Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 1 : 2023
ISSN : **1906-9685**

A NOVEL APPROACH FOR SECURE MEDICAL DATA SHARING IN CLOUD COMPUTING: PRIVACY PROTECTION MODEL AND SECURE SEARCH PROTOCOL

1.Uppala Vijay Kumar, 2.E.Madhusudhana Reddy
1.Research Scholar, Dept,. of Computer Science Engineering Career Point University, Kota.
dr.uppalavijaykumar@gmail.com
2. Professor of CSE & Principal Bhoj Reddy Engineering College for Women Hyderabad, Telangana, India.

Abstract

The medical field is producing and transmitting an increasing amount of multimedia data due to the rapid development of different multimedia technologies. Digital media data can also be widely distributed thanks to the internet. Editing, modifying, and duplicating digital material becomes considerably simpler. In addition, digital papers are vulnerable to numerous risks due to their ease of distribution and copying. maintaining Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM). In this suggested system, a novel secure search protocol is methodically constructed to allow cloud servers to execute secure search with no information of the real data of both trapdoors and keywords. We present a novel family of PP Functions and Additive Order to rank search outcomes while maintaining the secrecy of relevance scores among files and key phrases. We suggest a unique dynamic secret key generation mechanism and new data user authentication protocol to hinder the attackers from listening in on secret keys and posing as authorized data users submitting searches. Due to its ease of use, more and more people are utilizing cloud storage (CS) every day. There's a chance that cloud-stored data includes some confidential documents as well. For this reason, secure data retrieval and storage are required. There are numerous searchable cloud algorithms available. However, fewer of them offer adequate protection for data that is kept. When there are several data owners, a tree-based RMS technique might be utilized to increase confidentiality. The top search results are returned by TF-IDF method, which develops a multi-keyword search by taking into account a significant amount of cloud data. To locate the relevant file in the cloud, the cloud server additionally employs a depth first search technique.

Keyword: Ranking, Cloud Computing (CC), Secret Key generation, AES, Privacy Preserving (PP), Ranked Multi-keyword Search (RMS), inverse domain frequency (IDF), relevance scores (RSs)

INTRODUCTION

Using software and hardware resources that are provided as a service over a network, usually the Internet, is known as cloud computing. The term refers to how a cloud-shaped symbol is frequently used in system diagrams as abstraction for intricate infrastructure it comprises. Through cloud computing, user data, software, and computation are entrusted to remote services. Hardware and software resources that are controlled third-party services that are made available online comprise cloud computing. These services usually grant access to sophisticated software programs and top-tier server computer networks. The digital data is kept in logical pools in CS. The same data will have multiple owners in a multi-owner scenario.

To manage all of the data, there will be a core server. It is possible for the cloud to house numerous servers, each located in a different location. Data processing and protection will be under the purview of CS providers or primary servers. These CS providers offer storage space for purchase or lease to their customers. Network access to digital data may be spread and scaled to CS. Securing search over encrypted data is one of the issues with cloud storage. Encrypted cloud data (ECD) safe search is the hardest challenge in cloud storage. Different search techniques are available. Nevertheless, they either result in system overhead or make the methods extremely problematic to apply to many datasets. The data will be kept in the cloud in an encrypted format to avoid unauthenticated access.

A multi-keyword search system based on trees is developed to offer an effective search [1]. An index is formed by recognizing the words that appear to be keywords for given content. The resulting indexes are then combined into a single index. A depth first search is performed for each search request to find the user's relevant data file. To get the best results, the TF-IDF model is employed. An efficient search is carried out by using a depth first approach.

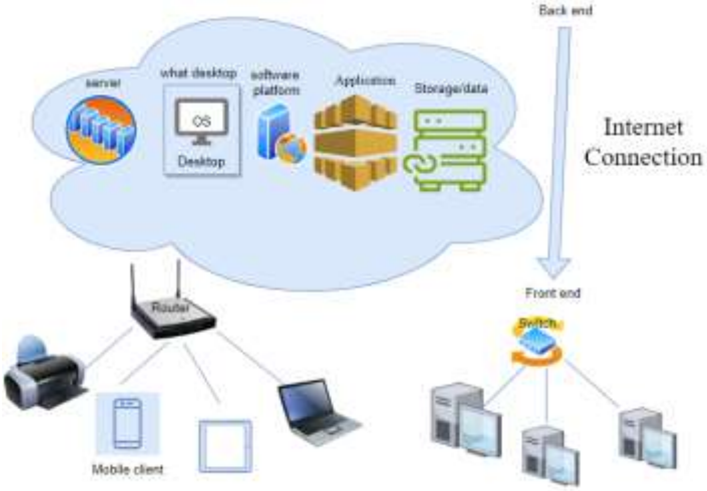

Figure.1.Cloud Computing

Numerous users share computing resources. Organizations can get benefits of cloud computing in addition to individual consumers. One of them is cloud-based data storage. The cost of maintaining and managing the data warehouse is eliminated by the virtualization of software and hardware resources in the cloud. Storage services are offered by numerous cloud computing systems, including Microsoft Azure, Drop Box, Amazon S3, Google Drive, and SkyDrive. As figure 1 illustrates, privacy and security issues are the main obstacles in cloud computing. Cloud providers have deployed security methods, such as firewalls, both software and hardware. Owing to the benefits of cloud services, a growing amount of private and sensitive data—including emails, government papers, private movies and images, medical records, and financial data for businesses—is being consolidated onto cloud servers.

The long-awaited goal of computing as a CC, utility, has potential to completely change the IT sector by changing how IT hardware is built and bought, as well as by making software even more alluring as a service. Innovative developers of new Internet services no longer need to invest significant sums of money in hardware or hire a staff of people to run their services. There are several ways to search data files in the cloud for scenarios involving multiple owners. Several of the topics covered below.

Practical Techniques for Searches on Encrypted Data

Authors in [2] suggested a way to search without jeopardizing the privacy of data. If a mobile user with a limited bandwidth wishes to recover documents from the mail storage server that include a specific term. The issue is that the content of the papers must be known by the server. Supporting search queries without disclosing all data is hence the challenge. The servers need to be reliable and not divulge information without the required authority. Applications run unwelcome security and privacy risks as a result of the unfrosted server. Instead of learning about the ciphertext, the untreated server shouldn't be informed about the plaintext. So that entrusted server cannot utilize controlled searching techniques to look up a word without the permission of the user. By allowing hidden queries, the user can instruct an untrusted server to look up a secret word without disclosing the term to the server. By enabling query isolation, the untrusted server only has knowledge about the plain text search result.

Initially, the issue of searching encrypted data is clarified. Let's say that user A has some documents that she keeps on server S, which is not trustworthy. A might be a mobile user, for instance, and keep her emails on an unreliable mail server. A wants to encrypt her documents and only keep the ciphertext on S because she doesn't trust S. One can separate each document into "words." Any token, including words and sentences, can be used for each "word." The user A wants to recover only the papers that comprise the word W from server S, even though they may only have a low-bandwidth network connection.

Secured RMS over ECD:

In CC, data owners are encouraged to outsource their intricate data management systems from their local locations to for-profit public clouds in order to benefit from increased flexibility and lower costs. Data must be encrypted before being stored to guarantee its security. It is significant to initiate a search using encrypted data as well.

Privacy Preserving Keyword Searches on Remote Encrypted Data:

Examine the following scenario: a user u wishes to save his files on a distant file server S in an encrypted format. Subsequently, user U wishes to quickly access a portion of encrypted files that comprise particular keywords while maintaining the confidentiality of keywords themselves and not jeopardizing the data security that are stored remotely.

Search on ECD:

The pay-per-use cloud model of today. The problem of efficient yet safe ranked keyword search over ECD has been defined and resolved in this study [4]. By providing matching files in ranked order based on certain related criteria (such as keyword frequency), ranked search significantly improves usability of the system and moves the implementation of PP cloud computing data hosting services closer to reality.

Secure Index for Resource-Constraint Mobile Devices in Cloud Computing

For RMS, the authors in [3] developed a secure index based on counting Bloom filters (CBF). More businesses and consumers are outsourcing their data to cloud servers these days. Sensitive data must be encrypted in order to safeguard data privacy, which adds significantly to the processing overhead and presents significant difficulties for devices with limited resources. This system uses a pruning technique to remove duplicate entries in order to save space, while numerous processes are developed to maintain and lookup CBF.

This article discusses the issue of safe ranked search over encrypted data on a cloud server. The suggested method creates a secure index for ranking multiple keyword searches by counting Bloom filters. Additionally, a pruning process is utilized to remove repeat items to save space, and other methods are devised to maintain and lookup CBF. The RSs are encrypted using Paillier cryptosystem. Even identical relevance scores are guaranteed to be encrypted into distinct bits that might aid in thwarting statistical analysis. The mobile devices may quickly search across encrypted data since the cloud server performs the majority of the computational work on encrypted RSs, which is the source constraint.

The RSs are encrypted using the Parlier cryptosystem. It will ensure that different bits of encryption are used for the same relevance scores. As a result, this can withstand statistical analysis using the relevance ratings' ciphertext. Additionally, Parlier cryptosystem allows for the homomorphism addition of encrypted text without requiring information of private key, which might transfer majority of the computing work associated with ranking from the user to cloud server. As a result, this approach can be used efficiently in mobile devices with limited resources, including 5G mobile terminals.

LITERATURE REVIEW

We build public-key systems that allow more general searches, like subset queries ($x \in S$), as well as comparison queries ($x \geq a$) on encrypted data. These systems can handle any kind of conjunctive query ($P1 \wedge \cdots \wedge P\ell$) without disclosing any information about specific conjuncts. Furthermore, we provide a generic framework for building and evaluating public-key systems that enable encrypted data queries [5].

In this work, we suggest a bed tree-based method that respects privacy and supports fuzzy multi-keyword features. With our system, incremental updates are simple to implement. We've put our solution into practice. According to the findings of our evaluation, our method is more economical in terms of building time and storage capacity. When it comes to multi-keyword queries, our search time typically outperforms the wildcard strategy, which returns a large number of encrypted files using single-word queries for approaches that do not enable multi-keyword queries [6].

We present the technique for fuzzy keyword search over ECD in this work. Fuzzy results are produced by using K-grams. We employ two different servers that are incompatible with one another for security reasons. The outcome of our experiment demonstrates the effectiveness and scalability of our system in handling a large number of encrypted data [7].

In order to avoid the necessity for real-time metadata generation for repaired data, this study offers an accurate repair method. In comparison to erasure code-based storage solutions, our developed service has a significantly lower computational cost during data retrieval, according to the performance analysis and testing results, but it has comparable storage and transmission costs. In comparison to distributed storage systems based on network coding, it results in lower storage costs, significantly faster data retrieval, and similar transmission costs [8].

In this article, we suggest a methodology for recognizing cloud computing's security and privacy issues. It identifies threats and attacks unique to clouds and provides a comprehensive explanation of how to mitigate and counteract them. Additionally, we provide a general cloud computing security architecture that helps meet cloud security and privacy needs while shielding the systems from a variety of threats [9].

Data is stored and retrieved via cloud computing. We need to be connected to the internet in order to access cloud data. This study presents the first data owner's analysis of commercially hosted public cloud data. Cloud data offers complete efficiency and privacy. This cloud management supplier offers ECD maintenance and multi-keyword ranked search. With the use of similarity measurement approaches, the MRSE keyword is used to access measurement of various individuals "coordinate matching" [10].

PROBLEM STATEMENT

Conventional authentication procedures typically consist of three stages. Initially, a secret key, let's say k0, is shared by the data authenticator and requester. Second, the requester delivers the encrypted data (d0)k0 to authenticator after encrypting his personally identifiable information (d0) using k0 [11]. Third, using k0 to decode the received data, the authenticator verifies the decrypted data.

This work proposes a safe, effective, and dynamic search method that allows dynamic document insertion and deletion in addition to precise multi-keyword ranked search.

In an effort to outperform linear search in efficiency, we suggest the 3DES method. To further cut down on time, the search procedure can also be done simultaneously. Using the secure method, the scheme's security is safeguarded against two threat models.

ALGORITHM

Data security and safety became more and more of a concern, necessitating data encryption. The data encryption techniques DES (Data Encryption Standard), 3DES (Triple DES), AES (Advanced Encryption Standard), and Blowfish (Best performance) are compared in this article [12].

DES, 3DES (Triple DES), AES, and Blowfish are among the several data encryption methods that are being compared.

DES

The first encryption method based on IBM's Lucifer algorithm was called DES. Being the original encryption standard, AES was extremely dangerous due to its numerous flaws and the discovery of multiple attacks.

3DES

An improvement on DES, which offered triple security compared to DES, is triple DES. To boost security, the same algorithm is used, but the encryption method is used three times [13].

AES

The National Institute of Standards and Technology (NIST) suggested the Advanced Encryption Standard as DES's replacement. The only known way to breach AES protection is by a brute force assault, which lets an attacker try different character combinations. Even with a supercomputer, Brute Force is a challenging task if number of combinations is unreasonably large.

Algorithm for Searchable Encryption [14]

An algorithm composed of randomized algorithms with polynomial time. They are as follows:

KeyGen(s): This security parameter is taken and utilized to create a public or private key pair. PEKS(Apub,w): This generates a searchable encryption by using A Pub, a public key, and w, a word. Trapdoor (Apriv, w): To create a trapdoor Tw, a private key called Apriv and a word called w are combined.

Ciphertext Security

This methodology is employed to ensure the security of the encrypted information. By simply rearranging the keywords and sending the resulting ciphertext for decoding, an attacker using cipher text might effortlessly breach semantic security. It is possible to crack this using a common method known as ciphertext security[15].

Private Key Searchable Encryption

Data encrypted with a private key is searched using a mechanism known as private key searchable encryption. Data is encrypted by the user in order to arrange it whatever they choose.

Public Key Searchable Encryption

Data might be encrypted and sent to the server utilizing the public key searchable encryption mechanism. The decryption key given by the owner could differ.

The framework does not display operations on data documents because the data owner might simply encrypt the data using conventional SK cryptography and then outsource the encryption process. Four algorithms make up the MRSE system, which focuses on the index and query:

a. Configuration (ℓ) the data owner outputs a symmetric key as SK after receiving a security parameter ℓ as input.
b. Construct Index (F, SK) The data owner creates a searchable index I using the dataset F. This index is then outsourced to the cloud server and encrypted using symmetric key SK. Following index construction, the document collecting can be contracted out and encrypted separately.
c. Trapdoor (fW): This algorithm creates a matching trapdoor TfW given t keywords of interest in fW as input.
d. Query (K, I, TfW) Upon receiving a query request with the format (TfW, k), the cloud server employs trapdoor TfW to do a ranked search on index I. Ultimately, it provides FfW, which is a ranked id list of the top k documents arranged as per how similar they are to f.W

Additional evidence from real-world data sets indicates that the suggested techniques undoubtedly result in minimal computational and communication overhead. In order to preserve security and search semantics in Figure 2 and the internal flow diagram in Figure 3, we suggest a hybrid technique in this study that combines AES, DES, and 3DES.

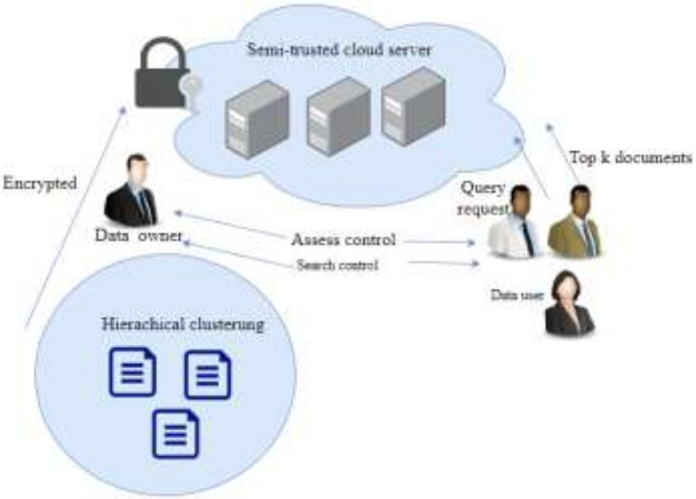

Figure.2: Architecture of multi-keyword synonym query over encrypted cloud data**.**

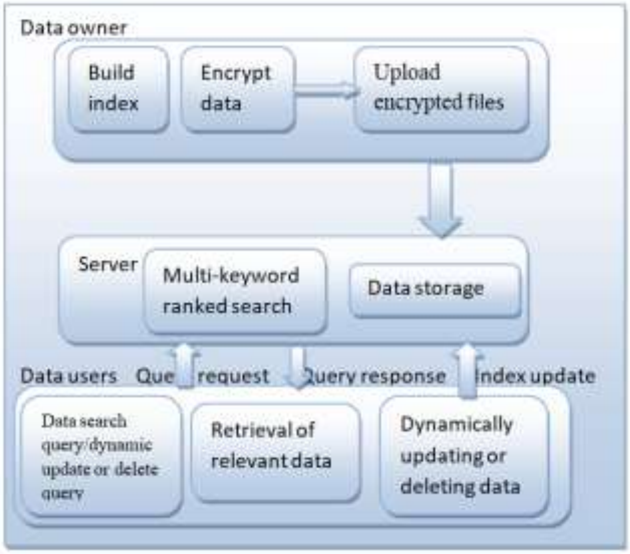
Data owner
Build index
Encrypt data
Upload encrypted files
Server
Multi-keyword ranked search
Data storage
Data users
Query request
Query response
Index update
Data search query/dynamic update or delete query
Retrieval of relevant data
Dynamically updating or deleting data

Figure.3. Internal Flow Diagram

Results & Analysis

a. Data User: This contains the login credentials for the user's registration.

b. Data Owner: This aids the owner in registering those particulars and includes login information.

c. File Upload: This allows file owners to upload an encrypted file that uses the AES and 3DES algorithms. By doing this, the data is guaranteed to be safe from unauthorized access.

d. Rank Search: This makes sure that users can look for files that are looked up a lot utilizing rank search.

e. File Download: Using his secret key, the user might download a file and use it to decrypt data that has been downloaded.

Examine both uploaded and downloaded files.

This makes it possible for the Owner to see the downloaded and uploaded files.

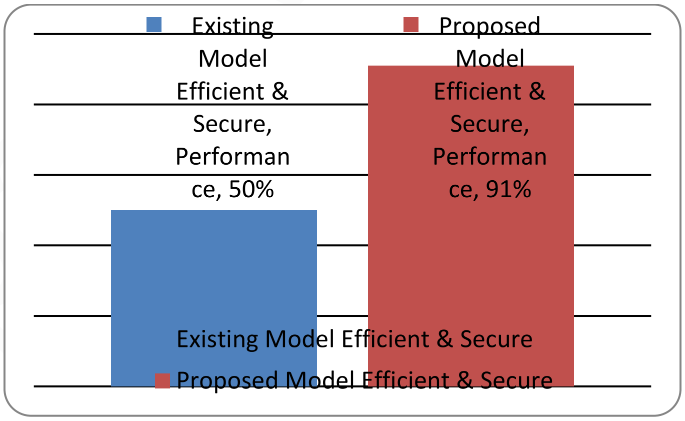

Figure.4. Analysis between Existing and present model

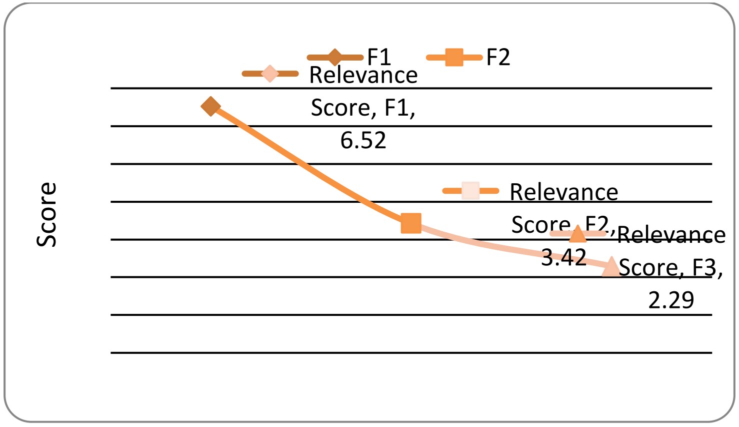

Figure.5.Graphicalrepresentationofrelevancescores

The collection of screenshots displayed in figure 4 can be used to describe the experimental outcomes.

Initially, a data owner uploads a file containing encrypted data using the AES algorithm [16].

By employing the SSE (secure symmetric encryption) technique, a secure search can be conducted on the uploaded file on the cloud server. The search results are presented using order-preserving symmetric encryption, or OPSE, in ranked form.

Using the formula TF×IDF, the algorithm determines the RS of files based on phrase frequency (TF) and IDF. where TF is defined as the quantity of a term or keyword that seems in the provided file. The IDF might be computed by separating total no. of files in the collection by total no. of files in Figure 5 that include that keyword. Files can therefore be ordered for greater symmetry based on relevance scores.

CONCLUSION

The need for quicker data retrieval from the cloud has grown in the contemporary context, as cloud infrastructure usage is increasing. Cloud providers must employ better algorithms that allow for quick retrieval without sacrificing user data security as more people save their information in the cloud. To create an index and perform searches within encrypted text, etc., numerous methods are employed. An effective RMS method over encrypted data is conducted, however, in a multiple data owner method that is taken into consideration for assessing data sharing in cloud computing. Every data file's index trees are combined into a single index tree. A DFS algorithm is used for the search. This secure search protocol enables various data owners to use diverse keys to encrypt their files and indexes. The cloud server may then combine encrypted indexes without having any knowledge of the individual data owners thanks to a tree-based index structure for every data owner. Compared to other current techniques, this tree-based search system is more effective at keyword mapping.

REFERENCES

1. M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, "Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions," J. Cryptol., vol. 21, no. 3, pp. 350–391, 2008.
2. P. Prasad, B. Ojha, R. R. Shahi, R. Lal, "3-dimensional security in cloud computing," in 3rd International Conference on Computer Research and Development (ICCRD, 2011), pp. 198-208, 2011.
3. L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proc. of ICICS, 2005.
4. C. Rong, S. T. Nguyen, M. G. Jaatun, "Beyond Lightning: A survey on security challenges in cloud computing," Computers and Electrical Engineering, vol. 39, no. 1, pp. 47-54, Jan. 2013.
5. D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. Of TCC,

2007, pp. 535–554.

6. M. Chuah and W. Hu, "Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data," in Distributed Computing Systems Workshops, 2011 31st International Conference, IEEE, (2011).
7. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, Mar. 2010.
8. N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.
9. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "A view of cloud computing," Communication of the ACM, Vol. 53, No. 4, pp. 50–58, 2010.
10. S. Deshpande, "Fuzzy keyword search over encrypted data in cloud computing," World Journal of Science and Technology, vol. 2, no. 10, (2013).
11. P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. of ACNS, 2004, pp. 31–45.
12. R. Brinkman, "Searching in encrypted data," PhD thesis, University of Twente, 2007.
13. Y. Hwang and P. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Pairing, 2007.
14. J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner products," in Proc. Of EUROCRYPT, 2008.
15. S. K. Sood, "A combined approach to ensure data security in cloud computing," Journal of Network and Computer Applications, vol. 35, no. 6, pp. 1831-1838, 2012.
16. Cong Wang, Qian Wang, and Kui Ren, Wenjing Lou, "Ensuring Data Storage Security in Cloud Computing," in IWQoS. IEEE International Workshop on Quality of Service, 2009. pp. 1-9.