International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

2023IJIEMR.Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 06th Feb 2023. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=ISSUE-02

DOI: 10.48047/IJIEMR/V12/ISSUE 02/29

Title **A Survey on Storage Distribution Approaches and Secure Data Storage for Protecting Data in Cloud Platform**

Volume 12, Issue 02, Pages: 180-187

Paper Authors
Prof R Ramesh, Dr Vekata Kishore Kumar Rajeti, Dr G Rajesh Chandra, N Hari Krishna

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

A Survey on Storage Distribution Approaches and Secure Data Storage for Protecting Data in Cloud Platform

Prof R Ramesh[1], Dr Vekata Kishore Kumar Rajeti[2], Dr G Rajesh Chandra[3], N Hari Krishna[4]
[1]Professor, CSE Dept,
KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.
Mail ID: repudiramesh@gmail.com
[2]Associate Professor, CSE Dept,
KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.
Mail ID: mail2rvkk@gmail.com
[3] Professor, CSE Dept,
KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India. Mail ID: grajeshchandra@gmail.com
[4]Asst Professor, CSE Dept,
KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India
harimtech2012@gmail.com

Abstract
At present real time requires the least amount of upfront capital investment, and has the greatest amount of scalability and many other benefits, the cloud service is being adopted by a huge number of academic institutions, government agencies, and business organisations. The cloud platform supports a variety of capabilities, but it also has a number of drawbacks. The main issue in the field of record security and cloud computing is data protection. Many solutions had been put out to address this issue. A need to study, classify, and examine the sizable current body of work for examining the applicability of those solutions to address the needs arises because many of the existing solutions need complete examination. This article offers a comparative and systematic analysis as well as a thorough examination of the top methods for distributing and safeguarding data on the cloud platform. Every committed strategy is discussed in terms of its ability to protect data, modern solutions in the field, medium and good records, such as workflow, successes, scope, gaps, and future recommendations. roughly every resolution. The described strategies are also given a thorough and comparative review. The applicability of the solutions is then discussed in relation to the requirements, and the research gaps along with potential future directions are highlighted.

Index Terms: Cloud computing, information privacy, data security, data safety, information storage, information sharing, IoT, ML, cryptography.

Introduction

As it determines the uniqueness of each agency, data is recognised as the most crucial corporate asset. It is the main basis for knowledge, skill, and eventually the knowledge for wise choices and actions. It is maybe aiding in the treatment of a disorder, boosting business sales, improving building efficiency, or being in charge of fulfilling the objectives and enhancing performance [1]. Additionally, any business must provide the essential services of data storage, analysis, and sharing in order to improve performance [2]. However, organisations are under significant pressure to store the massive amounts of data domestically as a result of the information's fast evolution [3]. Additionally, because of constrained sources, statistics exploration has become challenging [4].

Because of the cloud's many benefits, like on- call carrier, scalability, dependability, elasticity, measurable services, data recovery, accessibility, and a host of others, the majority of organisations have switched to it for those services [5]. Cloud computing is a

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

paradigm that allows for large amounts of memory, large computing capacity at a reasonable price. Customers can access the desired services across different domains regardless of time, which offers a significant level of ease to cloud users [6]. Clients can get cost reduction and productivity improvements to handle project and form collaborations by moving the local data into cloud by using cloud-based services [7]. As a result, for a few services, people and businesses are migrating to the cloud more frequently [8]. It is not difficult to anticipate that most businesses would migrate to the cloud in the near future given the rapid advancement of cloud computing technologies.

Although cloud computing offers many benefits, it faces a number of obstacles that, if no longer properly addressed, could impede its rapid expansion [9]. Think a real-world solution where a company permits its team of employees to store and distribute the information over the cloud. By utilising the cloud, the company can totally free itself from the burden of managing and keeping the data locally [10],[11]. But it also faces a variety of security risks, which are cloud customers' top worries [12]. First of all, outsourcing data to cloud servers means that users lose control of their data, which causes them grief because the outsourced data may also identify sensitive and priceless data. Second, information sharing is frequently implemented in an open, competitive environment, and the Attackers decided to target cloud servers. In the worst-case scenario, client data may be illegally protected through the cloud server itself [13], [14]. In order to upgrade the results of the business, the information must also be communicated across fantastic relevant stakeholders, such as business partners, employees, clients, and many others, both inside and beyond the business's premises. The receiver party may, nonetheless, mishandle these details and unintentionally or wilfully divulge them to some unapproved third party [15], [16].

Because of limited storage and computing storage of the companies and the numerous benefits of clouds, Fig. 1 depicts a distributed platform in which the statistics owners must provide the corporation's priceless records to the cloud platform. Furthermore, in accordance with specific requirements for its application purpose, the cloud data is shared with a select group of users.

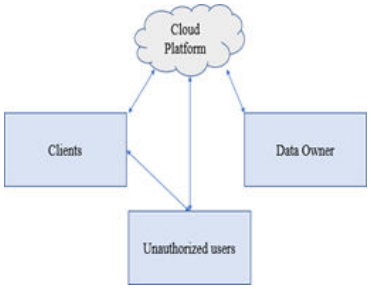

Fig 1: Data Sharing platform Block Diagram

However, after receiving the information, the recipient party could also reveal it. The information may also be misused by an unauthorised party through unauthorised access or may be released by the concerned parties or other third parties. Data loss or leakage could pose a serious threat to the confidentiality of the company. It can harm the reputation and goodwill of the company, lower the rank and position of the organisation, and lower shareholder prices [16]. Given the importance of data to a company, maintaining the security of this asset is crucial. There is a requirement for solutions that can effectively preserve the facts in a sharing setting.

For several packages, a variety of methods for statistics security in cloud environment have been investigated and developed. Typically, leakage prevention and leaker identification are used to ensure statistical safety, and this book focuses on achieving effective protection by halting leaking and identifying the hostile entity responsible for leakage as shown in Fig. 2. While leaker identification is specifically carried out using watermarking and probabilistic approaches, the major methods for preventing data leakage are tailored utilising encryption, access control mechanisms, and differential privacy with machine learning algorithms [17].

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

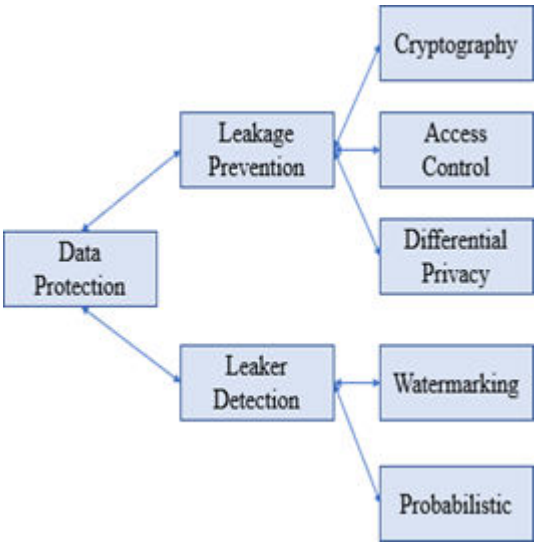

Fig 2: Types of Data Protection

Literature Review

[1]This study suggests a unique leaker identity technique that addresses the dynamic issue by managing the user demands in a common internet practice. The addition of a distribution plan lowers the risk of disclosing the information and increases the likelihood that the leaker will be found before the harmful person finds the statistics. The results show an improvement over the earlier work of up to 318%, 368% and 42% for average possibility, average fulfilment fee, average cutting charge, respectively. In addition, the suggested architecture completely reduces the possibility of data leakage by up to 88% and simultaneously gains 100% efficiency.

[2]Organizations now store, access, and share information differently thanks to cloud computing. Big data units are constantly shared among a hierarchy of several unique individuals with big data units being uploaded to the cloud on a regular basis. different privileges for access Finding a convenient and effective data access structure have become a major research challenge as more data storage needs migrate to the cloud. In order to solve the management and sharing of large data sets, the Privilege-based Multilevel Organizational Data-sharing scheme (P-MOD) is developed in this study. This system combines a privilege-based access form with a feature-based encryption method. As the number of users increases, the difficulty of designing hierarchies can be reduced thanks to our suggested privilege-based access structure, which makes it possible to manage healthcare information utilising mobile healthcare devices. Additionally, it can let organisations use extensive information analytics to roundly apprehend populations. Security analysis indicates that P-MOD is protected from an adaptively chosen plaintext attack, if the DBDH supposition is true. P-MOD is more efficient in computing complexity and storage space than the previous systems, according to comprehensive performance and simulation assessments using the genuine U.S. Census Income information set.

[3]On-demand data exchange amongst a few different entities is necessary to improve an agency's effectiveness. However, a small number of criminal actors have the ability to monitor this data to an uninvited third party, which could cause significant financial, reputational, and long-term balance losses for the establishments. A unique version of GUIM- SMD is provided by this study to help identify the guilty party who leaked information to an unauthorised party in a shared environment. An efficient This version proposes an efficient distribution technique to distribute the information to the users based on the access control mechanism. The approach introduces the summation matrix, which is computed using the D-rating and U-rating that are given to the classified facts and user, respectively. Additionally, D-rating and U-score, whose values range from 0 to 1, are entirely based on the file's data sensitivity and consumer responsibility. For the distribution of information among several clients, the evaluated summation matrix is used. In comparison to the prior work, the results demonstrate improvement up to 98.74, 236.38, and 252.39% for common possibility, average achievement fee, and detection charge, respectively.

[4] With the use of cloud storage options, consumers can upload their files remotely and comprehend information sharing. To ensure the integrity of the data saved in the cloud, remote facts integrity auditing is suggested. The cloud record might include certain sensitive information in

some common cloud storage systems, such as the digital fitness data machine. While the cloud document is shared, the sensitive data shouldn't be made public. Encrypting the entire shared record can decipher the sensitive information concealed, but it prevents other people from using it. It has not yet been investigated how to interpret data sharing with sensitive data concealing in distant data integrity audits. We suggest a remote data integrity auditing technique that realises data sharing with sensitive facts hidden on this article to address this issue. According to this approach, a sanitizer is utilised to sanitise information blocks that are identical to the report's sensitive records and converts their signatures into ones that are legitimate for the sanitised file. In the integrity auditing segment, these signatures are used to confirm the accuracy of the cleaned record. As a result, our system enables the cloud-saved document to be shared and utilised by others when the sensitive material is hidden, while still enabling efficient completion of the distant facts integrity auditing. The suggested approach, meanwhile, simplifies the challenging certificate management because it is based on identity-based cryptography. The suggested scheme is pleasant and environmentally friendly, as shown by the protection evaluation and overall performance assessment.

[5] One of the major challenges in the field of cloud computing, where information exchange is crucial between a few parties, is data leakage. An established Data Leaker Detection Model (DLDM) is provided in this study, allowing for the identification of the malevolent party responsible for information leaking. The method that is being suggested combines hashing, watermarking, and cryptography to secure the data. Additionally, the model identifies the poisonous customer by combining watermark extraction with possibility estimation. The findings show that when 200 documents totalling 20 MB are given to lone or wonderful people, it takes 3580 ms to find the malevolent person. For a burden cost of 2, the typical chance of identifying the poisonous user is 0:969518, which indicates an excessive likelihood of discovering the responsible agent. The performance of the suggested version is confirmed by the experimental findings.

[6] With the rise of cloud computing, mobile devices may access and retrieve personal data at any time and from any location. As a result, the problem with data security in cellular clouds is getting worse and preventing further development of cellular clouds. There has been a lot of research done to strengthen cloud protection. The majority of them, nevertheless, are irrelevant for mobile clouds because mobile devices only have a limited amount of computing power and electricity. For mobile cloud programmes, solutions with reduced computational overhead are incredibly necessary. We recommend a light-weight statistics sharing scheme (LDSS) for mobile cloud computing in this research. It adopts CP-ABE, an access control generation used in conventional cloud platforms, but modifies the access control tree's structure to make it suitable for mobile cloud environments.

[7] A challenge for the owner of the cloud is maintaining control over the facts while updating the security system. We recommend a tiered structure to decrease the overhead at the cloud provider associated with applying security to each report before sending it to the customer. With this technique, the sensitive document's protection and data privacy are maintained. The suggested technique classifies the information according to its sensitivity in order to strike a compromise between data protection and utility. Different algorithmic frameworks are required for the persistence of diverse categorizations. We set up a cloud-based platform where data is divided into four categories according to its level of sensitivity: public, exclusive, mystery, and top secret. A different method has been utilised to maintain security at each level. We added a mechanism to identify the problematic node that is responsible for information leaking at the levels that are the most sensitive, such as secret and top-mystery data. In order to evaluate the performance of the layered approach, an experimental evaluation is completed. The results of the experiment

demonstrate that the processing time (in ms) for 200 files with a 20 MB size is 437, 2239, 3142, and 3900 for When the documents are circulated to various users, they reveal public, exclusive, mysterious, and top-secret facts in that order, demonstrating the viability of the suggested strategy.

[8] Through the use of cloud technology, an increasing number of users and organisations have recently trusted proxy cloud carrier provider (PCSP) with the generation and storage of their data. According to this historical context, the characteristic-based encryption (ABE) mechanism is a replacement for traditional encryption that comfortably accommodates fine-grained policy and prevents collusion. When the get-in policy and report need to be revised in practical applications, there are some security issues that arise. Additionally, the ABE faces issues with excessive storage and calculation expenses. This paper suggests an effective cipher text-policy ABE method with file updates and coverage replaces in cloud. The first encryption's cipher text components may be shared while the coverage replacement and report update occur. It lowers the consumer's storage and communication costs as well as the PCSP's computing expenses. Additionally, it is demonstrated that the proposed system is agreeable below the notion of the decision q-parallel bilinear Diffie-Hellman exponent (BDHE). Last but not least, experimental modelling reveals that the suggested approach is significantly green in terms of record update and coverage replace.

[9] One of the most challenging issues that cloud computing faces is data protection. The cloud statistics are shared among many parties, and any agent can accidentally or purposefully reveal them to an unwanted recipient. Therefore, finding the malicious agent for hiding shared statistics has become necessary. In order to reduce the likelihood of further leakage, we provide a framework in this respect that is entirely based on probabilistic estimate and identifies the hostile agent. The distribution of the data among several sellers is accomplished in the suggested model by using 2-stage trees. When the records are leaked with the help of any agent, the parameters for malicious agent identification are computed entirely on the basis of chance. The experimental outcomes increased the average chance, average achievement charge, and average detection charge for the various types of agents to at least one, 0.98, and 0.76, respectively.

[10] Information integrity has drawn a lot of attention as a key security issue in reliable cloud storage. A verifier can accurately assess the accuracy of the outsourced records without downloading the data auditing methods. The complexity of key management is a significant research task related to present designs of records auditing techniques. In this research, we attempt to introduce fuzzy identification- based auditing—the first in such a technique, to the best of our knowledge—in order to address the challenging key management assignment in cloud information integrity checking. We focus more on the fundamentals of fuzzy identification-based information auditing, where a user's identity can be seen as a set of rigid descriptive characteristics.

[11] In today's rapidly evolving environment, there is a desire to exchange information with people inside or outside the firm, including sensitive financial information. For instance, businesses should share sensitive information with their partners, employees, and many other entities. The 0.33 party might be used to reveal this private information. Later, the distributor discovers the compromised data in a few illegal locations (such as via a criminal discovery process, on the user's hard drive, or online). We suggest a model that determines the likelihood that the data was stolen by one or more marketers or that it was independently gathered via some other way. The version's goal is to protect sensitive information by identifying the leaker responsible for data leakage and detecting leakage.

[12] Data sharing is a practical and affordable service made possible by cloud computing. Due to the fact that the information is outsourced to certain cloud servers, data contents privacy also

results from this. Different techniques are utilised to improve access control over the shared data in order to protect the sensitive and priceless information. Cipher text- policy attribute-based fully encryption (CP-ABE) could improve security and convenience in such methods. Traditional CP-ABE places a focus on data confidentiality, whereas at gift, protecting the customer's non-public privacy is a crucial issue. The use of CP-ABE with hidden access control provides both the security of the data and the privacy of the user. However, the majority of the currently used techniques are inefficient in terms of calculation cost and verbal interchange overhead. Furthermore, the majority of these works pay little to no attention to the issue of privacy leaking during the authorization verification phase. This study introduces a privacy- maintaining CP-ABE system with green authority verification to address the issues outlined above.

[13] This letter offers a fresh plan for identifying the agent responsible for data leakage. The method is based on an effective distribution strategy and threshold degree. Every record item's threshold price is determined using the concept of fact sensitivity as a means of controlling how those objects are allocated. Data is delivered by selecting the sellers in a round-robin fashion, and then the item with the best threshold price is chosen. The average possibility, average fulfilment rate, and average detection price are used to gauge the scheme's overall performance. In compared to past study, the results for the aforementioned parameters demonstrate an improvement of up to 75%.

[14] The challenging issue of secure data sharing in cloud computing has been addressed by the encryption technology known as cipher text-coverage characteristic-based encryption (CPABE). In general, shared information papers have a multilayered hierarchy, especially in the healthcare and military sectors. The hierarchy structure of shared documents hasn't been investigated in CP-ABE, though. A green document hierarchy attribute-based complete encryption technique for cloud computing is suggested in this study. The layered access control structures are immediately combined into a single access control structure, and the combined access control structure is then used to encrypt the hierarchical documents. The documents may allow the sharing of the attribute-related cipher text additions. As a result, the time value of encryption and each cipher text storage are saved. Additionally, it is demonstrated that the suggested system is secure based on the common assumption. The results of an experimental simulation show that the proposed approach is extremely effective at both encrypting and decrypting data. As the variety of documents expands, the benefits of our plan become more and more obvious.

[15] Data leaking has grown to be a serious problem for every organisation, especially when the 0.33 party is involved and has legal authorization to access the organization's private data. In this study, a statistical assessment-based leaker identity (SELI) scheme is presented for identifying the malicious party when data is released by any agent at a few unapproved sites. A new method of object and agent distribution that boosts the popularity of responsible agents is introduced. For the agent selection, the three algorithms SELI-first come first serve, SELI- spherical Robin (SELI-RR), and SELI-shortest request first (SELI-SRF) are provided. Additionally, the fictitious data objects are added even while the data sets are distributed, allowing some of the dealers to interpret the leaker in their own particular way. Based on the records that were allotted, the SELI scheme uses biograph to estimate the guilty party. In comparison to the previous work, the suggested solution improves average chance and average success charge by up to 25.37 and 58.18% for SELI- SRF and 163.68% for SELI-RR, respectively. Additionally, the SELI scheme demonstrated its effectiveness by achieving accuracy, precision, recall, and specificity scores of 99.82%, 99.92%, 99.4%, and 99.97%.

[16] With the rise in popularity of cloud computing, concerns over its security and privacy have grown. Data owners must

encrypt outsourced data to implement secrecy because the cloud computing platform is sent and un-trusted. Therefore, it is a pressing issue that needs to be resolved how to gain practical access control of encrypted statistics in an un-trusted environment. The promising method of attribute-based total encryption (ABE) is suitable for access control in cloud storage systems. In this paper, a hierarchical characteristic-based access control technique with constant-size cipher text is proposed. Because the range of bilinear pairing opinions to a consistent and the length of the cipher text are fixed, the approach is efficient. In encryption and decryption algorithms, it has a low computational value. The burden and threat of an unmarried authority scenario is also lessened by our scheme's hierarchical permission structure.

Conclusion

In cloud computing and statistics protection, data security is a challenging task. An abundance of labour is considered to lessen this mission. The thorough examination of the ongoing solutions, however, falls short. From this vantage point, this study provided a thorough analysis and investigated the most strategies about the ability and the pertinent solutions to share the information securely for data protection within the cloud platform. The key and appropriate information that is preferred to bring the method's focus, as well as the research gaps and recommended next steps for each presented response is highlighted.

Additionally, a thorough comparison and review of the majority of the refereed procedures is done. Each approach's applicability is evaluated in light of the context. No approach has been shown to be effective in ensuring that the information in the device is completely safe from every party who is directly or indirectly involved with it. By incorporating the techniques for providing the machine with total safety inside the sharing platform, the robust solution may be created. Furthermore, it is anticipated that the revealed analysis will serve as a landmark for the capability researchers operating in the area as well as other emerging applications requiring secure data exchange and storage. This is due to the set of highlights of the addressed fantastic solutions.

References

[1] A. K. Singh and I. Gupta, ``Online information leaker identification scheme for secure data sharing,'' Multimedia Tools Appl., vol. 79, no. 41, pp. 31165-31182, Nov. 2020.
[2] E. Zaghloul, K. Zhou, and J. Ren, ``P-MOD: Secure privilege-based multilevel organizational data- sharing in cloud computing,'' IEEE Trans. Big Data, vol. 6, no. 4, pp. 804-815, Dec. 2020.
[3] I. Gupta and A. K. Singh, ``GUIM-SMD: Guilty user identification model using summation matrix- based distribution,'' IET Inf. Security., vol. 14, no. 6, pp. 773- 782, Nov. 2020.
[4] W. Shen, J. Qin, J. Yu, R. Hao, and J. Hu, ``Enabling identity-based integrity auditing and data sharing with sensitive information hiding for secure cloud storage,'' IEEE Trans. Inf. Forensics Security, vol. 14, no. 2, pp. 331-346, Feb. 2019.
[5] I. Gupta and A. K. Singh, ``An integrated approach for data leaker detection in cloud environment,'' J. Inf. Sci. Eng., vol. 36, no. 5, pp. 993-1005, Sep. 2020.
[6] R. Li, C. Shen, H. He, X. Gu, Z. Xu, and C.-Z. Xu, ``A lightweight secure data sharing scheme for mobile cloud computing,'' IEEE Trans. Cloud Computer., vol. 6, no. 2, pp. 344-357, Apr. 2018.
[7] I. Gupta, N. Singh, and A. K. Singh, ``Layer- based privacy and security architecture for cloud data sharing,'' J. Communication. Software. Syst., vol. 15, no. 2, pp. 173-185, Apr. 2019.
[8] J. Li, S. Wang, Y. Li, H. Wang, H. Wang, H. Wang, J. Chen, and Z. You, ``An efficient attribute- based encryption scheme with policy update and le update in cloud computing,'' IEEE Trans. Ind. Information., vol. 15, no. 12, pp. 6500-6509, Dec. 2019.
[9] I. Gupta and A. K. Singh, ``A

framework for malicious agent detection in cloud computing environment,'' Int. J. Adv. Sci. Technol., vol. 135, pp. 49-62, Feb. 2020.
[10] Y. Li, Y. Yu, G. Min, W. Susilo, J. Ni, and K.-R. Choo, ``Fuzzy identity-based data integrity auditing for reliable cloud storage systems,'' IEEE Trans. Dependable Secure Computer., vol. 16, no. 1, pp. 72- 83, Jan./Feb. 2019.
[11] I. Gupta and A. K. Singh, ``A probabilistic approach for guilty agent detection using bi- graph after distribution of sample data,'' Proc. Computer. Sci., vol. 125, pp. 662-668, Jan. 2018.
[12] L. Zhang, Y. Cui, and Y. Mu, ``Improving security and privacy attribute-based data sharing in cloud computing,'' IEEE Syst. J., vol. 14, no. 1, pp. 387-397, Mar. 2020.
[13] I. Gupta and A. K. Singh, ``Dynamic threshold- based information leaker identification scheme,'' Inf. Process. Lett., vol. 147, pp. 69-73, Jul. 2019.
[14] S. Wang, J. Zhou, J. K. Liu, J. Yu, J. Chen, and W. Xie, ``An efficient le hierarchy attribute-based encryption scheme in cloud computing,'' IEEE Trans. Inf. Forensics Security, vol. 11, no. 6, pp. 1265-1277, Jun. 2016.
[15] I. Gupta and A. K. Singh, ``SELI: Statistical evaluation-based leaker identification stochastic scheme for secure data sharing,'' IET Communication., vol. 14, no. 20, pp. 3607-3618, Dec. 2020.
[16] W. Teng, G. Yang, Y. Xiang, T. Zhang, and D. Wang, ``Attribute-based access control with constant- size ciphertext in cloud computing,'' IEEE Trans. Cloud Computer., vol. 5, no. 4, pp. 617-627, Oct./Dec. 2017.
[17] I. Gupta, A. K. Singh, C. -N. Lee and R. Buyya, "Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, and Future Directions," in IEEE Access, vol. 10, pp. 71247-71277, 2022, doi: 10.1109/ACCESS.2022.3188110.