ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

A Survey on Improved Privacy Security over Cloud using Multi-Keyword Graded Exploration

Rohini R. Mahajan, Ganesh N. Dhanokar

M.E.2nd Year, Dept. of Computer Science Engineering and Information Technology, G.H.R.I.E.M., Jalgaon, Maharashtra, India

Assist. Professor, Dept. of Computer Science Engineering and Information Technology, G.H.R.I.E.M., Jalgaon, Maharashtra, India

ABSTRACT: Despitetaking into the consideration of benefits of cloud computing data, proprietors are developing more priority to put their data in the cloud. Records in the cloud should be encrypted before outsourcing. Considering the large network traffic and size of the data it is necessary to allow multiple keywords in the search request to get appropriate results. Currently, the mechanism of search encryption is mainly focusing on the limited keywords. In the present paper, for the first time, we are introducing multi-keyword ranked search over encrypted data in the cloud computing (MRSE) by considering privacy. For multi-keyword search mechanism on the cloud, we choose the most similar "Coordinate matching" i.e., the maximum similar results to capture the search query. Initially, the idea proposed for the MRSE based on secure inner product computation, it gives couple of significantly improved MRSE schemes to achieve various stringent privacy requirements in two distinct threat models.

KEYWORDS: Accuracy, Cloud computing, Multi-keyword, Privacy preserving

I.INTRODUCTION

1.1. Data Storage in Cloud Computing

Distributed computing is a field of computer science that studies distributed systems. A distributed system is a software system in which components located on networked computers communicate and coordinate their actions by passing messages. There are many alternatives for the message passing mechanism, including RPC-like connectors and message queues. Three significant characteristics of distributed systems are: concurrency of components, lack of a global clock, and independent failure of components. An important goal and challenge of distributed systems is location transparency. Examples of distributed systems vary from SOA-based systems to massively multiplayer online games to peer-to-peer applications. Distributed computing also refers to the use of distributed systems to solve computational problems. In this survey paper, the computational entities are called computers or nodes.

Distributed system may have a common goal, such as solving a large computational problem. Alternatively, each computer may have its own user with individual needs, and the purpose of the distributed system is to coordinate the use of shared resources or provide communication services to the users. Other typical properties of distributed systems include the following (1) The system has to tolerate failures in individual computers. (2) The structure of the system (network topology, network latency, number of computers) is not known in advance, the system may consist of different kinds of computers and network links, and the system may change during the execution of a distributed program. (3) Each computer has only a limited, incomplete view of the system. Each computer may know only one part of the input.

The same system may be characterized both as "parallel" and "distributed"; the processors in a typical distributed system run concurrently in parallel. Parallel computing may be seen as a particular tightly coupled form of distributed

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

computing, and distributed computing may be seen as a loosely coupled form of parallel computing. Nevertheless, it is possible to roughly classify concurrent systems as "parallel" or "distributed" using the following criteria:(1) In parallel computing, all processors may have access to a shared memory to exchange information between processors. (2)In distributed computing, each processor has its own private memory (distributed memory). Information is exchanged by passing messages between the processors.

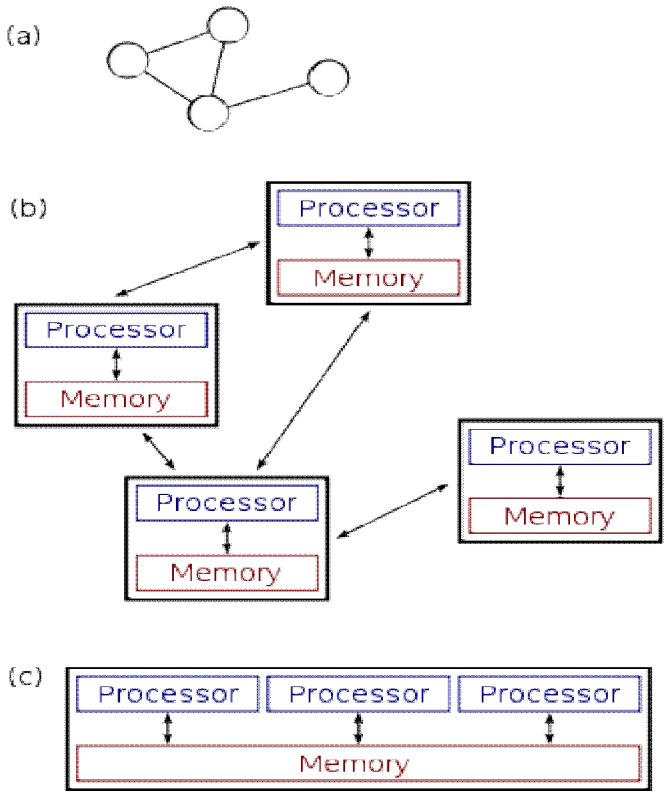


Figure.1(a) is a schematic view of a typical distributed system; as usual, the system is represented as a network topology in which each node is a computer and each line connecting the nodes is a communication link. Figure (b) shows the same distributed system in more detail: each computer has its own local memory, and information can be exchanged only by passing messages from one node to another by using the available communication links. Figure (c) shows a parallel system in which each processor has a direct access to a shared memory.

The situation is further complicated by the traditional uses of the terms parallel and distributed algorithm that do not quite match the above definitions of parallel and distributed systems

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

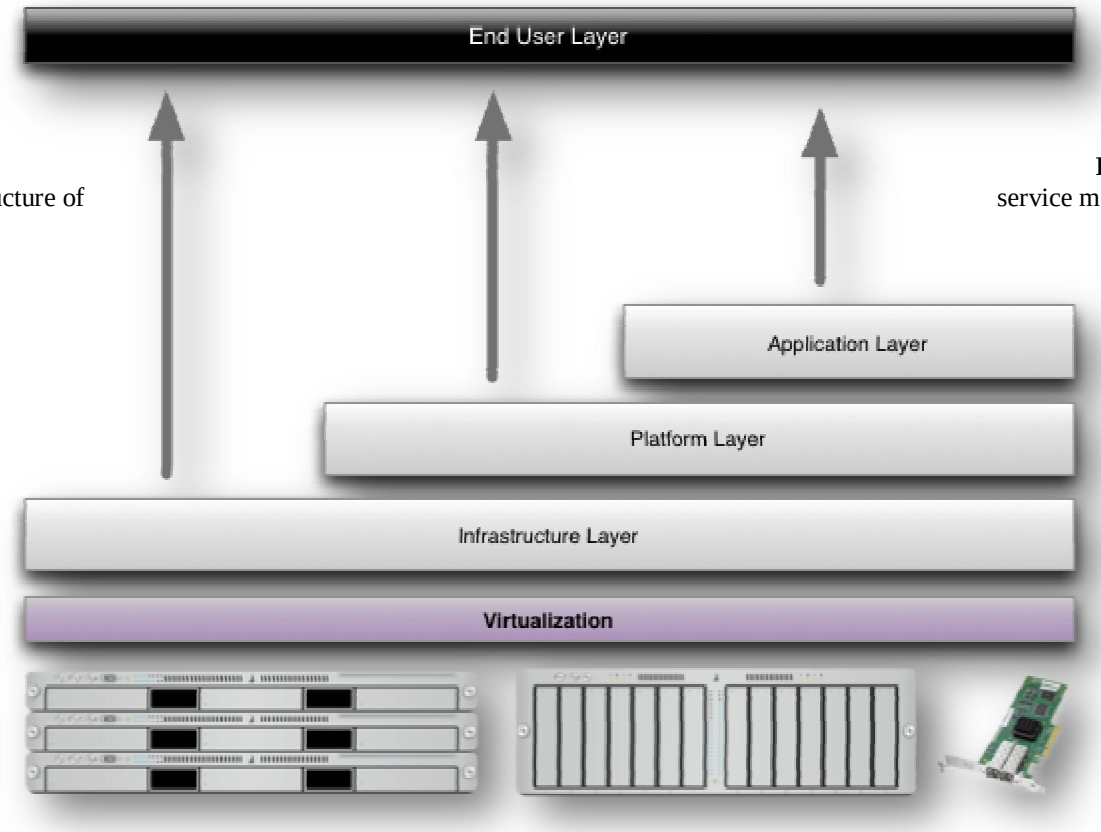
(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

1.2. Services Models

Cloud Computing comprises three different service models, namely Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). The three service models or layer are completed by an end user layer that encapsulates the end user perspective on cloud services. The model is shown in figure below. If a cloud user accesses services on the infrastructure layer, for instance, she can run her own applications on the resources of a cloud infrastructure and remain responsible for the support, maintenance, and security of these applications herself. If she accesses a service on the application layer, these tasks are normally taken care of by the cloud service provider.



Structure of **Figure.2**. service models

II.LITERATURE SURVEY

Cao, N., Wang, C., Li, M., Ren, K and Lou W.

With the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. Considering a large number of data users and documents in the cloud, it is necessary to allow multiple keywords in the search request and return documents in the order of their relevance to these keywords. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely categorizethe search results. In the present work, for the first time, we define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted data in cloud computing (MRSE).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Kamara,S., andLauter, K.
We consider the problem of building a secure cloud storage service on top of a public cloud infrastructure where the service provider is not completely trusted by the customer. We describe, at a high level, several architectures that combine recent and non-standard cryptographic primitives in order to achieve our goal. We survey the benefits such architecture would provide to both customers and service providers and give an overview of recent advances in cryptography motivated specifically by cloud storage.

Chang, Y.C.,and Mitzenmacher,M.
We consider the following problem: a user $\mathcal{U}$ wants to store his files in an encrypted form on a remote file server $\mathcal{S}$. Later the user $\mathcal{U}$ wants to efficiently retrieve some of the encrypted files containing (or indexed by) specific keywords, keeping the keywords themselves secret and not jeopardizing the security of the remotely stored files. For example, a user may want to store old e-mail messages encrypted on a server managed by Yahoo or another large vendor, and later retrieve certain messages while travelling with a mobile device. In our studies, we offer solutions for this issueunder well-defined security requirements.

III. PROPOSED WORK AND METHDOLOGY

3.1. Architecture of search encryption data over cloud

Here, we define the problem of secure ranked keyword search over encrypted data on cloud, and provide such an effective protocol, which fulfils the secure ranked search functionality with little relevance score information leakage against keyword privacy.Thorough security analysis shows that our ranked searchable symmetric encryption scheme indeed enjoys “as-strong as- possible” security guarantee compared to previous schemes.

3.2. Design

3.2.1 Input Design: The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system.Design considered the following criteria:

a. What data should be given as input?
b. How the data should be arranged or coded?
c. The dialog to guide the operating personnel in providing input.
d. Methods for preparing input validations and steps to follow when error occur.

3.2.3 Output Design

A quality output is one, which meets the requirements of the end user and presents the information clearly. Efficient and intelligent output design improves the system’s relationship to help user decision-making.The output form of an information system should accomplish one or more of the following objectives:

a. Convey information about past activities, current status or projections of the Future.
b. Signal important events, opportunities, problems, or warnings.
c. Trigger an action.
d. Confirm an action.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

IV. RESULTS

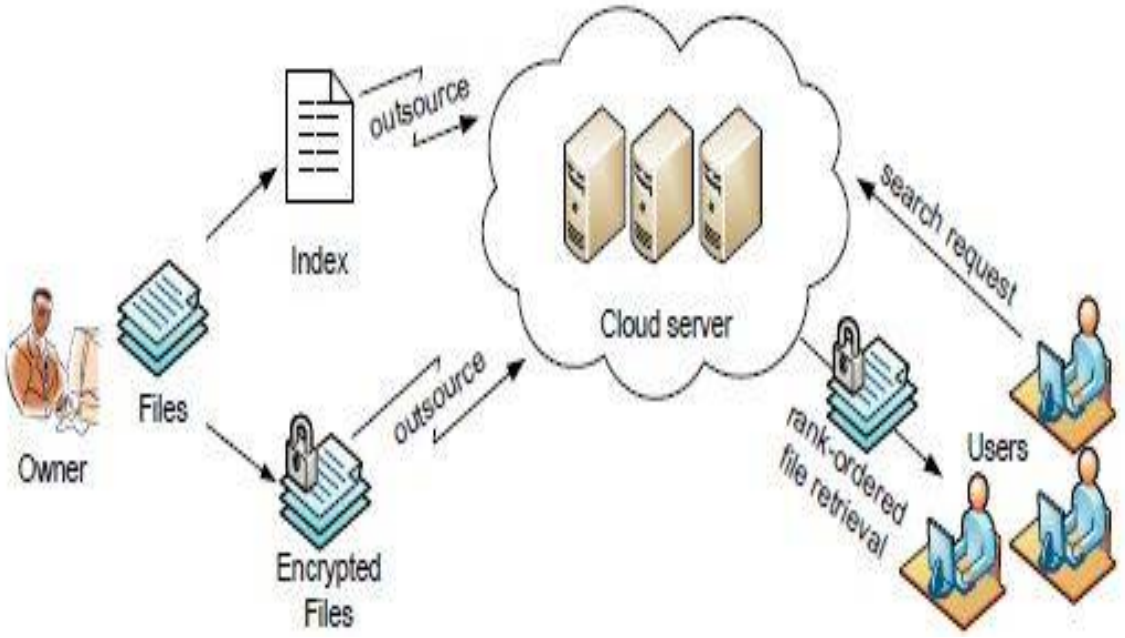


Figure 3. Architecture of search encrypted data over cloud

4.1 Architecture of search encrypted data over cloud

As illustrated in Figure 3 we consider a cloud data hosting service involving three different entities, data owner(O), Owner outsource Files outsource Encrypted Files search request Rank-ordered file retrieval Index Users Cloud server data user (U), and cloud server (CS). Data owner has a collection of *n* data files *C* = (*F*1, *F*2, . . ., *Fn*) that he wants to outsource on the cloud server in encrypted form while still keeping the capability to search through them for effective data utilization reasons. To do so, before outsourcing, data owner will first build a secure searchable index *I* from a set of *m* distinct keywords *W* = (*w*1, *w*2, ..., *wm*) extracted2 from the file collection *C*, and store both the index *I* and the encrypted file collection *C* on the cloud server.

We assume the authorization between the data owner and users is appropriately done. To search the file collection for a given keyword *w*, an authorized user generates and Submits a search request in a secret form a trapdoor *Tw* of the keyword *w*—to the cloud server. Upon receiving the search request *Tw*, the cloud server is responsible to search The index *I* and return the corresponding set of files to the user.

4.2 Merits of Input Design

a. Input design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
b. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

c. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow.

4.3 Merits of Output Design

Encryption	Decryption
• Through Public & Private key generation • Convert message to Unicode • Generate SHA instance • Convert message using SHA • Generate Blind values • Signed • Unbind and verify	• Generate RSA key • Convert into bytes • Generate SHA • Get Blind values • Signed retrieved values • Unbind retrieved values • Verify the retrieved values

4.5 Basic Privacy Protection

To enable efficient similarity search, data owner builds a secure index and outsources it to the cloud server along with the encrypted data items. Server performs search on the index according to the queries of the data users without learning anything about the data other than what data owner allows an adversary to learn. In Phase-I, I present the index structure. In Phase-II, Describe the search scheme that is built on top of the index.

4.5.1 Phase I: The Index Structure

Searchable Symmetric Encryption (SSE) scheme is based on a secure index structure that is built through locality sensitive hashing (LSH). LSH maps objects into several buckets such that similar objects collide in some buckets while dissimilar ones do not with high probability. Index structure is constructed on top of this property.

4.5.2 Phase II: Basic Secure Search Scheme

In this part, I describe the basic protocol for similarity SSE scheme, overview of which is presented in Figure 3 Initially data owner gets private keys and then he creates the index for the data collection. Alice encrypts the items with key to form the encrypted collection. Suppose a user is interested in retrieving the items, the user generates search query. Once the user receives file, user decrypts encrypted file. Once the encrypted items corresponding to the search request are retrieved, user decrypts them with the key to obtain their plain versions.

IV. CONCLUSION AND FUTURE SCOPE

Altogether, we describe, confirmed and solve the problem of multi-keyword ranked search over encrypted cloud data, and recognised a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of “coordinate matching,” i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use “inner product similarity” to quantitatively evaluate such similarity measure. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, we proposed a basic idea of MRSE using secure inner product computation. Additionally, we provided two improved MRSE schemes to achieve various stringent privacy requirements in two distinctive threat models. Further we also investigated some auxiliary enhancements of our ranked search mechanisms, including supporting more search semantics, i.e., TF_IDF, and dynamic data operations. Complete analysis investigating privacy and efficiency securities of proposed schemes is given, and experiments on the real-world data set show our proposed schemes introduces low overhead on both computation and communication. In future investigation, we will explore examining the integrity of the rank order in the search result assuming the cloud server is untrusted.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

REFERENCES

1. Cao, N.,Wang,C. M., Ren, Li. K. and Lou,W. "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829-837, Apr, 2011.
2. L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50-55, 2009.
3. Cao,N. Yu, S., Yang, Z., Lou, W.,and Hou,Y. "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.
4. Kamara,S. and Lauter, K."Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptography and Data Security, 2010.
5. Singhal, A. "Modern Information Retrieval: A Brief Overview," IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, 2001.
6. WittenI.H., Moffat, A. and Bell, T.C. "Managing Gigabytes: Compressing and Indexing Documents and Images". Morgan Kaufmann Publishing ,1999.
7. Song, D., Wagner, D., and Perrig, A., "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.
8. Goh, E.H., "Secure Indexes," Cryptology ePrint Archive, http:// eprint.iacr.org/2003/216, 2003.
9. Chang,Y.C., and Mitzenmacher, M., "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.
10. Curtmola, R., Garay, J.A., Kamara, S., and Ostrovsky, R., "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. 13th ACM Conf. Computer and Comm. Security (CCS '06), 2006.
11. Boneh, D.,Crescenzo, G.D.,Ostrovsky, R., and Persiano, G., "Public Key Encryption with Keyword Search," Proc. Int'l Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2004.
12. Bellare, M., Boldyreva, A., and ONeill, A., "Deterministic and Efficiently Searchable Encryption," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.
13. Abdalla, M.,Bellare, M.,Catalano, D.,Kiltz, E.,Kohno, T., Lange,T., Malone-Lee, J.,Neven, G.,Paillier, P.,and Shi, H., "Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous Ibe, and Extensions," J. Cryptology, vol. 21, no. 3, pp350- 391, 2008.
14. Li, J., Wang, Q.,Wang,C., Cao, N.,Ren, K., and Lou, W., "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, Mar. 2010.
15. Boneh, D.,Kushilevitz, E.,Ostrovsky, R., and W.E.S. III, "Public Key Encryption That Allows PIR Queries," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.
16. Golle, P., Staddon, J., and Waters, B., "Secure Conjunctive Keyword Search over Encrypted Data," Proc. Applied Cryptography and Network Security, pp. 31-45, 2004.
17. Ballard, L.,Kamara, S., and Monrose, F.,"Achieving Efficient Conjunctive Keyword Searches over Encrypted Data," Proc. Seventh Int'l Conf. Information and Comm. Security (ICICS '05), 2005.
18. Boneh,D. and Waters, B., "Conjunctive, Subset, and Range Queries on Encrypted Data," Proc. Fourth Conf. Theory Cryptography (TCC), pp. 535-554, 2007.