ARTIFICIAL INTELLIGENCE FOR HYBRID CLOUD SECURITY APPLICATIONS

Abstract

One of the newest study topics today is security in cloud computing. They fall into two categories: private cloud and public cloud, depending on their features and services. In this situation, these two categories of cloud services are necessary for a company to offer society a better service. In order to do this, a brand-new hybrid cloud solution that combines both private and public clouds was introduced. In the modern world, security in the cloud environment is a difficult problem, especially for the hybrid cloud because of the integration of both. Artificial intelligence (AI), which is a set of computer algorithms that improves natural reality, serves this function by securing hybrid cloud networks while keeping data there and accessing it from the cloud. For redundant or duplicate-free secure storage and retrieval, a Support Vector Machine (SVM) deduplication processing algorithm is also described. A new access control system and the use of RSA (Rivest–Shamir–Adleman) encryption are also present. When storing, accessing, and extracting data from the cloud database, the hybrid cloud's security has been tested while the suggested security architecture has been implemented.

Keywords: AI, SVM Deduplication, RSA, Hybrid Cloud, Data Collection

Authors

Srinivas Dava
Associate Professor
Department of Computer Science and Engineering
Jyothishmathi institute of technology and science
karimnagar 505481, Telangana, India.
srinivasdava450@gmail.com

D. F. Jingle Jabha
Professor
Department of Electrical and Electronics Engineering
SRM TRP Engineering College
Tiruchirapalli, Tamilnadu, India.
jinglejabha@gmail.com

Shaji. K. A. Theodore
Faculty of Information Technology - Networking
Department of Information Technology
UTAS MUSANNA, Oman.
theodore7733@hotmail.com

Arul V.H
Associate professor
Department of Electronics and Computer Engineering
Thejus Engineering College,
Kerala-680584. vharul@gmail.com

I. INTRODUCTION

In a variety of industries, including medicine, commerce, and education, the cloud computing environment offers a tremendous amount of assistance to the world [1]. The hybrid cloud is a distinct sort of cloud computing technique [2] because it is possible to access the capabilities of both private and public clouds. Industries today are gradually becoming more and more dependent on hybrid clouds in order to provide their employees with a flexible solution for accessing data and seeing information by individuals. Organisations can use the hybrid computing environment to increase their flexibility and safeguard their on-site corporate data from firewalls. This hybrid cloud is enabling businesses to increase their computing capabilities while also removing the need for exorbitant outlays to handle demand spikes that last only a short while [3]. Additionally, the hybrid cloud offers all of the advantages of the cloud, including flexibility, scalability, and cost savings, with the least amount of danger to your data [4]. Despite the numerous dangers involved in offering all of these amenities to an organization's employees and end consumers.

The services that are provided around the world must have security. Data security is a key factor in cloud network environments [5]. Data encryption and user access control are two of the six security threats that are most harmful. A variety of encryption methods can be used in this situation to encrypt data and convert text into cipher-text, an encrypted version of the input that is inaccessible to unauthorised users. The decryption technique for encrypted data additionally utilises a separate key, which activates it and grants the approved user access to the original text [6, 7]. One of the major elements of hybrid cloud security is the ability to control users through knowledge of their login credentials and the users who have permission to access the cloud data that is kept in the database in the cloud.

A well-known method for scalable and efficient data management in cloud computing is the deduplication process [8]. Recently, this strategy has drawn increased interest from both cloud users and the general public. Data duplication is the removal of unnecessary copies of data from cloud storage by the use of data compression [9]. It also improves storage utilisation. When ensuring data secrecy, the common encryption method clashes with the data deduplication method.

In order to make the deduplication process infeasible, the convergent encryption approach is used [10]. By boosting the performance of the processors, the memory and, hybrid cloud also offers a variety of services to the cloud applications. Basic networks, traditional databases, and operating systems with adaptable operating systems for hybrid clouds all used data storage and retrieval algorithms. In this case, modifying the data access restriction is required to protect the data from malicious users. Even still, these fundamental procedures are insufficient for storing and retrieving data in cloud applications that have a sizable amount of data scattered across numerous locations. This research introduces artificial intelligence for hybrid data cloud security to address these drawbacks. The AI will be useful in further strengthening cloud data security.

In this study, it is suggested that AI be used to secure the hybrid cloud while data is being stored, retrieved, or accessed from cloud databases. Additionally, a deduplication procedure based on SVM is developed to prevent data duplication during retrieval as well as safe data storage. For its secure encryption process, RSA is employed the suggested security architecture also uses a newly developed dynamic access control technique.

The remainder of the paper is structured as follows: Section 2 contains the thorough survey. The overall architecture of the suggested system is described in Section 3. The suggested security framework is described in Section 4. Section 5 displays the experimental findings and how well they performed. The conclusion appears in Section 6.

II. RELATED WORKS

Mohamed Goudjil *et al* (2018) suggested a brand-new active learning technique for classifying texts. The basic goal of active learning is to intelligently choose which samples should be labelled in order to decrease the labelling effort without sacrificing classification accuracy. Using a collection of multi-class SVM classifiers to produce posterior probabilities, the suggested technique chooses a batch of informative samples, which are then manually labelled by a subject matter expert. According to experimental findings, the suggested active learning strategy greatly decreases the labelling effort while also improving classification accuracy.

Meng Hao *et al* (2019) presented a privacy improved and effective federated learning (PEFL) system for commercial AI. In comparison to current methods, the proposed PEFL can stop privacy leaking from both local gradients and shared parameters. PEFL offers a high level of safeguarding our privacy even with numerous trustworthy entities because it is non-interactive in each aggregation. Performance analysis shows that PEFL has real-world accuracy, and in the near future, research is anticipated on high-dimensional datasets and complicated neural networks.

Al-Juaid *et al* (2019) proposed an improved solution for protecting sensitive text data on home computers that would use both steganography and cryptography. The system security is produced by using audio-based steganography and RSA cryptography as two sequential layers, ensuring the best security possible while obtaining the benefits of both. To examine the relationship between security, capacity, and data reliance, the study modelled the system and put it to the test. Future work will involve modifying the crypto layer to test other symmetric methods.

O. F. A. Wahab *et al* (2021) employed a hybrid data compression algorithm that may be used for both lossy and lossless compacting Steganography methods. By reducing the amount of data sent each time, this strategy can help with quick transmission over a sluggish internet connection or take up little room on various storage devices. Future implementations of this concept could improve cloud service security through the use of artificial intelligence approaches.

Jinbo Xiong *et al* (2019) provided a solution that used there-encryption algorithm to successfully complete permitted deduplication and the convergent encryption algorithm to prevent confidentiality data spilling. The management centre is also introduced in order to lower the client's computation costs and management overhead and to implement the dynamic updating of the authorised user's privilege. Finally, the security analysis shows how secure our suggested scheme is, and the performance evaluation proves how effective and efficient the proposed system is. However, there is a need to boost deduplication efficiency.

III. PROPOSED SYSTEM

Fig. 1 depicts the proposed system's general architecture. Documents, a cloud database, a user interface module, deduplication, and access control are among the five main parts.

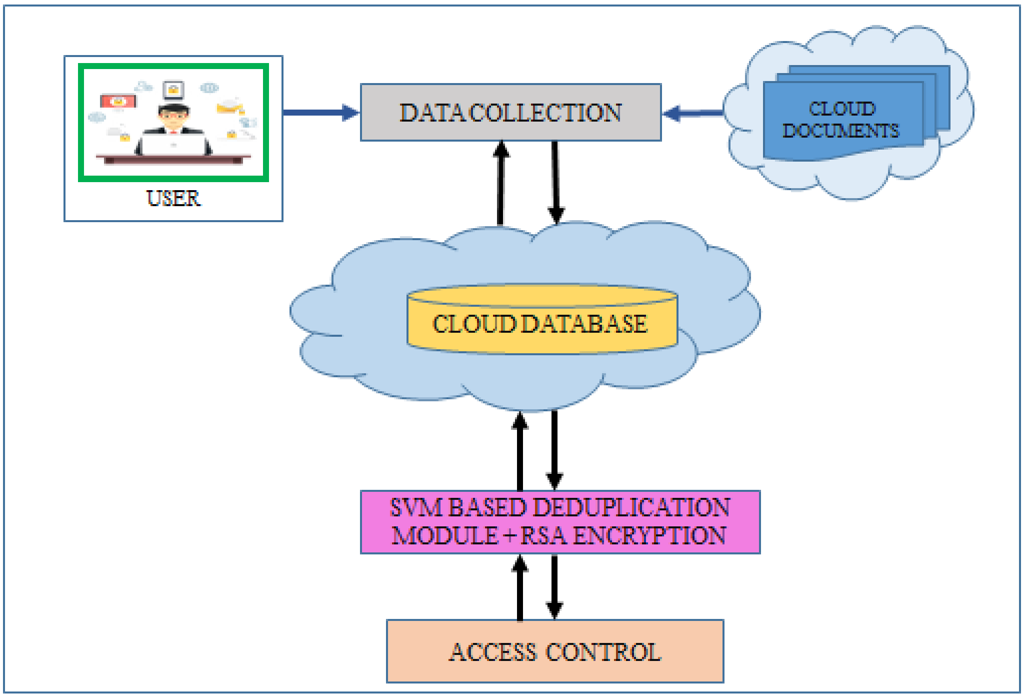

Figure 1: Proposed Hybrid Cloud System

The data collection module has gathered and stored in cloud databases both the organization's documents and the data of cloud users. The encrypted de-duplication module reduces data duplication and secures the information that is already there and kept in the cloud database. The user interface module will store the encrypted material into the cloud database without redundant storage. The access control system is in charge of limiting cloud users.

1. **Hybrid Cloud :** A hybrid cloud consists of two or more cloud deployment models that are connected so that data can be transported between them without interfering with one another. These clouds are commonly developed by corporations, and both the organisation and the cloud provider share management responsibilities. A business can use this model to describe the needs and goals for the services it provides. A well-designed hybrid cloud can be advantageous for both security-related operations, like processing client payments, and business-related operations, such processing employee paychecks. Interactions between private and public components could impede implementation. It is necessary to aggregate and use services from various sources as

though they were from a single source. These clouds, which could be private, communal, or public, are connected by a specialised or widely used technology that enables the transfer of data and applications between them.

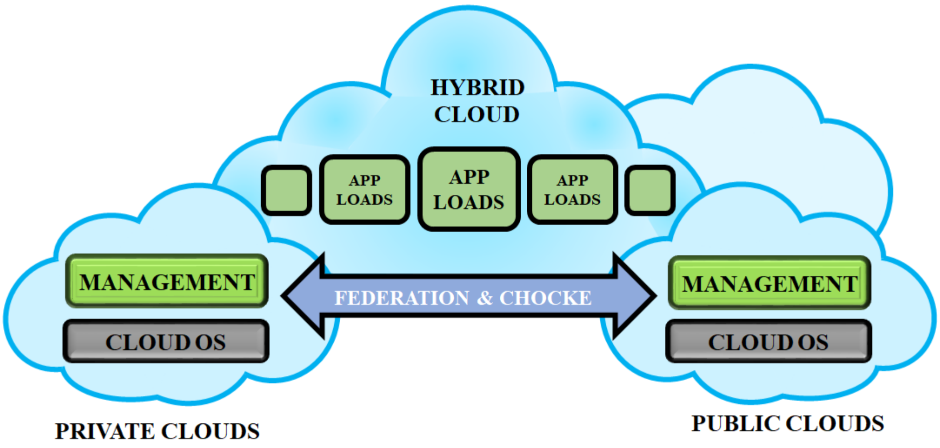

Figure 2: Hybrid Cloud

Figure 1 depicts a hybrid cloud as a solution to the deduplication problem in distributed computing, with varying advantages. As a result, a protected duplication check is completed with a variety of advantages using symmetric encryption, proof of possession, and combination approach to handle. In order to prevent key sharing among clients, we managed the keys in a private cloud. Keeping a safe distance from beast power assaults in daylight mists and performing a safe duplicate check were both accomplished using the Novel encryption key. To effectively address the deduplication issue in cloud computing with diverse benefits.

2. **SVM Based Deduplication:** A specific data compression technique called data deduplication is used to eliminate redundant copies of data from storage. The technique can be used to reduce the quantity of data that needs to be transported over networks and improve storage efficiency. Either at the file level or the block level, deduplication is possible. It gets rid of multiple copies of the same file when it comes to file base deduplication. Deduplication at the block level is a different option that removes duplicate chunks of data from non-identical files. An SVM Classifier for deduplication is part of the system under examination.

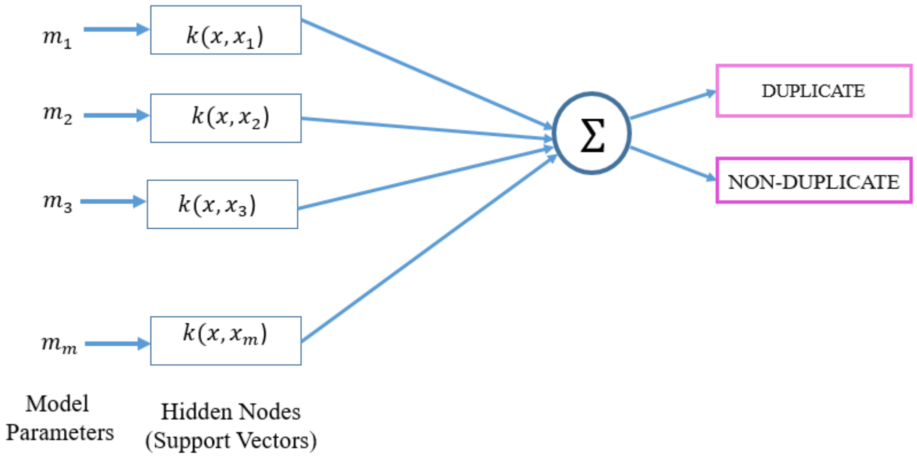

Figure 3: SVM for Deduplication

The SVM's design for deduplication purposes is as stated above. The SVM of the proposed deduplication method will yield the two outcomes Non-Duplicate and Duplicate. As seen in the graphic, the SVM model created for the suggested deduplication method. In order to identify the duplicate and deduplicated records in datasets and train the SVM classifier, we require specific data features. Following training with the data features, the classifier will determine if the provided data are duplication. The classifier has to be given the values once we have computed all the data features. These findings allow us to train the classifier to distinguish between duplicate and non-duplicate records in the sample. We can provide a fresh record after the SVM classifier has been trained to determine whether it contains duplicate or non-duplicate records. Following comparison, the SVM classifier will determine whether the provided MRI image falls within the duplication category or not, and it will inform us of its findings.

3. **RSA Encryption:** Traditional encryption does not work with data deduplication while maintaining data confidentiality. In particular, traditional encryption requires that many users encrypt their data using different keys. Therefore, deduplication is not possible since copies of the same data made by different users would have different ciphertexts. It has been proposed to enforce data privacy while enabling deduplication using convergent encryption. It generates a data copy by computing the value of the data copy's content and then encrypts or decrypts it using a convergent key. Identical data copies will produce the same overlapping key and ciphertext since the encryption process is predictable and depends on the data content. After the proof, users who possess the same file will not need to upload the exact same file, but rather will receive a pointer from the server. Only the owners of the pertinent data may decode an encrypted file with a pointer that a user can download from the server using their convergent keys.

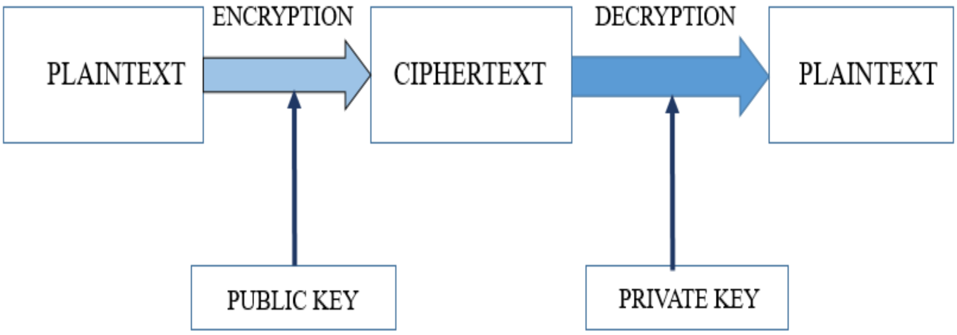

Figure 4: RSA Algorithm

The algorithm builds the publickey and the privatekey using two enormous prime values. The RSA algorithm has been considered as a potential form of authentication. The working public key of the RSA algorithm is made available to everyone, but the private key is kept confidential. For its secure encryption process, RSA is well-known. Positive integer prime integers employed in exponential fashion for encryption and decryption are the foundation of RSA's operation.

4. **Cloud Database:** The data on several data centres situated at various locations is stored in the cloud database. A cloud database has several nodes spread across it that are intended for query services for corporate data centres as well as data centres located in various geological areas. This linkage is necessary for the database to be easily and completely accessible via cloud services. There are various ways to access the database through cloud services; a user can do so using a computer and the internet, or they can use a mobile phone and 3G or 4G services to do so.

IV. RESULTS AND DISCUSSION

This section contains the findings of the experiments conducted to assess the effectiveness of the record deduplication method we suggested. There is a demonstration of the encryption and decryption times.

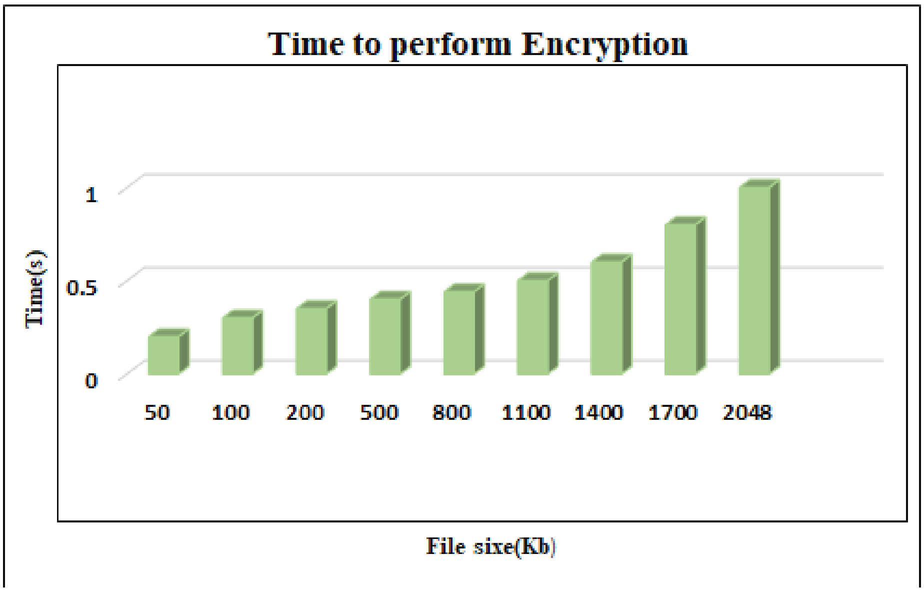

Figure 5: Encryption Time

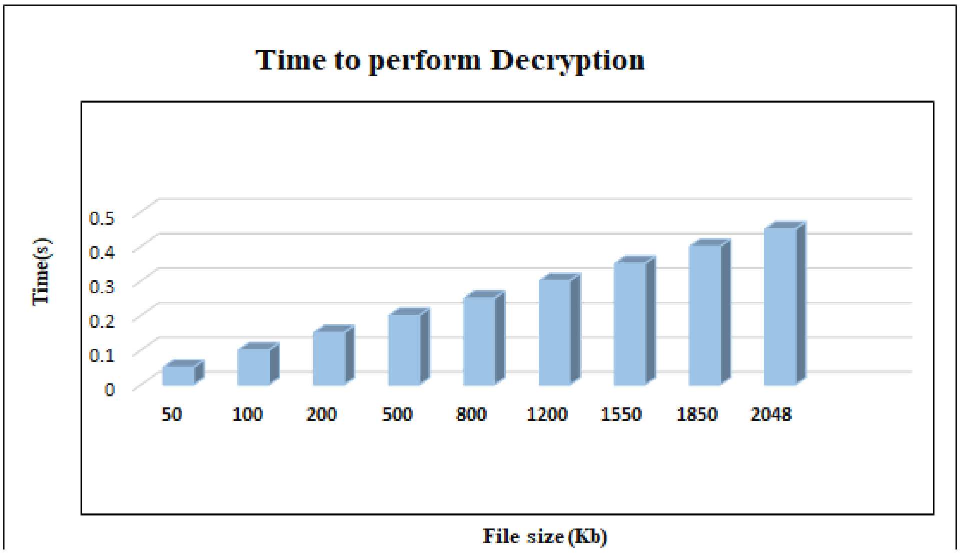

Figure 6: Decryption Time

Figure 5 and 6 depicts the amount of time necessary to encrypt and decrypt data, demonstrating that the amount of time required grows gradually with file size and reduces as the process progresses. The Encryption time and Decrytion time for the maximum file size 2048kb is 1s and 0.45s.

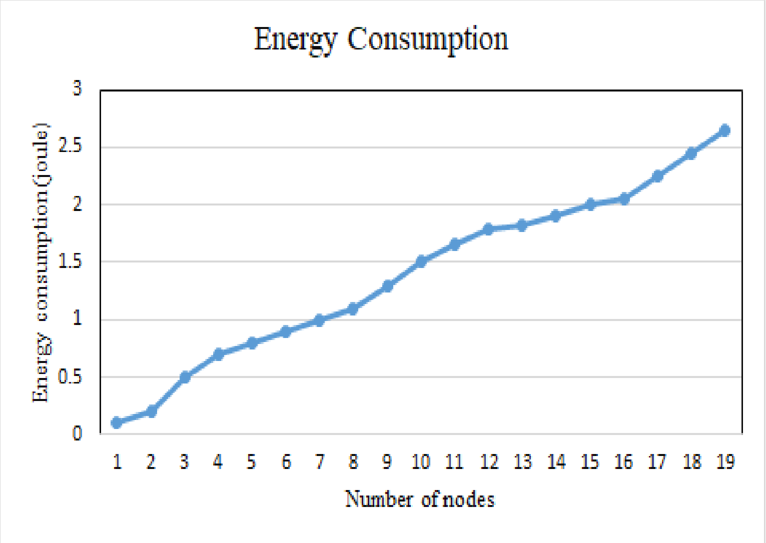

Figure 7: Energy Consumption Ratio

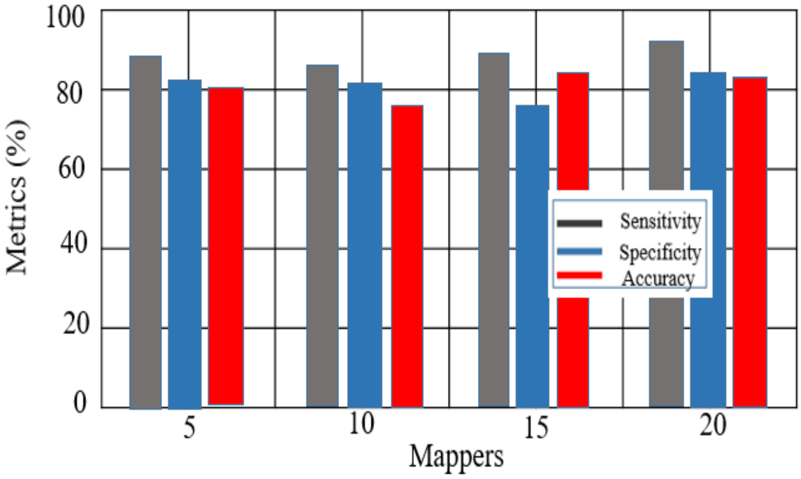

Figure 8: Comparison of SVM Deduplication Performance Metrics

Figure 7 shows the energy consumption ratio, which compares the number of nodes to the ratio's rising energy consumption. Figure 8 depicts the data deduplication process using mappers versus SVM. To evaluate a system's effectiveness, one might distinguish between sensitivity, specificity, and accuracy. The accuracy, sensitivity, and specificity of the suggested approach for mappers were 80.02%, 84.0%, and 82.0%, respectively.

V. CONCLUSION

With the storage and access of data from cloud databases, this work presents a novel AI application for safeguarding hybrid cloud networks. They incorporated a recommended SVM-based deduplication procedure approach to avoid redundant data and to offer safe storage and retrieval in their application. The suggested AI programme took user privacy into account while designing it and assessing the security of hybrid cloud networks used to store, retrieve, and access data from cloud databases. By offering data protection and limiting data access, this proposed application's key benefit is that it lowers security concerns in hybrid clouds. The accuracy, sensitivity, and specificity of the suggested approach for mappers were 80.02%, 84.0%, and 82.0%, respectively. The encryption and decryption times for files up to 2048 kilobytes are respectively 1 second and 0.45 second.

REFERENCES

[1] L. Zhang, Y. Cui and Y. Mu, "Improving Security and Privacy Attribute Based Data Sharing in Cloud Computing," in IEEE Systems Journal, vol. 14, no. 1, pp. 387-397, March 2020.

[2] Yu, Han, Xiuqing Lu, and Zhenkuan Pan. "An authorized public auditing scheme for dynamic big data storage in cloud computing." *IEEE Access* 8 (2020).

[3] Asif, Muhammad, Sagheer Abbas, M. A. Khan, Areej Fatima, Muhammad Adnan Khan, and Sang-Woong Lee. "MapReduce based intelligent model for intrusion detection using machine learning technique." *Journal of King Saud University-Computer and Information Sciences* (2021).

[4] J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun and Y. Xiang, "Block Design-Based Key Agreement for Group Data Sharing in Cloud Computing," in IEEE Transactions on Dependable and Secure Computing, vol. 16, no. 6, pp. 996-1010, 1 Nov.-Dec. 2019.

[5] Y. Hou, S. Garg, L. Hui, D. N. K. Jayakody, R. Jin and M. S. Hossain, "A Data Security Enhanced Access Control Mechanism in Mobile Edge Computing," in IEEE Access, vol. 8, pp. 136119-136130, 2020.

[6] Saračević, Muzafer H., Saša Z. Adamović, Vladislav A. Miškovic, Mohamed Elhoseny, Nemanja D. Maček, Mahmoud Mohamed Selim, and K. Shankar. "Data encryption for Internet of Things applications based on catalan objects and two combinatorial structures." *IEEE Transactions on Reliability* 70, no. 2 (2020): 819-830.

[7] Abd El-Latif, Ahmed A., Bassem Abd-El-Atty, Wojciech Mazurczyk, Carol Fung, and Salvador E. Venegas-Andraca. "Secure data encryption based on quantum walks for 5G Internet of Things scenario." *IEEE Transactions on Network and Service Management* 17, no. 1 (2020): 118-131.

[8] Tyj, Naga Malleswari. "Adaptive deduplication of virtual machine images using AKKA stream to accelerate live migration process in cloud environment." *Journal of Cloud Computing* 8, no. 1 (2019): 3.

[9] Y. Fu, N. Xiao, H. Jiang, G. Hu and W. Chen, "Application-Aware Big Data Deduplication in Cloud Environment," in IEEE Transactions on Cloud Computing, vol. 7, no. 4, pp. 921-934, 1 Oct.-Dec. 2019, doi: 10.1109/TCC.2017.2710043.

[10] X. Yang, R. Lu, J. Shao, X. Tang and A. A. Ghorbani, "Achieving Efficient and Privacy-Preserving Multi-Domain Big Data Deduplication in Cloud," in IEEE Transactions on Services Computing, vol. 14, no. 5, pp. 1292-1305, 1 Sept.-Oct. 2021, doi: 10.1109/TSC.2018.2881147.

[11] Goudjil, Mohamed, Mouloud Koudil, Mouldi Bedda, and Noureddine Ghoggali. "A novel active learning method using SVM for text classification." *International Journal of Automation and Computing* 15 (2018): 290-298.

[12] M. Hao, H. Li, X. Luo, G. Xu, H. Yang and S. Liu, "Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence," in IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6532-6542, Oct. 2020, doi: 10.1109/TII.2019.2945367.

[13] Al-Juaid, Nouf, and Adnan Gutub. "Combining RSA and audio steganography on personal computers for enhancing security." *SN Applied Sciences* 1 (2019): 1-11.

[14] O. F. A. Wahab, A. A. M. Khalaf, A. I. Hussein and H. F. A. Hamed, "Hiding Data Using Efficient Combination of RSA Cryptography, and Compression Steganography Techniques," in IEEE Access, vol. 9, pp. 31805-31815, 2021, doi: 10.1109/ACCESS.2021.3060317.

[15] J. Xiong, Y. Zhang, S. Tang, X. Liu and Z. Yao, "Secure Encrypted Data With Authorized Deduplication in Cloud," in IEEE Access, vol. 7, pp. 75090-75104, 2019, doi: 10.1109/ACCESS.2019.2920998.