International Journal of Innovative Research in Computerand Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 7, July 2020||

Searchable Symmetric Encryption and Data Search in Cloud Services

Vaishali R. Kandekar, Dr. Sangve Sunil M.

PG Student, Department of Computer Engineering, Zeal College of Engineering and Research, Narhe, Pune, India

HOD, Department of Computer Engineering, Zeal College of Engineering and Research, Narhe, Pune, India

ABSTRACT: As cloud storage expands, more and more data owners will probably outsource their data to cloud providers. For privacy reasons, confidential data must be encrypted before subcontracting. This paper proposed a symmetric search encryption scheme (SSE) that is safe and easy to implement when looking for strings. The communication process and calculation time in n documents are included. The proposed system uses a hash string to generate indices instead of encryption so that it is appropriate for light applications. This framework presents a new concept of privacy for search patterns that provides a measure of security against harm. It is supported by a research model which defines skills. Integrity checks are carried out to inform the stored file to be tempered or not by checking the previous one.

KEYWORDS: Cloud storage, chain search, SSE, hash-chain, lightweight cryptography, Symmetric key, Searchable encryption, Security,

I.INTRODUCTION

Cloud storage is a computer data storage model that stores the digital data in logical pools. Physical storage involves multiple servers (sometimes in multiple locations), and the physical environment is typically owned and managed by the hosting company. These cloud storage providers are responsible for keeping the data available and accessible and protecting and running the physical environment. People and organizations purchase or lease storage capacity from providers to store user, organization or application data.A wide range of documents are stored on a cloud server, the search for a keyword leads to a great many documents that are not related to the subject.This motivates the idea of a chain search, which allows a more specific search.Chain searches are a search with multiple keywords that store the order of the keywords. Then their order and adjacencies must be considered during the search, as well as the presence of all such keywords in a document. A simple but powerful way to search the encrypted data while maintaining the privacy of the data is provided by the SSE. The reason for selecting a primitive private / symmetric key for our implementation lies primarily in the fact that it results in significantly lower overhead computing compared to its public / asymmetric key counterpart, making it more suitable for mobile devices.

MOTIVATION

In order to solve confidentiality problems, data encryption systems can be useful for users to be able to control the secrecy of their stored data. Users encrypt their data before cloud outsourcing. However, standard encryption schemes will reduce the ability of users to search for stored data, since after encryption, the user simply cannot use a plain-text keyword to perform a search and therefore cannot retrieve content in an efficient manner. In order to ensure a secure and efficient data retrieval, it is necessary to ensure that the user can search the encrypted data without disclosing the contents and the searched keyword on the server.

A. *OBJECTIVE*

The aim of this paper is to study the searchable encryption schemes in detail and to implement a solution that allows data storage and recovery in cloud computing to be protected from privacy.

II.REVIEW OF LITERATURE

1. In this paper, wedesign GSSE, a dynamically verifiable SSE scheme, which can be applied to any SSE schemes with a three-party model and does not require modifications on them. By building authenticators and a proof index,

GSSE provides efficient search result verification, while preventing data freshness attacks and data integrity attacks in SSE. The experimental results demonstrate that GSSE introduces acceptable overhead in verifying search results.[1]

2. In this paper, we focus onhow to improve the performance of the SSE scheme and reduce its leakages. Based on non-colluding clouds, we propose the M-SSE scheme, which achieves both forward and backward security. Apart from the good performance, M-SSE can protect the size pattern. Distributing the leakages to different clouds to reduce the information leakage may be a new idea to protect users' privacy. In future work, we plan to design more secure searchable encryption with better performance and the forward secure order-preserving encryption scheme.[2]

3. present a series of attacks that recover rawtext from database columns encrypted with DTE and OPE using only the encrypted column and the publicly available auxiliary information. The concrete security provided by these systems is studied. This system presents a series of attacks that recover plain text from columns of databases encrypted with DTE and OPE using only the encrypted column and the auxiliary information publicly available. This system takes into account known attacks, including frequency analysis and classification, as well as new attacks based on combinatorial optimization.[3].

4. Present the fundamental principles of modern cryptography, with emphasis on formal definitions, clear hypotheses and rigorous security tests. The book begins by focusing on private-key cryptography, including the exhaustive treatment of private-key cryptography, message authentication codes and hash functions. The authors also present design principles for widely used block ciphers and flow ciphers, including RC4, DES, and AES, in addition to providing interim constructions of flow codes and block ciphers of lower- level primitives. The second half of the book deals with public key cryptography, beginning with an autonomous introduction to the numerical theory needed to understand thecryptosystems RSA, Diffie-Hellman and El Gamal (and others), followed by an exhaustive treatment of several standardized cryptography public key and digital signature[4].

5. SSE scheme proposal to satisfy all the properties described above. Our construction expands the focus of the inverted Index in several non-trivial ways and introduces new techniques for designing SSE[5].

6. Transformation of an encryption scheme based on anonymous identity (IBE) into a secure PEKS scheme that, unlike the previous one,guarantees consistency. Finally, this system suggests three extensions of the basic concepts considered here, that is, cryptography based on anonymous hierarchical identity, public key cryptography with search of temporary keywords and identity-based cryptography with keywordresearch[6].

7. One of our constructions, called RSA-DOAEP, has the additional feature of preserving length, so it is the first example of a public key code. This system generalizes this to obtain a notion of cryptographic schemes that can efficiently search for a more flexible privacy to the exchanges in terms of searchthroughatechniquecalledgrouping[7].

8.This document presented the first SSE scheme that supports the search for an arbitrary string. The idea of OXT is to retrieve the documents corresponding to aselected token from a joint query and then make an intersection with the other tokens. This system has modified this idea to adapt a chain search configuration in the following way. The main idea of our scheme is to recover the documents that correspond to the first character of a query string and then verify if the subsequent characters of the retrieved documents are the same as the characters of the query string. To implement this, this system uses the fact that you can verify if a text contains a string as a sub string using a multipleset[8].

9. This system present two constructions that this system shows safe under our new definition. It is interesting to note that besides complying with greater security guarantees, our buildings are more efficient than all previous constructions [9].

10.propose a basic ideafor MRSE based on a secure internal calculation of the product, and then provide two MRSE schemes significantly improved to achieve various privacy requirements for the generation of strings in two different threat models. To improve the search experience of the data search service, this system further extends these two schemes to support more search semantics. An in-depth analysis is provided to investigate the privacy and efficiency guarantees of the proposed schemes. Solve the complex problem of multiple keyword classified search that preserves the privacy of encrypted data in cloud computing. Classify the document according to the corresponding result.[10].

III. SYSTEM ARCHITECTURE/ SYSTEM OVERVIEW

The proposed system will provide data security. Proposes a secure search protocol where the cloud service can perform secure search trapdoors.Provide an efficient and easy-to-use symmetric searchable encryption scheme (SSE) for string search. Four entities are involved in the proposed system computing model, such as data owners, data users, cloud servers and TPA.

Data owners have collection of files. The data proprietors upload the file to create the indexes. Data owners encrypt encrypted cloud storage files and outsource them.When a data customer wishes to search through cloud files, the search string is entered. The system gives files that are matched. Then customer sends decryption request and customer receives this via mail.Only file is downloaded to the customer if key and trapdoor matches. Then customers have to join trapdoor and customers can access and decrypt files.The third-party auditor checks for data integrity and informs the proprietor of this detail. The program can pick the hash value from a position and substitute it with a repique of the previously stored hash value.

Advantages

1. It does not just look for the keywords but also find the order in the way suggested string search.
2. Provide multi keyword searching in secure way.
3. The implemented scheme enables a user to store data securely in the cloud by encrypting it before outsourcing and also provides user capability to search over the encrypted data without revealing any information about the data or the query.

A. Algorithms

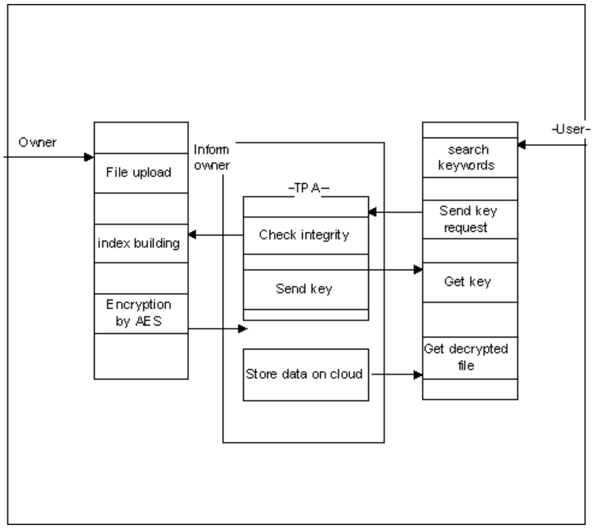

Fig. 1. Proposed System Architecture

Stage 2: Generate a SecretKey, utilizing generateKey() API technique for KeyGenerator

Step 3: Create a Mac object.

Step 4: Initialize the MAC with the above key, using init(Key key) API method of Mac.

Step 5: Create a new String message and get its byte array.

3) SHA-1(Secure Hash Algorithm1)

This algorithm is used for the development of a file hash and is used to test integrity to ensure that no file is tempered. SHA-1 is a cryptographic hash function that accepts an entry and creates a 160-piece esteem appreciation called a process of message which is normally prepared as a hexadecimal number. SHA, a family of encryption functions designed to protect data, is 40 digits. Secure Hashing Algorithms, also known as SHA. It works by transforming data using a Hash function, a bit-by-bit algorithm, modular add-ons and compression. The Hash function creates a string that doesn't match the original.

1) Advanced Encryption Standard (AES) Algorithm: It's a symmetrical algorithm for encryption.
It is used to convert plain text to cipher text.
Steps Encryption
1: 128-bit data block
2:key expansion
3: add round key
4: sub byte, shiftrow, mixcolumns, add round key 5: subbyte, shift rows, add round key
6:128 but encrypted block
Decryption
1:128 but encrypted block
2:key expansion
3:add round keys, shiftrows, sub bytes
4:add round keys, mix columns, shift rows, sub bytes
5:add round key
6: 128-bit block

2) HMAC-MD5:keyed-hashing algorithm which is a cryptographic technique for authenticating messages.
The sender and receiver share the symmetric key K for establishing the MAC process. Essentially, the MAC is an encrypted checksum generated on the underlying message that is sent along with a message to ensure the authentication of the message. Generate a Code for Message Authentication (MAC). This framework uses the Mac class that provides the usefulness of the "Message Authentication Code" (MAC) calculation. To put it plainly, the following steps are used to create a Message Authentication Code: compression functions. The hash function produces a string of fixed dimensions that does not resemble the original.

These algorithms function in one way, which means that it is virtually impossible to convert them back into the original data when they've transformed themselves into their hash values. Password encryption is a common SHA application, because the server-side must only trace the hash value of specific users instead of the actual password.

B. Mathematical Model

A block cipher is specified by an encryption function

$$E_K(P)=E\,(K,P):\{0,1\}^k\times\{0,1\}^n\rightarrow\{0,1\}^n$$

which takes as info a key K of bit length k, called the key size, and a bit string P of length n, called the square size, and returns a string C of n bits. P is known as the plaintext, and C is named the ciphertext. For every K, the capacity EK(P) is required to be an invertible mapping on 0,1n. The converse for E is characterized as a function

$E{-1}\ K\ (C) = DK(C) = D\ (K, C)$: $\{0,1\}\ k\times\ \{0,1\}\ n \rightarrow \{0,1\}\ n$
$\{0,1\}\ n$

Taking a key K and a ciphertext C to return a plaintext value P, suchthat

$$\forall K{:}D_K\,(E_K\,(P)) = PallK{:}D_K\,(E_K\,(P)) = P.$$

Step 1: Create a new Key Generator for the HmacMD5 algorithm.

C. HARDWARE AND SOFTWAREREQUIREMENTS

Hardware Requirements

1) Processor - Intel i5core

2) Speed - 1.1GHz
3) RAM -2GB
4) Hard Disk - 40GB
5) Key Board - Standard WindowsKeyboard
6) Mouse - Two or Three ButtonMouse
7) Monitor -SVGA

Software Requirements

1) Operating System - XP,Windows7/8/10
2) Codinglanguage-Java,MVC,JSP,HTML,CSSetc
3) Software -JDK1.7
4) Tool - Eclipse
5) Server - Apache Tomcat7.0
6) Database - MySQL5.0

IV. SYSTEM ANALYSIS AND RESULT

Experimental setup Table 1-shows the execution time for searching the entered keywords in no. of documents. The pro-installed search for the system string that takes one round of communication, O(n) times of computation over n documents. It provides an efficient matched document in accordance with the entry string, as it considers keywords adjacent. Fig.2-Shows a pictorial representation of no. of-documents search time. X — Axis includes no. of document and y-axis with (ms) time.

Graph shows how the time of search varies with respect to the number of documents in the proposed system. In our implementation, the time of search depends not only on the number of documents returned, but also on the number of documents in which the keywords to be searched are present. This is due to the fact that, for efficient implementation, this system will first check if all keywords are present in a document. If all keywords are present in a document, then and only then this system will check the adjacency. It can be noted that the most important part of the search time is taken by checking.

TABLE I
EXECUTION TIME FOR SEARCHING THE ENTERED KEYWORDS IN NO. OF DOCUMENTS.

Index	No. of documents	position integers	Hash chaining
1	1000	850	1010
2	2000	1050	1400
3	3000	1400	1900

V.CONCLUSION

The proposed system proposes a new secure search protocol. It introduces a new security scheme in SSE, named, search pattern in distinguishability. It can be observed that with security, although keywords are guaranteed to be secure against possible leakage from the index, they do not guarantee security against possible leakage from the trapdoor. To this end, the system introduces probabilistic trapdoor for the first time and proves that our scheme is safe under this criterion.

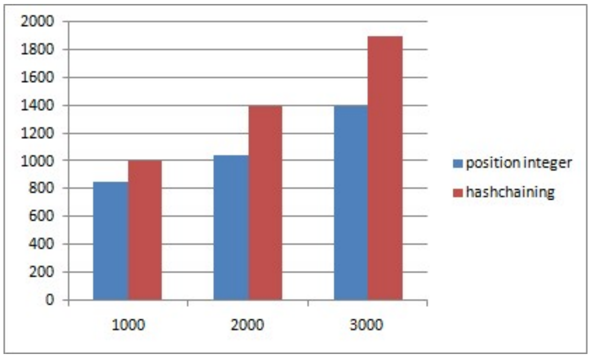

Fig. 2. searching time for no. of documents.

VI. ACKNOWLEDGMENT

The authors would like to thank the researchers as well as publishers for making their resources available and teachers for their guidance. We are thankful to the authorities of Savitribai Phule University of Pune and concern members of cPGCON2020 conference, organized by, for their constant guidelines and support. We are also thankful to the reviewer for their valuable suggestions. We also thank the college authorities for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

REFERENCES

[1] Enabling Generic, Verifiable, and Secure Data Search in Cloud Services Jie Zhu, Qi Li, Senior Member, IEEE, Cong Wang, Senior Member, IEEE, Xingliang Yuan, Member, IEEE, Qian Wang, Member, IEEE, Kui Ren, Fellow, IEEE (2018).

[2] M-SSE: An Effective Searchable Symmetric Encryption with Enhanced Security for Mobile Devices CHONGZHI GAO1,4, SIYI LV2 , YU WEI2 , ZHI WANG2 , ZHELI LIU2 , AND XIAOCHUN CHENG3 IEEE(2018)

[3] Muhammad Naveed, Seny Kamara, and Charles V Wright. Inference Attacks on Property-Preserving Encrypted Databases. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, pages 644655. ACM,2015.

[4] Jonathan Katz and Yehuda Lindell. Introduction to ModernCryptography. CRC press,2014.

[5] Seny Kamara, Charalampos Papamanthou, and Tom Roeder. Dynamic Searchable Symmetric Encryption. In Proceedings of the 2012 ACM conference on Computer and communications security, pages 965976. ACM,2012.

[6] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi. Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions. volume 21, pages 350391. Springer,2008.

[7] Mihir Bellare, Alexandra Boldyreva, and Adam ONeill. Deterministicand Efficiently Searchable Encryption. In Annual International Cryptology Conference, pages 535552. Springer,2007.

[8] Ning Cao, Cong Wang, Ming Li, Kui Ren, and Wenjing Lou. Privacy- Preserving Multi-Keyword Ranked Search Over Encrypted Cloud Data. volume25,pages222233.IEEE,2014.

[9] David Cash, Joseph Jaeger, Stanislaw Jarecki, Charanjit S Jutla, Hugo Krawczyk, Marcel-Catalin Rosu, and Michael Steiner. Dynamic Search- able Encryption in Very-Large Databases: Data Structures and Implementation.volume2014,page853.Citeseer,2014.

[10] Yoshinao Uchide and Noboru Kunihiro. Searchable symmetric encryption capableofsearchingforanarbitrarystring.WileyOnlineLibrary,2016.Jonathan Katz and Yehuda Lindell. Introduction to ModernCryptography. CRC press,2014.