ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 3, Issue 4, April 2018

An Economical multi - keyword seek for conserving energy over Mobile Cloud: TEES

[1] P. Sreevenkatramana, [2] Dr.AR.Arunachalam
[2] Professor&Head, Department Of Computer Science And Engineering, BIHER
[1] PG student, Department Of Computer Science And Engineering , BIHER

Abstract: ***Due to the increasing popularity of cloud computing, more and more sensitive or private information has been outsourced onto the cloud server. However, sensitive data should be encrypted before outsourcing for privacy requirements. The amount of protection needed to secure data in cloud is directly proportional to the value of the data stored. Security of the Cloud can be improved by trusted computing and cryptography. Data security is an important aspect of quality of service. Data encryption is a heavy overhead for the mobile devices, and data retrieval process incurs a complicated communication between the data user and cloud. This cause over consuming the mobile resource usage normally with limited bandwidth capacity and limited battery life, these issues introduce heavy overhead to computing and communication as well as a higher power consumption for mobile device users, which makes the encrypted search over mobile cloud very challenging. In this project, propose TEES (Traffic and Energy saving Encrypted Search), architecture for mobile cloud storage applications. TEES achieves the efficiencies through employing multi-keyword search scheme.***

Keywords: Cloud Computing,Conservation,TEES,Mobile Cloud.

I. INTRODUCTION

Cloud storage system is a service model in which data are maintained, managed and backuped remotely on the cloud side, and meanwhile data keeps available to the users over a network. Mobile Cloud Storage (MCS) denotes a family of increasingly popular on-line services, and even acts as the primary file storage for the mobile devices. MCS enables the mobile device users to store and retrieve files or data on the cloud through wireless communication, which improves the data availability and facilitates the file sharing process without draining the local mobile device resources.

Cloud Server:
The cloud server hosts third-party data storage and retrieve services.Since data may contain sensitive information, the cloud servers cannot be fully entrusted in protecting data. For this reason, outsourced files must be Encrypted. Any kind of information leakage that would affect data privacy are regarded as unacceptable.

Data Owner:
The data owner has a collection of n files C = {f1,f2,…fn} to outsource onto the cloud server in encrypted form and expects the cloud server to provide keyword retrieval service to data owner himself or other authorized users. To achieve this, the data owner needs to build a searchable index I from a collection of l keywords W = {w1,w2,…wl} extracted out of C, and then outsources both the encrypted index I' and encrypted files onto the cloud server. The data user is authorized to process multi keyword retrieval over the outsourced data. The computing power on user side is limited, which means that operations on user side should be simplified.

Data User:
The data user encrypts the query and sends it to the cloud server that returns the relevant files to the data user. Afterwards, the data user can decrypt and make use of the files.This Project introduces TEES (Traffic and Energy saving Encrypted Search) architecture for mobile cloud storage applications. TEES achieves the efficiencies through employing and modifying the ranked keyword search as the encrypted search platform basis, which has been widely employed in cloud storage systems. Traditionally, two categories of encrypted search methods exit that can enable the cloud server to perform the search over the encrypted data: ranked keyword search and Boolean keyword search. The ranked keyword search adopts the relevance scores to represent the relevance of a file to the searched keyword and sends the top-k relevant files to the client. It is more suitable for cloud storage than the boolean keyword search approaches, since boolean keyword search approaches need to send all the matching files to the clients, and therefore incur a larger amount of network traffic and a heavier post-processing overhead for the mobile devices.TEES offloads the security calculation to the cloud side to save the energy consumption of mobile

devices, and TEES also simplify the encrypted search procedure to reduce the traffic amount for retrieving data from encrypted cloud storage.

II. LITERATURE SURVEY

Wide range of studies has been done for efficient multi keyword search scheme over mobile cloud. One of the well-known algorithms for efficient multi keyword search is RSA, which is widely used cryptosystem in the world. In the paper, Dawn Xiao dong Song [1] described some cryptographic schemes for the problem of searching on encrypted data and provided proofs of security for the resulting crypto systems. These techniques have a number of crucial advantages. They are provably secure, they provide provable secrecy for encryption, meaning that the untrusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching, so that the untrusted server cannot search for an arbitrary word without the user's authorization.Dan Boneh [2] performed for the problem of searching on data that is encrypted using a public key system. Consider user Bob who sends email to user Alice encrypted under Alice's public key. An email gateway wants to test whether the email contains the keyword \urgent" so that it could route the email accordingly. Alice, on the other hand does not wish to give the gateway the ability to decrypt all her messages. This construct a mechanism that enables Alice to provide a key to the gateway that enables the gateway to test whether the word \urgent" is a keyword in the email without learning anything else about the email. It refers to this mechanism as Public Key Encryption with keyword Search. Using this mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some special keyword, but learn nothing else.
In the paper [3] Wang defined and solved the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy. Specifically, we explore the statistical measure approach. In the paper [4] proposed by Cong Wang, proposed a definition for ranked searchable symmetric encryption, and give an efficient design by properly utilizing the existing cryptographic primitive, order-preserving symmetric encryption (OPSE). Thorough analysis shows that the proposed solution enjoys "as-strong-as possible" security guarantee compared to previous SSE schemes, while correctly realizing the goal of ranked keyword search.In the paper [5], Shucheng Yu proposed a novel multi keyword fuzzy search scheme by exploiting the locality-sensitive hashing technique. Proposed scheme achieves fuzzy matching through algorithmic design rather than expanding the index file. It also eliminates the need of a predefined dictionary and effectively supports multiple keyword fuzzy searches without increasing the index or search complexity. Extensive analysis and experiments on real-world data show that the proposed scheme is secure, efficient and accurate. In the paper [6], Juan Ramos examined the results of applying Term Frequency Inverse Document Frequency (TF-IDF) to determine what words in a corpus of documents might be more favorable to use in a query. As the term implies, TF-IDF calculates the values for each word in a document through an inverse proportion of the frequency of the word in a particular document to the percentage of documents the word appears in. Words with high TF-IDF numbers imply a strong relationship with the document they appear in,suggesting that if that word were to appear in a query, the document could be of interest to the user.

III. OBJECTIVES OF THE RESEARCH

The objective is to develop a new architecture, TEES as an initial attempt to create a traffic and energy efficient encrypted multi keyword search tool over mobile cloud storages.Order Preserving Encryption algorithm leaks data privacy. So this project makes use of another powerful RSA algorithm which will not harm the efficiency.This project also introduce the concept of keyword buffer controller that allows for quick search of documents and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality.The mobile client has a heavy workload for decrypting the selected index, calculating and ranking the relevance scores. It will take more time when comes to the mobile client since the computing capacity of a mobile device is limited. This is also clearly inappropriate when the battery of a mobile device is taken into account. Second, the one round-trip for each file search and retrieval request, is not capable in fetching exact result when compared to two round – trips. This is a heavy burden for mobile devices with limited bandwidth and traffic fees.

IV. EXISTING TECHNIQUE

The framework of the existing method consists of three steps:
1) Construction of stem word.

2) Order Preserving Encryption (OPE) to encrypt the file index.
3) Compose Term Frequency Table

The existing system developed a new architecture, TEES (Traffic and Energy saving Encrypted Search) as an initial attempt to create a traffic and energy efficient encrypted keyword search tool over mobile cloud storages with single keyword search scheme.It is slightly more time and energy consuming than keyword search over plain-text, but at the same time it saves significant energy compared to traditional strategies.Searchable symmetric encryption (SSE) retrieves encrypted data over cloud. It allows a party to outsource the storage of its data to another party (a server) in a private manner, while maintaining the ability to selectively search over it. Searchable encryption focuses on single keyword search or Boolean keyword search, and rarely differentiates the search results. This implements server side ranking based on order-preserving encryption (OPE), OPE leaks data privacy.

An order-preserving symmetric encryption (or OPE) scheme is a deterministic symmetric encryption scheme whose encryption algorithm produces cipher texts that pre-serve numerical ordering of the plaintexts

V. PROPOSED TECHNIQUE

The proposed method can be broadly classified into three stages.

1. Process of Preprocessing and Indexing
2. Apply RSA algorithm (Ron Rivest, Adi Shamir and Leonard Adleman) encryption and MD5 (Message Digest) hash function.
3. Data Search and Retrieval after Authentication

In Stage I, The data owner first executes the preprocessing and Indexing work. He should invert files that are selected to store on the cloud, for text search engines. Every word in these files undergoes stemming to retain the word stem. After this step, the data owner encrypts and hashes every term (word stem) to fix its entry in the index. The index is then created by the data owner. Finally, the data owner encrypts the index and stores it into the cloud server, together with the encrypted file set. As energy consumption becoming important, a complicated algorithm is not suitable in mobile devices. Therefore we choose a simple RSA Algorithm in TEES. In Stage III, A data user can only access a file after being authenticated by the data owner. In the process of authentication, the data user sends his identity to the data owner. The data owner sends the encrypted keys back if the user is a legal user. In the process of search and retrieval, the cloud server helps the users to find the top-k relevant files for a given keyword without decrypting it.

VI. ALGOIRTHM SPECIFICATION

RSA Algorithm

RSA is a widely used cryptosystem in the world. It is a public key cryptosystem which uses two kinds of key, private key and public key. Every user has both of the keys, a private one and a public one. If user A wants to send a message to B, he need B's public key to encrypt the message. After encrypted, the message is received by B, then B uses his private key to decrypt the message.

RSA algorithm can be classified as three algorithms, the key generation algorithm, encryption algorithm, and decryption algorithm. RSA key generation algorithm can be described as follows,

- Step 1. Generate two large random and distinct primes P and Q
- Step 2. Calculate $N = P.Q$ and $\varphi = (P - 1)(Q - 1)$
- Step 3. Choose a random integer E, $1 < E < \varphi$, such that $\gcd(E,\varphi) = 1$ The rules to select E are:
    a. E is positive integer
    b. $0 < E < M$

c. GCD (M, E) = 1 ... (GCD = Greater Common Divisor) NOTE: It is recommended to use E = 65537 (17 bits).

- Step 4. Compute the unique integer D, $1 < D < \varphi$, such that $ED \equiv 1 \pmod{\varphi}$
- Step 5. Public key is (N, E) and private key is (N, D)

RSA encryption algorithm can be described as follows,

C = ME mod N,

RSA decryption algorithm can be described as follows, M = CD mod N,

Which C represents cipher text and M represents message.

A SIMPLE pseudo code for this algorithm is below:

E = exponential value,

M = message to encrypt,

N = modulo value

C = cipher text value

C = 1 ... set it as default value

While E

If E is odd

C = C * M

C = C % N

E = E / 2

M = M * M

M = M % N

While Loop

Message Digest Algorithm

MD5 Message Digest is a widely used hash technique, such that it will produce 128-bit hash value. We need to convert the input data into bytes in order to convert it to hash value. This is useful in many security applications and it ensures data integrity.

ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 3, Issue 4, April 2018

No one should be able to produce input for given pre-specified output. No one should be able to produce two different inputs for which the transformation function returns the same output. Message-Digest (Fingerprint) algorithms are special functions which transform input of (usually) arbitrary length into output (so-called "message digest") of constant length.

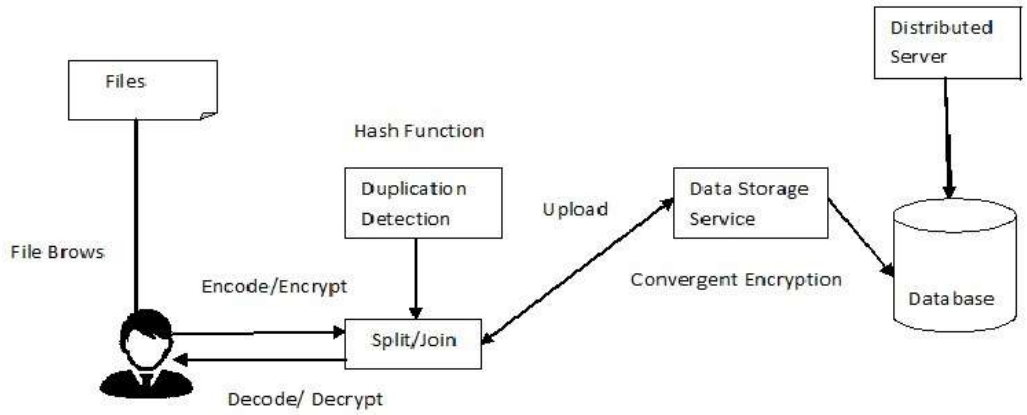
MD5 algorithm takes input message of arbitrary length and generates 128-bit long output hash.

MD5 hash algorithm consists of 5 steps.

- Step 1. Append Padding Bits
- Step 2. Append Length
- Step 3. Initialize MD Buffer
- Step 4. Process Message in 16-Word Blocks
- Step 5. Output

V. SECURE DEDUPLICATION

Deduplication systems with distributed cloud servers allow more error prevention. Additional security can be achieved by implementing a sharing strategy using secrecy.In more details, a file is first split and encoded into fragments by using the technique of secret sharing, instead of encryption mechanisms. Sharing distribution is done in multiple servers which are independent. Additional deduplication is done using systems that provide efficiency and reliability for block and file levels. Analysis of security shows us how secure these systems are in the defines way of the structure.We implement our deduplication systems using the Public secret sharing scheme that enables high reliability and confidentiality levels.This design fixes the issue of previous work that the computational load at user or auditor is too huge for tag generation. Finishing the fine grains, is done by Sec cloud designed by audit in sector and block levels. This also helps with security. The complication arises when preventing dictionary-attacks. This system proposes two things- secure auditing and deduplication of files.



VI. CONCLUSION

The way to solve this is by saving the space in storage and reducing the need of band-widths. Maintaining cloud data is eased by this. User retrieves data and files barring any loss of data. Segmenting-Binning modules add to the UI.The time taken for the user to interact with the cloud reduces considerably as bandwidth is a vital resource for the user. Testing was carried out thoroughly and the results suggest a considerable saving in the storage space and bandwidth requirements. Deduplication is implemented on the storage space of the cloud controller.

VII. CONCLUSION

This project, proposed a secured way of accessing files from cloud. This project proposed a secured and reliable scheme for data owner to provide better services to the users. The owner side encryption scheme and index file generation helps the data user to get secure and protected data with better QOS.

To improve the QOS a client side ranking process has been adopted. Searching the query in the index file rather than the file system cloud server can give quick response.

Proposed scheme fulfils the security requirements of multi keyword top-k retrieval over the encrypted cloud data.

This project devise a server-side ranking SSE scheme. It propose a two-round searchable encryption (TRSE) scheme employing the fully homomorphic encryption, which fulfills the security requirements of multi-keyword top- k retrieval over the encrypted cloud data. By security analysis, this shows that the proposed scheme guarantees data privacy. According to the efficiency evaluation of the proposed scheme over real dataset, extensive experimental results demonstrate that our scheme ensures practical efficiency.

REFERENCES

[1] TEES: An Efficient Search Scheme over Encrypted Data on Mobile Cloud, Volume: PP, Issue: 99. 2-2-2015.

[2] D. Huang, “Mobile cloud computing,” IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, 2011.

[3] D. Niyato and E. Hossain, “Integration of WiMAX and WiFi: Optimal pricing for bandwidth sharing,” IEEE Commun. Mag., vol. 45, no. 5, pp.140–146, May 2007.

[4]. C.-Y. Chang, T.-Y. Wu, C.-C. Huang, A. J.-W. Whang, and H.-C.Chao, “Robust header compression with load balance and dynamic bandwidth aggregation capabilities in WLAN,” J. Internet Technol., vol.8, no. 3, pp. 365–372, 2007.

[5] H. Son, S. Lee, S.-C. Kim, and Y.-S. Shin, “Soft load balancing over heterogeneous wireless networks,” IEEE Trans. Vehic. Technol., vol. 57, no. 4, pp. 2632–2638, Jul. 2008.

[6]. L. Zhou, H.-C. Chao, and A. V. Vasilakos, “Joint forensics-scheduling strategy for delay-sensitive multimedia applications over heterogeneous networks,” IEEE J. Selected Areas Commun., vol. 29, no. 7, pp. 1358–1367,Aug 2011