ARTIFICIAL INTELLIGENCE IN CLOUD DATA INTEGRITY ASSURANCE

Dhanya R[1], S.Mythili[2]*

Abstract

Cloud computing/cloud storage has emerged as a widely embraced solution to address the escalating storage expenses faced by IT enterprises. The exorbitant costs associated with data storage devices and the rapid generation of information make it financially burdensome for companies or individual users to update their hardware frequently. In addition to cost reduction, outsourcing data to the cloud contributes to decreased maintenance efforts. Cloud storage enables users to transfer data to expansive data centers situated remotely, over which they have no direct authority or control. While cloud-based data storage is an efficient solution but introduces new security challenges requiring resolution. The main goal is to develop a procedure that guarantees the integrity of data stored in the cloud, enabling users to confirm that their data has not been tampered with.This proof of integrity is adaptable to both the cloud provider and the customer. Despite the cost-effectiveness of data storage, security remains a critical concern. Although various encryption algorithms can enhance security, users cannot guarantee the absolute security of their data. AI, a field that merges computer science with reliable datasets, facilitates efficient problem-solving. Within AI, machine learning and deep learning sub-fields generate specific algorithms. Machine learning and deep learning, being integral to AI, contribute to the development of expert systems that make predictions based on input data. The storage and retrieval of data in the cloud rely on encryption/decryption keys. Instead of generating keys through mathematical functions, biometrics or digital signatures can serve as the key, allowing for encryption/decryption processes.

Keywords: Data Integrity, Cloud storage, Artificial Intelligence.

I. INTRODUCTION

Storing and retrieving large volumes of data through cloud storage is an efficient method employed by both individuals and enterprises. An essential aspect of cloud storage is ensuring the integrity of the data for all users, a critical feature that the cloud service must provide. The challenge is to implement a method or algorithm that provides evidence of data integrity in the cloud. This issue involves examining and obtaining proof that user data stored in a remote cloud storage location remains unchanged during various processes such as uploading, downloading, or archival, thus ensuring the integrity of the data.

This type of integrity assurance, where no modifications can occur without the owner's consent, proves highly beneficial in various scenarios such as storage systems that operate on a peer-to-peer basis, network file systems, long-term archives, web-service object stores, and database systems.Authentication systems of this nature serve as a deterrent against unauthorized alterations or deletions of data stored in the cloud, employing regular authentication checks on the storage archives. This authentication mechanism the data owner can efficiently, quickly, and securely confirm that the integrity of the data is being maintained in the cloud storage preventing any unauthorized actions. The primary goal is to provide users with the utmost assurance of integrity for the data they deem valuable.

The primary concept involves delegating the responsibility of ensuring data security to AI technology. Rather than manually implementing encryption algorithms with specific private keys, the cloud can employ biometric authentication or digital signatures that have already been verified and assigned to a biometric reader. Upon registering with the cloud, users are required to provide biometric information, which is then linked to an AI tool. Subsequently, when a user intends to upload data to the cloud, they must provide the associated biometric information, which is compared to the pre-stored biometric data. If a match is

[1,2]Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
* Corresponding Author

confirmed, the user is permitted to upload data, and the information undergoes encryption using an encryption algorithm with the biometric data serving as the key.

II. RELATED WORK

Lately, there has been an increasing interest in validating.remotely stored data. Pioneering data ownership verification (PDP) methods like Atenize, which validates file ownership, have explored the concept of publicly auditing untrusted repositories. These methods utilize homomorphic tags based on RSA for outsourcing audits, enabling public verification of the system. Atenize et al., however, did not address the dynamic data storage scenario. Directly extending the approach from static to dynamic storage could result in design and security challenges. In response, Atenese et al. introduced a dynamic version of the PDP method in their subsequent work. Nevertheless, this system imposes limitations on the number of pre-defined requests and supports only basic block operations with very limited functionality, excluding support for block insertions.

My tec Technologies has pioneered innovative biometric key protection technology. Unlike traditional two-stage processes involving user authentication and key release, Mytec's solution employs biometrics directly as the key for cryptography. After saving the key in cloud storage, it is associated with biometrics at a basic level and retrieved using biometrics for subsequent verification. Notably, the key remains independent of biometric data. This approach nsuring that the biometric application remains uncompromised even when the key is tampered with, and allowing for the easy replacement or update of the key later. Known as Biometric Encryption, the process developed by Mytec Technologies serves as a key for encryption/decryption. During user registration, the Biometric Encryption process, a secure data piece called a Bioscrypt is generated by combining the biometric image with a key generated using biometrics., enhancing file transfer security in the cloud. The Bioscrypt, acting as a cryptographic key, remains secure as it cannot independently reveal a fingerprint or key. In the verification process, the biometric-based security algorithm combines the biometric representation with the BIOS password to derive an encryption key. So rather than simply providing a key that will be forwarded to the user as authentication, biometric encryption requires combining a biometric image with Bioscrypt to extract a key that can be reproduced.

III. METHODOLOGY

Commonly utilized cloud security algorithms include RSA, AES, DES, among others. These algorithms employ private/public keys for the encryption/decryption process. DSA, alongside RSA, is widely recognized as one of the most utilized digital signature algorithms. DSA operates differently from RSA; it doesn't directly encrypt messages with private keys or decrypt messages with public keys. Instead, it utilizes a specialized algorithm To produce a digital signature, two 160-bit numbers are generated from the message digest and the private key. DSA utilizes public keys in its process for signature authentication, the authentication process is more intricate compared to RSA.

The proposed approach suggests employing biometrics for data encryption instead of using a specific key when a user intends to store or retrieve data. To access cloud storage, clients must be authorized users, and the client can log in by providing either biometrics or a digital signature for authentication. When uploading a file, AI technology will utilize either face recognition or biometric concepts as an encryption key. Similarly, when downloading the file, the same biometric information or facial recognition can serve as the key. Cloud storage's role is to convert the biometric information into a key format and provide it to the algorithm. While any of the previously mentioned algorithms can be applied, additional statements are needed for converting the biometric information into a key. The innovative idea here is to use biometric information converted into a key, feeding it into the algorithm instead of the conventional approach of providing a key as a public/private key pair.

Fig a. Model of Data Integrity Schemes

IV. BIOMETRIC ENCRYPTION ALGORITHM

The goal of the biometric-based encryption algorithm is to define a method for linking and later retrieving a digital-based key on a biometric characteristic, such as a biometric. This key functions as the encryption key.

Biometric encryption algorithms analyze the entire fingerprint image using a correlation mechanism. In this context, the correlation technique assesses the proximity of biometric details entered during registration and data transfer to the cloud.

In image processing terminology, correlation is applied to determine the response of the image mask, where the biometric information is treated as an image processed by an AI-driven application. The mask is applied to the matrix from left to right. The input images are represented as a two-dimensional array denoted by f(y), and theirtransform is denoted as F(v). Here, y denotes the spatial domain. The capital F indicates an array of transform domains.The correlation function c(y) between the extended input f(y) and the initial input f0(y), obtained during the proof, is defined as the integral from minus infinity to infinity of f1(v) times the complex conjugate of f0(y-v) dv, where * denotes the complex conjugate. In practical correlation systems, the system output is computed as the inverse Fourier transform (FT-1) of the product of F1(u) and the complex conjugate of , expressed as the integral over x of the inverse Fourier transform of {F1(u) * F0(u)} du, denoted as F, with F0*(u) typically represented by the filter function H(u) derived from f0(y). In correlation systems, the filter function H(u) is designed to generate a correlation peak, typically resembling a delta function, at the system output, a characteristic often employed for tracking objects of interest.The correlation function c(x) between the extended input f1(x) and the initial input f0(x), obtained during proof, is defined as the integral from minus infinity to infinity of f1(v) times the complex conjugate of f0(x-v) dv, where * denotes the complex conjugate. In correlation systems, the filter function H(u) is designed to generate a correlation peak, typically resembling a delta function, at the system output, a characteristic often employed for tracking objects of interest.

V. EXPERIMENTAL RESULTS

Fig (a) The cloud homepage serves as a representation of the authorized individuals who have access to the information stored in the cloud. The data maintained in the cloud remains secure and unaltered, as only authorized persons can view the information. Users are required to log in using the biometric information provided during the registration process before being directed to their home pages.

On the Key Request page, each time an authorized user logs in to access their cloud data using biometric information, a key or one-time password (OTP) is sent to their email to unlock their homepage. The owner of the cloud storage undergoes biometric or facial authentication whenever attempting to log in, ensuring an additional layer of security. This approach mandates authentication using biometric data each time, preventing key generation without the corresponding biometric data.

When a user wishes to upload data, their biometric information is utilized to encrypt the uploaded data. Similarly, for file downloads, the user must provide biometric data, which serves as the key for decryption. This process ensures integrity assurance each time a file is either uploaded or downloaded, contributing to the overall security of the system.

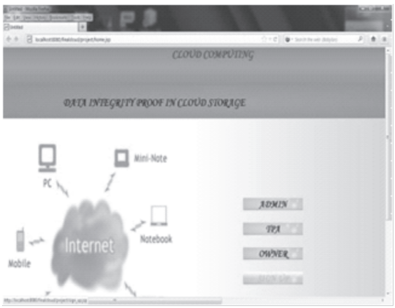
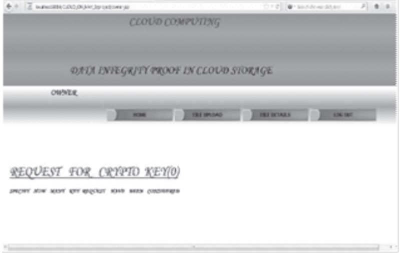


Fig (a) Cloud Homepage



REQUEST FOR CRYPTO KEY(0)

Fig (b) Authorised user Key Request page

VI. CONCLUSION

The implementation of "AI IN CLOUD DATA ASSURANCE" aids clients in ensuring the integrity of data stored on cloud storage servers with minimal amount and effort. This methodology was developed to alleviate the computation and storage burdens for both the customer and the cloud storage server. The client only needs a biometric detector, the key generator function , and the encryption/decryption function . Consequently, the client's data storage needs are kept to a minimum. As a result, the data storage requirements for the client are kept to a minimum making this scheme particularly advantageous for thin clients. As a result, the integration of AI into cloud storage facilitates users in verifying and maintaining data integrity.

REFERENCES

[1] CongWang; Chow,S.S.M. ; QianWang ; KuiRen ;WenjingLou "Privacy_preserving Public Auditing for Secure CloudStorage", IEEE Transactions on Computers Volume: 62, Issue: 2 2013 ,PP no : 362–375

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Storage Security in Cloud Computing," Proc. IEEE INFOCOM '10, Mar. 2010.

[3] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing," http://csrc.nist.gov/groups/SNS/cloudcomputing/index.html, June 2009.

[4] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the Clouds: A Berkeley View of Cloud Computing," Technical Report UCB-EECS- 2009-28, Univ. of California, Berkeley, Feb. 2009.

[5] R. Sravankumar and Saxena," Data integrity proofs in cloud storage" in IEEE 2011.

[6] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in SP '00: Proceedings of the 2000 IEEE Symposium on Security and Privacy. Washington, DC, USA: IEEE Computer Society, 2000, p.

[7] A. Juels and B.S. Kaliski, Jr., "Pors: proofs of retrievability for large files," in CCS"07: Proceedings of the 14th ACM conference on Computer and communications security.

[8] H. Shacham and B. Waters, "Compact Proofs of Retrievability," In Proceedings.ofAsiacrypt "08, Dec. 2008.

[9] Russell, S/ Norvig, P.," Artificial Intelligence: A Modern Approach",

[10] Patterson, Dan W."Introduction to Artificial Intelligence & Expert Systems "

[11] Omaima AL-Allaf,Abdelfatah A Tamini, "Face Recognition System Based on Different Artificial Neural Networks Models and Training Algorithms"

[12] Albert Bodo, "Method for producing a digital signature with aid of a biometric feature", German
patent DE 42 43 908 A1, (1994).

[13] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence", IEEE Trans. on Pattern Analysis and Machine Intelligence 15, 1148-1161, (1993)