DS Journal of Multidisciplinary
Volume 1 Issue 1, 23-30, Jan - Mar 2024
ISSN: XXXX-XXXX

Original Article

IoT-Centric Data Protection Using Deep Learning Technique for Preserving Security and Privacy in Cloud

[1]J. Abitha, [2]R. Sadhana

[1,2] *Department of Information Technology, St. Joseph's College (Autonomous), Tamil Nadu, India.*

[1]abithajesuraj2000@gmail.com

Received: 21 October 2023; Revised: 20 November 2023; Accepted: 13 December 2023; Published: 22 January 2024;

Abstract - A system of interconnected, the Internet of Things (IoT) is a term that refers to physical objects that may be connected online. Concerns over user privacy on the Internet of Things are growing as a result of the large amounts of personal data being gathered and exchanged there. IoT devices may increase productivity, accuracy, and financial gain in addition to reducing human intrusion, giving Internet of Things applications the most flexibility and convenience. Overhead of communications, security, and privacy, IoT is experiencing issues as well. As a result, protecting data has grown to be a difficult undertaking that must be handled carefully. This study offers a secure data security solution for preserving privacy in the cloud environment to address this crucial and difficult topic. It does this by efficiently separating the data by separating sensitive data from non-sensitive data with an SVM classifier and then using the data to partially decrypt and analyze, which increases the effectiveness of the model while ensuring security. The sensitive data was protected using Okamato Uchiyama encryption. The model safely stores, analyzes, and shares data to ensure the system's safety and privacy. The novel method was compared to existing methodologies regarding particular parameters like precision, accuracy, F1 score, and recall.

Keywords - Cloud computing, Internet of Things, Support Vector Machine, Okamato Uchiyama.

1. Introduction

Internet of Things (IoT) devices, which generate enormous amounts of data but lack both the storage and the processing capability to process it, are the foundation of the cyber-physical system. As a result, for various services, data must be moved from the on-premises platform to the cloud platform [1, 2]. The cloud platform has established itself as a top-tier method of storing, analyzing, and exchanging data with numerous stakeholders for the best results [3, 4]. However, it is not advisable to put your faith in a third-party cloud platform, particularly for preserving private information, as renting data to the cloud results in the gadget losing ownership of it [5].

An average data loss expense worldwide will be $4.35 million in 2022, up 2.6% and 12.7% from 2021 and 2020, respectively, according to research by IBM and Ponemon Institute [6]. These factors have made data protection a major problem, inspiring researchers to provide strategies for maintaining data privacy [7]. Most of the models that are now in use are based on encryption techniques [8-10] and differential privacy [11-14], but they

have a limit. According to research by IBM and the Ponemon Institute [6], the average cost of a data breach globally will be $4.35 million in 2022, an increase of 2.6% and 12.7% from 2021 and 2020, respectively.

These reasons have made data protection a significant issue, prompting researchers to offer techniques for preserving data privacy [7]. The majority of models now in use are based on encryption methods [8-11] as well as differential privacy [12-15], although their accuracy, utility, and efficiency might be improved. As far as the authors are aware, there are currently no models that adequately balance the accuracy and privacy of the data being outsourced. To address the aforementioned problem, this study presents a revolutionary safe data protection system known as SP-DPM, which allows secure data analysis, storage, and sharing in a cloud context. IoT devices create data, divide it up, and only encrypt the sensitive portions. The fog node partially decrypts the encrypted data provided by IoT devices. The Cloud Service Provider (CSP) classifies the acquired data before sharing the resulting data with the utility suppliers for decision-making. The following are the paper's main contributions:

- To provide better privacy and security for data generated by IoT devices, a novel data protection approach has been developed.
- In order to maintain privacy and speed up computing, a data partition method is implemented that divides the data into different groups depending on how sensitive it is.
- By applying partial decryption and data encryption techniques before classifying the data using a voting classifier, the created data analysis strategy increases efficiency while ensuring data privacy.
- Security analysis is done to assess and validate the effectiveness of the suggested model, which has been tested using a variety of data sets. The outcomes are compared to state-of-the-art works of art through several performance indicators.

Section 3 presents the proposed framework. The experimental setting is demonstrated in Section 4. The collected results are then detailed in Section 5. With a conclusion and future work, the paper is completed.

2. Literature Survey

Zhang et al. [15] proposed the HP-CP-ABE hidden access policy. The CP-ABE scheme is used to protect and ensure data security and verify authorized users. The writers also developed an identification strategy to confirm the legitimacy of the users and carry out the decryption operation. Regardless of the user's characteristics, this technique guarantees a private key of a fixed size and reduces both transmission and storage costs.

J. Li et al., [16], the proposed method solely backs the "AND" policy, making it a lax security plan. A strategy for outsourcing that updates the matrix accessibility structure of the linear secret-sharing system is the CP-ABE method. It improves the effectiveness of the policy and dynamically updates the files in the cloud computing environment. The suggested plan lowers proxy CSPs' compute costs and the owners' storage and communication expenses.

Guo et al. [17] suggested the MA-CP-ABE method for the purpose of encrypting the hierarchical PHR, the suggested approach combines a number of several entry points into a single structure. The collusion attack hinders (N-1) vulnerable organizations out of N authorities even though it lacks a single trustworthy source of information.

Chai et al. [18] presented a privacy-protecting outsourcing classification system. The authors employed an alternative OU cryptosystem to reduce bandwidth consumption. The suggested system has limited data sharing and is immune to an attack using substitution and comparison.

Wei et al. [19], The privacy of learning model parameters is protected via a system for Notifying before model Aggregation Federated Learning (NbAFL). A differential privacy approach was used by the authors to add noise to the clients' local settings. Insufficient data exchange, however, plagues this paradigm.

3. Proposed Methodology

A malicious utility provider might take data that has been outsourced from the cloud, store it, analyze it, and share it with the parties involved to gather private information that could be abused. As a result, protecting data has grown to be a difficult undertaking that must be handled carefully. This paper provides a safe data protection technique for preserving confidentiality in a cloud context to address this crucial and difficult problem. It does so by effectively separating the data into sensitive and non-sensitive categories using an SVM classifier, partially decrypting the data, and performing data analysis that increases the model's effectiveness while maintaining security. Okamato Uchiyama encryption has been used to protect sensitive data. By carrying out safe data storage, analysis, and exchange, the model guarantees the security and privacy of the system. Specific criteria, including precision, accuracy, F1-score and recall, have been used to compare the suggested method to the existing methodologies. The general block diagram for the suggested work is displayed in Figure 1.

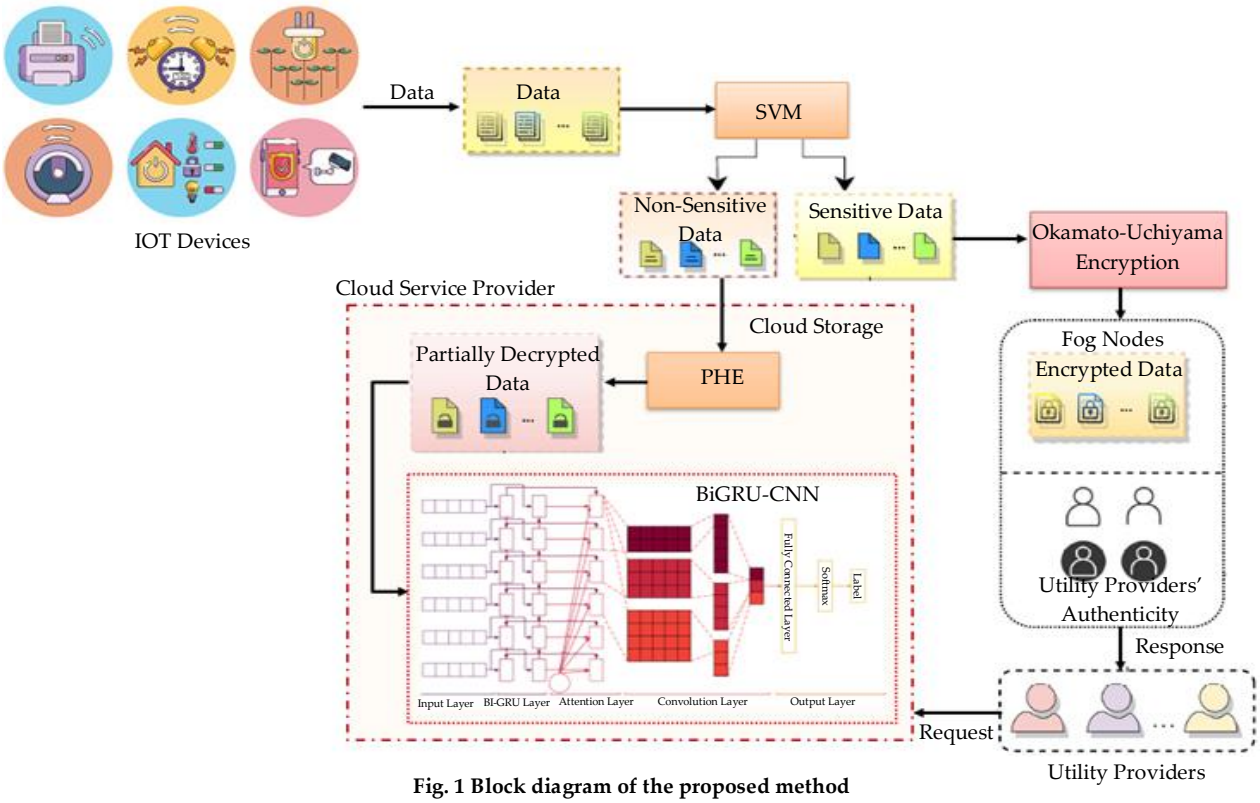

Fig. 1 Block diagram of the proposed method

3.1. Support Vector Machine

Support Vector Machines (SVMs) are a subclass of supervised machine learning methods that convert difficult, highly non-linear situations into binary classification models [5]. The SVM must construct the decision surface, which is a hyperplane, utilizing data samples to maximize the margin around it. During the training stage of the algorithm, each data sample is assigned a class designation and the projected value.

The data sample contains what are referred to called characteristics; these are the variables in the data that specify the data sampling vector's activity. The weights applied to each input feature and a collection of support vectors that construct the ideal hyperplane are used to forecast the results of the SVM training phase. In contrast

to other neural networks, the SVM maximizes the number of nonzero weights while lowering the overall number of nonzero weights by maximizing the margin. These only match the important traits that provide information useful for selecting the hyperplane.

The kernel function modifies the data dimensions to define the hyperplane's form, which is a crucial step in the SVM. Simply put, the kernel function increases the hyperplane's size to help distinguish between the classes. Employing many kernel types, including the polynomial, linear, and both sigmoid and Gaussian radial basis functions, is possible. The type of data sample affects how each kernel performs. The simplest kernel, the linear kernel, performs better when applied to linear problems. The supplied characteristics are combined by the polynomial, RBF, and sigmoid kernels to produce support vectors. They work best with non-linear data, but how complex they are depends on how many additional features they find. In Figure 2, the Support Vector Machine (SVM) flowchart is displayed.

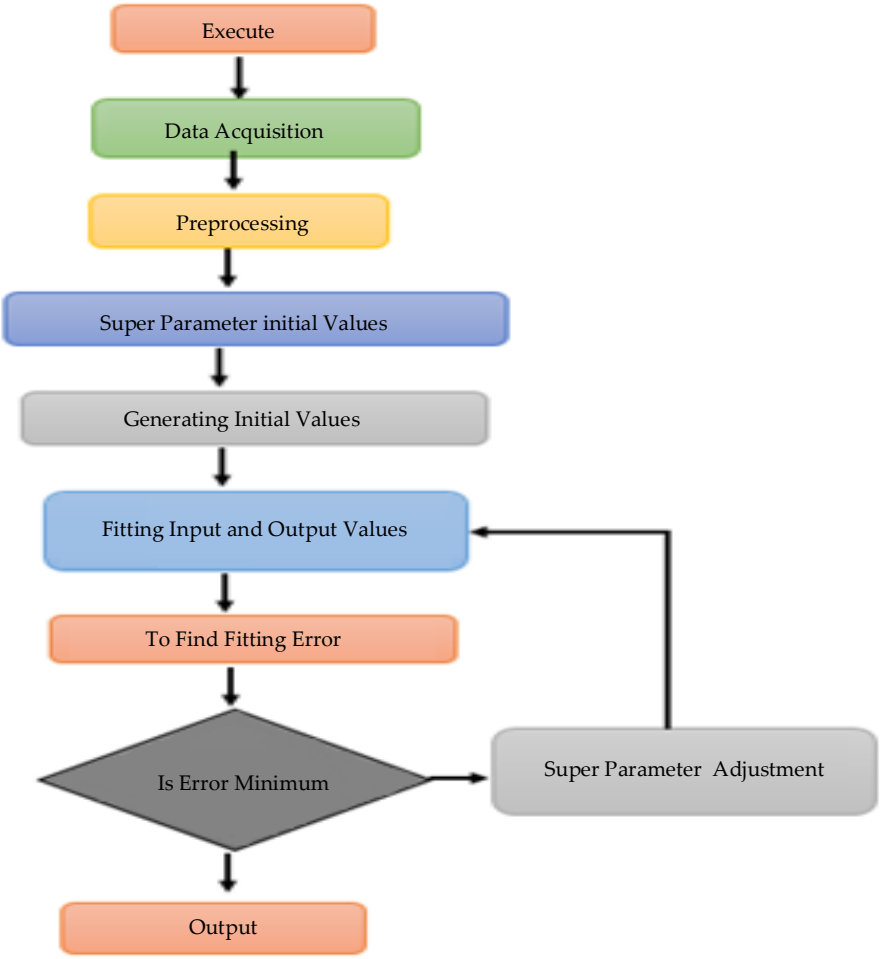

Fig. 2 Flow chart of Support Vector Machine (SVM)

3.2. Okamato Uchiyama Encryption

Numerous concepts from number theory, discrete mathematics, and abstract algebra are used in the Okamato-Uchiyama cryptosystem. Numerous of these ideas are fundamental and applied in various cryptography areas. However, despite the fact that beyond integer computations, there is no need to provide detailed or rigorous treatments, those essential notions are still important but are not sufficiently covered in mathematics curricula in underdeveloped and developing countries; by examining the fundamental concepts and mathematics used in the Okamoto-Uchiyama algorithm.

3.2.1. Key Generation

The following procedure generates a public/private key:

1. Create the two big primes, A and B.
2. Compute N=B^2A
3. Select an integer with a random value G2 n-1 such that G(B-1)
4. Computes H=GN mod N.

Next, we have (N, G, H) as the public key and (B, A) as the private key.

3.2.2. Encryption

Using the public key (B, A), the following can be done to encrypt a message M < B.

1. A random number R1n-1 should be chosen so that G(B-1)
2. Compute C=GN HR mod N.

The value C is the encryption of M.

3.2.3. Decryption

An encrypted message C can be decrypted with the private key (B, A) as follows:

1. Compute P= l (C^{B-1}mod B^2)
2. Compute Q= l (G^{B-1}mod B^2), P and Q will be integers.
3. Using the extended Euclidean technique, compute the opposite value of Q modulo B. $Q' = Q^{-1}\ mod\ B$
4. Compute M=$PQ'mod\ B$

The value M is the decryption of C.

4. Result and Discussion

As shown in the next sections, we tried a number of experiments in this work to address the privacy issue using deep learning algorithms. By carrying out safe data storage, analysis, and exchange, the model guarantees the security and privacy of the system. Specific criteria, including precision, accuracy, F1-score and recall, have been used to compare the suggested method to the existing methodologies.

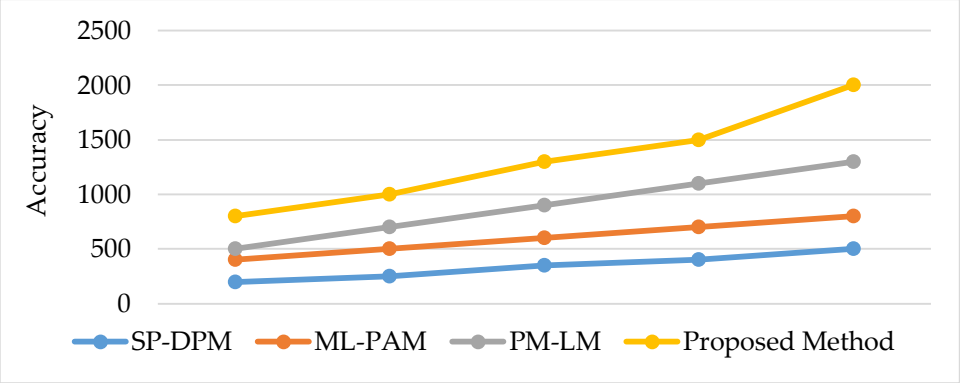

Fig. 3(a) Accuracy graph

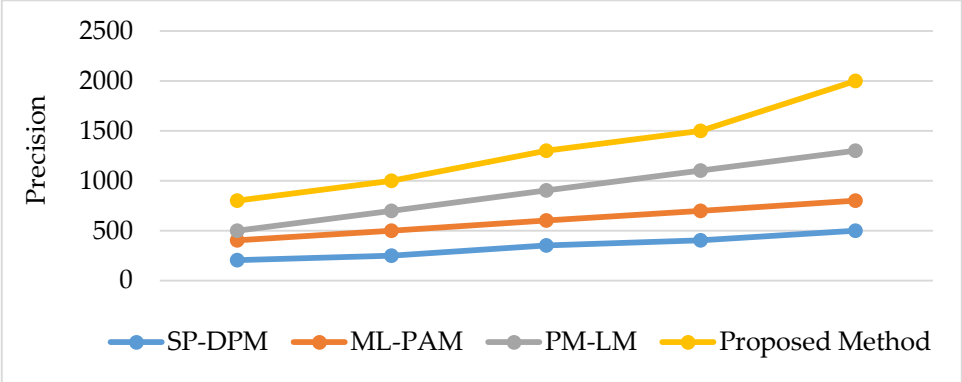

Fig. 3(b) Precision graph

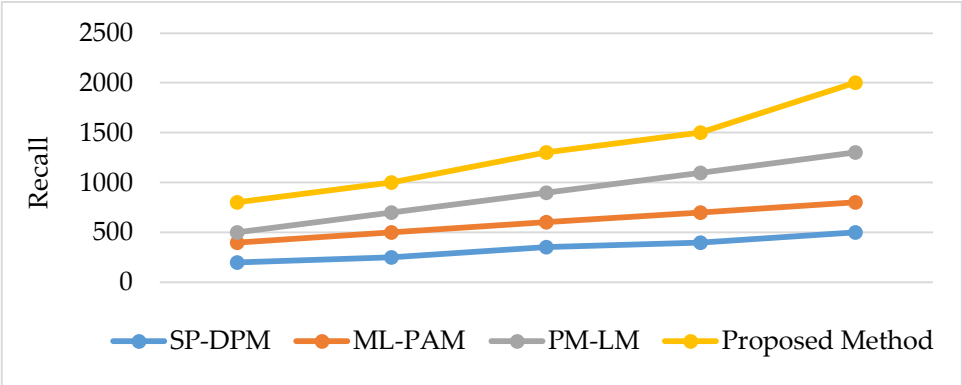

Fig. 3(c) Recall graph

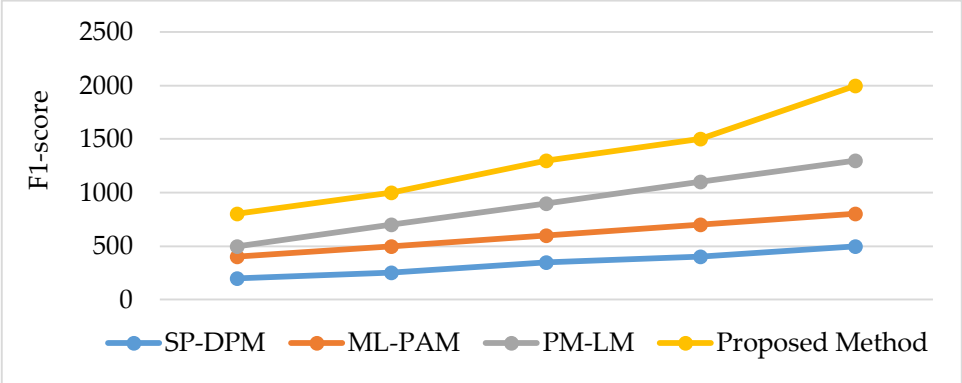

Fig. 3(d) F1-score graph

Figure 3 (a,b,c,d) illustrates the findings, demonstrating that all deep learning methods offer high evaluation metrics. In contrast, the averages for benign are 99.92%, 98.85%, and 99.90% for precision, recall, and F1-score.

5. Conclusion

This paper provides a safe data security method for protecting privacy in cloud computing settings to address this crucial and difficult topic. It does this by efficiently separating the data by separating sensitive data from non-sensitive data with an SVM classifier and then using the data to partially decrypt and analyze, which increases the effectiveness of the model while ensuring security. The sensitive data was protected using Okamato Uchiyama encryption. The model safely stores, analyzes, and shares data to ensure the system's safety and

privacy. The novel method was compared to existing methodologies in terms of particular parameters like precision, accuracy, F1 score, and recall.

References

[1] Ashutosh Kumar Singh, and Ishu Gupta, "Online Information Leaker Identification Scheme for Secure Data Sharing," *Multimedia Tools and Applications*, vol. 79, pp. 31165–31182, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ishu Gupta et al., "Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, and Future Directions," *IEEE Access*, vol. 10, pp. 71247–71277, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Deepika Saxena et al., "A Fault Tolerant Elastic Resource Management Framework toward High Availability of Cloud Services," *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 3048–3061, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Rishabh Gupta et al., "Quantum Machine Learning Driven Malicious User Prediction for Cloud Network Communications," *IEEE Networking Letters*, vol. 4, no. 4, pp. 174-178, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ishu Gupta, and Ashutosh Kumar Singh, "SELI: Statistical Evaluation Based Leaker Identification Stochastic Scheme for Secure Data Sharing," *IET Communications*, vol. 14, no. 20, pp. 3607–3618, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] IBM Ponemon, Cost of a Data Breach Study, 2022. [Online] Available: https://www.ibm.com/security/data-breach

[7] Ishu Gupta, and Ashutosh Kumar Singh, "Dynamic Threshold Based Information Leaker Identification Scheme," *Information Processing Letters*, vol. 147, pp. 69–73, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Xindi Ma et al., "PDLM: Privacy-Preserving Deep Learning Model on Cloud with Multiple Keys," *IEEE Transactions on Services Computing*, vol. 14, no. 4, pp. 1251–1263, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Ping Li et al., "Privacy-Preserving Outsourced Classification in Cloud Computing," *Cluster Computing*, vol. 21, no. 1, pp. 277–286, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Maoguo Gong, Jialun Feng, and Yu Xie, "Privacy-Enhanced Multi-Party Deep Learning," *Neural Networks*, vol. 121, pp. 484–496, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Rishabh Gupta et al., "Differential and Triphase Adaptive Learning-Based Privacy-Preserving Model for Medical Data in Cloud Environment," *IEEE Networking Letters*, vol. 4, no. 4, pp. 217-221, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ping Li et al., "Privacy-Preserving Machine Learning with Multiple Data Providers," *Future Generation Computer Systems*, vol. 87, pp. 341–350, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[13] Rishabh Gupta et al., "A Differential Approach and Deep Neural Network Based Data Privacy-Preserving Model in Cloud Environment," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 4659-4674, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ishu Gupta et al., "MLPAM: A Machine Learning and Probabilistic Analysis Based Model for Preserving Security and Privacy in Cloud Environment," *IEEE Systems Journal*, vol. 15, no. 3, pp. 4248–4259, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Shulan Wang et al., "An Efficient File Hierarchy Attribute-Based Encryption Scheme in Cloud Computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1265–1277, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[16] Hai Liu et al., "A Fair Data Access Control Towards Rational Users in Cloud Storage," *Information Sciences*, vol. 418, pp. 258–271, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Zechao Liu et al., "Practical Attribute-Based Encryption: Outsourcing Decryption, Attribute Revocation and Policy Updating," *Journal of Network and Computer Applications*, vol. 108, pp. 112–123, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[18] Hong Zhong et al., "Multi-Authority Attribute-Based Encryption Access Control Scheme with Policy Hidden for Cloud Storage," *Soft Computing*, vol. 22, no. 1, pp. 243–251, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Leyou Zhang, Yilei Cui, and Yi Mu, "Improving Security and Privacy Attribute-Based Data Sharing in Cloud Computing," *IEEE Systems Journal*, vol. 14, no. 1, pp. 387–397, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jianqiang Li et al., "An Efficient Attribute-Based Encryption Scheme with Policy Update and File Update in Cloud Computing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6500–6509, 2019. [CrossRef] [Google Scholar] [Publisher Link]