

ENHANCING CLOUD SECURITY: A MULTI-FACTOR AUTHENTICATION AND ADAPTIVE CRYPTOGRAPHY APPROACH USING MACHINE LEARNING TECHNIQUES

Prof. Rakesh Ramesh Tannu*¹

*¹Head Of Department, Information Technology, JCEI's Jaihind Polytechnic, Kuran, Maharashtra, India.

ABSTRACT

The necessity for sophisticated security measures to safeguard private information on distant servers is highlighted by the cloud computing industry's explosive growth. To protect these data, authentication is essential. Vulnerabilities continue despite the different approaches that have been suggested. Using machine learning approaches based on an intrusion-detection system, this research presents a revolutionary multi-factor authentication system combined with a hybrid cryptographic framework that dynamically changes encryption algorithms. The suggested method uses fingerprint authentication, conditional characteristics, and passwords to extract the encryption key from fingerprint information. It uses a dual-encryption strategy that combines five algorithm pairs: AES + HMAC (SHA-256), ECC + HMAC (SHA-512), HMAC-MD5 + PBKDF2, Twofish + Argon2, and Blowfish + HMACSHA3-256. A Hybrid CNN-transformer model predicts and classifies attacks by dynamically adjusting an encryption algorithm to secure the data. The framework exhibited strong resilience against brute force, spoofing, phishing, guessing, and impersonation attacks. The proposed model achieved a commendable accuracy of 96.8%, outperforming other models. Implementing this framework in a cloud authentication environment significantly enhances data confidentiality and protects against unauthorized access. This study highlights the potential of combining multi-factor authentication and adaptive cryptography to obtain robust cloud security solutions.

Keywords: Adaptive Cryptography, Attack Prediction, Cloud Security, Dual Encryption, Dynamic Encryption Algorithms, Hybrid CNN-Transformer Model, Machine Learning Techniques, Multi-Factor Authentication.

I. INTRODUCTION

The introduction of cloud computing [CC] has drastically altered how people and companies handle and store data. CC provides on-demand services through the internet and makes it possible to store data and application software with less administrative work. Any platform can be transformed into a cloud-based infrastructure by integrating these four crucial components: Rapid scalability, resource pooling, extensive network accessibility, and self-service on demand. CC offers excellent services to individuals through the Internet, and is essential for everyday software solutions. Many cloud-supported applications rely on user, personal, and location information that can pose security and privacy risks. Research on cloud services mainly concentrates on creating safe user authorization techniques to shield private information from possible dangers in cloud computing settings. Four well-known techniques are available for user authentication in cloud systems: knowledge-based (based on the user's knowledge), ownership-based (based on the user's possessions), characteristic-based (based on the user's identity), and location-based (based on the user's location). As shown in Fig. 1, each of these techniques is known as an authentication factor. In order to prevent unauthorized access to applications, data, goods, services, and resources, common authentication mechanisms such as password-based, certificate-based, token-based, and multi-factor approaches are essential. Finding the most secure authentication method that is widely accepted by users is a major challenge in cloud computing because of various threats that can compromise the login process. Developing a trustworthy authentication method for cloud services requires detailed awareness of potential threats and techniques to prevent them. Over the last several decades, many user authentication systems and cryptographic techniques have been proposed for securing personal information in the cloud. However, recent research indicates that existing cloud computing systems lack adequate security measures for user authentication and management. The primary findings of this study are as follows.

II. METHODOLOGY

The classifiers included K-Nearest Neighbors, Random Forest, Logistic Regression, Decision Tree, XGBoost,

CatBoost, Support Vector Machine, and Artificial Neural Networks. Each classifier was trained using various techniques depending on the design specifications. However, this study focuses on specific ensemble methods and algorithm pairings, potentially overlooking other effective strategies. Further exploration of the classifications based on different attack types is required.

The current literature suggests that both multi-factor authentication (MFA) and anomaly detection systems require significant enhancement. MFA systems require improvements in user-friendliness, security optimization, and comprehensive method analysis, with further research on advanced techniques, biometrics, and cryptographic methods. Anomaly detection systems should focus on refining hybrid methods, enhancing attack-specific classification, boosting real-time performance, and conducting thorough security analysis. Addressing the challenges of various attack types and ensuring cost efficiency are critical. Resolving these issues can significantly improve the security, accuracy, and performance of cloud-based anomaly detection and authentication systems.

In this section, we present an improved multi-factor authentication framework designed for cloud environments. The proposed scheme involves three key phases: (1) registration, (2) login and authentication, and (3) identity update

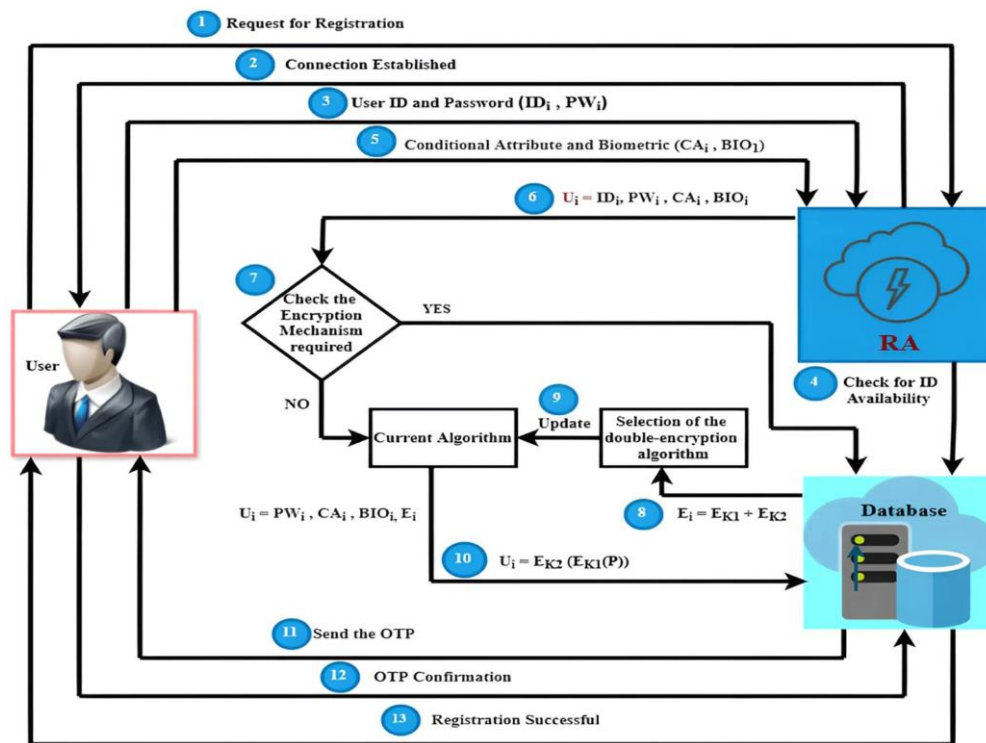


FIGURE 1. Registration phase.

A. USER REGISTRATION PHASE

The user initiates a request for a new registration. Once a request is received, the Registration Authority (RA) establishes a secure communication channel. User (U_i) then proceeds with the registration process using RA, as shown in Fig. 2.

To begin the registration process, user (U_i) submits its

selected user ID (UID_i) to the Registration Authority (RA). The RA then verifies the availability of UID_i by checking it against the existing User List. If UID_i is already obtained, the user is prompted to select a different UID_i . Once a unique UID_i is confirmed, the user is instructed to create a strong password (PW_i). After successfully setting the password, the registration proceeds to the next stage of the process. The Registration Authority (RA) initiates the creation of conditional attributes (CA_i) for each user by administering a series of security questions. Users are required to select more than one conditional attribute, defined according to the following formula:

$$ID_i \rightarrow (CA_1, CA_2, \dots, CA_n)$$

Next, users submit their biometric information (BIO_i)

during this phase, which involves feature extraction techniques [31], [32] to derive relevant features. From these extracted features, Biometric Master Key (BMK_i) was generated [33]. Finally, the RA obtains all the information from the user.

$U_i \rightarrow (UID_i, PW_i, CA_i, BIO_i, BMK_i)$

· During the encryption algorithm selection phase, RA

verifies the status of the encryption machine to determine the appropriate encryption algorithm. This selection process adheres to a specific rule in which cryptographic mechanisms are rotated periodically for groups of N users.

· The value of N is chosen to ensure that the computational

overhead is optimal and the security is not compromised. The database already contained a list of five pairs of encryption mechanisms.

· Based on the chosen encryption mechanisms, a dual

encryption approach was applied to encrypt specific U_i details such as (PW_i, CA_i, BIO_i, E_i) . These encrypted details, along with information about the encryption techniques (E_i) are stored securely in a database. The Biometric Master Key (BMK_i) generated from the user's biometric data (BMK_i) is split into two keys: K_1 for the first encryption mechanism (E_1) and K_2 for the second encryption mechanism (E_2).

Subsequently the OTP was sent to the user's mobile number and email account. Once the OTP is confirmed by the user during registration, the registration process is completed.

In this study, five sets of dual encryption algorithms were employed to encrypt user information. A hybrid encryption algorithm was chosen for each user based on specific conditions and then used to encrypt the information. The dual encryption algorithms used were AES + HMAC (SHA-256), ECC

+ HMAC (SHA-512), HMAC-MD5 + PBKDF2, Twofish +

Argon2, and Blowfish + HMAC SHA3-2.

B. AUTHENTICATION PHASE

During the authentication phase, the client's identity is confirmed before they can access the cloud data services. This process occurs when clients submit requests for cloud services.

· Upon receiving the login request, the proposed frame-

work initiates a multi-factor authentication process. Initially, the users were prompted to enter their UID_i and PW_i . Subsequently, the RA verifies the existence of the UID_i . If a UID_i exists, the PW_i is encrypted and compared with the encrypted password in the database. If a match is found, then the process proceeds to the next authentication step.

· At this stage, the user is prompted to provide necessary

CA_i information after verification by the server. If the verification is successful, the user is permitted to input the BIO_i information. These data were processed and compared with server records. If a match is found, the login will be successful; otherwise, it will fail.

Although fingerprint biometric authentication is widely recognized for its convenience and effectiveness, it is essential to acknowledge its limitations. Various factors, including skin conditions, moisture, quality of the fingerprint reader, unavailability of the scanner, and difficulties with fingerprint recognition owing to issues such as worn or damaged skin, can impede its overall performance. To address these challenges, we incorporate the concept of "skip" as a fallback mechanism when biometric authentication fails or is unavailable. This allows us to explain how the skip concept can be utilized in MFA systems with biometric authentication. There are several types of skip concepts available, such as one-time tokens (OTT), hardware tokens, and push notifications, and we chose

the OTT method for our implementation.

When the user is directed to fingerprint authentication,

users are presented with the option of skipping the connection if they are unable to provide necessary the

biometric details. The skip concept can only be used by users under unavoidable circumstances. When the User selects the skip connection option, the server generates a new OTT. The OTT password was randomly generated by the system using the formulation provided below.

$$UID = OTT \text{ where } OTT = \{r\}$$

· The OTT for a specific UID_i is generated by randomly choosing a large prime number 'r' using the method described above. Every time a new OTT is generated for UID_i , it is sent to the user's mobile number and email account via a secure channel. This OTT is intended for single login use, ensuring that only authorized users with an OTT can access the authentication process

· In our system, we implemented various approaches to address the issue of the excessive use of the skip option. We closely monitored its usage and established specific boundaries to deter its overreliance. If someone exceeds the limit, we promptly notify both the individual and administrator. In addition, we introduced a system that restricts users to a set number of skips per month, typically three, in order to strike a balance between convenience and security. A visual representation of the authentication phase IDENTITY UPDATE PHASE

The Identity Update Phase is an essential part of the user management system responsible for managing modifications to a user's identity details while safeguarding the security and accuracy of the data stored. In this phase, users (U_i) have the option to update various aspects of their identity information, namely, their password (PW), Conditional Attributes (CA), or into the key-generation algorithm to produce a biometric key. Subsequently, this key was compared with to existing key. If there are any differences, then the PWD, CA, and BIO of the respective users are encrypted using the newly generated key and stored securely in the database. For each update type, the system maintains an audit log containing the user ID, timestamp, and authentication method to ensure traceability and accountability.

III. HYBRID CNN-TRANSFORMER MODEL FOR ATTACK PREDICTION

The UNSW-NB15 dataset, developed by the Australian Center for Cyber Security (ACCS) using the IXIA PerfectStorm tool, is a valuable resource for testing network intrusion detection systems (NIDS). It includes contemporary attack scenarios and regular activities, making it ideal for evaluating and comparing machine-learning methods for network anomaly detection. The dataset consists of raw network packets captured using Wireshark, extracted features, and labeled data, with 49 attributes categorized into various types Table 1. lists these data fields and their descriptions. This data set consist of

Feature Categories:

- 1) **Flow Features:** Basic features derived from network flows. Examples: Duration, protocol type, service, source and destination bytes.
- 2) **Basic Features:** Features related to TCP/IP connections. Examples: Source and destination IP addresses, port numbers, and timestamps.
- 3) **Content Features:** Information extracted from the packet payload. Examples: Number of failed login attempts, and shell prompts.
- 4) **Time-based Features:** Information based on time-based patterns. Examples: Number of connections to the same host within a certain time frame.
- 5) **Additional Generated Features:** Features created by analyzing traffic patterns. Examples: Number of packets, number of bytes, and average packet size.

TABLE 1. Examples of Data Fields and Their Descriptions in the UNSW-NB15 Dataset

Features	Description	Features	Description
srcip	Source IP address	sttl	Source to dest. time to live value
sport	Source port number	dttl	Dest. to source time to live value
dstip	Destination (Dest.) IP address	sloss	Source packets retransmitted or dropped
dsport	Destination port number	dloss	Dest. packets retransmitted or dropped

proto	Protocol type (e.g., TCP, UDP)	service	http, ftp, smtp, ssh, dns, ftp-data ,irc
state	State of the connection (e.g., FIN, SYN)	Sload	Source bits per second
dur	Duration of the connection	Dload	Dest. bits per second
sbytes	Source-to-destination bytes	Spkts	Source to destination packet count
dbytes	Dest.-to-source bytes	Dpkts	Dest. to source packet count
label	Attack type or normal	swin	Source TCP window advertisement value

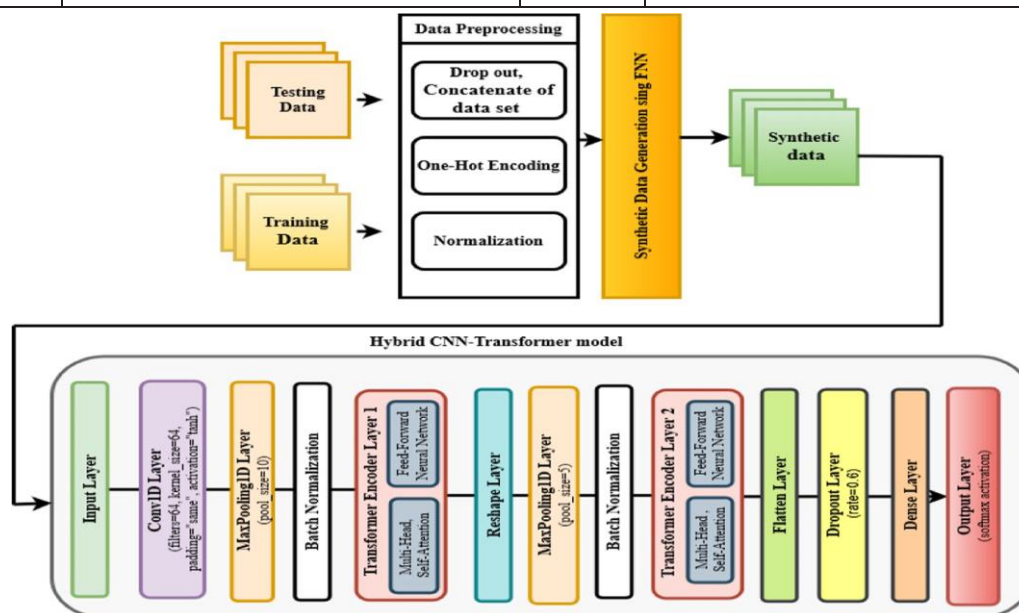


FIGURE 2. Hybrid CNN - transformer model.

A. DATA PREPROCESSING

The UNSW-NB15 dataset underwent thorough preprocessing to enhance its quality and usability for training machine learning models. The process involved importing the datasets into Pandas DataFrames, conducting exploratory data analysis to understand the attack category distribution, addressing class imbalances, removing irrelevant columns, and ensuring no missing values. Categorical features were transformed into a numerical format using the one-hot encoding method, and both the training and testing datasets were merged for consistent preprocessing. Feature normalization was applied through min-max scaling to ensure equal contributions of all features during model training. To address class imbalances, a Feedforward Neural Network (FNN) was used to generate synthetic samples for minority classes, resulting in a balanced dataset. The final step involved separating the class labels from the features and preparing the dataset for machine learning tasks. These preprocessing steps transformed the original UNSW-NB15 dataset into a structured, balanced, and normalized format, thereby improving the learning capabilities and prediction accuracy of the model.

B. IMPLEMENTATION OF THE PROPOSED MODEL

Various machine and deep learning techniques have been employed to create effective attack classification systems. We propose a hybrid Convolutional Neural Network (CNN) transformer model, as depicted in Fig. 2. This model leverages the strengths of both CNNs and transformers to enhance classification accuracy.

The proposed model integrates an FNN for synthetic data generation. These synthetic data, produced by the FNN, augment the training set and improve the ability of the model to generalize and accurately classify attacks. Once generated, the synthetic data are passed to the hybrid CNN transformer model for classification. The CNN component effectively captured the spatial hierarchies in the data, whereas the transformer component captured long-range dependencies, thereby boosting the overall performance of the model. This hybrid model was implemented using the Jupyter Notebook platform, which provides an interactive environment for the coding, visualization, and testing of the model.

The effectiveness of the proposed model was evaluated by comparing it with other models, including artificial

neural networks (ANN), Gradient Boosting Classifier (GBC), and Random Forest Classifier (RFC). Each of these models represents a different approach to attack classification and provides a comprehensive benchmark for the proposed model. The results demonstrate that the proposed hybrid CNN-transformer model outperforms the other models, achieving the highest accuracy rate of 96.8%. This indicates that the integration of the CNN and transformer architectures, along with the use of synthetic data generation, significantly enhances the ability of the model to accurately classify attacks, making it a robust tool for cybersecurity applications.

C. TRAINING OF MODELS AND EVALUATION OF SYSTEMS

The process of model training involves utilizing 5-fold cross-validation with stratified KFold to ensure a balanced representation of classes. SMOTE was employed to address the class imbalance. Each training and validation fold included compiling and fitting of the model using the Adam optimizer and categorical cross-entropy loss. The accuracy and loss metrics were monitored throughout the epochs. After training, the performance of the model was evaluated using the out-of-sample data for each fold. This evaluation includes computing metrics such as accuracy and confusion matrices to assess the effectiveness of classification across different attack categories. In addition, ROC curves were constructed to graphically examine the model's multiple-class classification performance in terms of true-positive and false-positive rates, thereby providing a comprehensive understanding of its predictive capabilities.

IV. DYNAMIC ENCRYPTION ALGORITHM SELECTION BASED ON PREDICTED ATTACKS

After predicting the likelihood of an attack using UNSW-NB1, an optimal encryption algorithm based on anticipated attack types was dynamically selected. Refer to Fig. 3(a) for details on the encryption algorithms employed and their respective resistance to the predicted attack scenarios.

The process commences with the utilization of a trained Feedforward Neural Network (FNN) and scaled input features

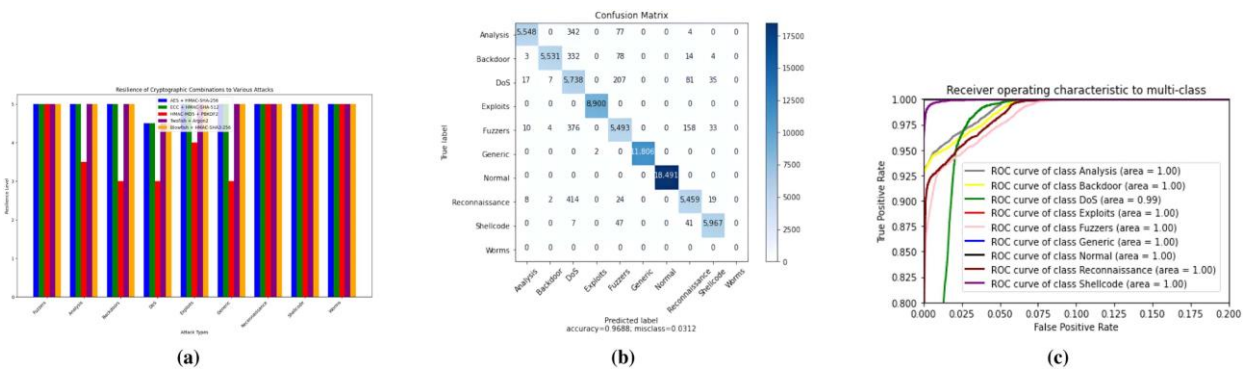


FIGURE 3. (a) Resilience of cryptographic combinations to various attacks, (b) confusion matrix, and (c) ROC curve for the proposed model.

To predict attack types. Subsequently, the predicted attack type was used to determine a suitable encryption algorithm. This selection is based on a predefined mapping that associates each attack type with specific algorithms, such as AES

+ HMAC (SHA 256), ECC + HMAC (SHA512), HMAC-MD5 + PBKDF2, Twofish + Argon2, and Blowfish + HMAC SHA3-256. Data encryption was implemented accordingly, and the encrypted credentials were updated in the database using SQL queries. This automated workflow guarantees the security of sensitive information against anticipated threats, and dynamically adjusts encryption methods based on real-time threat predictions.

V. RESULT AND DISCUSSION

It is crucial to compute a range of evaluation metrics, including accuracy, precision, recall, and F1-score, to evaluate and compare the performance of the machine learning model utilized in this research. The effectiveness of the models can be evaluated by analyzing these metrics based on specific criteria.

VI. CONCLUSION

Finally, this study provides a complete method for improving cloud security using a revolutionary multi-factor

authentication system and adaptive cryptography framework. The system solves significant weaknesses in cloud computing environments by merging passwords, conditional attributes, fingerprint authentication mechanisms and machine learning-driven dynamic encryption algorithm updates. The dual encryption technique enhances data safety by utilizing combinations such as AES + HMAC (SHA-256) and ECC + HMAC (SHA-512), which are dynamically customized based on the discovered attack type. Furthermore, the use of a hybrid CNN-transformer based model for attack prediction resulted in good accuracy, precision, recall, and F1-Score values of 96.8% and 97.2%, respectively. 96.8% and 96.9% outperformed the standard models. This predictive capacity enables the system to preemptively alter security measures, assuring strong defense against brute force, guesses, phishing, spoofing, and impersonation assaults. Implementing this advanced approach allows enterprises to improve the privacy and integrity of the sensitive information stored in cloud settings. The framework not only meets strict security standards but also includes adaptive capabilities to effectively combat evolving threats. This study demonstrates the efficacy of combining multi-factor authentication and adaptive cryptography, opening the way for resilient and secure cloud computing systems in the face of growing cybersecurity threats. Future improvements could leverage real-time datasets for more accurate attack predictions. Deploying this framework in live cloud environments enhances security and readiness against evolving cyber threats. By employing behavioral analysis techniques, abnormal usage patterns of the skip mechanism can be detected and mitigated, thereby preventing potential misuse.

VII. REFERENCES

- [1] G. Ramesh, J. Logeshwaran, and V. Aravindarajan, "A secured database monitoring method to improve data backup and recovery operations in cloud computing," *BOHR Int. J. Comput. Sci.*, vol. 2, no. 1, pp. 1–7, 2022, doi: 10.54646/bijcs.019.
- [2] B. T. Rao, "A study on data storage security issues in cloud computing," *Procedia Comput. Sci.*, vol. 92, pp. 128–135, 2016, doi: 10.1016/j.procs.2016.07.335.
- [3] K. Latha and T. Sheela, "Block based data security and data distribution on multi cloud environment," *J. Ambient Intell. Humanized Comput.*, vol. 15, 2024, Art. no. 53, doi: 10.1007/s12652-019-01395-y.
- [4] K. Raju and M. Chinnadurai, "An identity-based secure and optimal authentication scheme for the cloud computing environment," *Comput. Mater. Continua*, vol. 69, no. 1, pp. 1057–1072, 2021, doi: 10.32604/cmc.2021.016068.
- [5] A. Ometov, S. Bezzateev, N. Mäkitalo, S. Andreev, T. Mikkonen, and Y. Koucheryavy, "Multi-factor authentication: A survey," *Cryptogr.*, vol. 2, no. 1, pp. 1–31, 2018, doi: 10.3390/cryptography2010001.
- [6] S. Sudha and S. S. Manikandasaran, "A survey on different authentication schemes in cloud computing environment," *Int. J. Manage., IT Eng.*, vol. 9, no. 1, pp. 359–375, 2019.
- [7] A. Alhothaily, C. Hu, A. Alrawais, T. Song, X. Cheng, and D. Chen, "A secure and practical authentication scheme using personal devices," *IEEE Access*, vol. 5, pp. 11677–11687, 2017, doi: 10.1109/ACCESS.2017.2717862.
- [8] N. Anusha and N. R. Suma, "A review on secured file system using multi-factor authentication with visual cryptography for cloud environment," *Int. Res. J. Modernization Eng. Technol. Sci.*, vol. 4, no. 6, pp. 4433–4436, 2012.
- [9] Cybersecurity, "What is password encryption and how does it work," Accessed: Jun. 15, 2023. [Online]. Available: <https://teampassword.com/blog/what-is-password-encryption-and-how-much-is-enough>
- [10] M. Hazratifard, F. Gebali, and M. Mamun, "Using machine learning for dynamic authentication in telehealth: A tutorial," *Sensors*, vol. 22, no. 19, 2022, Art. no. 7655, doi: 10.3390/s22197655.
- [11] N. Siddiqui, L. Pryor, and R. Dave, "User authentication schemes using machine learning methods—A review," in *Proc. Int. Conf. Commun. Comput. Technol.*, Singapore, 2021, pp. 703–723, doi: 10.1007/978-981-16-3246-4_54.
- [12] T. N. Tan and H. Lee, "High-secure fingerprint authentication system using ring-LWE cryptography," *IEEE Access*, vol. 7, pp. 23379–23387, 2019, doi: 10.1109/ACCESS.2019.2899359.
- [13] C. Singh and D. Singh, "A 3-level multifactor authentication scheme for cloud computing," *Int. J. Comput. Eng. Technol.*, vol. 10, no. 1,

- [15] pp. 184–195, 2019. [Online]. Available: <https://ssrn.com/abstract=3537621>
- [16] S. A. Sagar, O. Bhat, M. Raina, and S. Patil, "Authentication system using cryptographic secure password storage," *Int. J. Innov. Res. Eng. Multidisciplinary Phys. Sci.*, vol. 6, no. 6, pp. 76–78, 2018.
- [17] V. Kakkad, M. Patel, and M. Shah, "Biometric authentication and image encryption for image security in cloud framework," *Multiscale Multidisciplinary Model., Exp. Des.*, vol. 2, no. 4, pp. 233–248, 2019, doi: 10.1007/s41939-019-00049-y.
- [18] M. Obaidat, J. Brown, S. Obeidat, and M. Rawashdeh, "A hybrid dynamic encryption scheme for multi-factor verification: A novel paradigm for remote authentication," *Sensors*, vol. 20, no. 5, 2020, Art. no. 4212, doi: 10.3390/s20154212.
- [19] K. DeviPriya and S. Lingamgunta, "Multi factor two-way hash-based authentication in cloud computing," *Int. J. Cloud Appl. Comput.*, vol. 10, no. 2, 2020, Art. no. 21, doi: 10.4018/IJCAC.2020040104.
- [20] K. M. Prabha and P. V. Saraswathi, "Suppressed K-anonymity multi-factor authentication based Schmidt-Samoa cryptography for privacy preserved data access in cloud computing," *Comput. Commun.*, vol. 158, pp. 85–94, 2020, doi: 10.1016/j.comcom.2020.04.057.
- [21] B. O. ALSaleem and A. I. Alshoshan, "Multi-factor authentication to systems login," in *Proc. Nat. Comput. Colleges Conf.*, 2021, pp. 1–4, doi: 10.1109/NCCC49330.2021.9428806.
- [22] K. Venkatachalam, P. Prabu, A. Almutairi, and M. Abouhawwash, "Secure biometric authentication with de-duplication on distributed cloud storage," *PeerJ Comput. Sci.*, vol. 7, 2021, Art. no. e569, doi: 10.7717/peerj-cs.569.
- [23] J. Mohammed Ubada and M. Mohamed Surputheen, "Evaluation of multifactor user security through multi authentication verifiable hybrid revert encryption for cloud computing environment," *Int. J. Comput. Sci. Netw. Secur.*, vol. 22, no. 9, pp. 481–488, 2022, doi: 10.22937/IJC-SNS.2022.22.9.62.
- [24] S. Kaur, G. Kaur, and M. Shabaz, "A secure two-factor authentication framework in cloud computing," *Secur. Commun. Netw.*, vol. 2022, no. 1, 2022, Art. no. 7540891, doi: 10.1155/2022/7540891.
- [25] D. Carrillo-Torres, J. A. Pérez-Díaz, J. A. Cantoral-Ceballos, and C. Vargas-Rosales, "A novel multi-factor authentication algorithm based on image recognition and user established relations," *Appl. Sci.*, vol. 13, no. 3, 2023, Art. no. 1374, doi: 10.3390/app13031374.
- [26] D. V. K. Vengala, D. Kavitha, and A. S. Kumar, "Three factor authentication system with modified ECC based secured data transfer: Untrusted cloud environment," *Complex Intell. Syst.*, vol. 9, no. 3, pp. 2915–2928, 2023, doi: 10.1007/s40747-021-00305-0.