IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

MULTI-LAYERED CLOUD DATA LEAKAGE DETECTION

Dhrithi Shetty[1], Nidhi Shetty[2], Sneha Shinde[3], Prof. Monica Charate[4]

[1234]Department of Computer Science and Technology, SNDT Women's University, Mumbai, India

Abstract: Cloud computing has revolutionized the way data is stored and accessed, offering scalable, flexible, and cost-efficient solutions. However, it also brings with it significant data security concerns, especially data leakage, which exposes sensitive information to unauthorized entities. This paper proposes a robust, multi-layered framework for detecting data leakage in cloud environments, integrating AES-256 encryption, watermarking for traceability, and fake object insertion for intrusion detection. Each layer offers a distinct form of security—confidentiality, accountability, and threat identification. The proposed system is validated against existing techniques and proves superior in terms of real-time breach detection and insider threat prevention.

I. INTRODUCTION

Cloud computing has transformed data storage and access, providing affordable and flexible solutions that effortlessly scale to meet any need. It removes all physical storage infrastructure, making online data sharing and collaboration simple. Cloud computing has attractive features but also important risks. Sensitive data stored in the cloud is at risk from forbidden access, accidental breaches, and malicious attacks, thereby compromising privacy and security [3].

Older security methods are frequently insufficient to address insider threats, along with locating the source of data breaches, even with improved encryption. Encryption protects all data from forbidden access; however, it does not monitor every instance of access or sharing without permission. This important limitation poses a challenge. It is particularly problematic in multi-user settings with frequent data sharing among multiple stakeholders.

Furthermore, advanced cyber-attacks are increasingly targeting cloud systems. Attackers access sensitive data by exploiting vulnerabilities in shared resources and they also gain access by employing social engineering tactics or bypassing weak authentication measures [4]. Such breaches can severely affect organizations, resulting in financial losses, reputational harm, and legal issues from data protection regulation non-compliance.

A thorough, multi-faceted security approach is necessary to address these challenges [1]. This paper proposes a novel system for detecting data leakage in cloud environments. This system uses three key components: AES-256 encryption, watermarking techniques to embed several traceable identifiers and the insertion of many fake objects to detect forbidden access attempts. By integrating these techniques, the proposed system offers considerably improved protection, strong accountability, and highly effective real-time threat detection for modern cloud security challenges.

II. LITERATURE REVIEW

Data security in cloud computing has been studied in detail, and various approaches have been suggested for secure storage, access control, and leakage detection. However, existing works are incomplete in practical implementation, real-time monitoring, adaptability, and flexibility to new threats. These gaps draw attention to a need for adaptive, real-time, and feasible security solutions in cloud computing.

Table 1: Literature Review

No	Paper Name	Author(s)	Remarks
1	Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, and Future Directions (2022).	Ishu Gupta Ashutosh Kumar Singh Chung-Nan Lee Rajkumar Buyya	A systematic review of secure data storage and sharing distinguishes security models into cryptographic techniques, differential privacy, probability and watermarking, but does not consider real-world scenarios and active breach detection.
2	Avoiding the data leakage and providing privacy to data in Networking (2016)	Madhavi Suryawanshi, Prof . Sarita Patil	Similarly, research on data privacy in networking examines encryption and access control but provides generic solutions that have not been tested in practice or integrated with new technologies.
3	A Survey on Data Breach Challenges in Cloud Computing Security: Issues and Threats(2017)	R.Barona, E.A.Mary Anita	Further works on cloud security threats categorize breaches into integrity, confidentiality, and compliance risks but without proposing specific mitigation strategies.
4	An Analysis of Data Leakage and Prevention Techniques in Cloud Environments (2015)	Govinda. K, Divya Joseph	The latest study on guilty agent detection is based on allocation strategy, but it is difficult to implement and causes computational overhead.
5	An Analysis of Data Leakage and Prevention Techniques in Cloud Environments (2015)	T. Brindha, R. S. Shaji	A related work on data leakage prevention reviews encryption and monitoring techniques but lacks in real-world implementation.
6	Data Leakage Detection (2011)	Panagiotis Papadimitriou, Hector Garcia-Molina	Another approach based on smart data allocation improves leak detection but functions based on probabilistic models, which may result in potential false alarms.

III. PROPOSED SYSTEM

The proposed system improves cloud security by adding multiple protection mechanisms, guaranteeing data confidentiality, traceability, and intrusion detection. It employs AES-256 encryption to secure stored files, making decryption nearly impossible without proper authorization. This encryption method protects against brute-force attacks and unauthorized modifications, ensuring that sensitive information remains intact. Additionally, the system incorporates watermarking, embedding unique identifiers within files to track user access and detect unauthorized sharing. These watermarks act as digital fingerprints, allowing administrators to trace data leaks back to their source, thereby discouraging malicious activities.

To further strengthen security, the system utilizes fake object insertion, where decoy files resembling real data are strategically placed within the storage environment [6]. These fake files serve as traps for unauthorized users, and any interaction with them triggers alerts and logs the activity for administrative review. This proactive detection mechanism helps identify potential threats and improves security protocols based on observed attack patterns. Furthermore, an OTP-based authentication process is implemented for both user registration and file downloads, ensuring that only verified users can create accounts and access stored files. This additional layer of protection mitigates risks associated with compromised credentials and unauthorized access.

Another key feature of the system is its strong password enforcement, which guides users to create complex and secure passwords. By integrating encryption, watermarking, fake object insertion, OTP-based authentication, and secure password enforcement, the proposed system establishes a multi-layered defense strategy for cloud storage. These combined measures not only enhance data security but also promote accountability, making the system a comprehensive solution for preventing data leakage and unauthorized access.

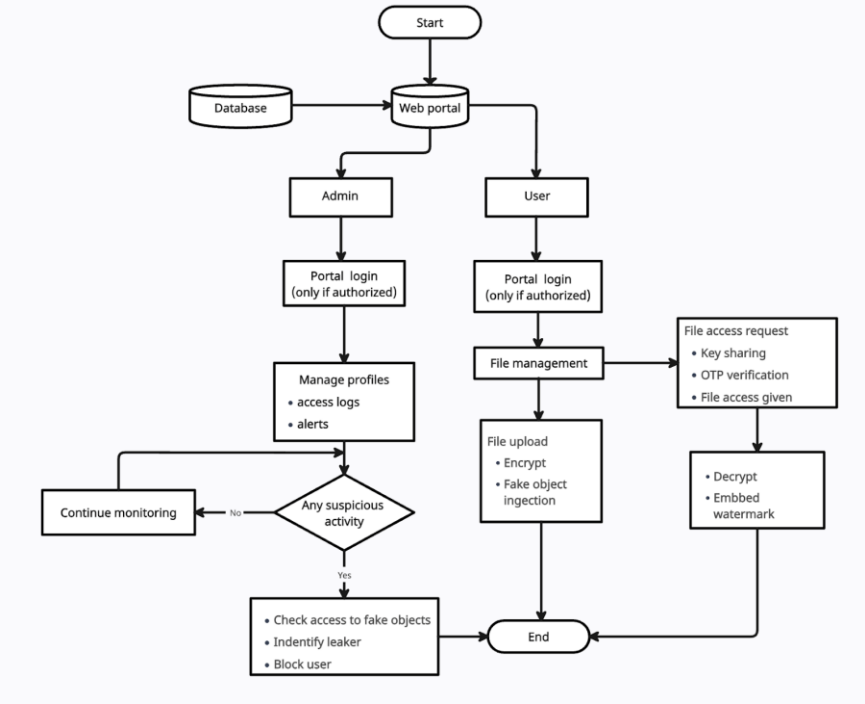

Fig 1: Workflow of the Proposed System.

IV. System Architecture

The architecture of the proposed system integrates AES-256 encryption, watermarking, and fake object insertion into a cohesive framework, providing a multi-layered defense mechanism against data leaks in cloud environments. This comprehensive approach ensures that data remains secure while enabling traceability and intrusion detection, thereby reinforcing the overall security posture of the cloud storage system.

The process begins with user authentication, where the system verifies the identity of the user before granting access to the portal. Once authenticated, users can upload, download, or access files stored in the cloud. This step ensures that only legitimate users can interact with the system, reducing the risk of unauthorized access.

During file upload, the system applies AES-256 encryption to secure the data. AES-256 is a symmetric encryption algorithm that encrypts data using a 256-bit key, which is highly resistant to brute-force attacks due to its large key size and complex encryption process. The encrypted data is then stored securely in the cloud, ensuring that even if the storage is compromised, the data remains inaccessible without the correct decryption key. When a user attempts to access or download a file, the system decrypts the data using the same key, ensuring authorized users can view the content.

Upon successful encryption, the system embeds a unique watermark into the file. This watermark contains metadata such as the uploader's ID, timestamp, and access permissions. If a data breach occurs, the embedded

watermark helps trace the origin of the leak, allowing the system to identify the user or source responsible for the unauthorized activity. Watermarking ensures that any leaked data can be traced back to the source, strengthening accountability and data traceability.

To enhance intrusion detection, the system generates and inserts fake objects into the storage. These fake objects mimic legit files but contain no sensitive information. If an unauthorized user attempts to access or manipulate a fake object, the system flags the activity as suspicious. This serves as an early indicator of a potential breach, triggering further analysis and security measures. Fake object insertion creates an additional layer of defense by luring malicious actors into revealing their presence.

All user interactions, including file uploads, downloads, and access attempts, are logged in a secure database. When suspicious activities are detected—such as multiple failed login attempts, unusual access patterns, or tampering with fake objects—the system generates an alert for the administrator. This enables immediate action to prevent data loss or compromise.

Finally, the system maintains detailed logs of file access and user activity, which aids in post-incident analysis and recovery. Watermarked files ensure that even if data is leaked, the source can be traced, facilitating accountability and reinforcing security protocols. The combination of AES-256 encryption for data confidentiality, watermarking for traceability, and fake object insertion for intrusion detection creates a robust, multi-layered defense strategy that effectively secures cloud-based data and minimizes the risk of unauthorized access or leakage.

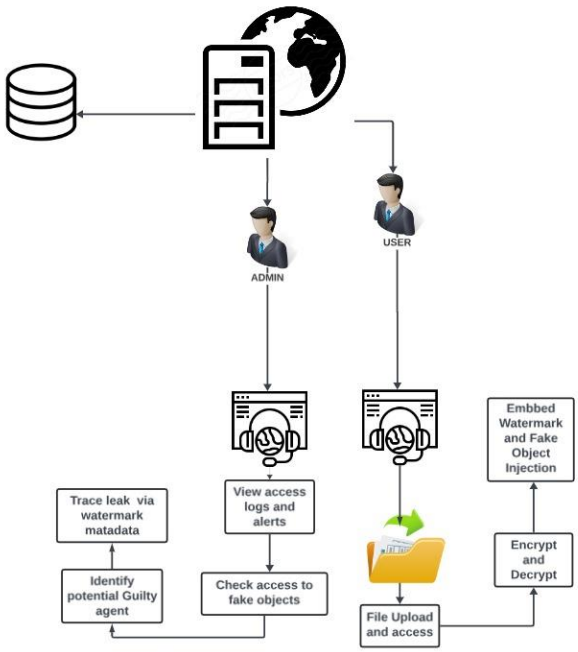

Fig 2: System Architecture of the Proposed System.

V. ALGORITHM

3.1 AES -256 ENCRYPTION

1. Input: Plaintext file and 256-bit encryption key.
2. Initial Round: XOR plaintext with the round key.
3. Main Rounds: Execute 14 rounds, each involving SubBytes, ShiftRows, MixColumns, and AddRoundKey operations.
4. Final Round: Perform SubBytes, ShiftRows, and AddRoundKey operations without MixColumns.
5. Output: Encrypted ciphertext.

3.2 WATERMARKING

1. Convert the unique identifier (UID) to binary format.
2. Embed the binary UID into the file's metadata without altering its primary content.
3. Extract and decode the embedded UID during an investigation to trace the source of unauthorized access.

3.3 FAKE OBJECT INSERTION

1. Create realistic decoy files that mimic genuine data.
2. Monitor access patterns and interactions with these decoys.
3. Generate alerts for any suspicious behavior involving fake objects.

IV. RESULT AND ANALYSIS

In our analysis, we studied data from over 30,000 leaked records belonging to various organizations across different sectors. The bar graph below represents the frequency of data leaks over four-year intervals, highlighting trends in organizational data leakages. These insights provide a clearer understanding of how data leaks have increased over time and emphasize the need for stronger security measures to protect sensitive organizational data.

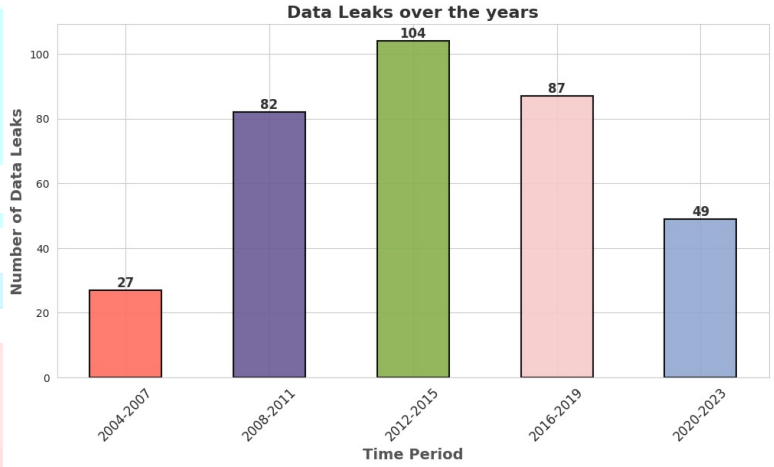

Fig 3: Frequency of Data Leaks in different Organizations across year.

V. CONCLUSION

The implemented system successfully enhances cloud data security by integrating AES-256 encryption, watermarking, and fake object insertion, offering a multi-layered defense against unauthorized access and data leakage. Adding OTP-based authentication during user registration and file downloads further strengthens access control, ensuring that only legitimate users can retrieve sensitive data. The inclusion of a dedicated fake files log enables better tracking of unauthorized access attempts, reinforcing system accountability. This security framework presents a reliable and effective approach to protecting cloud storage. Future work can focus on optimizing performance, improving scalability, and expanding compatibility with diverse cloud environments.

REFERENCES

[1] Gupta, I., Singh A.K., Lee, C.N., Buyya, R., "Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments," IEEE, 2022.

[2] Suryawanshi, M., Patil, S., "Avoiding the Data Leakage and Providing Privacy to Data in Networking," IEEE, 2016.

[3] Barona, R., Anita, E.A.M., "A Survey on Data Breach Challenges in Cloud Computing Security," IEEE, 2017.

[4] Govinda, K., Joseph, D., "Dynamic Data Leakage using Guilty Agent Detection over Cloud," IEEE, 2017.

[5] Brindha, T., Shahaji, R. S.,"An Analysis of Data Leakage and Prevention Techniques in Cloud Environment," IEEE, 2015.

[6] Papadimitriou, P., Garcia-Molina, H., "Data Leakage Detection," IEEE, 2011.

[7] Kumar, N., Katta, V., Mishra, H., Garg, H., "Detection of Data Leakage in Cloud Computing Environment", IEEE, 2014.

[8] Zhang, L., Cui, Y., Mu, Y., " Improving Security and Privacy Attribute Based Data Sharing in Cloud Computing", IEEE, 2019.

[9] D. Chandramohan, T. Vengattaraman, D. Rajaguru, R. Baskaran, and P. Dhavachelvan,"A Privacy Breach Preventing and Mitigation Methodology for Cloud Service Data Storage", IEEE, 2013.

[10] S. Haddad, G. Coatrieux, A. Moreau-Gaudry, and M.Cozic,"Joint watermarking-encryption-JPEG-LS for medical image reliability control in encrypted and compressed domains," IEEE, 2020.

[11] Holambe, S., Shinde, U., Bhosale, A., "Data Leakage Detection Using Cloud Computing", ISSN, 2229-5518.

[12]Jonnadula, V., Kommana, S., Panidapu, N., Rokkam, Y.,"Data Leakage Detection in Cloud Computing Environment", IJIRSET, 2024.

[13] Zuo, C., Lin, Z., Zhang, Y., "Why Does Your Data Leak? Uncovering the Data Leakage in Cloud from Mobile Apps", IEEE, 2019.