International Journal of Safety and Security Engineering
Vol. 15, No. 4, April, 2025, pp. 653-675
Journal homepage: http://iieta.org/journals/ijsse

Advances and Challenges in Cloud Data Storage Security: A Systematic Review

Mohammed El Moudni*, Elhoussaine Ziyati

C3S Research Laboratory, High School of Technology, Hassan II University, Casablanca 20000, Morocco

Corresponding Author Email: mohammed.elmoudni1-etu@etu.univh2c.ma

https://doi.org/10.18280/ijsse.150403

Received: 30 January 2025
Revised: 10 April 2025
Accepted: 22 April 2025
Available online: 30 April 2025

Keywords:
systematic review, cloud computing, data storage, data privacy and security, data integrity, data availability, encryption, access control

ABSTRACT

Cloud computing has significantly changed how data is stored by offering enhanced flexibility and scalability. However, its rapid growth has introduced serious security challenges, particularly concerning data integrity, confidentiality, and availability. This systematic review investigates recent research in cloud data storage security, focusing on research published between 2020 and 2024. A structured selection process led to the inclusion of 77 relevant studies that addressed key research questions. The review synthesizes current knowledge, identifies ongoing challenges, and evaluates six main security techniques, including, encryption, access control, data loss prevention (DLP), blockchain, machine learning, and data redundancy. Each method is analyzed based on its principles, application context, advantages, and limitations, along with a comparative assessment. Encryption is widely adopted and offers strong confidentiality but may reduce system performance. Access control enables accurate access management but is often complex to implement. DLP helps prevent sensitive data leaks but can result false positives. Blockchain improves transparency and trust but introduces latency and integration challenges. Machine learning enhances anomaly detection but depends on large datasets and computational resources. Data redundancy supports data availability but increases storage costs. The findings show that relying on a single method is not sufficient to ensure a complete data protection in cloud environments. A multi-layered approach, integrating various techniques, is necessary, particularly with the increased reliance on cloud services due to the expansion of the Internet of Things and the impact of the COVID-19 pandemic. This review contributes to the field by offering a comprehensive comparison of modern security models and provides direction for future research.

1. INTRODUCTION

Cloud computing has changed how we use technology by making data storage and computing power more flexible and easier to scale. Instead of depending on one system, tasks can be shared across many resources, which improves efficiency and makes access easier. This has led to big improvements in information technology operations and global connectivity, making cloud services a key part of modern infrastructures. However, as more people and businesses use the cloud, security concerns have also grown. Protecting cloud data means ensuring it stays private, accurate, and available, while preventing unauthorized access and data loss [1].

A key advantage of cloud computing is its potential to manage large amounts of data in a scalable and cost-effective way. It also supports advanced data analysis, real-time teamwork, and smooth integration with other digital services. However, the complexity of cloud systems brings security challenges that need careful management [2]. A major risk is sensitive information breaches, leading to leaks and privacy violations [3]. Insider threats are another issue, as both intentional misuse and accidental mistakes can harm security.

Many organizations use multi-cloud and hybrid cloud strategies to gain flexibility and backup options. While these approaches have benefits, they also make it harder to maintain consistent security across different cloud providers and platforms [4]. To manage these challenges, various security methods are used, including encryption, access control, and data redundancy techniques [5]. Additionally, new technologies like machine learning, blockchain, and zero-trust security models are being explored to improve cloud security further [6].

The shared-responsibility concept represents a key idea in cloud security. It explains the positions of both cloud service providers and users [7]. While cloud providers concentrate on the cloud infrastructure, users are responsible for managing their data, setting up access controls, and conducting regular security checks [8].

Despite progress in cloud security, cyber threats are always changing. Research shows that current security measures still have weaknesses, especially against advanced attacks and new technologies [9]. This study aims to analyze the current cloud storage security literature, identify key research gaps, and suggest new perspectives [10].

Although there is a wide range of security techniques available, they are often applied independently and may not

work effectively when combined. Some methods are complex to implement or can negatively affect system performance. Furthermore, much of the existing research focuses on individual techniques without offering a comprehensive comparison, making it difficult to assess which approaches are most effective.

This paper addresses these gaps by conducting a systematic review of cloud data storage security methods published between 2020 and 2024. Through a detailed comparison of key techniques, the study highlights effective practices, identifies persistent challenges, and provides insights into how cloud storage security can be strengthened in the future.

1.1 Motivation

The use of cloud services is growing very quickly. Almost 95% of businesses now depend on cloud platforms for their operations, especially as they adapt to the consequences of remote work due to the covid-19 pandemic [11]. This new paradigm has accelerated digital transformation, leading to more remote work, increased use of mobile devices, and a higher demand for cloud computing solutions [12].

However, this rapid growth, along with the expected rise of over 80 billion Internet of Things devices by 2025 [13]. The increase in remote work has led to more data being stored in the cloud, making systems more vulnerable to cyberattacks. In 2024, more than 27 billion records were exposed, and the mean cost of a data breach hit $4.88 million [14]. These numbers show how important it is to ensure the security of cloud as its use continues to grow. At the same time, it remains difficult for researchers and organizations to clearly understand the strengths and weaknesses of the different approachs, due to the lack of a comprehensive and updated comparison.

1.2 Our contribution

This study focuses on two main areas:

- *Analysis of Security Mechanisms:* This part gives a detailed look at the principles, how they are implemented, their benefits, and their limitations. It also includes a table that summarizes key metrics, making it easier to compare different security mechanisms.
- *Comparative Analysis:* This part highlights a detailed analysis of the evaluated techniques. It highlights at the same time their advantages and weaknesses.

1.3 Organization of the paper

The structure of the paper is as follows: Section 2 provides the background and context of the study. Section 3 explains the research methodology used for this systematic review. Section 4 reviews the literature and existing models. Section 5 presents the results of the study. Section 6 highlights the limitations of this research. Lastly, section 7 concludes the study by summarizing the main findings and potential perspectives.

2. CONTEXT OF THE STUDY

Before diving into the specific ideas behind this systematic review, it is important to first build a basic understanding of the cloud landscape. This foundation will help make the concepts clearer.

2.1 Cloud computing

Cloud computing is a new paradigm that provides a set of resources hosted on the Internet, users can use them with a pay as you go pricing model. The cloud paradigm enables users in the same time to exploit any service without needing to own or manage the physical infrastructure [15]. Its scalability and flexibility help organizations also to allocate resources more efficiently, and to improve performance and reduce costs [16].

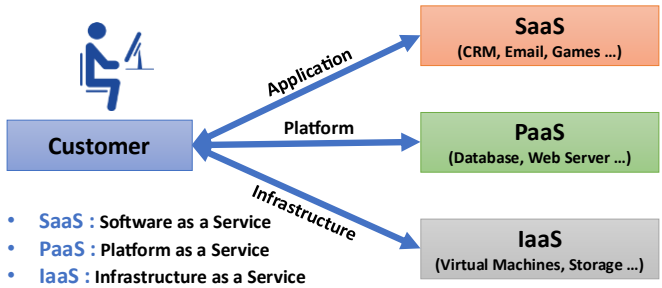

Figure 1. Main categories of cloud services

Cloud computing services are typically classified into three main categories, as depicted in Figure 1, Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). These distinct models offer varying degrees of user control and management [17].

2.2 Cloud data storage

Cloud storage refers to the storage and management of data in the cloud. Each approach differs in the way it organizes and processes data to meet specific requirements. The main categories of cloud storage are as follows [18]:

- *Object-based storage:* Information is stored as distinct objects, which may include documents, multimedia files, or other forms of unstructured data.
- *Block-level storage:* Data is fragmented into separate units (blocks) and stored independently, making it ideal for structured data applications.
- *File-based storage:* Provides a hierarchical file system that enables data organization, storage, and retrieval through a directory structure.

Table 1. Data storage offers per cloud service provider

Type	Microsoft Azure	Amazon AWS	Oracle OCI	Google GCP
File	Azure File Storage	Amazon EFS	OCI File Storage	Google Cloud Filestore
Block	Azure Blob Storage	Amazon EBS	OCI Block Volume	Cloud Persistent Disk
Object	Azure Blob Storage	Amazon S3	OCI Object Storage	Google Cloud Storage

Cloud service providers offer tailored data storage solutions, categorized as file, block, and object storage, each category is suited to specific needs. Table 1 lists storage services from major providers [19].

2.3 Shared responsibility model

The shared responsibility model allows to separate the security perimiter of both cloud service providers (CSPs) and users. CSPs protect cloud infrastructure [20], whereas users responsible of protecting their data through encryption, access controls, adequat configuration, and vigilance against breaches and internal threats [21].

Table 2. Separation of responsabilities in the cloud

Layer	IaaS	PaaS	SaaS
User Access	USER	USER	USER
Data	USER	USER	USER
Application	USER	USER	CSP
Operating System	USER	CSP	CSP
Virtualization	CSP	CSP	CSP
Servers	CSP	CSP	CSP
Storage	CSP	CSP	CSP
Network	CSP	CSP	CSP
Physical	CSP	CSP	CSP

Table 2 illustrates the cooperative aspect of cloud security. A secure cloud environment can be achieved when both parties clearly understand their respective responsibilities [22].

2.4 Data storage security threats

2.4.1 Identified threats

Although cloud data storage offers flexibility and scalability, it faces threats, such as data breaches and insider attacks, where authorized individuals act maliciously or unintentionally. Malware attacks, misconfigurations, and derial of service attacks that can compromise the data integrity and availability. Additionally, risks such as data loss or corruption can lead to permanent loss or damage to information [23].

Table 3. Key security threats in cloud data storage

Threat Type	Potential Impact	Exploitation Techniques
Data Breaches	Exposure of sensitive or confidential data.	Exploiting vulnerabilities, phishing attacks, credential theft.
Insider Threats	Unauthorized access or data leakage by authorized users.	Misuse of access privileges, intentional or accidental data theft.
Misconfiguration	Data exposure, loss, or unintentional disclosure due to incorrect settings.	Incorrect cloud storage configurations, improper access controls.
Data Loss/Corruption	Permanent loss or corruption of data.	Hardware failures, software bugs, or human error.

2.4.2 Mitigation techniques

To overcome the security challenges posed by storing data in the cloud, it is essential to employ effective mitigation techniques that are tailored to specific threats. Table 3 provides an overview of these techniques [24].

Table 4 presents a short identification of the main mitigation techniques, including the essential methods used to improve security in the cloud.

Table 4. Mitigation techniques for identified threats

Threat Type	Mitigation Technique
Data Breaches	Encryption / DLP
Insider Threats	Access Control
Malware Attacks	Intrusion Detection / Backup
Misconfiguration	Configuration Management / Access Control
Denial of Service	Firewall Rules / Intrusion Detection
Storage	Data Redundancy and Backup

3. METHODOLOGY

The literature survey is an essential part the systematic review and serves as the basis for this investigation. This comprehensive process involves a methodical examination of existing studies, allowing to gather and synthesize relevant research on used security techniques.

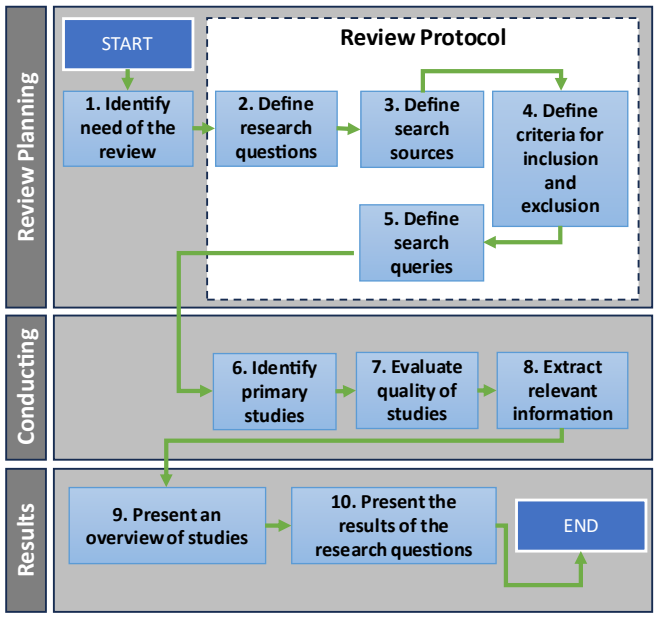

Figure 2. Literature survey process

As illustrated in Figure 2, the literature survey process starts with thorough review planning, including defining research goals, identifying key themes, and selecting relevant databases and sources. It also involves setting criteria for study inclusion and exclusion. The next phase is the review execution, which includes systematic searching and critical evaluation of the literature based on relevance, methodology, and outcomes. The process concludes with result synthesis, where insights are consolidated, patterns are identified, and research gaps are mentioned to offer a general overview of the context.

3.1 Research question

The primary focus of the paper is to present the critical elements of data protection in cloud storage. Specifically, our study aims to identify the most used techniques and to evaluate their effectiveness, implementation challenges, advantages and disadvantages.

The review question shown in Table 5 was formulated to guide a comprehensive and organized analysis of the literature, ensuring that our findings are pertinent and significant.

Table 5. Research questions

Question	Sub-Questions	Purpose
What are the most effective techniques for securing data in cloud environments?	1. What are the main data protection techniques currently employed in cloud storage? 2. How do these techniques address various data security concerns? 3. What are the comparative advantages and limitations of these techniques? 4. What are the primary threats and vulnerabilities associated with each technique?	To identify, analyze, and evaluate different data protection methods, their effectiveness, and associated threats.
How do these data protection techniques compare in terms of efficacy, implementation challenges?	1. What are the success rates and performance metrics of these techniques? 2. What implementation challenges are associated with each technique?	To evaluate the practical impact, versatility, and threats related to data protection techniques.

3.2 Search strategy

In order to conduct an extensive review of the literature, we developed a search strategy that adhered to the PRISMA guidelines. Our methodology involves identifying and selecting relevant academic sources in the domains of cloud, data, and security.

Table 6. Study journals and sources

Type	Name	Quartile	Impact Factor
Journal	IEEE Access	Q1	3.993
Journal	IEEE Internet of Things Journal	Q1	8.408
Journal	IEEE Transactions on Cloud Computing	Q1	4.075
Journal	IEEE Transactions on Dependable and Secure Computing	Q1	4.717
Journal	Journal of Cloud Computing	Q2	2.787
Journal	Journal of Information Security and Applications	Q2	2.152
Journal	SN Computer Science	Q2	1.374

Table 6 lists the academic journals and databases used as a source of articles. By focusing on Q1 and Q2 journals, such as IEEE Access and the Journal of Cloud Computing, we ensured that the included literature had a high impact and relevance.

This selection process provides a solid foundation for peer-reviewed studies, which are crucial for a comprehensive review of cloud data protection techniques. After identifying the sources, we have applied a range of inclusion and exclusion rules to refine the search results.

As illustrated in Table 7, to ensure that the literature is both recent and of high quality, we only included peer-reviewed studies published between 2020 and 2024. Non-peer-reviewed articles and articles that did not directly address cloud data protection were excluded. This step is crucial to reduce the large volume of initial search results to a more manageable and relevant subset of studies. Finally, we executed a detailed search strategy using specific queries from various databases to identify the most relevant literature.

Table 7. Inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
Journal Type	Peer-reviewed journals, conference papers	Non-peer-reviewed sources, editorials, opinion pieces
Publication Year	2020 - 2024	Articles published before 2020
Language	English	Non-English publications
Focus	Techniques for data protection in cloud storage, security measures, and effectiveness	Studies not directly related to data protection or cloud security
Paper Type	Full-text articles, review articles	Abstracts only, conference posters, presentations
Methodology	Empirical studies, systematic reviews, meta-analyses	Theoretical papers without empirical data

Table 8. Search queries

Search Query	(PUBYEAR > 2019 AND PUBYEAR < 2025) AND (SUBJAREA = "COMPUTER SCIENCE") AND (DOCTYPE = "article") AND (EXACTKEYWORD("Cloud Computing") OR EXACTKEYWORD("Security") OR EXACTKEYWORD("Cryptography") OR EXACTKEYWORD("Data Privacy") OR EXACTKEYWORD("Privacy") OR EXACTKEYWORD("Digital Storage") OR EXACTKEYWORD("Access Control") OR EXACTKEYWORD("Blockchain") OR EXACTKEYWORD("Authentication") OR EXACTKEYWORD("Cloud Storage"))		
Database	IEEE Xplore	Scopus	Google Scholar
Results	55	60	38
Filters	Q1 and Q2 journals peer-reviewed 2020-2024, English	Q1 and Q2 journals, peer-reviewed, 2020-2024, English	Q1 and Q2 Peer-reviewed, 2020-2024, English

Table 8 details the practical applications of our search queries across IEEE Xplore, Scopus, and Google Scholar databases. By using targeted queries and applying filters, we gathered a current collection of articles.

3.3 Study selection

The study selection process was conducted in two stages to ensure rigor, transparency, and relevance. In the initial screening stage, we reviewed the titles and abstracts of articles to identify studies aligned with our research objectives and inclusion criteria. Duplicate records were automatically removed using reference management software prior to screening to avoid redundancy.

To ensure the comprehensiveness of the search, we used multiple academic databases and applied a broad set of keywords related to cloud data storage security. The search results were cross-verified to reduce the risk of missing relevant studies.

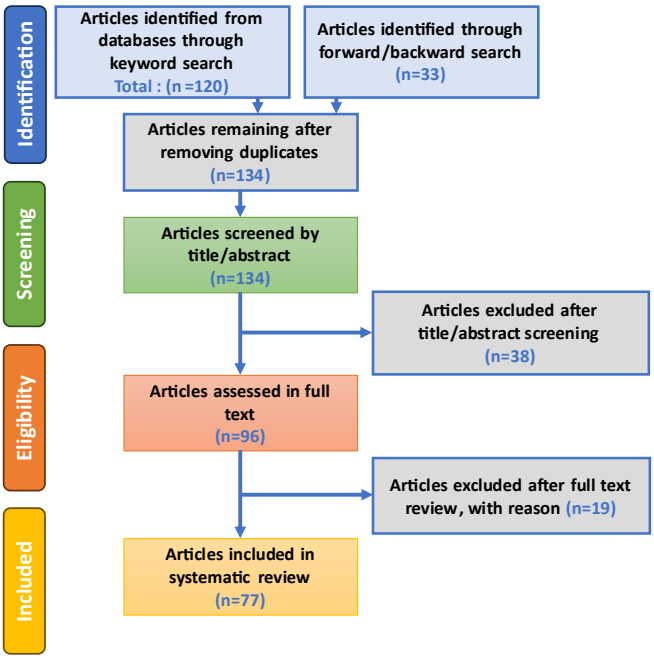

Figure 3. Study collection process

As illustrated in Figure 3, the initial review identified 153 potentially relevant studies. In the eligibility stage, we conducted a detailed full-text review of these articles, evaluating their methodology, quality, adherence to research standards, and contribution to cloud data protection. This rigorous process led to the selection of 77 high-quality studies for inclusion in our systematic review.

4. LITERATURE SURVEY

4.1 General overview

Over the past four years, research on cloud data storage security has grown significantly, as shown by the increasing number of conferences, workshops, and publications focused on this topic. Following the explanation of our research method, this section presents a comprehensive review of the selected studies. It highlights key contributions, explains the main security techniques, evaluates their strengths and limitations, and outlines directions for future research.

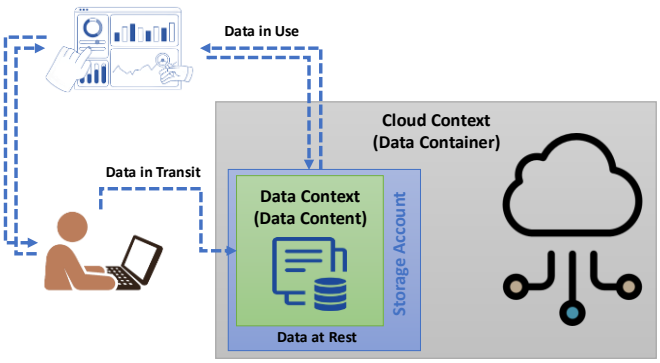

Figure 4. Macro view of the cloud data storage workflow

As shown in Figure 4, the literature survey on cloud data storage security is systematically divided into two main areas.

The first area focuses on securing the data container or context, which involves ensuring the protection of the cloud infrastructure.

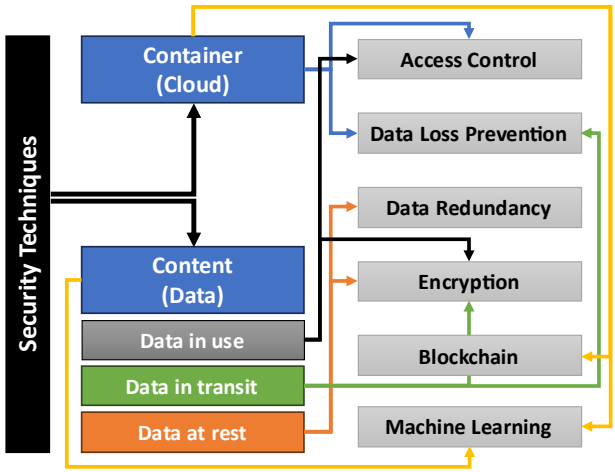

Figure 5. Main cloud data-storage security techniques

The second area as shown in both Figure 4 and Figure 5, addresses the security of data content across its three critical states: data at rest, which point to stored data, data in transit, which pertains to data being transferred to the cloud storage, and data in use, which involves data actively being used or processed.

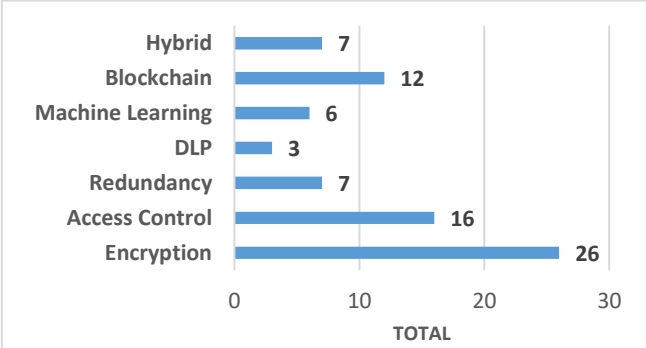

Figure 6. Number of articles per proposed technique

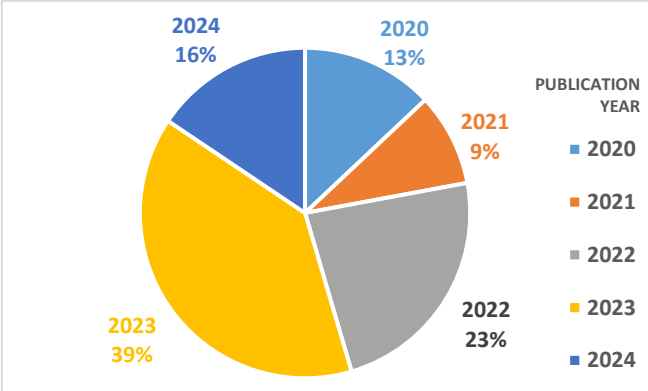

Figure 7. Number of articles per proposed technique

Figure 6 shows that encryption is the most used technique, featured in 26 studies, underscoring its critical role in securing cloud storage. Access control is highlighted in 16 studies, emphasizing its importance for managing data access.

Blockchain appears in 12 papers, reflecting growing interest in decentralized security methods. Redundancy and hybrid approaches are discussed in seven studies, indicating their value in enhancing fault tolerance. Machine learning is covered in six papers, suggesting its emerging role in threats classification and detection, while Data Loss Prevention is noted in only three papers, marking it as a niche area within cloud storage protection.

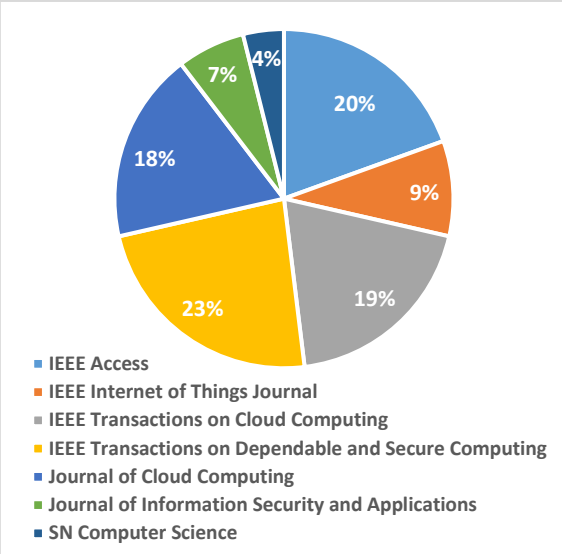

Figure 8. Number of articles per journal

The publication year trend illustrated in Figure 7 reveals a notable increase in research on cloud data protection in recent years. 2023 stands out, with 30 papers representing the peak of research activity and a significant surge in focus on this topic, followed by 2022, which saw 18 papers, demonstrating sustained interest. Publications for 2024 show 12 papers, reflecting ongoing research efforts. In contrast, 2020 and 2021 had fewer publications, with 10 and seven papers, respectively, suggesting lower levels of research activity during those years.

The journal distribution shown in Figure 8 highlights the focus of research on cloud storage security. IEEE Transactions on Dependable and Secure Computing leads with 18 papers, followed by IEEE Access and IEEE Transactions on Cloud Computing, each with 15 papers, reflecting significant contributions to the field from these sources. The Journal of Cloud Computing also shows notable contribution with 14 papers. In contrast, the Journal of Information Security and Applications has five papers and SN Computer Science has three papers, indicating a smaller volume of research on this topic within these journals.

4.2 Related work

4.2.1 Encryption

Background. Encryption is an essential security measure that converts data into an unreadable format, thereby ensuring confidentiality and integrity by preventing access by unauthorized users.

Table 9. Encryption techniques and use case

Technique	Principle	Use Case
Symmetric Encryption	The same key is used for encryption and decryption; needs secure key management.	Efficient for large-scale data encryption.
Asymmetric Encryption	Two keys are used (public and private) for encryption and decryption.	Ideal for secure communication and digital signatures.
Homomorphic Encryption	Enables handling encrypted data without decryption; preserves data privacy.	Ideal for sensitive and confidential information.

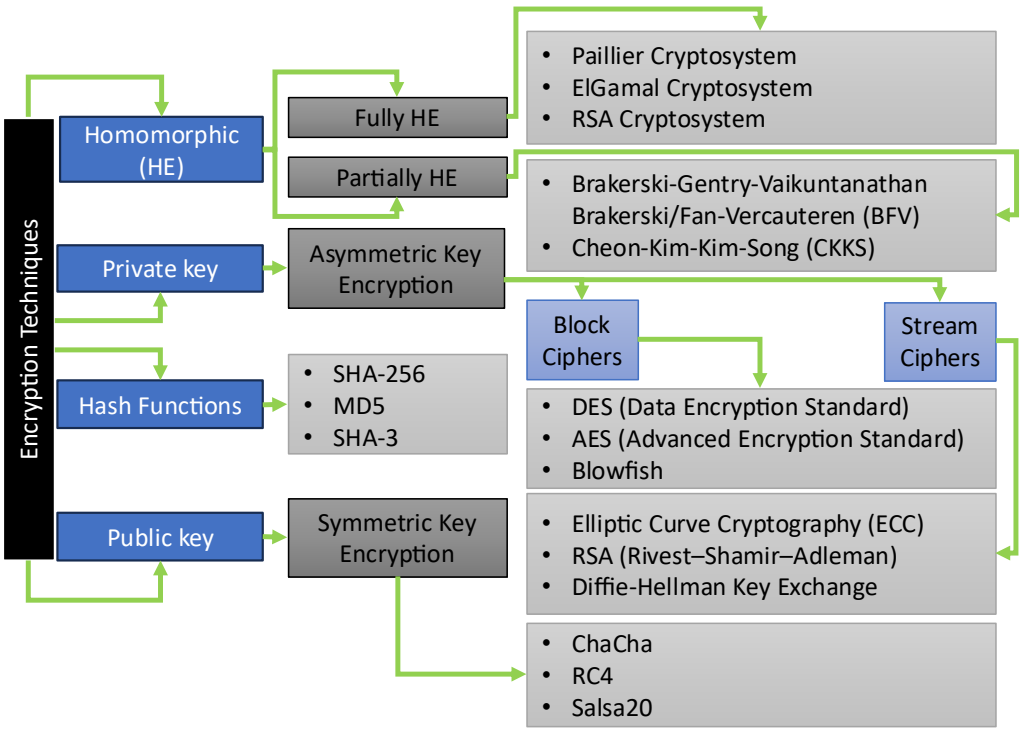

Figure 9. Encryption techniques and algorithms

Table 10. Comparative analysis of proposed encryption models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[25]	Attribute-Based Encryption (ABE), Public Traceability Mechanism	High	Moderate	Fine-grained access control and improved confidentiality.	Complex setup and management of large-universe attribute-based encryption schemes.
[26]	Advanced Encryption Standard (AES), Public Key Cryptography	Medium	Low	Increased data integrity and reduced unauthorized exposure.	Increased computational overhead due to data hiding techniques and encryption.
[27]	Digital Signatures, Cryptographic Hash Functions	Medium	Moderate	Efficient data integrity auditing and tampering detection.	Potential performance degradation from additional auditing mechanisms.
[28]	Advanced Encryption Standard (AES), Data Dispersion Techniques	Medium	Moderate	Improved data integrity and availability.	Overhead from data dispersion and encryption methods.
[29]	Attribute-Based Encryption (ABE)	High	Moderate	Dynamic key management and enhanced data confidentiality.	Complexity in managing multiple authorities and encryption keys.
[30]	Attribute-Based Encryption (ABE), Elliptic Curve Cryptography (ECC)	High	Moderate	Fine-grained access control and secure data transactions.	Performance impact from fine-grained access control mechanisms.
[31]	Homomorphic Encryption	Medium	Low	Secure data manipulation with confidentiality preservation.	Computationally intensive data obfuscation techniques.
[32]	Ciphertext-Policy Attribute-Based Encryption (CP-ABE), Keyword-Based Search	High	Moderate	Verifiable search capabilities and enhanced data confidentiality.	Potential latency in keyword-based searchable encryption.
[33]	Homomorphic Encryption (HE)	High	Moderate	Dynamic encryption key management and improved access control.	Complexity in integrating revocability with encryption and data integrity.
[34]	Attribute-Based Encryption (ABE), Cryptographic Hashes, Digital Signatures	High	Moderate	Effective protection against decryption key exposure.	Performance overhead associated with privacy-preserving searchable encryption.
[35]	Homomorphic Encryption (HE)	High	Low	Privacy-preserving searchable encryption and secure data access.	Increased computational cost for handling homomorphic encryption operations.
[36]	Cryptographic Techniques for Data Obfuscation	Medium	Moderate	Support for secure computations and encrypted data analytics.	Complexity in verifying data integrity and managing encrypted searches.
[37]	Attribute-Based Encryption (ABE), Advanced Encryption Algorithms	High	Low	Enhanced data correctness and confidentiality with integrity verification.	High computational requirements for joint data and function homomorphic encryption.
[38]	Attribute-Based Encryption (ABE), Multi-Keyword Search	High	Moderate	Privacy-preserving and serverless searchable encryption.	Complexity and potential performance issues with dynamic encryption.
[39]	Hierarchical Block Variable Length Coding, Advanced Encryption Schemes	Medium	Low	Obscured data access patterns for improved confidentiality.	Reversible data hiding techniques may have limitations in security strength.
[40]	Multi-Proxy Assisted Encryption	Medium	Low	Optimal traceability and accountability in decentralized systems.	Challenges in achieving optimal traceability and maintaining system efficiency.
[41]	Stochastic Gradient Descent, Long Short-Term Memory (LSTM) Networks	High	Moderate	Secure multi-key searchable encryption for complex queries.	Performance overhead associated with multi-key searchable encryption.
[42]	Lattice-Based Cryptography, Attribute-Based Encryption (ABE)	High	Moderate	Reversible data obfuscation and secure data hiding.	Potential performance issues with reversible data hiding techniques.
[43]	Martino Homomorphic Encryption	Medium	Low	Enhanced confidentiality with dynamic access control in mobile cloud environments.	Overhead due to multi-proxy assisted encryption mechanisms.
[44]	Advanced Encryption Standard (AES), Rivest–Shamir–Adleman (RSA)	Medium	Moderate	Efficient encryption function for secure data storage and retrieval.	Computationally intensive secure encryption algorithms.

	Encryption				
[45]	Multi-Key Homomorphic Encryption	High	Moderate	Robust protection against decryption key exposure.	Potential performance impact of lattice-based encryption schemes.
[46]	Advanced Encryption Standard (AES), Blockchain Key Management	Medium	Moderate	Secure data protection with homomorphic encryption.	Complexity in implementing and managing homomorphic encryption schemes.
[47]	Attribute-Based Encryption (ABE), Searchable Encryption	High	Moderate	Hybrid cryptography for medical data security and isolation.	Hybrid cryptography models may face integration and performance challenges.
[48]	Verified Public Key Encryption, Equality Test	Medium	Moderate	Multi-key encryption for enhanced privacy in cloud computing.	Performance overhead associated with multi-key homomorphic encryption.
[49]	Advanced Encryption Standard (AES), Rivest–Shamir–Adleman (RSA) Encryption	Medium	Moderate	Dynamic key management and improved data security.	Potential complexity in dynamic encryption and blockchain key management.
[50]	Homomorphic Encryption, Data Obfuscation Techniques	High	Moderate	Enhanced privacy protection for mobile cloud storage.	Complexity in implementing privacy-preserving mobile cloud storage solutions.

As illustrated in Figure 9 and Table 9, Encryption techniques are used to secure data through mathematical transformations. Symmetric encryption Employs the same key for encryption and decryption, expressed as $C = Enc(Key, Data)$ where Enc is the encryption function, Key represent the encryption key, $Data$ is the plaintext, and C is the ciphertext. Asymmetric encryption Employs a set of two keys (one public, one private) for data encryption and decryption, described by $C = Enc(K_{public\_key}, Data)$ and $Data = Dec(K_{private\_key}, C)$ where K_{public_key} is the public key, $K_{private\_key}$ is the private key, Enc is the encryption function, and Dec is the decryption function. Homomorphic encryption Enables processing of encrypted data without decryption, represented as $E(K, f(P1, P2))$ where f is a function such as addition or multiplication, and E is the encryption function. Hashing involves converting data into a fixed-length hash, via $Hash(P)$ where $Hash$ is the hash function and P is the plaintext. Hashing is a one-way function designed to be irreversible.

Existing studies. Table 10 introduces an analysis of the suggested encryption models. Each paper is examinated based on the used method, implementation complexity, and performance impact. In addition, it highlights the advantages and disadvantages of each approach. Various encryption models have been developed, focusing on confidentiality, integrity, and access control. Attribute-Based Encryption (ABE), including both revocable and multi-authority variants, plays a key role in providing fine-grained access control, allowing for flexible management of encryption keys and user permissions. Homomorphic encryption is another critical technique, enabling secure computations on encrypted data without compromising privacy, which is particularly valuable for sensitive data storage. Advanced Encryption Standard (AES) is frequently integrated with other cryptographic methods, such as digital signatures and cryptographic hashes. Searchable encryption models, combining Ciphertext-Policy ABE with keyword-based search capabilities, address the need for secure data retrieval while maintaining confidentiality. Additionally, hybrid cryptographic approaches that combine symmetric and asymmetric encryption, alongside blockchain-based key management, further enhance the security of cloud data storage, particularly healthcare sector.

Discussion. As shown in Figure 10, the distribution of encryption techniques reveals a clear focus on homomorphic encryption with then papers. Symmetric and asymmetric encryption methods each account for five studies, illustrating their sustained importance. Mixed encryption approaches, involving combinations of different techniques, are also notable, with six papers reflecting a trend towards hybrid solutions. This balance indicates a robust exploration of both traditional and advanced encryption methods in this study.

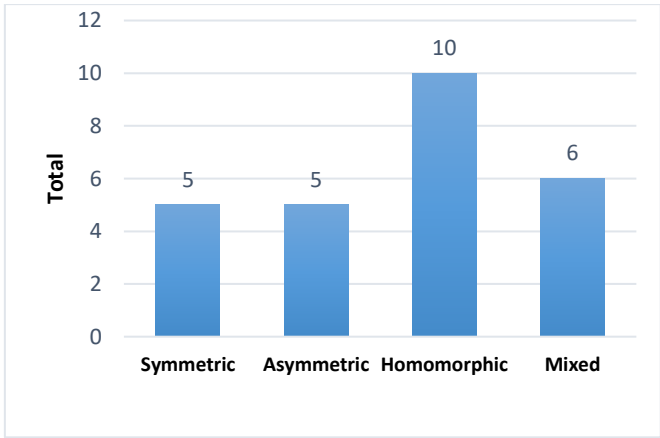

Figure 10. Distribution of articles per encryption type

Table 11. Comparative analysis of proposed access control models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[51]	Multi-Authority Access Control (MAAC)	Moderate to High	High	Provides efficient multi-authority management and access control, allowing scalable and secure cloud storage.	Can involve high complexity in managing multiple authorities, potentially increasing operational overhead.

[52]	Dual Access Control	Moderate	Moderate	Combines dual access control layers, improving the security of data sharing and minimizing unauthorized access.	Dual access controls might introduce additional configuration challenges, impacting ease of use and deployment.
[53]	Multi-Keyword Ranked Search	High	Low	Supports fine-grained search capabilities with multi-key access, enhancing flexibility and precision in data retrieval.	May require significant computational resources for multi-key searches, potentially affecting performance.
[54]	Biometric-Based Access Mechanism	Moderate	Moderate	Integrates biometric authentication, offering a strong, user-friendly method for verifying identities and securing access.	Biometric systems can be prone to false negatives or positives, possibly leading to access issues or security vulnerabilities.
[55]	Least-Privilege Model	Low	Low	Implements least-privilege principles effectively, ensuring that users have only the necessary permissions.	Least-privilege implementation might be complex to configure, requiring careful management of permissions.
[56]	Attribute-Based Access Control (ABAC)	Moderate	Moderate	Provides secure attribute-based access with consistent policies, ensuring robust protection of sensitive data.	Policy consistency can be challenging to maintain, particularly in dynamic environments with frequent changes.
[57]	Editable Data Sharing with Accountability	High	High	Allows for controlled data sharing with high accountability, making it easier to manage and audit data access.	Controlled data sharing mechanisms might limit flexibility, potentially complicating user access scenarios.
[58]	Distributed Data Access Control	Moderate	Moderate	Ensures privacy preservation in distributed environments, with strong protection against unauthorized data access.	Privacy-preserving techniques can impact performance, potentially leading to slower data access or increased latency.
[59]	Sanitizable Access Control	Moderate	Moderate	Protects against malicious data publishers by allowing data sanitization, enhancing data integrity.	Data sanitization processes might introduce delays or complexity in ensuring the integrity of shared data.
[60]	Optimized Role-Based Access Control	Moderate to High	Moderate to High	Optimized for e-health environments, incorporating trust mechanisms to improve data access and security.	Trust-based mechanisms in e-health environments may require extensive validation, complicating implementation.
[36]	Oblivious Random Data Access	High	Low	Offers oblivious access control to prevent data leakage, safeguarding user privacy in cloud environments.	Oblivious access control methods can be complex to implement and may require significant resources for effective operation.
[61]	Attribute-Based Access Control with Performance Optimization	Moderate	Moderate	Features performance-optimized remote file sharing, combining security with efficient access to cloud-stored data.	Performance optimization might be limited by the complexity of attribute-based access controls, affecting efficiency.
[62]	Admission Control and Key Agreement	Moderate	Moderate	Employs anonymous identity-based controls, improving security and privacy through advanced key agreement techniques.	Anonymous identity techniques might face challenges in maintaining user convenience while ensuring security.
[63]	Verifiable Data Storage and Retrieval	High	Low	Ensures reliable data storage and access with verifiable attributes, enhancing data integrity.	Verifiable attribute-based methods can be highly resource-intensive, which can affect overall system performance.
[64]	Role-Based Encrypted Keyword Search	High	Low	Enhances access control through encrypted keyword search, providing robust protection for outsourced cloud data.	Encrypted keyword search mechanisms can introduce additional overhead, affecting the speed of data retrieval.
[65]	Privacy-Preserving Attribute-Based Access Control with Data Duplication	Moderate to High	Moderate	Provides efficient multi-authority management and access control, allowing scalable and secure cloud storage.	Can involve high complexity in managing multiple authorities, potentially increasing operational overhead.

The reviewed studies, as summarized in Table 11, showcase various encryption techniques tailored to different security needs. Attribute-Based Encryption (ABE), noted in [25, 29, 30, 37], excels in fine-grained access control and confidentiality but may impact performance due to its complexity. Advanced Encryption Standard (AES), used in [26, 28, 39, 40, 46], offers effective symmetric encryption with strong security and efficiency but lacks advanced access control features. Homomorphic Encryption, featured in [31-33, 35, 43-47, 50], supports secure computations on encrypted data but often incurs high computational costs. Hybrid Encryption methods, explored in [34, 40, 41, 44, 48], blend symmetric and asymmetric techniques to balance security and manageability but can be complex to implement. These studies highlight the importance of choosing encryption methods based on specific use cases and balancing security, efficiency, and computational demands.

4.2.2 Access control

Background. The principle of access control in cloud computing consists in controlling who may access resources, in which conditions, and what operations they may carry out. This control is crucial to ensure data protection, compliance and security.

Table 12. Articles per used technique

Type	Total	Used algorithms	Papers
Symmetric	5	- Advanced Encryption Standard (AES) - Data Encryption Standard (DES)	[26, 28, 39, 40, 46]
Asymmetric	5	- RSA - Elliptic Curve Cryptography (ECC) - Attribute-Based Encryption (ABE)	[25, 29, 30, 37, 50]
Homomorphic	10	- Fully Homomorphic Encryption (FHE) - Partially Homomorphic Encryption (PHE) - Privacy-Preserving Encryption	[31, 32, 33, 35, 43-47, 50]
Mixed	5	- Hybrid Encryption (combining symmetric and asymmetric methods) - Multiple Encryption Techniques	[34, 40, 41, 44, 48]

Table 13. Comparative analysis of proposed data redundancy models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[66]	Demand-Aware Erasure Coding	Moderate	Moderate	Balances redundancy and storage efficiency, adapting to data demands and failures.	Complexity in encoding and decoding operations, potential for high overhead.
[67]	Secure and Distributed Data Storage	High	Moderate	Ensures secure and resilient data storage across adversarial networks, addressing challenges in data distribution.	High complexity in maintaining security and consistency in distributed settings.
[68]	Identity-Based Provable Multi-Copy Data Possession	High	Moderate	Verifies data redundancy and integrity in the multi-cloud context, enhancing data protection.	Intensive computational requirements for verification, complex cryptographic protocols.
[69]	Secure Distributed Storage Orchestration	High	High	Manages data distribution and redundancy effectively across heterogeneous cloud-edge infrastructures.	High complexity in orchestration and increased resource demands.
[70]	Prediction-Based Replica Selection	Moderate	Low	Optimizes data placement and reduces latency by predicting optimal replica locations.	Complexity in predictive modeling and potential suboptimal placement.
[71]	Cost-Effective Consistency Model	Moderate	Moderate	Maintains strong consistency and redundancy for geo-diverse data replicas, balancing cost and reliability.	Performance impact due to maintaining consistency across geographically dispersed locations.
[72]	Dynamic Replication and Placement	Moderate	Moderate	Enhances data redundancy and availability in multi-cloud environments through dynamic replication strategies.	Additional complexity in managing dynamic replication and distribution.

Table 14. Access control models used in the cloud

Model	Description	Key Features
Role-Based Access Control (RBAC)	Access is granted based on roles assigned to users, with permissions tied to these roles.	Simplifies management, supports hierarchical roles, easy to audit, scalable for large organizations.
Discretionary Access Control (DAC)	Access is controlled by the resource owner, who can grant or revoke permissions.	Flexible permissions, user-controlled access, potential security risks if mismanaged.
Mandatory Access Control (MAC)	Access is enforced by a central authority based on security labels and classifications; users cannot change rights.	Enforced policies, suitable for environments with strict security needs.
Attribute-Based Access Control (ABAC)	Access is determined by user attributes, resource attributes, and contextual factors.	Fine-grained, dynamic, context-aware, supports complex policies, adaptable and scalable.

As illustrated in Table 12, access control models are often used in the cloud, in Role-Based Access Control, access is defined as:

$$Access(U,R) = U_{i=1}^{n} Permissions\ (R_i) \quad (1)$$

where, U is the user, R is the resource, and $Roles(U)$ represents the roles assigned to the user. Discretionary Access Control (DAC) can be represented as:

$$Access(U,R) = Owner\ (R) \cup DelegAccess(U,R) \quad (2)$$

where, the resource owner or delegated users determine access. Mandatory Access Control (MAC) follows:

$$Access(U,R) = Sec_{Level}(U) \geq Sec_{Level(R)} \quad (3)$$

Ensuring that access relies on security classifications. Attribute-Based Access Control (ABAC) is modeled as:

$$Access(U,R) = f\ (A_U, A_R, A_E) \quad (4)$$

where, A_U, A_R and A_E are attributes of the user, resource, and environment, respectively, and f is a policy function determining access.

Existing studies. Table 13 offers a comparative analysis of various access control models proposed for cloud data-storage security. Xiong et al. [51] introduced SEM-ACSIT, a multi-authority framework using attribute-based encryption (ABE) for IoT cloud storage. Ning et al. [52] combined role-based access control (RBAC) with attribute-based policies to enhance flexibility and security. Li et al. [53] developed a multi-keyword ranked search mechanism with access control through searchable encryption. Panchal et al. [54] used biometric authentication, including fingerprint and facial recognition, for securing cloud services. Gill et al. [55] applied least-privilege access control in AWS with granular policies. Xue et al. [56] proposed a secure attribute-based model with hybrid encryption. Hou et al. [57] created a fine-grained model with editability features using cryptographic proofs. Nasiraee and Ashouri-Talouki [58] focused on privacy-preserving distributed access control with advanced encryption standards. Susilo et al. [59] developed a sanitizable access control system to guard against data tampering. Butt et al. [60] optimized RBAC in e-health with trust mechanisms. Liu et al. [36] introduced an oblivious random data access scheme for privacy. Chen et al. [61] developed a secure remote file sharing system with attribute-based control. Paulraj et al. [62] designed an anonymous identity-based admission control policy. Bera et al. [63] integrated integrity verification into attribute-based encryption for verifiable data storage. Miao et al. [64] introduced REKS, a role-based encrypted keyword search model with enhanced control. Pavithra et al. [65] proposed a privacy-preserving model with data duplication for maintaining confidentiality.

Discussion. The reviewed studies revealed a diverse range of access control models, each addressing specific security and performance needs. Multi-Authority Access Control (MAAC) [51] offers scalable management but is complex to implement. Dual-Access Control [52] combines role-based and attribute-based policies to enhance security, though it may involve configuration challenges. The Multi-Keyword Ranked Search model [53] provides precise data retrieval but demands high computational resources. Biometric-Based Access Mechanisms [54] deliver strong authentication but require specialized hardware. The Least-Privilege Model [55] simplifies permissions but can be complex to manage. Attribute-Based Access Control (ABAC) [56, 57] supports dynamic policies but may struggle with policy consistency. Techniques like Fine-Grained Access Control with Editability [58] and Privacy-Preserving Access Control [59] improve data handling but may affect performance. These models highlight the need to balance security needs with implementation and performance considerations.

4.2.3 Data redundancy

Background. In cloud storage, maintaining multiple copies of data across multiple regions and availability zones is a crucial security strategy known as data redundancy.

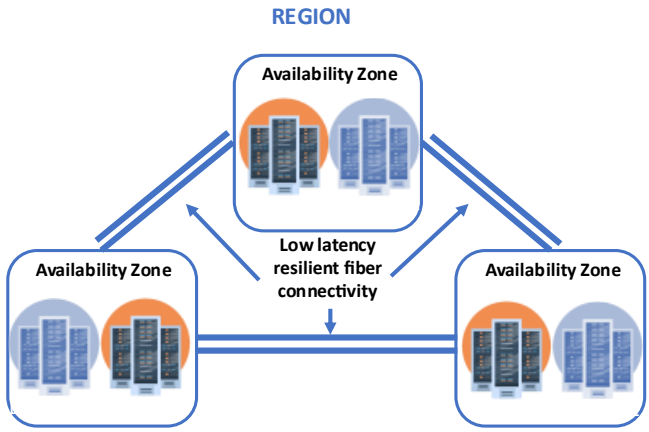

Figure 11. Cloud region concept

As shown in Figure 11, a cloud region is a specific geographic area with multiple datacenters, while availability zones are isolated locations within a region. Distributing data across these regions and zones enhances redundancy and helps to prevent data loss. Let $R(C)$ be the set of regions for a CSP C .

$$R(C) = \{R_i, R_2, R_3, \dots, R_n\}\ with\ i \in N, i \geq 1 \quad (5)$$

Each region R_i contains a set of availability zones, denoted as $AZ(R_i)$, where each AZ has a minimum of 2 data centers.

$$AZ(R_i) = \{AZ_{ij} \mid j = 1, \dots, m_i\}, m_i \in N, m_i \geq 2 \quad (6)$$

Each availability zone AZ_{ij} contains a set of data centers $DC(AZ_{ij})$ with at least 1 data center per AZ.

$$DC(AZ_{ij}) = \{DC_{ijk} \mid k = 1,2, \dots, k_{ij}\}\ with\ k_{ij} \in N \quad (7)$$

For a given region R_i, the set of services $S(R_i)$ available in that region is a subset of all services offered by the CSP. Thus, if a CSP C offers a variety of services, the specific set of services available in each region R_i might differ. Let $S(C)$ denote the set of all services provided by CSP C, For each region R_i, the set of available services $S(R_i)$ is a subset of $S(C)$.

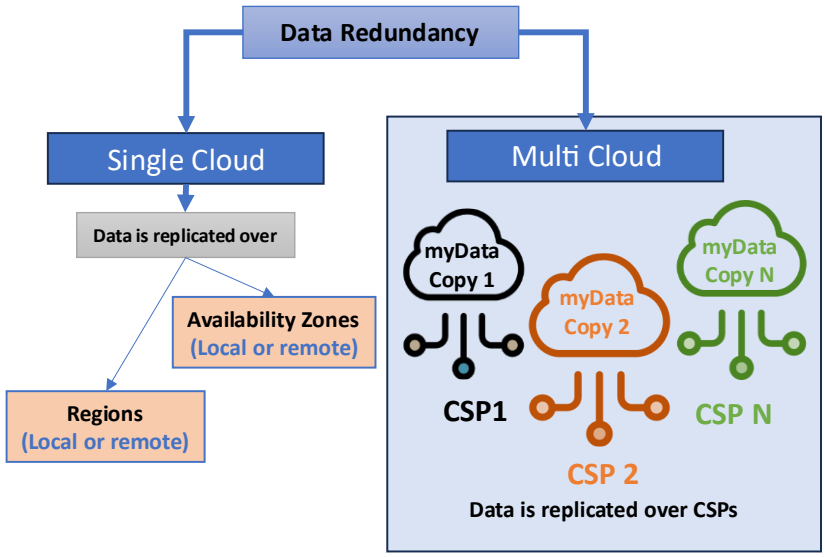

Figure 12. Data redundancy over cloud models

Table 15. Comparative analysis of proposed DLP models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[74]	CloudDLP (Data Sanitization)	Moderate	Potentially minimal	Transparent data sanitization, effective leakage reduction	May complicate user experience in shared environments
[75]	Loco-Store (Locality-Based Oblivious Storage)	High	Possible performance degradation	Enhanced protection against data leakage by hiding access patterns	Frequent data re-shuffling can impact system performance
[76]	Process Mining (Security Detection in Multi-Cloud)	High	High computational demands	Strengthens data loss prevention through anomaly detection	May hinder real-time detection due to high computational needs

Data redundancy in cloud storage is often represented by the replication factor R, which denotes the number of copies of data stored across different regions or availability zones as shown in Figure 12, if data is replicated in three regions, $R = 3$. The overall availability of the data P_{total}, can be calculated by considering the availability P_i of each individual copy. Assuming independent failure events, the availability is given by:

$$P_{total} = 1 - \prod_{i=1}^{R}(1 - P_i) \quad (8)$$

Eq. (8) highlights how increasing the replication factor enhances the data availability and reliability in cloud environments.

Existing studies. Table 14 summarizes recent approaches to data redundancy and distribution in cloud storage systems. Li and Li [66] proposed a demand-aware erasure coding scheme to optimize redundancy and fault tolerance. Ren et al. [67] explored secure distributed storage in adversarial networks. Li et al. [68] introduced a cryptographic mechanism for verifying data redundancy in multi-cloud environments. Kontodimas et al. [69] developed a framework for managing data distribution across cloud-edge infrastructures. Shithil and Adnan [70] proposed a strategy for replica selection to improve data retrieval in geo-distributed systems. Du et al. [71] presented a consistency model for maintaining data replication across geographically diverse nodes. Aldailamy et al. [72] focused on dynamic replication in multi-cloud environments for online social networks.

Discussion. The reviewed studies reveal several critical challenges in managing data redundancy and distribution.

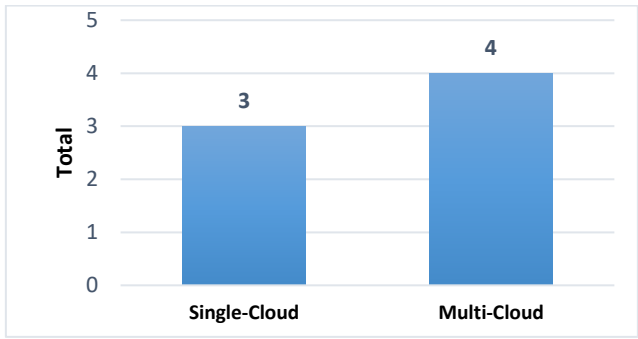

Figure 13. Distribution of papers per cloud model

This section examines various approaches related to data redundancy and distribution in cloud storage systems, and hightlights several challenges. Li and Li [66] proposed a Demand-Aware Erasure Coding method that optimizes redundancy, though its deployment can be complex. Ren et al. [67] addressed secure data storage in adversarial networks, with the challenge of maintaining consistency across nodes. Li et al. [68] introduced an Identity-Based Multi-Copy Data Possession mechanism that ensures data integrity but requires significant computational resources. Kontodimas et al. [69] developed a framework for secure distributed storage, enhancing data distribution while being resource-intensive. Shithil and Adnan [70] presented a Prediction-Based Replica

Selection strategy to improve retrieval efficiency, although it depends on predictive models that may not always be accurate. Du et al. [71] proposed a Cost-Effective Consistency Model, which may face performance issues due to the complexities of replication. Aldailamy et al. [72] focused on Dynamic Replication and Placement in multi-cloud environments, increasing complexity in dynamic scenarios. Figure 13 illustrates the distribution of these studies between single- and multi-cloud environments.

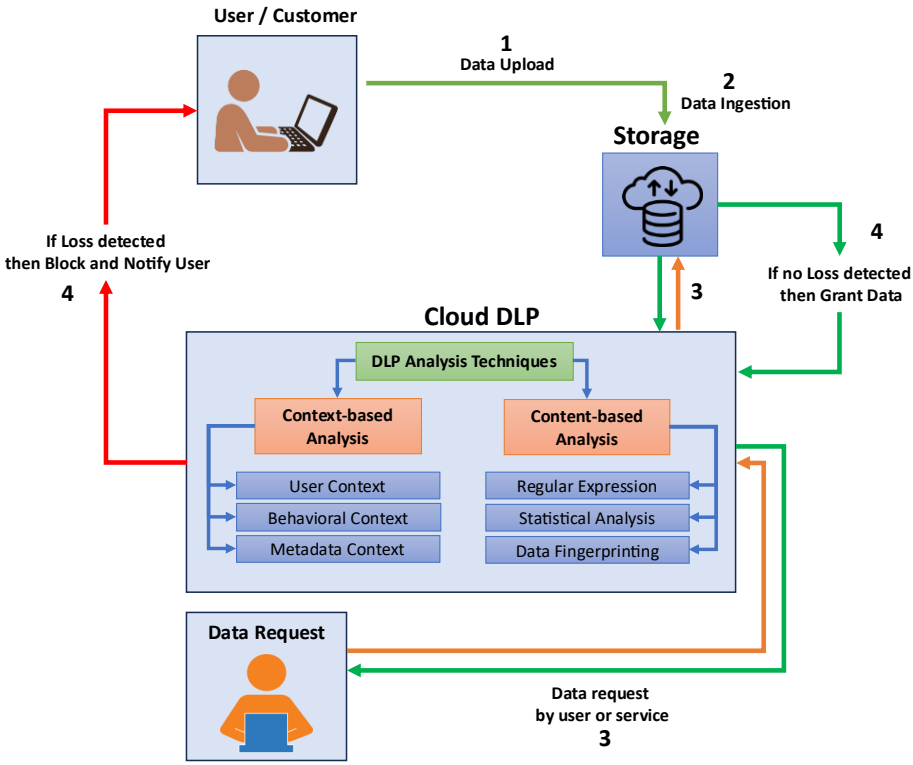

Figure 14. Concept of cloud DLP

Table 16. Comparative analysis of proposed machine learning models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[76]	Ensemble Learning with Feature Selection (Random Forest and PCA) for Network Intrusion Detection Systems (NIDS)	High	Moderate	Enhances detection accuracy by combining multiple classifiers and reducing dimensionality	High computational cost due to ensemble model complexity, potential delays in real-time applications
[77]	Cloud-Assisted Secure Data Classification using K-Nearest Neighbors (KNN) and Homomorphic Encryption	Moderate	High	Maintains data confidentiality and integrity during classification in smart city environments	Increased latency and reliance on robust cloud infrastructure
[78]	Time Series Anomaly Detection via Long Short-Term Memory (LSTM) Networks for Intrusion Detection	Moderate	Moderate	Enables accurate detection of temporal anomalies in cloud environments, enhancing security	Limited generalization to unseen attacks, moderate computational demands
[79]	Privacy-Preserving Cross-Media Retrieval using Searchable Encryption and Secure Indexing	High	High	Ensures data privacy during cross-media retrieval operations in cloud systems	High computational overhead, making it challenging for low-resource environments
[80]	Proactive Drive Failure Prediction for Cloud Storage System Through Semi-Supervised Learning with Label Propagation and Self-Training	Moderate	Moderate	Improves reliability by predicting drive failures proactively using semi-supervised learning techniques	May require extensive historical data for accurate predictions, and could be less effective with limited data
[81]	Hybrid Intrusion Detection Enhancement using Machine Learning and Deep Learning models	High	High	Combines both deep learning and machine learning models for improved detection accuracy	High computational requirements and potential complexity in tuning and integration

4.2.4 Data loss prevention

Background. Data Loss Prevention models are playing a key role in ensuring sensitive information detection and protection. It utilizes content-based techniques, such as pattern recognition and keyword matching, to identify sensitive data. Additionally, context-based methods assess data sensitivity by analyzing usage patterns and user roles. These combined techniques enhance data security and ensure compliance with regulatory standards in the cloud storage.

Figure 14 categorizes data protection approaches into content-based and context-based methods. Content-based techniques, like keyword matching and regular expressions, detect sensitive information directly within the data. Context-based methods, such as contextual analysis and behavioral monitoring, evaluate data sensitivity based on its environment and usage patterns. In keyword matching, the detection score S_D is calculated as:

$$S_D = \sum_{i=1}^{n} w_i \,.\, K_i \tag{9}$$

where, K_i indicates the presence of keyword k_i and w_i represents its weight. In regular expressions, the score is given by:

$$S_D = \sum_{i=1}^{m} w_i \,.\, P_i(D) \tag{10}$$

where, $P_i(D)$ is 1 if pattern p_i matches data D. Context-based analysis integrates content with contextual factors to assess sensitivity using:

$$S = \alpha \,.\, ContentScore(D) + \beta . ContextScore(C) \tag{11}$$

where, α and β are weights. Behavioral analysis involves calculating an anomaly score A as:

$$S = \frac{B_i - \mu}{\sigma} \tag{12}$$

where, B_i represents behavioral metrics, the average and standard deviation of typical behavior are represented by μ and σ, respectively.

Existing studies. Table 15 reviews recent data loss prevention approaches for cloud storage. Han et al. [73] developed CloudDLP, which sanitizes data during transfers to prevent leakage but may affect user experience. Tian et al. [74] introduced Loco-Store, which hides access patterns to protect data but may slow performance due to frequent data reshuffling. Zhang et al. [75] proposed a security detection framework for multi-cloud environments that uses process mining to detect anomalies, though its high computational demands may limit real-time threat detection.

Discussion. Preventing data loss and leakage in cloud storage is complex, as shown by recent studies. Han et al. [73] developed CloudDLP, a data sanitization tool that reduces leakage risks but may complicate user experience in shared settings. Tian et al. [74] introduced Loco-Store, which hides access patterns but suffers from performance issues due to frequent data reshuffling. Zhang et al. [75] created a process mining-based framework for multi-cloud environments to improve data loss prevention, though it faces challenges with high computational demands that can affect real-time detection. These studies underscore the trade-offs between security, usability, and performance in cloud storage solutions.

4.2.5 Machine learning

Background. Machine learning (ML) has significantly advanced data protection in cloud storage by utilizing algorithms that analyze both data context and content. As shown in Figure 15, ML can detect anomalies by examining user behavior and access patterns, identifying potential security issues like unauthorized access or data breaches. Figure 15 illustrates that while each ML model addresses specific security problems independently, their combined use enhances overall performance.

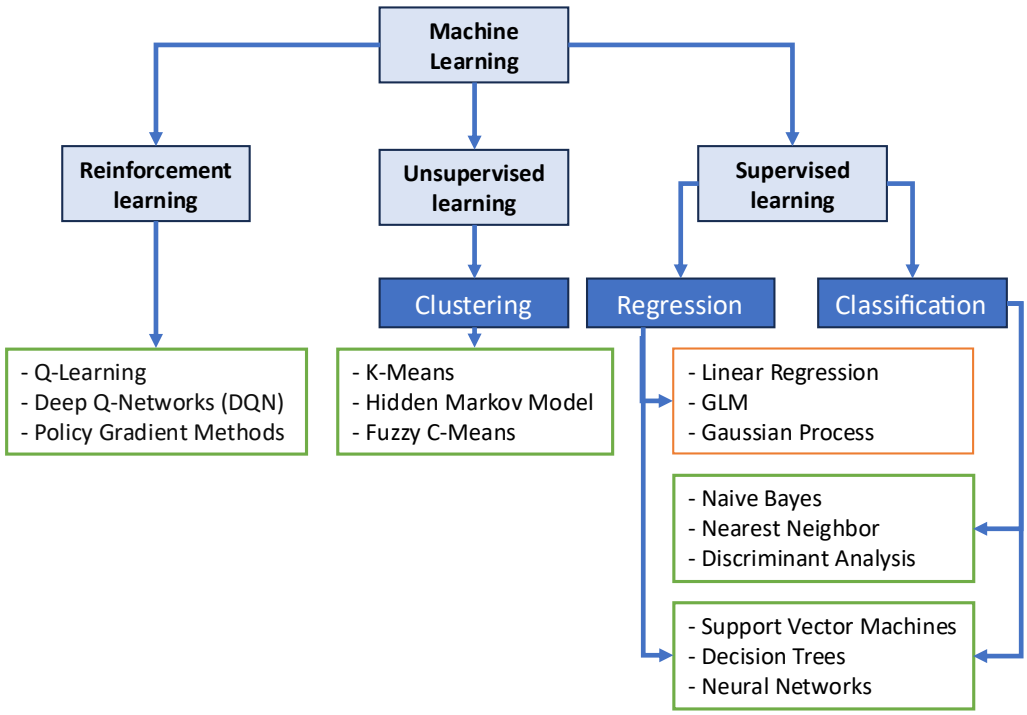

Figure 15. Machine Learning Models

Table 17. Machine learning applied in cloud security

Model	Category	Usage	Focus
Supervised Learning	Data Classification	Identifies and protects sensitive data at rest by learning from labeled examples.	Data Content
Unsupervised Learning	Anomaly Detection	Monitors data in transit to detect unusual patterns or threats without prior labels.	Data Context
Reinforcement Learning	Dynamic Security Policies	Optimizes security policies and access controls in real-time based on interactions with data.	Data Context

Machine learning enhances cloud data protection through the various techniques listed in Table 16, each of which utilizes specific mathematical models. Supervised learning classifies data using models like decision trees, where the classification function $f(x)$ is trained with labeled data (x, y) optimizing the objective function.

$$\min \sum_{i=1}^{n} L(y_i, f(x_i)) \tag{13}$$

where, L is a loss function. Unsupervised learning detects anomalies by learning data distribution. Reinforcement learning optimizes security policies based on a cumulative reward function, where represents the reward at each step and is the discount factor.

Existing studies. Table 17 reviews recent machine learning approaches for cloud data storage security. Khan and Haroon [76] developed a network intrusion detection system that uses ensemble learning to improve accuracy but may increase computational demands. Kumar et al. [77] proposed a cloud-based classification method for secure data storage in smart cities, effective for large-scale data but potentially affected by cloud dependency. Al-Ghuwairi et al. [78] created an intrusion detection system that identifies time-series anomalies, though it relies heavily on historical data, which might limit its adaptability. Wang et al. [79] introduced a privacy-preserving retrieval framework that combines cryptographic techniques with machine learning, ensuring data confidentiality and efficient retrieval. Zhou et al. [80] developed a proactive drive failure prediction system using semi-supervised learning, improving accuracy and reliability. Sajid et al. [81] presented a hybrid machine and deep learning approach to enhance intrusion detection.

Discussion. Recent machine learning-based approaches have demonstrated both advancements and limitations. Khan and Haroon's [76] network intrusion detection system (NIDS) offers high accuracy but suffers from high computational complexity, impacting real-time performance. Kumar et al.'s [77] cloud-supported classification method effectively safeguards data in smart cities but faces scalability and latency issues due to its reliance on cloud infrastructure. Al-Ghuwairi et al.'s [78] time-series anomaly detection system enhances data protection but is limited by its dependence on historical data, affecting its adaptability to new threats. Wang et al.'s [79] privacy-preserving retrieval framework maintains data confidentiality but may encounter efficiency challenges across diverse data types. Zhou et al.'s [80] proactive drive failure prediction system improves prediction accuracy but may struggle with hardware generalization. Sajid et al.'s [81] hybrid machine and deep learning approach improves intrusion detection but involves higher computational costs and integration complexities.

4.2.6 Blockchain

Background. Blockchain has emerged as a valuable promising option in the age of cloud security. Its decentralized structure helps reduce key risks in traditional cloud systems by storing data transactions across multiple nodes, making unauthorized changes more difficult. The immutable ledger of blockchain ensures that transation records cannot be altered once logged, which provide a strong protection against data breaches and fraud.

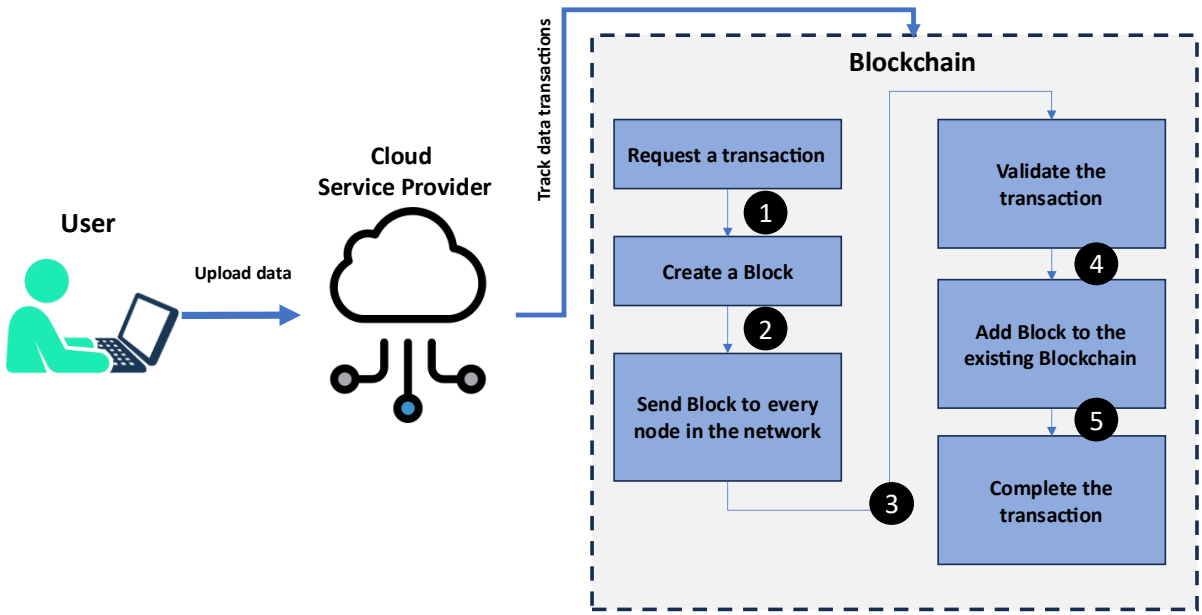

Figure 16. Usage of blockchain in cloud storage

Table 18. Comparative analysis of proposed blockchain models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[82]	Collaborative Auditing with Blockchain (Merkle Trees and Smart Contracts)	Moderate	Moderate	Enhances data integrity verification through blockchain's immutable ledger and Merkle Tree structures; smart contracts automate audit processes.	Complexity in integrating blockchain with existing cloud systems; potential latency in smart contract execution and Merkle Tree updates.
[83]	Multi-Replica Public Auditing (Proof-of-Replication and Cryptographic Proofs)	High	Moderate	Improves fault tolerance and data resilience by utilizing Proof-of-Replication for data verification across multiple cloud environments, supported by cryptographic proofs.	High complexity in synchronizing and managing data replicas; increased overhead due to cryptographic operations.
[84]	Decentralized Privacy-Preserving Auditing (Zero-Knowledge Proofs and Blockchain)	Moderate	Moderate	Safeguards data confidentiality during auditing by integrating Zero-Knowledge Proofs with blockchain's decentralized ledger.	Computational overhead from Zero-Knowledge Proofs may impact system performance.
[85]	AuthPrivacyChain (Decentralized Identity Management and Privacy-Enhanced Access Control)	High	Low to Moderate	Provides robust access control and privacy protection through decentralized identity management and privacy-enhancing access control protocols using blockchain.	High implementation complexity; potential performance trade-offs in privacy-enhanced access controls.
[86]	Selective Sharing of Outsourced Encrypted Data (Encryption and Access Control Policies)	Moderate	Moderate	Facilitates the safe exchange of encrypted data with a flexible access control policy, ensuring data confidentiality in the context of cloud	Complexity in managing encryption keys and access control policies; potential performance impact due to encryption overhead.
[87]	Efficient Data Integrity Verification (Blockchain and Hashing Algorithms)	Moderate	Moderate	Enhances data integrity checks across multiple cloud platforms by leveraging blockchain's decentralized ledger technology and sophisticated hashing techniques for rapid authentication.	Challenges in scaling and managing large volumes of data across multiple cloud providers.
[88]	Decentralized Public Auditing (Blockchain and Cryptographic Proofs)	High	Moderate	Enhances transparency and integrity of cloud storage through decentralized auditing mechanisms supported by cryptographic proofs.	High computational requirements for cryptographic proofs; potential delays in audit results.
[89]	Secure Deduplication and Shared Auditing (Blockchain and Deduplication Techniques)	High	Moderate	Combines secure deduplication with shared auditing using blockchain to minimize redundant data storage and ensure data integrity.	Complexity in implementing secure deduplication processes; increased overhead from shared auditing.
[90]	Decentralized Storage Auditing (Blockchain and Proof-of-Ownership)	High	Moderate	Utilizes blockchain for decentralized storage auditing and Proof-of-Ownership to enhance data security and auditing efficiency.	High complexity in maintaining and verifying Proof-of-Ownership; potential performance trade-offs.
[91]	Certificateless Public Cloud Data Integrity Auditing (Blockchain and Certificateless Cryptography)	High	Moderate	Integrates blockchain with certificateless cryptography for secure and efficient public cloud data integrity auditing.	Complexity in implementing certificateless cryptography; potential performance impact due to additional cryptographic operations.
[92]	Deduplication and Integrity Auditing (Blockchain and Encryption)	High	Moderate	Integrates blockchain technology with cryptographic methods to enable deduplication and verify integrity in encrypted cloud storage systems.	High implementation complexity; potential performance degradation due to encryption and deduplication processes.
[93]	Key Security Management with Blockchain and Digital Twins (Blockchain and Digital Twins)	High	High	Enhances key security management in cloud storage through integration with digital twins and blockchain technology.	High complexity in managing digital twins and blockchain integration; potential performance impact.

As shown in Figure 16, blockchain technology operates based on the principle of decentralized consensus and cryptographic security. At its core, blockchain uses hash functions such as SHA-256, where a given input M is transformed into a fixed-size output H via:

$$H = SHA - 256(M) \quad (14)$$

To ensure data integrity and enable efficient verification, blockchain employs Merkle trees, where each parent node hash H_{parent} is computed as:

$$H_{parent} = Hash(H_{left} \parallel H_{right}) \quad (15)$$

Digital signatures, created using asymmetric cryptography, secure transactions with equations such as:

$$Signature = Sign(H_{transaction}, PrivateKey) \quad (16)$$

which can be verified with:

$$Verify(Signature, H_{transaction}, PublicKey) \quad (17)$$

In cloud data storage, blockchain can be used to track who accessed data, check if data have been changed, and create a secure record of storage activities. This is especially helpful in environments where accountability is important, such as healthcare or financial contexts. However, using blockchain in cloud systems also has some challenges. These include high computing costs and difficulty connecting with existing cloud service providers.

Existing studies. Table 18 provides a comparative overview of recent models that implements blockchain. Each study is assessed based on the used method, implementation complexity, and performance impact. Blockchain technology is increasingly utilized to address security concerns in cloud storage by offering decentralized and immutable data protection solutions. Huang et al. [82] developed a collaborative auditing framework leveraging blockchain and smart contracts for automated audit processes. Yang et al. [83] introduced a public audit scheme for multi-cloud environments, incorporating blockchain to ensure data synchronization and verification. Miao et al. [84] created a privacy-preserving auditing approach using blockchain and Zero-Knowledge Proofs to maintain data confidentiality. Yang et al. [85] proposed AuthPrivacyChain, combining blockchain with decentralized identity management for enhanced access control. Other studies, such as those by Sifah et al. [86], Zhang et al. [87], and Shu et al. [88], focused on improving data integrity, transparency, and deduplication in cloud storage using blockchain technologies. These advancements reflect a growing emphasis on integrating blockchain to bolster cloud data security, balancing enhanced protection with the challenges of implementation and performance.

Discussion. Recent studies have explored the use of blockchain to improve cloud data storage security, showing both promising outcomes and notable challenges. Huang et al. [82] and Yang et al. [83] used blockchain and Zero-Knowledge Proofs to enhance data integrity and privacy auditing, though their methods face integration and performance issues. Miao et al. [84] and Yang et al. [85] applied Proof-of-Replication for data resilience but introduced added complexity. Sifah et al. [86] and Zhang et al. [87] improved transparency with decentralized auditing, yet scalability remains a concern. Shu et al. [88] and Tian et al. [89] used blockchain for access control and identity management, but their models may reduce efficiency. Du et al. [90] and Du et al. [91] combined encryption with blockchain for secure sharing and deduplication, facing trade-offs in processing speed. Song et al. [92] and Huang and Yi [93] proposed decentralized auditing and key management, but these also increased implementation complexity.

Table 19. Comparative analysis of proposed hybrid models

Ref.	Method	Implementation Complexity	Performance Impact	Advantages	Disadvantages
[94]	Identity-Based Encryption (IBE) + Access Control Mechanism	Moderate	Moderate	Strong identity-based access control, enhanced security for sensitive data sharing in cloud environments.	Complex key management and potential performance issues with large datasets.
[95]	Machine Learning-Based Trust Management + Blockchain Integration	High	High	Robust security through dynamic trust management and immutable blockchain records.	High computational cost and resource-intensive implementation.
[96]	Role-Based Access Control (RBAC) + Authorized Keyword Search	Moderate	Low	Efficient data retrieval with role-based access, enhanced data security through keyword-based encryption.	Limited scalability and potential complexity in managing user roles and permissions.
[97]	Artificial Bee Colony Algorithm + Adaptive Data Security	High	Moderate	Improved data security and encryption strength using adaptive algorithms.	Complexity in implementation and tuning of the algorithm parameters.
[98]	Attribute-Based Encryption (ABE) + Blockchain for Public Traceability	High	Moderate	Enhanced data confidentiality, fine-grained access control, and secure traceability of access events.	Complex setup and management of large-universe attribute-based encryption schemes.
[99]	Digital Twin-Based Security + Privacy Enhancements	High	High	Enhanced security and privacy for medical records in cloud storage via digital twin models.	High implementation complexity and resource requirements for maintaining digital twins.
[46]	Dynamic AES Encryption + Blockchain Key Management	High	Moderate	Strong encryption with dynamic AES and secure key management through blockchain.	Potential latency issues and high resource consumption during key management operations.

Table 20. Comparative analysis of proposed approaches

Approach	Real-World Use Cases	Benefits	Limitations	Considerations
Encryption	Protecting stored health records, financial data, and legal files in cloud storage	Ensures data confidentiality and protection	Resource-intensive, potential performance impact	Key management, encryption algorithm strength
DLP	Blocking sensitive data uploads to cloud storage	Prevents data leakage and unauthorized access	May produce false positives, complex configuration	Integration with existing systems, scalability
Access Control	Managing user access to cloud storage buckets, collaboration tools, and shared company drives	Restricts access based on user roles and permissions	Complexity in managing permissions and roles	Policy enforcement, granularity of access control
Data Redundancy	Backing up data across cloud regions, maintaining mirrored storage for disaster recovery	Enhances data availability and disaster recovery	Increased storage costs, potential redundancy overhead	Storage costs, recovery time
Machine Learning	Detecting unusual file access, insider threats, and ransomware in cloud storage	Improves threat detection and response	Requires large datasets for training, potential false positives	Model accuracy, training data quality
Blockchain	Recording file changes, verifying data integrity, and auditing access history in decentralized cloud logs	Provides immutability and transparency	Complex integration with existing systems	Integration with cloud infrastructure, scalability

4.2.7 Hybrid models

Existing studies. Table 19 highlights papers proposing hybrid and mixed- approaches for cloud data security. Yang et al. [94] proposed an Identity-Based Encryption (IBE) system with access control for secure data sharing. Franklin et al. [95] combined Machine Learning with blockchain for improved trust management. Sultan et al. [96] developed an RBAC scheme with keyword search to enhance data retrieval. Geetha [97] introduced an Adaptive Artificial Bee Colony Algorithm for dynamic encryption strengthening. Yan et al. [98] integrated Attribute-Based Encryption (ABE) with blockchain for fine-grained access control. Yi [99] proposed a Digital Twin-Based Security framework for improved security and privacy. Shakor et al. [46] created a Dynamic AES Encryption method with blockchain key management for robust encryption. These models blend advanced technologies to enhance data security while considering practical implementation challenges.

Discussion. Yang et al. [94] and Franklin et al. [95] proposed an advanced identity-based encryption and machine learning-based trust management, integrating these with blockchain for enhanced security. However, both face issues with complex key management and high computational demands. Sultan et al. [96] and Geetha [97] introduced role-based access control with keyword search and adaptive algorithms, respectively, which improve security but may struggle with scalability and implementation complexities. Yan et al. [98] and Yi [99] proposed attribute-based encryption and digital twin-based frameworks for robust access control and privacy, but they face challenges in managing extensive attributes and integrating with cloud environments. Shakor et al. [46] developed dynamic AES encryption with blockchain key management, enhancing encryption and key security while potentially increasing latency and resource consumption.

5. FINDINGS

5.1 General overview

Table 20 provides an overview of various cloud data storage security approaches, each with distinct benefits and challenges. Encryption ensures data confidentiality and compliance but can impact performance. It is the most frequently used method in the reviewed studies due to its fundamental role in securing data at rest and in transit, ease of integration with existing systems, and broad regulatory acceptance. Its mature implementation in both public and private cloud services also contributes to its widespread adoption. Data Loss Prevention models help to protect sensitive information but may produce false positives and be complex to configure. This complexity, along with the need for constant policy updates and fine-tuning, may explain its less frequent use. Access Control methods secure cloud environments based on user roles but can be difficult to manage. Data Redundancy enhances availability and disaster recovery but can lead to higher storage costs. Machine Learning offers advanced threat detection but needs extensive training data and may have false positives. Blockchain provides immutability and transparency for records but can be complex to integrate. Each method has unique challenges, highlighting the need for a balanced security strategy.

5.2 Performance metrics

Studies in the current literature review have employed a range of performance metrics to assess each model's performance. As shown in Figure 17, the metrics for evaluating cloud data security approaches include encryption throughput and key management overhead for encryption methods, data classification accuracy and false positive rate for DLP systems, and policy application time and role management efficiency for access control mechanisms. Data redundancy is assessed by replication speed and failover time, machine learning models by anomaly detection rate and training time complexity, and blockchain solutions by transaction processing speed and block verification time. This comprehensive evaluation highlights the performance strengths and trade-offs of each approach in enhancing data protection.

5.3 Key challenges

Cloud data storage protection faces significant challenges across the key areas studied in this systematic review:

- **Encryption:** This introduces computational overhead and complex key management issues, particularly in distributed environments, where rapid key updates are necessary.

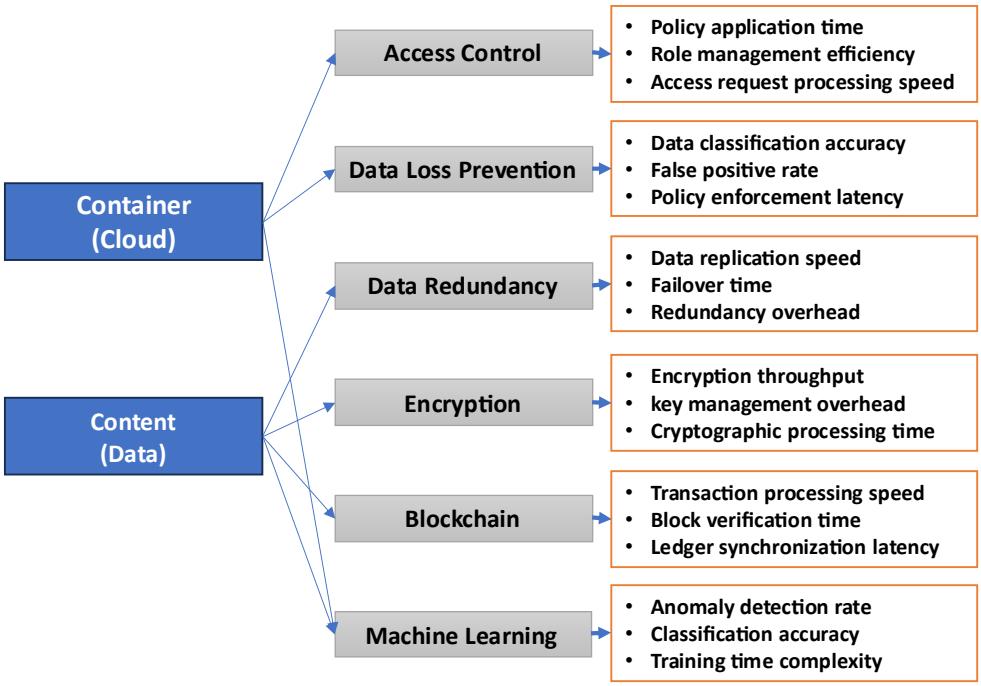

Figure 17. Metrics used in the current literature survey

- **Access Control Mechanisms:** They struggle with scalability, leading to performance bottlenecks and potential security gaps as the number of users and roles grows.
- **DLP Models:** They often deal with significant false-positive ratios and the difficulty of real-time data inspection, which can hinder operational efficiency.
- **Blockchain:** While enhancing data integrity, blockchain faces challenges such as latency, high resource consumption, and scalability issues owing to decentralized consensus protocols.
- **Machine Learning:** Models used for threat detection need large volumes of training data and significant computational resources, and are vulnerable to adversarial attacks.
- **Data Redundancy:** Models must balance the need for fault tolerance with the challenges of managing data consistency, storage costs, and compliance with regulations in a distributed cloud environment.

Furthermore, as complex security threats continue to advance, it has become evident that existing methods must be improved to adequately secure the data in the cloud.

6. LIMITATION OF THIS REVIEW

This systematic review of cloud data storage security techniques acknowledges several inherent limitations. These include potential biases in literature selection and the fast-paced evolution of technology, which may affect the applicability and relevance of the findings. The following section details these key constraints to consider when interpreting the review's conclusions.

Table 21 summarizes the key constraints of this study, clarifying the limitations discussed in the systematic review. It categorizes these limitations, highlights their potential impacts on the findings, and provides insights into how they might affect the interpretation and generalizability of the results.

Table 21. Limitations of the review

Limitation	Description	Potential Impact
Literature Scope	Focuses on English-language studies and major databases, possibly missing relevant work in other languages or sources.	May exclude significant findings and perspectives.
Evolution of Technology	Rapid changes in cloud storage security could render some discussions outdated.	Limits long-term relevance of the review's conclusions.
Sectors Considerations	Does not extensively cover sector-specific challenges and solutions.	Might not fully address unique security needs of specific sectors.

7. CONCLUSION

This paper presented a systematic review of cloud data storage security techniques, based on 77 studies published between 2020 and 2024. The findings reveal that no single method is sufficient to fully protect cloud data from evolving security threats. A multi-layered approach is therefore essential to address the key principles of data security: confidentiality, integrity, and availability.

The review confirms that encryption is the most widely adopted technique, offering strong protection for data confidentiality. Access control and data redundancy also play vital roles in preventing unauthorized access and ensuring data availability. Blockchain enhances auditability and data integrity, while machine learning contributes to real-time threat detection through classification techniques. Despite these advantages, each method faces practical challenges, such as computational overhead, integration complexity, and regulatory compliance issues.

The main contribution of this paper is a structured and comparative analysis of the major security techniques used in

cloud storage. By highlighting their strengths, limitations, and application scenarios, the study provides meaningful insights for researchers aiming to develop more secure, scalable, and efficient cloud data protection models.

In particular, the integration of blockchain and machine learning presents a promising direction for future research. Blockchain can serve as a tamper-proof ledger for recording data access and system activities, while machine learning can analyze this data to detect anomalies, predict threats, and automate response mechanisms. Together, these technologies offer a foundation for intelligent, adaptive, and transparent cloud storage security approachs. This review underscores the importance of combining multiple techniques to address modern cloud storage security challenges and supports the development of innovative models in both academic and industrial contexts.

ACKNOWLEDGMENT

The C3S Research Laboratory generously supported this research. We sincerely thank everyone whose ideas, concepts, and methodologies, shared through academic articles and publications, have inspired and enriched our work.

REFERENCES

[1] Shen, J., Liu, D., Shen, H., Chen, X., Susilo, W. (2019). Cloud computing security: Fundamental challenges and future research. IEEE Access, 7: 22328-22340. https://doi.org/10.1109/ACCESS.2019.2908331

[2] Tabrizchi, H., Kuchaki Rafsanjani, M. (2020). A survey on security challenges in cloud computing: Issues, threats, and solutions. The Journal of Supercomputing, 76(12): 9493-9532.

[3] Singh, S., Jeong, Y.S., Park, J.H. (2016). A survey on cloud computing security: Issues, threats, and solutions. Journal of Network and Computer Applications, 75: 200-222. https://doi.org/10.1016/j.jnca.2016.09.002

[4] Yang, H., Fu, X., Guo, Y., Li, H. (2020). Secure data sharing and searching at the edge of cloud-assisted Internet of Things. IEEE Access, 8: 27373-27384. https://doi.org/10.1109/ACCESS.2020.3035350

[5] Das, A.K., Ma, P.M.W., Lo, R.P.W., Rodrigues, J.J.P.C. (2020). Privacy-preserving cloud-based personal health record system using attribute-based encryption and blockchain. IEEE Access, 8: 182922-182933. https://doi.org/10.1109/ACCESS.2020.3033210

[6] Ruj, S., Basu, S., Sakurai, K. (2021). Cloud computing security: Fundamentals, technologies, and applications. IEEE Access, 9: 157401-157418. https://doi.org/10.1109/ACCESS.2021.3075727

[7] Sun, P.J. (2019). Privacy protection and data security in cloud computing: A survey, challenges, and solutions. IEEE Access, 7: 147420-147452. https://doi.org/10.1109/ACCESS.2019.2935273

[8] Zhao, M., Liu, W., He, K. (2022). Research on data security model of environmental monitoring based on blockchain. IEEE Access, 10: 120168-120180. https://doi.org/10.1109/ACCESS.2022.3228365

[9] Khashan, O.A. (2020). Secure outsourcing and sharing of cloud data using a user-side encrypted file system. IEEE Access, 8: 210855-210867. https://doi.org/10.1109/ACCESS.2020.3035451

[10] Sahi, A., Lai, D., Li, Y. (2021). A review of the state of the art in privacy and security in the eHealth cloud. IEEE Access, 9: 104127-104141. https://doi.org/10.1109/ACCESS.2021.3085432

[11] Flexera. State of the Cloud. http://info.flexera.com/CM-REPORT-State-of-the-Cloud, accessed on Sep. 5, 2024.

[12] IDC. IDC Report. https://www.idc.com/getdoc.jsp?containerId=prUS52343224, accessed on Sep. 5, 2024.

[13] IoT Analytics. State of IoT – Spring 2023. http://iot-analytics.com/product/state-of-iot-spring-2023, accessed on Sep. 5, 2024.

[14] IBM. Data Breach Report. http://www.ibm.com/reports/data-breach, accessed on Sep. 5, 2024.

[15] He, J., Zhang, Z., Li, M., Zhu, L., Hu, J. (2019). Provable data integrity of cloud storage service with enhanced security in the Internet of Things. IEEE Access, 7: 6226-6239. https://doi.org/10.1109/ACCESS.2019.2912345

[16] Albshaier, L., Budokhi, A., Aljughaiman, A. (2024). A review of security issues when integrating IoT with cloud computing and blockchain. IEEE Access, 12: 109560-109595. https://doi.org/10.1109/ACCESS.2024.3176894

[17] Alouffi, B., Hasnain, M., Alharbi, A.S., Alosaimi, W., Alyami, H., Ayaz, M. (2021). A systematic literature review on cloud computing security: Threats and mitigation strategies. IEEE Access, 9: 57792-57807. https://doi.org/10.1109/ACCESS.2021.3076791

[18] Amazon Web Services. What is cloud storage? https://aws.amazon.com/what-is/cloud-storage/, accessed on Sep. 5, 2024.

[19] Spiceworks. What is cloud storage? https://www.spiceworks.com/tech/cloud/articles/what-is-cloud-storage/, accessed on Sep. 5, 2024.

[20] McGowan, J.S.P., Shukla, M.S. (2020). Understanding the shared responsibility model in cloud security. IEEE Access, 8: 151328-151339. https://doi.org/10.1109/ACCESS.2020.3019815

[21] Hu, S.J., Wu, M.L., Chang, K.C. (2020). Cloud security and compliance: The role of the shared responsibility model. IEEE Cloud Computing, 7(3): 50-58. https://doi.org/10.1109/MCC.2020.3019381

[22] Yang, R.Y., Islam, M.S., Islam, N.M.S. (2021). A survey on cloud security and the shared responsibility model. IEEE Access, 9: 127482-127496. https://doi.org/10.1109/ACCESS.2021.3095128

[23] Al-Kahtani, M.M., Kim, Y., Kumar, M.M.S. (2021). Cloud data security threats and solutions: A survey. IEEE Access, 9: 82188-82205. https://doi.org/10.1109/ACCESS.2021.3085450

[24] Liu, R., Liu, Y., Liu, X. (2023). Secure data storage in cloud computing: Challenges and mitigation techniques. IEEE Access, 11: 97542-97558. https://doi.org/10.1109/ACCESS.2023.3238651

[25] Zhang, Z., Zeng, P., Pan, B., Choo, K.K.R. (2020). Large-universe attribute-based encryption with public traceability for cloud storage. IEEE Internet of Things Journal, 7(10): 10314-10323. https://doi.org/10.1109/JIOT.2020.2986303

[26] Moyou Metcheka, L., Ndoundam, R. (2020). Distributed data hiding in multi-cloud storage environment. Journal of Cloud Computing, 9(1): 68.

https://doi.org/10.1186/s13677-020-00208-4
[27] Zhang, X., Si, W. (2021). Efficient auditing scheme for secure data storage in fog-to-cloud computing. IEEE Access, 9: 37951-37960. https://doi.org/10.1109/ACCESS.2020.2971630
[28] Song, H., Li, J., Li, H. (2021). A cloud secure storage mechanism based on data dispersion and encryption. IEEE Access, 9: 63745-63751. https://doi.org/10.1109/ACCESS.2021.3075340
[29] Ming, Y., He, B., Wang, C. (2021). Efficient revocable multi-authority attribute-based encryption for cloud storage. IEEE Access, 9: 42593-42603. https://doi.org/10.1109/ACCESS.2021.3066212
[30] Qi, S., Lu, Y., Wei, W. Chen, X. (2021). Efficient data access control with fine-grained data protection in cloud-assisted IIoT. IEEE Internet of Things Journal, 8(4): 2886-2899. https://doi.org/10.1109/JIOT.2020.3020979
[31] Mossebo Tcheunteu, S.W., Moyou Metcheka, L., Ndoundam, R. (2021). Distributed data hiding in a single cloud storage environment. Journal of Cloud Computing, 10(1): 43. https://doi.org/10.1186/s13677-021-00258-2
[32] Miao, Y., Tong, Q., Deng, R.H., Choo, K.K.R., Liu, X., Li, H. (2022). Verifiable searchable encryption framework against insider keyword-guessing attack in cloud storage. IEEE Transactions on Cloud Computing, 10(2): 835-848. https://doi.org/10.1109/TCC.2020.2989296
[33] Liu, J.N., Luo, X.A., Wang, J., Yang, A.J., Wang, X.A., Li, M. (2022). Enabling efficient, secure and privacy-preserving mobile cloud storage. IEEE Transactions on Dependable and Secure Computing, 19(3): 1518-1531. https://doi.org/10.1109/TDSC.2020.3027579
[34] Ge, C., Susilo, W., Baek, J., Liu, Z., Xia, J., Fang, L. (2022). Revocable attribute-based encryption with data integrity in clouds. IEEE Transactions on Dependable and Secure Computing, 19(5): 2864-2872. https://doi.org/10.1109/TDSC.2021.3065999
[35] Ihtesham, M., Tahir, S., Tahir; H., Hasan, A., Sultan, A., Saeed, S. (2023). Privacy preserving and serverless homomorphic-based searchable encryption as a service (SEaaS). IEEE Access, 11: 115204-115218. https://doi.org/10.1109/ACCESS.2023.3324817
[36] Liu, H., Lu, X., Duan, S., Zhang, Y., Xiang, Y. (2023). An efficient oblivious random data access scheme in cloud computing. IEEE Transactions on Cloud Computing, 11(2): 1940-1953. https://doi.org/10.1109/TCC.2022.3173260
[37] Nasiraee, H., Ashouri-Talouki, M., Liu, X. (2023). Optimal black-box traceability in decentralized attribute-based encryption. IEEE Transactions on Cloud Computing, 11(3): 2459-2472. https://doi.org/10.1109/TCC.2022.3210137
[38] Zhang, Y., Zhu, T., Guo, R., Xu, S., Cui, H., Cao, J. (2023). Multi-keyword searchable and verifiable attribute-based encryption over cloud data. IEEE Transactions on Cloud Computing, 11(1): 971-983. https://doi.org/10.1109/TCC.2021.3119407
[39] Xu, S., Horng, J.H., Chang, C.C., Chang, C.C. (2023). Reversible data hiding with hierarchical block variable length coding for cloud security. IEEE Transactions on Dependable and Secure Computing, 20(5): 4199-4213. https://doi.org/10.1109/TDSC.2022.3219843
[40] Cui, J., Li, B., Zhong, H., Xu, Y., Liu, L. (2023). Achieving revocable attribute group-based encryption for mobile cloud data: A multi-proxy assisted approach. IEEE Transactions on Dependable and Secure Computing, 20(4): 2988-3001. https://doi.org/10.1109/TDSC.2022.3204549
[41] Suganya, M., Sasipraba, T. (2023). Stochastic gradient descent long short-term memory based secure encryption algorithm for cloud data storage and retrieval in cloud computing environment. Journal of Cloud Computing, 12(1): 74. https://doi.org/10.1186/s13677-023-00442-6
[42] Huang, B., Gao, J., Li, X. (2023). Efficient lattice-based revocable attribute-based encryption against decryption key exposure for cloud file sharing. Journal of Cloud Computing, 12(1): 37. https://doi.org/10.1186/s13677-023-00414-w
[43] Rupa, C., Greeshmanth, Shah, M.A. (2023). Novel secure data protection scheme using Martino homomorphic encryption. Journal of Cloud Computing, 12(1): 47. https://doi.org/10.1186/s13677-023-00425-7
[44] Gadde, S., Amutharaj, J., Usha, S. (2023). A security model to protect the isolation of medical data in the cloud using hybrid cryptography. Journal of Information Security and Applications, 73: 103412. https://doi.org/10.1016/j.jisa.2022.103412
[45] Li, X., Li, H., Gao, J., Wang, R. (2023). Privacy preserving via multi-key homomorphic encryption in cloud computing. Journal of Information Security and Applications, 74: 103463. https://doi.org/10.1016/j.jisa.2023.103463
[46] Shakor, M.Y., Khaleel, M.I., Safran, M., Alfarhood, S., Zhu, M. (2024). Dynamic AES encryption and blockchain key management: A novel solution for cloud data security. IEEE Access, 12: 26334-26343. https://doi.org/10.1109/ACCESS.2024.3351119
[47] Hosseingholizadeh, A., Rahmati, F., Ali, M., Damadi, H., Liu, X. (2024). Privacy-preserving joint data and function homomorphic encryption for cloud software services. IEEE Internet of Things Journal, 11(1): 728-741. https://doi.org/10.1109/JIOT.2023.3286508
[48] Zhang, B., Yang, W., Zhang, F., Ning, J. (2024). Efficient attribute-based searchable encryption with policy hiding over personal health records. IEEE Transactions on Dependable and Secure Computing, 22(2): 1299-1312. https://doi.org/10.1109/TDSC.2024.3432769
[49] Li, W., Susilo, W., Xia, C., Huang, L., Guo, F., Wang, T. (2024). Secure data integrity check based on verified public key encryption with equality test for multi-cloud storage. IEEE Transactions on Dependable and Secure Computing, 21(6): 5359-5373. https://doi.org/10.1109/TDSC.2024.3375369
[50] Baseri, Y., Hafid, A., Firoozjaei, M.D., Cherkaoui, S., Ray, I. (2024). Statistical privacy protection for secure data access control in cloud. Journal of Information Security and Applications, 84: 103823. https://doi.org/10.1016/j.jisa.2024.103823
[51] Xiong, S., Ni, Q., Wang, L., Wang, Q. (2020). SEM-ACSIT: Secure and efficient multiauthority access control for IoT cloud storage. IEEE Internet of Things Journal, 7(4): 2914-2927. https://doi.org/10.1109/JIOT.2020.2963899
[52] Ning, J., Huang, X., Susilo, W., Liang, K., Liu, X., Zhang, Y. (2020). Dual access control for cloud-based data storage and sharing. IEEE Transactions on Dependable and Secure Computing, 19(2): 1036-1048.

https://doi.org/10.1109/TDSC.2020.3011525
[53] Li, J., Ma, J., Miao, Y., Yang, R., Liu, X., Choo, K.K.R. (2022). Practical multi-keyword ranked search with access control over encrypted cloud data. IEEE Transactions on Cloud Computing, 10(3): 2005-2019. https://doi.org/10.1109/TCC.2020.3024226
[54] Panchal, G., Samanta, D., Das, A.K., Kumar, N., Choo, K.K.R. (2022). Designing secure and efficient biometric-based access mechanism for cloud services. IEEE Transactions on Cloud Computing, 10(2): 749-761. https://doi.org/10.1109/TCC.2020.2987564
[55] Gill, P., Dietl, W., Tripunitara, M.V. (2022). Least-privilege calls to Amazon Web Services. IEEE Transactions on Dependable and Secure Computing, 20(3): 2085-2096. https://doi.org/10.1109/TDSC.2022.3171740
[56] Xue, K., Gai, N., Hong, J., Wei, D.S.L., Hong, P., Yu, N. (2022). Efficient and secure attribute-based access control with identical sub-policies frequently used in cloud storage. IEEE Transactions on Dependable and Secure Computing, 19(1): 635-646. https://doi.org/10.1109/TDSC.2020.2987903
[57] Hou, H., Ning, J., Zhao, Y., Deng, R.H. (2022). Fine-grained and controllably editable data sharing with accountability in cloud storage. IEEE Transactions on Dependable and Secure Computing, 19(5): 3448-3463. https://doi.org/10.1109/TDSC.2021.3100401
[58] Nasiraee, H., Ashouri-Talouki, M. (2022). Privacy-preserving distributed data access control for CloudIoT. IEEE Transactions on Dependable and Secure Computing, 19(4): 2476-2487. https://doi.org/10.1109/TDSC.2021.3060337
[59] Susilo, W., Jiang, P., Lai, J., Guo, F., Yang, G., Deng, R.H. (2022). Sanitizable access control system for secure cloud storage against malicious data publishers. IEEE Transactions on Dependable and Secure Computing, 19(3): 2138-2148. https://doi.org/10.1109/TDSC.2021.3058132
[60] Butt, A.U.R., Mahmood, T., Saba, T., Bahaj, S.A.O., Alamri, F.S., Iqbal, M.W. (2023). An optimized role-based access control using trust mechanism in E-health cloud environment. IEEE Access, 11: 138813-138826. https://doi.org/10.1109/ACCESS.2023.3335984
[61] Chen, E., Zhu, Y., Liang, K., Yin, H. (2023). Secure remote cloud file sharing with attribute-based access control and performance optimization. IEEE Transactions on Cloud Computing, 11(1): 579-594. https://doi.org/10.1109/TCC.2021.3104323
[62] Paulraj, D., Neelakandan, S., Prakash, M., Baburaj, E. (2023). Admission control policy and key agreement based on anonymous identity in cloud computing. Journal of Cloud Computing, 12(1): 71. https://doi.org/10.1186/s13677-023-00446-2
[63] Bera, S., Prasad, S., Rao, Y.S., Das, A.K., Park, Y. (2023). Designing attribute-based verifiable data storage and retrieval scheme in cloud computing environment. Journal of Information Security and Applications, 75: 103482. https://doi.org/10.1016/j.jisa.2023.103482
[64] Miao, Y., Li, F., Jia, X.H., Wang, H.X., Liu, X.M., Choo, K.R. (2024). REKS: Role-based encrypted keyword search with enhanced access control for outsourced cloud data. IEEE Transactions on Dependable and Secure Computing, 21(4): 3247-3261. https://doi.org/10.1109/TDSC.2023.3324640
[65] Pavithra, M., Prakash, M., Vennila, V. (2024). BGNBA-OCO based privacy preserving attribute based access control with data duplication for secure storage in cloud. Journal of Cloud Computing, 13(1): 8. https://doi.org/10.1186/s13677-023-00544-1
[66] Li, J., Li, B. (2021). Demand-aware erasure coding for distributed storage systems. IEEE Transactions on Cloud Computing, 9(2): 532-545. https://doi.org/10.1109/TCC.2018.2885306
[67] Ren, J., Li, J., Li, T., Mutka, M.W. (2022). Feasible region of secure and distributed data storage in adversarial networks. IEEE Internet of Things Journal, 9(11): 8980-8988. https://doi.org/10.1109/JIOT.2021.3119031
[68] Li, J., Yan, H., Zhang, Y. (2022). Efficient identity-based provable multi-copy data possession in multi-cloud storage. IEEE Transactions on Cloud Computing, 10(1): 356-365. https://doi.org/10.1109/TCC.2019.2929045
[69] Kontodimas, K., Soumplis, P., Kretsis, A., Kokkinos, P., Fehér, M., Lucani, D.E., Varvarigos, E. (2023). Secure distributed storage orchestration on heterogeneous cloud-edge infrastructures. IEEE Transactions on Cloud Computing, 11(4): 3407-3425. https://doi.org/10.1109/TCC.2023.3287653
[70] Shithil, S.M., Adnan, M.A. (2023). A prediction based replica selection strategy for reducing tail latency in geo-distributed systems. IEEE Transactions on Cloud Computing, 11(3): 2954-2965. https://doi.org/10.1109/TCC.2023.3244203
[71] Du, Y., Xu, Z., Zhang, K., Liu, J., Stewart, C., Huang, J. (2023). Cost-effective strong consistency on scalable geo-diverse data replicas. IEEE Transactions on Cloud Computing, 11(2): 1764-1776. https://doi.org/10.1109/TCC.2022.3161297
[72] Aldailamy, A.Y., Muhammed, A., Hamid, N.A.W.A., Latip, R., Ismail, W. (2024). Efficient multi-cloud storage using online dynamic replication and placement algorithms for online social networks. IEEE Access, 12: 20409-20425. https://doi.org/10.1109/ACCESS.2024.3361748
[73] Han, P., Wang, C., Chen, J., Zhao, S. (2020). CloudDLP: Transparent and scalable data sanitization for browser-based cloud storage. IEEE Access, 8: 68449-68459. https://doi.org/10.1109/ACCESS.2020.2985870
[74] Tian, W., Li, R., Xu, Z., Xiao, W. (2020). Loco-Store: Locality-based oblivious data storage. IEEE Transactions on Dependable and Secure Computing, 17(5): 1085-1095. https://doi.org/10.1109/TDSC.2020.3009428
[75] Zhang, X., Zhang, Y., Yang, H., Zhang, L. (2023). File processing security detection in multi-cloud environments: A process mining approach. Journal of Cloud Computing, 12(1): 100. https://doi.org/10.1186/s13677-023-00474-y
[76] Khan, M., Haroon, M. (2023). Detecting network intrusion in cloud environment through ensemble learning and feature selection approach. SN Computer Science, 5(1): 84. https://doi.org/10.1007/s42979-023-02390-z
[77] Kumar, A., Khan, S.B., Pandey, S.K., Shankar, A., Maple, C., Mashat, A., Malibari, A.A. (2023). Development of a cloud-assisted classification technique for the preservation of secure data storage in smart cities. Journal of Cloud Computing, 12(1): 92.

https://doi.org/10.1186/s13677-023-00469-9
[78] Al-Ghuwairi, A.R., Sharrab, Y., Al-Fraihat, D., AlElaimat, M., Alsarhan, A., Algarni, A. (2023). Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. Journal of Cloud Computing, 12(1): 127. https://doi.org/10.1186/s13677-023-00491-x
[79] Wang, Z., Qin, J., Xiang, X., Tan, Y., Peng, J. (2023). A privacy-preserving cross-media retrieval on encrypted data in cloud computing. Journal of Information Security and Applications, 73: 103440. https://doi.org/10.1016/j.jisa.2023.103440
[80] Zhou, H., Niu, Z.H., Wang, G., Liu, X.G., Liu, D.S., Kang, B.N., Hu, Z., Zhang, Y. (2024). Proactive drive failure prediction for cloud storage system through semi-supervised learning. IEEE Transactions on Dependable and Secure Computing, 21(4): 1528-1543. https://doi.org/10.1109/TDSC.2023.3286093
[81] Sajid, M., Malik, K.R., Almogren, A., Malik, T.S., Khan, A.H., Tanveer, J., Ur Rehman, A. (2024). Enhancing intrusion detection: A hybrid machine and deep learning approach. Journal of Cloud Computing, 13(1): 123. https://doi.org/10.1186/s13677-024-00685-x
[82] Huang, P., Fan, K., Yang, H., Zhang, K., Li, H., Yang, Y. (2020). A collaborative auditing blockchain for trustworthy data integrity in cloud storage system. IEEE Access, 8: 94780-94794. https://doi.org/10.1109/ACCESS.2020.2993606
[83] Yang, X., Pei, X., Wang, M., Li, T., Wang, C. (2020). Multi-replica and multi-cloud data public audit scheme based on blockchain. IEEE Access, 8: 144809-144822. https://doi.org/10.1109/ACCESS.2020.3014510
[84] Miao, Y., Huang, Q., Xiao, M., Li, H. (2020). Decentralized and privacy-preserving public auditing for cloud storage based on blockchain. IEEE Access, 8: 139813-139826. https://doi.org/10.1109/ACCESS.2020.3013153
[85] Yang, C., Tan, L., Shi, N., Xu, B., Cao, Y., Yu, K. (2020). AuthPrivacyChain: A blockchain-based access control framework with privacy protection in cloud. IEEE Access, 8: 70604-70615. https://doi.org/10.1109/ACCESS.2020.2985762
[86] Sifah, E.B., Qi, X., Hu, X., Agyekum, K.O.O., Acheampong, K.N., Cobblah, C.N.A., Gao, J. (2021). Selective sharing of outsourced encrypted data in cloud environments. IEEE Internet of Things Journal, 8(18): 14141-14155. https://doi.org/10.1109/JIOT.2021.3068226
[87] Zhang, Y., Geng, H., Su, L., Lu, L. (2022). A blockchain-based efficient data integrity verification scheme in multi-cloud storage. IEEE Access, 10: 105920-105929. https://doi.org/10.1109/ACCESS.2022.3211391
[88] Shu, J., Zou, X., Jia, X., Zhang, W., Xie, R. (2022). Blockchain-based decentralized public auditing for cloud storage. IEEE Transactions on Cloud Computing, 10(4): 2366-2380. https://doi.org/10.1109/TCC.2021.3051622
[89] Tian, G., Zhang, X., Zhang, Q., Li, Y., Wang, H. (2022). Blockchain-based secure deduplication and shared auditing in decentralized storage. IEEE Transactions on Dependable and Secure Computing, 19(6): 3941-3954. https://doi.org/10.1109/TDSC.2021.3114160
[90] Du, Y., Duan, H., Zhou, A., Wang, C., Au, M.H., Wang, Q. (2022). Enabling secure and efficient decentralized storage auditing with blockchain. IEEE Transactions on Dependable and Secure Computing, 19(5): 3038-3054. https://doi.org/10.1109/TDSC.2021.3081826
[91] Du, J., Dong, G., Ning, J., Xu, Z., Yang, R. (2023). A blockchain-assisted certificateless public cloud data integrity auditing scheme. IEEE Access, 11: 123018-123029. https://doi.org/10.1109/ACCESS.2023.3329558
[92] Song, M., Hua, Z., Zheng, Y., Huang, H., Jia, X. (2023). Blockchain-based deduplication and integrity auditing over encrypted cloud storage. IEEE Transactions on Dependable and Secure Computing, 20(6): 4928-4945. https://doi.org/10.1109/TDSC.2023.3237221
[93] Huang, J., Yi, J. (2024). The key security management scheme of cloud storage based on blockchain and digital twins. Journal of Cloud Computing, 13(1): 15. https://doi.org/10.1186/s13677-023-00587-4
[94] Yang, Y., Chen, Y., Chen, F., Chen, J. (2022). Identity-based cloud storage auditing for data sharing with access control of sensitive information. IEEE Internet of Things Journal, 9(13): 10434-10445. https://doi.org/10.1109/JIOT.2021.3121678
[95] Franklin, I.B., Arokiadass Jerald, M.P., Bhuvaneswari, R. (2022). Machine learning-based trust management in cloud using blockchain technology. SN Computer Science, 3(6): 429. https://doi.org/10.1007/s42979-022-01337-0
[96] Sultan, N.H., Laurent, M., Varadharajan, V. (2023). Securing organization's data: A role-based authorized keyword search scheme with efficient decryption. IEEE Transactions on Cloud Computing, 11(1): 25-43. https://doi.org/10.1109/TCC.2021.3071304
[97] Geetha, J.S. (2023). Adaptive artificial bee colony algorithm-based enhancement of data security in cloud computing. SN Computer Science, 5(1): 98. https://doi.org/10.1007/s42979-023-02419-3
[98] Yan, L., Ge, L., Wang, Z., Zhang, G., Xu, J., Hu, Z. (2023). Access control scheme based on blockchain and attribute-based searchable encryption in cloud environment. Journal of Cloud Computing, 12(1): 61. https://doi.org/10.1186/s13677-023-00444-4
[99] Yi, H. (2023). Improving cloud storage and privacy security for digital twin based medical records. Journal of Cloud Computing, 12(1): 151. https://doi.org/10.1186/s13677-023-00523-6