this outcome energy efficiency is increased for 10

3 PROPOSED METHODOLOGY: BIO KEY WITH GABOR –XOR

In distributed cloud computing, biometric-based authentication plays a vital role in current research. Distributed denial of service is the main threat in the cloud nowadays: several users trying to access a single cloud server leads to an increase in response time and complicates security. There are several methods to solve these issues in the cloud, even though there is a lack of confidentiality, reliability, and consistency of the data. To resolve that, we propose an approach called secure biometric authentication on distributed storage in the cloud. Figure 3 shows the overview of the proposed architecture. The owner's biometric information is recorded for authentication. Once the owner registration is complete, the data is encrypted using a distributed approach and is stored in the cloud. While the user tries to access the content, a cloud server authenticates the user validity and contacts the data owner for the bio key to access the data. In the proposed architecture, the data is stored securely and with the owner's reference, users can access the content avoiding duplicate copies of the same content and enhancing security through the bio key. The flow of the proposed design can be described as:

- The owner of the data can upload content onto the cloud that is encrypted through a distributed model.

- The duplication of the content can be avoided by the cloud server. The features of the fingerprint of the owner are extracted and stored in a database so that the original content can be referred by the user with the owner's permission.

- Finger print of the user is converted in to bio key using edge detection and hash functions.

3.1 Biometric based authentication with de-duplication

In this work, a fingerprint-based authentication along with a hash-based deduplication approach is used. Among the various biometric techniques, fingerprints have been widely accepted and implemented for secure authentication. The input fingerprint image of the owner is normalized before processing and then the features of the fingerprint image are extracted using an optimized self-learning method and are stored in a database for authentication. Fig 3 shows the biometric authentication process. A Gabor filter-based technique is used to enhance image equality by removing the noise. Subsequently, the enhanced image is
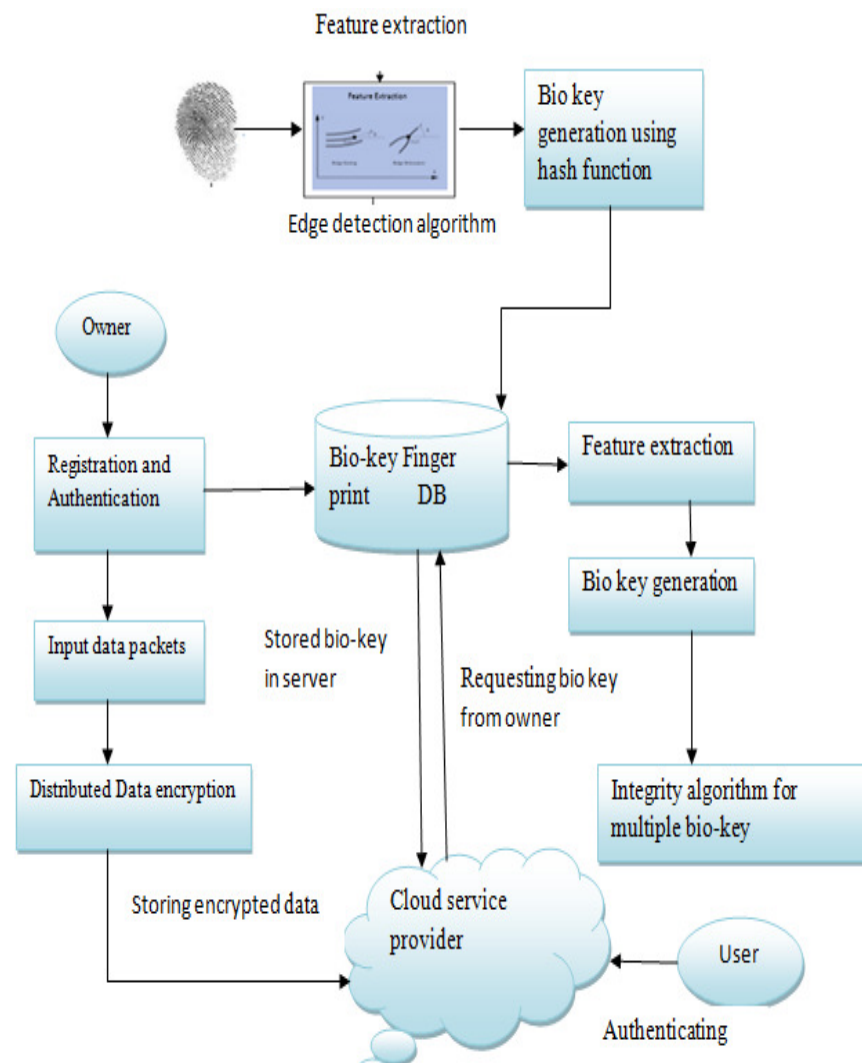
Figure 3. Overview of the proposed architecture

242 made ready for feature extraction Hur et al. (2016). The ridge endings and ridge bifurcation are extracted
243 as features using an edge detection algorithm then the hash functions of these features are considered as
244 bio-keys, which are stored in a database for authentication and de-duplication.

Algorithm 1: Biometric authentication with de-duplication

Result: Bio key (QR-code)
Input: Input fingerprint image;

Step 1: read the input image with the pixels as I(x,y);
Step 2: The input image can be preprocessed (e.g., binarization, thinning) and enhanced using the
following equations. During binarization, the grey level image is converted into a black and
white image, where black represents the ridges and white represents the valleys. Then the
Binarized image is thinned using three conditions. Thinning is the process of transforming the
ridge pixels into one pixel:;

$$I(x,y) = \frac{1}{|w||h|} \sum_{i=0} \sum_{j=0} (x,y) \tag{1}$$

$$g(x,y;\theta,f) = exp\left\{-\frac{1}{2}\left[\frac{x\theta}{\sigma^2}\right] cos(2\pi fx^2))\right\} \tag{2}$$

Where, $x_\theta = xcos_\theta + ysin_\theta$ and $y_\theta = -xsin_\theta + ycos_\theta$, $\theta = Orientation$, f=frequency and
σ_x, σ_y=standard deviation of gausian envelop;
Step 3: The features like ridge end and bifurcation are extracted using the following equation;

245

$$\sum_{i=0}^{7} N_i = thenridgeend \tag{3}$$

$$\sum_{i=0}^{7} N_i > 2thenbifurcation \tag{4}$$

Where $N_0, N_1..N_7$ are the eight neighbors of the pixel (x, y) of Image I. ;
Step 4: the binary pattern code of the processed image is generated as follows;

$$BC(g(x_c,y_c)) = \sum_{i=0}^{7} f(gx_i) - g(x_c)x2^n \tag{5}$$

Where, f (m) = 1 for $m >= 0$ and f (m) =0 for $m < 0$;
Step 5: The bio key for each binary pattern is generated with QR decomposition formula [2];
Step 6: De-duplication using a hash-based technique of the key as;

$$KH_i = hash(KH_i|G_k), KH_j \tag{6}$$

Where G_k- Gaussian random number.;
Step 7: Store the bio key of each owner into the cloud data base and content storage using
algorithm 2.

---

246 ## 3.2 Security based – distributed storage encryption
247 This proposed algorithm is taken from Kumar and Begum (2011), where the input data packet is di-
248 vided into two substrings. The substrings are further processed and then merged to store in the cloud

---

249 server. The two components of the data packets are considered as X , Y and Z is the random number.
250 To encrypt the data packet, an data packet an XOR operation is performed and the data is encrypted
251 that is sent to the cloud server. This approach is shown in fig 4. Hence, our proposed biometric
252 security-based- distributed storage encryption with a de-duplication approach can obtain secure data
253 transfer between the user and the cloud without duplication. This proposed approach can avoid du-
254 plication of content by sharing the key to the user for access. This will leads to reduce the storage
255 of the cloud. One method can prove the secured authentication using biometric and de-duplication.

Algorithm 2: P

Result: Encrypted data packet
Input: Data packet with name label NL and pre-defined label PL;

Step 1: The name label of the data packet is { n_1 ,n_2, , n_l } and pre-defined label are declared as {
p_1, p_2,..., p_l };
Step 2;
while *NL* **do**
    while *each packet of the input data* **do**
        if *lable <> PL* **then**
            Initialized and J,γ, δ = 0;
            Generate key *k* at random;
            if *(I&&Z! = 0)* **then**
                J=Z-I;
                γ=I XOR k;
                δ=J XOR k
            end
        else
            Encrypt the data packet using XOR operation;
        end
    end
end
Step 3: Output the encrypted data packet. That will store into the cloud server;
Step 4: De-duplication: algorithm 1 and algorithm 2 are combined.;
Note: The step 4 combines two algorithms and computes when user request for data, cloud server
  authenticate it and search for owner of the data. Finally it shared the bio key to the user to access
  the content. ;
seudo code : 2 (Distribute storage – security based encryption)

256

4 EXPERIMENTAL RESULTS AND DISCUSSIONS

258 This proposed work experimented in AWS cloud services. Different type of services of AWS is initiated
259 to each owner and user for a secure authentication process. The proposed work is implemented in three
260 stages. The first stage is to generate the bio key of the owner of the data before upload. In the second
261 stage, the hash function of this key is used for the de-duplication process. The third stage is the encryption
262 of the data that are uploaded to the cloud server. Fig. 5 represents the first stage of the proposed work. In
263 figure 5, (a) represents the original input fingerprint image of the owner and (b) and (c) represents the
264 equivalent orientation and frequency images of the input image. This is shown in Figure 6.
265   **Illustration of Figure 6:** It illustrates the Binarized form of the input image to identify the ridges
266 and valleys. Then the Binarized image is thinned into a single pixel to identify the features such as ridge
267 ending and bifurcation which is shown in fig 6 (c). These variations of the fingerprint patterns are used to
268 build the bio key with QR code as shown in Fig 6 (d).
269   The comparative analysis of time is shown in Fig 7. The data redundancy makes memory engaged
270 and takes more time to upload, download the data. While we upload, redundancy makes no space for
271 new data and resource allocation takes more time. Next, while we download the data, due to redundancy
272 system confuses which data to access. Memory space also very less. This makes more time to download
273 the data from a cloud server.Also, Figure 8 shows time taken by proposed model in data encryption and
274 decryption. To evaluate the performance of this fingerprint authentication de-duplication, the proposed
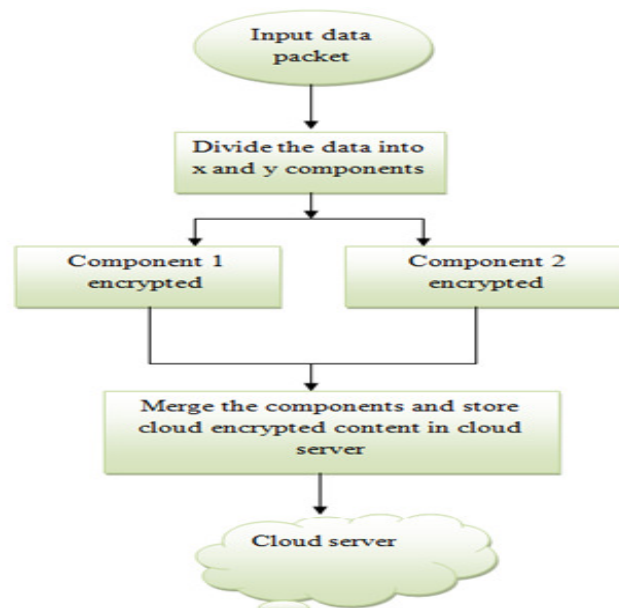
Figure 4. Security based distributed storage encryption

(a) 1a   **(b)** 1b   **(c)** 1b
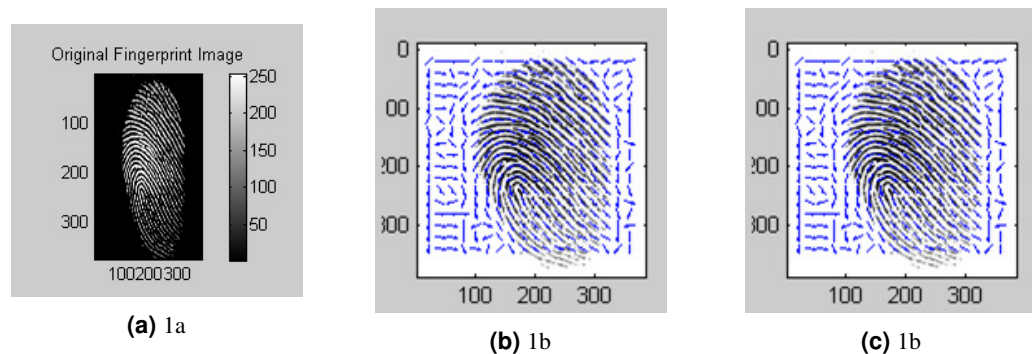
Figure 5. (a) Original fingerprint image (b) ridge orientation (c) ridge frequency

275   work is compared with existing algorithms such as Local Binary Pattern (LBP), Scale Invariant Feature
276   Transform (SIFT), Manifold Learning, and cloud user to service authentication (CU-SA) on different data
277   size. The evaluated results are shown in Figure 8. Proposed approach consumes only 890ms for 100 DB
278   size. Whereas other approach exceeds nearly 1000ms and above. From the resultant observation, the
279   proposed biometric authentication de-duplication is best in terms of computational time by having low
280   time compare to other standard existing algorithms.

281   To evaluate the proposed biometric security-based distributed storage authentication de-duplication
282   approach, the encryption and decryption using the bio key are calculate with communication cost for
283   different data sizes. The evaluated results are shown in Figures 9 and 10. In the encryption process, our
284   proposed algorithm takes 5ms, 8ms, 13ms, 23ms, and 26ms for encrypting data with the size of 10MB,
285   30MB, 50MB, 80MB, and 100MB. Similarly, for the decryption process, our proposed algorithm takes
286   4ms, 7ms, 12ms, 20ms, and 33ms for encrypting data with the size of 10MB, 30MB, 50MB, 80MB, and
287   100MB.

288   From all observations of the result experiments, the proposed bio key de-duplicating with security-
289   based distributed storage approach performs better encryption and decryption time and communication
290   cost with respect to different data size. The proposed approach is compared with the existing algorithms
291   such as convergent encryption (CE) Liu et al. (2017), leakage resilient (LR) Liu et al. (2017), randomized
292   convergent encryption (RCE) Liu et al. (2017), secure de-duplication scheme (SDS) Liu et al. (2017) to
293   prove the de-duplication performance. The figure 9 describes the communication cost of proposed system.

(a) 1a   **(b)** 1b   **(c)** 1b   **(d)** 1b
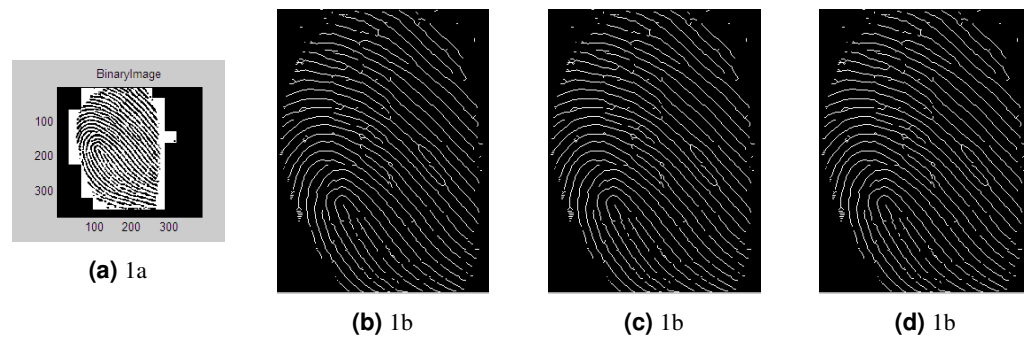
Figure 6. (a) binarization (b) thinned image (c) feature (Ridge & bifurcation) extraction (d) QR code of the fingerprint input image
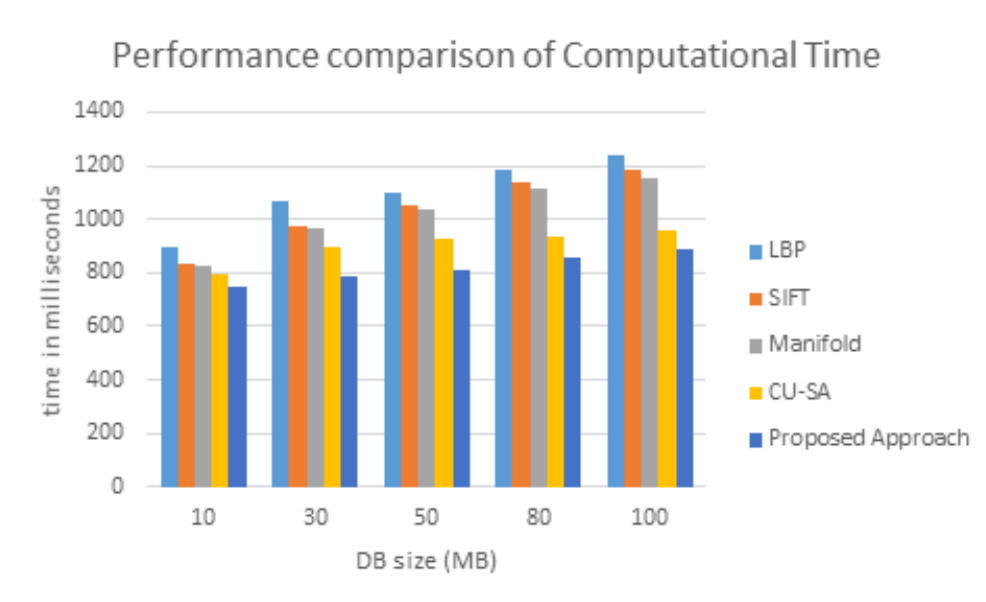
Figure 7. Evaluation of Computational Time of various algorithms

²⁹⁴   The evaluated experiment in terms of the data upload to the cloud server and the data download from the
²⁹⁵   server of various algorithms are shown in figure 10 and figure 11. The obtained result proved that our
²⁹⁶   proposed bio key-based encryption and de-duplication techniques secure minimum upload and download
²⁹⁷   time on various sizes of the data compared to other algorithms. Hence our proposed scheme achieves
²⁹⁸   better communication cost, better encryption and decryption time, and better data upload and downloads
²⁹⁹   time. Which proved that one method can be used for secure authentication, de-duplication with excellent
³⁰⁰   reliability, response time, and security.

5 CONCLUSION

³⁰²   Cloud storage has abundance data at every second for processing and storing in server by CSP. Data
³⁰³   deduplication in the cloud brings concern on security of stored data. There is lot of possibilities for
³⁰⁴   unauthorized access. In our proposed work, security using bio key generation makes users access data
³⁰⁵   securely. Here, the main task of accomplishment is avoiding redundancy in a cloud server and biometric
³⁰⁶   authentication using Gabor filters with XOR operation. This is a very complex fact due to biometric
³⁰⁷   scanning and matching for authentication. The evolution of high technologies over time assures the
³⁰⁸   security and integrity of the system. Also, data de-duplication is reducing multiple storages of the same
³⁰⁹   data over the cloud network. In this article, we have initiated security with the de-duplication of the data
³¹⁰   in the cloud servers. Security of the data is computed using the user's biometric parameters and a bio
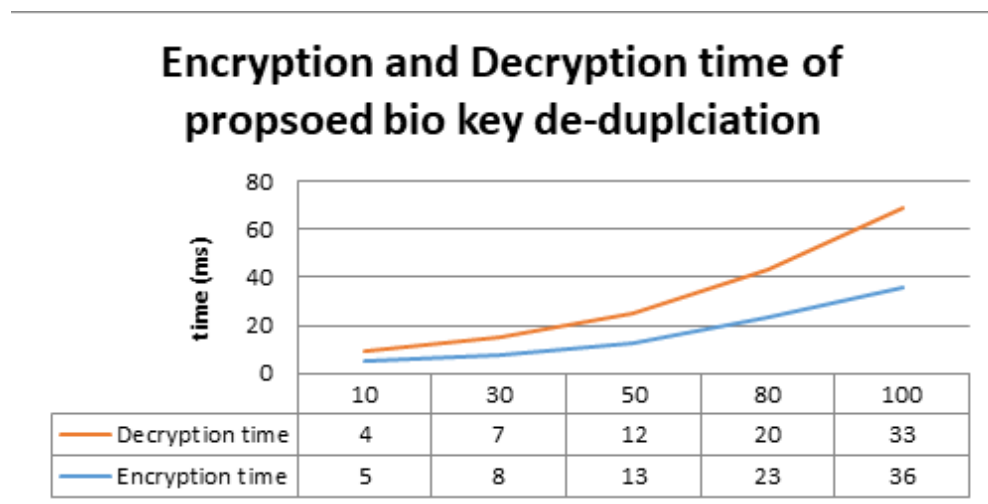
Figure 8. proposed method evaluation in terms of encryption and decryption time
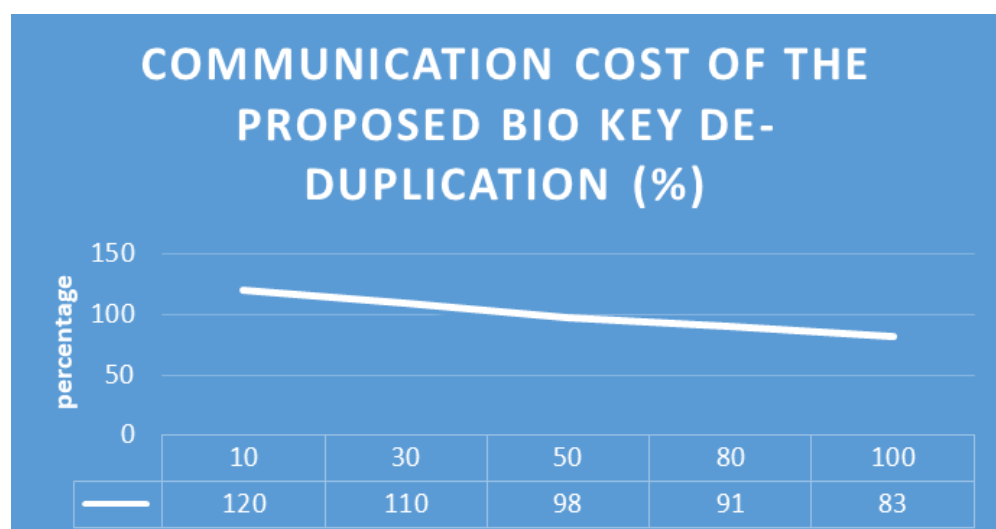
Figure 9. proposed method communication cost
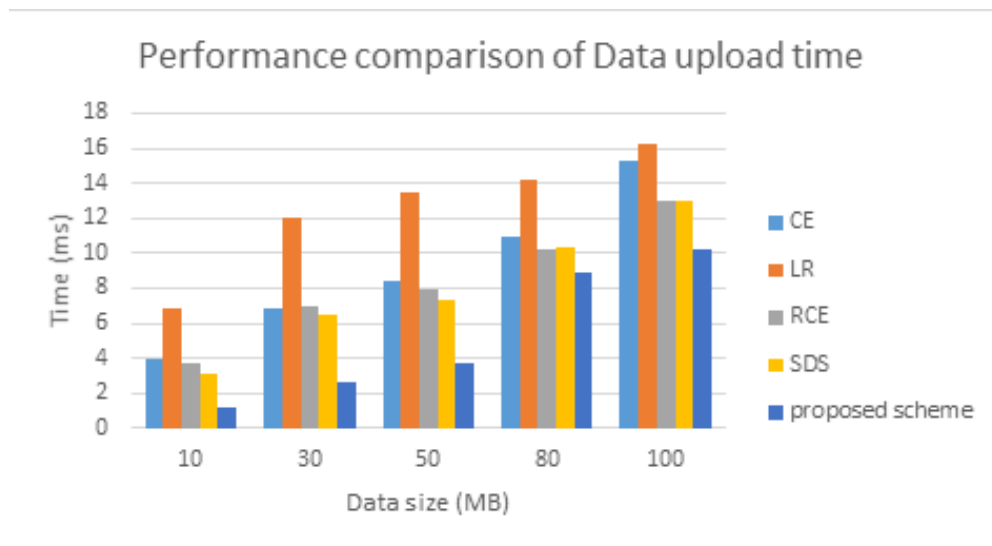
Figure 10. Performance comparison in terms of Data uploads time
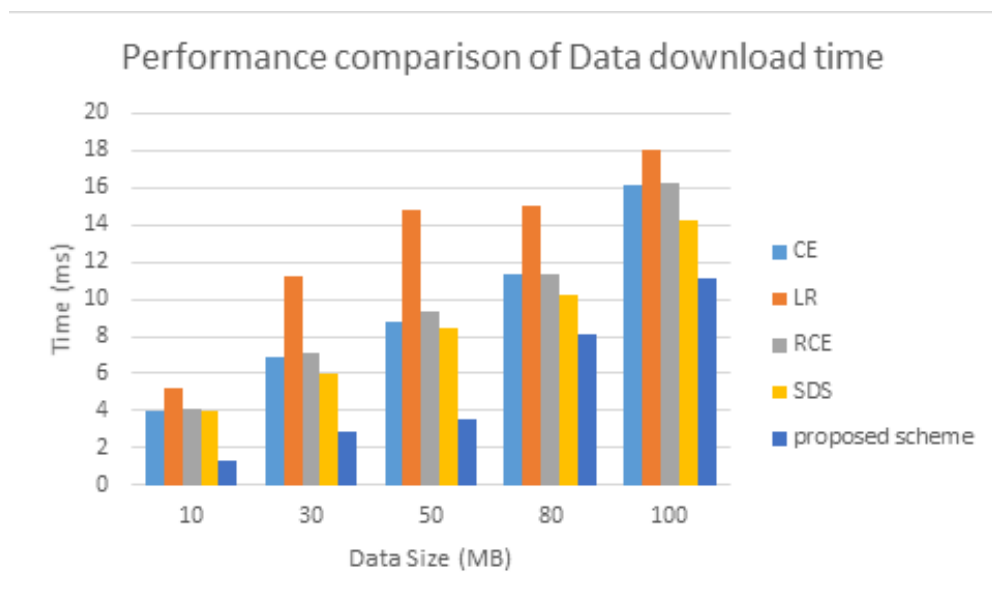
Figure 11. Performance comparison of Data downloads time

key is generated. The fingerprint of the owner is used to generate the bio key by QR code conversion. User is authenticated by transferring the bio key to users. Then the data is transferred to authenticated users. The most significant task carried out in this research work is biometric cryptographic security and reducing the de-duplication of data in cloud storage. The algorithm and its processing time are improved in our implementation process. The proposed methodology provides more reliability and fast encryption techniques. In the future, different biometric techniques can be expanded for user authentication processes. Standardized intelligent algorithms can be further used for checking data deduplication.

REFERENCES

Ah Kioon, M. C., Wang, Z. S., and Deb Das, S. (2013). Security analysis of md5 algorithm in password storage. In *Applied Mechanics and Materials*, volume 347, pages 2706–2711. Trans Tech Publ.

Al-Assam, H., Hassan, W., and Zeadally, S. (2019). Automated biometric authentication with cloud computing. In *Biometric-Based Physical and Cybersecurity Systems*, pages 455–475. Springer.

Alomar, N., Alsaleh, M., and Alarifi, A. (2017). Social authentication applications, attacks, defense strategies and future research directions: a systematic review. *IEEE Communications Surveys & Tutorials*, 19(2):1080–1111.

Aloul, F., Zahidi, S., and El-Hajj, W. (2009). Two factor authentication using mobile phones. In *2009 IEEE/ACS International Conference on Computer Systems and Applications*, pages 641–644. IEEE.

Amin, R., Kumar, N., Biswas, G., Iqbal, R., and Chang, V. (2018). A light weight authentication protocol for iot-enabled devices in distributed cloud computing environment. *Future Generation Computer Systems*, 78:1005–1019.

Balloon, A. M. (2001). From wax seals to hypertext: Electronic signatures, contract formation, and a new model for consumer protection in internet transactions. *Emory LJ*, 50:905.

Banyal, R., Jain, V., and Jain, P. (2014). Dynamic trust based access control framework for securing multi-cloud environment. In *Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies*, pages 1–8.

Benarous, L., Kadri, B., and Bouridane, A. (2017). A survey on cyber security evolution and threats: biometric authentication solutions. In *Biometric security and privacy*, pages 371–411. Springer.

Bhargav-Spantzel, A., Squicciarini, A. C., Modi, S., Young, M., Bertino, E., and Elliott, S. J. (2007). Privacy preserving multi-factor authentication with biometrics. *Journal of Computer Security*, 15(5):529–560.

Bonneau, J., Herley, C., Van Oorschot, P. C., and Stajano, F. (2015). Passwords and the evolution of imperfect authentication. *Communications of the ACM*, 58(7):78–87.

Bruun, A., Jensen, K., and Kristensen, D. (2014). Usability of single-and multi-factor authentication methods on tabletops: a comparative study. In *International Conference on Human-Centred Software Engineering*, pages 299–306. Springer.

Chandramohan, D., Vengattaraman, T., and Dhavachelvan, P. (2017). A secure data privacy preservation for on-demand cloud service. *Journal of King Saud University-Engineering Sciences*, 29(2):144–150.

Council, N. R. and Committee, W. B. (2010). Biometric recognition: Challenges and opportunities.

Coventry, L., De Angeli, A., and Johnson, G. (2003). Usability and biometric verification at the atm interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 153–160.

Dasgupta, D., Roy, A., and Nag, A. (2016). Toward the design of adaptive selection strategies for multi-factor authentication. *computers & security*, 63:85–116.

Dulari, P. and Bhushan, B. (2019). A novel approach for cloud data security enhancement through cryptography and biometric in the government cloud environment. *Int. J. Comp. Sci. Mobile Computing*, 8(12):59–63.

Grassi, P. A., Fenton, J. L., Newton, E. M., Perlner, R. A., Regenscheid, A. R., Burr, W. E., Richer, J. P., Lefkovitz, N. B., Danker, J. M., and Choong, Y. (2016). Draft nist special publication 800-63b digital identity guidelines. *National Institute of Standards and Technology (NIST)*, 27.

Grigoras, C. (2009). Applications of enf analysis in forensic authentication of digital audio and video recordings. *Journal of the Audio Engineering Society*, 57(9):643–661.

Gunson, N., Marshall, D., Morton, H., and Jack, M. (2011). User perceptions of security and usability of single-factor and two-factor authentication in automated telephone banking. *Computers & Security*, 30(4):208–220.

Harini, N. and Padmanabhan, T. (2013). 2cauth: A new two factor authentication scheme using qr-code. *International Journal of Engineering and Technology*, 5(2):1087–1094.

Heartfield, R. and Loukas, G. (2015). A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys (CSUR)*, 48(3):1–39.

Huang, X., Xiang, Y., Bertino, E., Zhou, J., and Xu, L. (2014). Robust multi-factor authentication for fragile communications. *IEEE Transactions on Dependable and Secure Computing*, 11(6):568–581.

Hur, J., Koo, D., Shin, Y., and Kang, K. (2016). Secure data deduplication with dynamic ownership management in cloud storage. *IEEE Transactions on Knowledge and Data Engineering*, 28(11):3113–3125.

Ibrokhimov, S., Hui, K. L., Al-Absi, A. A., and Sain, M. (2019). Multi-factor authentication in cyber physical system: A state of art survey. In *2019 21st International Conference on Advanced Communication Technology (ICACT)*, pages 279–284. IEEE.

Indu, I., Anand, P. R., and Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. *Engineering science and technology, an international journal*, 21(4):574–588.

Kim, J.-J. and Hong, S.-P. (2011). A method of risk assessment for multi-factor authentication. *Journal of Information Processing Systems*, 7(1):187–198.

Konoth, R. K., van der Veen, V., and Bos, H. (2016). How anywhere computing just killed your phone-based two-factor authentication. In *International Conference on Financial Cryptography and Data Security*, pages 405–421. Springer.

Kumar, D. A. and Begum, T. U. S. (2011). A novel design of electronic voting system using fingerprint. *International Journal of Innovative Technology & Creative Engineering*, 1(1):12–19.

Liu, X., Li, G., Zhang, S., and Liu, A. (2018). Big program code dissemination scheme for emergency software-define wireless sensor networks. *Peer-to-Peer networking and applications*, 11(5):1038–1059.

Liu, X., Liu, Q., Peng, T., and Wu, J. (2017). Dynamic access policy in cloud-based personal health record (phr) systems. *Information Sciences*, 379:62–81.

Malathi, R. and Raj.R, J. R. (2016). An integrated approach of physical biometric authentication system. *Procedia Computer Science*, 85:820–826.

Mohsin, J., Han, L., Hammoudeh, M., and Hegarty, R. (2017). Two factor vs multi-factor, an authentication battle in mobile cloud computing environments. In *Proceedings of the International Conference on Future Networks and Distributed Systems*, pages 1–10.

Naeem, A., Javed, A. R., Rizwan, M., Abbas, S., Lin, J. C.-W., and Gadekallu, T. R. (2021). Dare-sep: A hybrid approach of distance aware residual energy-efficient sep for wsn. *IEEE Transactions on Green Communications and Networking*.

Nor, N. A., Narayana Samy, G., Ahmad, R., Ibrahim, R., and Maarop, N. (2015). The proposed public key infrastructure authentication framework (pkiaf) for malaysian government agencies. *Advanced Science Letters*, 21(10):3161–3164.

Obergrusberger, F., Baloglu, B., Sänger, J., and Senk, C. (2012). Biometric identity trust: toward secure biometric enrollment in web environments. In *International Conference on Cloud Computing*, pages 124–133. Springer.

Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., and Koucheryavy, Y. (2018). Multi-factor authentication: A survey. *Cryptography*, 2(1):1.

Rehman, A., Rehman, S. U., Khan, M., Alazab, M., and Reddy, T. (2021). Canintelliids: Detecting in-vehicle intrusion attacks on a controller area network using cnn and attention-based gru. *IEEE Transactions on Network Science and Engineering*.

Scheidt, E. M. and Domangue, E. (2006). Multiple factor-based user identification and authentication. US Patent 7,131,009.

Shabbir, M., Shabbir, A., Iwendi, C., Javed, A. R., Rizwan, M., Herencsar, N., and Lin, J. C.-W. (2021). Enhancing security of health information using modular encryption standard in mobile cloud computing. *IEEE Access*, 9:8820–8834.

Sun, J., Zhang, R., Zhang, J., and Zhang, Y. (2014). Touchin: Sightless two-factor authentication on multi-touch mobile devices. In *2014 IEEE conference on communications and network security*, pages 436–444. IEEE.

Symeonidis, I., Mustafa, M. A., and Preneel, B. (2016). Keyless car sharing system: A security and privacy analysis. In *2016 IEEE International Smart Cities Conference (ISC2)*, pages 1–7. IEEE.

420 Tahir, H. and Tahir, R. (2008). Biofim: Multifactor authentication for defeating vehicle theft. In
421   *Proceedings of the World Congress on Engineering, London, UK*, pages 2–4. Citeseer.

422 Wang, C., Ding, K., Li, B., Zhao, Y., Xu, G., Guo, Y., and Wang, P. (2018). An enhanced user
423   authentication protocol based on elliptic curve cryptosystem in cloud computing environment. *Wireless*
424   *Communications and Mobile Computing*, 2018.

425 Wang, D. and Wang, P. (2015). Offline dictionary attack on password authentication schemes using smart
426   cards. In *Information security*, pages 221–237. Springer.

427 Wong, K.-S. and Kim, M. H. (2012). Secure biometric-based authentication for cloud computing. In
428   *International Conference on Cloud Computing and Services Science*, pages 86–101. Springer.

429 Zahrouni, L., Blackwood, D., Rizvi, S., Gualdoni, J., and Almiani, M. (2017). Preventing identity theft
430   using biometrics based authentication system. In *2017 IEEE Jordan Conference on Applied Electrical*
431   *Engineering and Computing Technologies (AEECT)*, pages 1–6. IEEE.

432 Ziyad, S. and Kannammal, A. (2014). A multifactor biometric authentication for the cloud. In *Computa-*
433   *tional Intelligence, Cyber Security and Computational Models*, pages 395–403. Springer.