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 6, June 2017

Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data

Ghansham R. Rathod, Prof. SonaliPatil

M. E. Student, Dept. of Computer, JSPM's BSIOTR, Wagholi, Pune, Maharashtra, India

Asst. Professor, Dept. of Computer, JSPM's BSIOTR, Wagholi, Pune, Maharashtra, India

ABSTRACT: As cloud computing is more flexible & effective in terms of economy, data owners are encouraged tooutsource their complex data systems from local sites to commercial public cloud. But for security of data, sensitive data has to be encrypted before outsourcing, which overcomes method of traditional data utilization based on plaintext keyword search. Considering the large number of data users and documents in cloud, it is necessaryfor the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval .Retrieving of all the files having queried keyword will not be affordable in pay as per use cloud .In this paper, we proposethe problem of SecuredMulti-keyword search over encrypted cloud data, and construct a group ofprivacy policiesfor such a secure cloud data utilization system. Fromnumber ofmulti-keyword semantics, we selectthe highly efficient rule of coordinate matching, i.e., as many matches as possible, to identifythe similarity between search query and data , and for further matching we useinner datacorrespondence to quantitatively formalize such principle for similarity measurement. We first propose a basic Secured multi keyword ranked search scheme using secure inner product computation, and then improve it to meet different privacy requirements. The Ranked result provides top k retrieval results.Also we propose an alert system which will generate alerts when un-authorized user tries to access the data from cloud, the alert will generate in the form of mail and message

KEYWORDS:multi-keyword ranked search, Searchable encryption,and cloud computing, dynamic update

I. INTRODUCTION

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g.,networks,servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud Computing means a remote server that access through the internet which helps inbusiness applications and functionality along with the usage of computer software. Cloud computing saves moneythat users spend on annualor monthly subscription.Due to advantage of cloud services, more and more sensitive information are being centralized into the cloud servers, such as emails, personal health records, private videos and photos, company finance data, government documents, etc. .

multiple users but as well as dynamically re-allocated as per demand. This can perform for assigning resources to users in dissimilar time zones. For example, a cloud computing service which serves American usersduring American business timings with a specific application (e.g. email) while the same resources are getting reallocated and serve Indian users during Indian business timings with another application (e.g. web server).

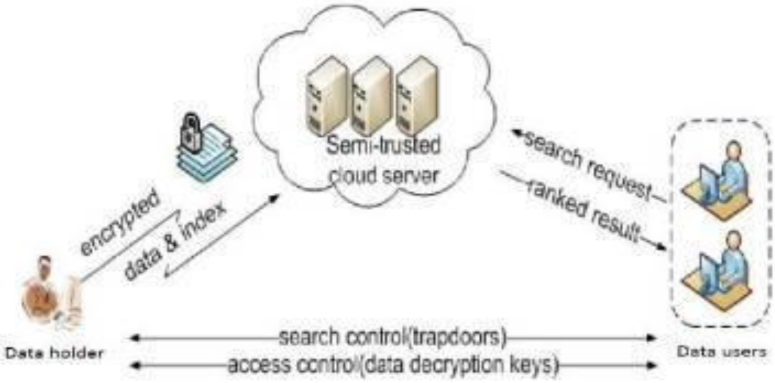
Semi-trusted cloud server
search request
ranked result
encrypted data & index
Data holder
Data users
search control(trapdoors)
access control(data decryption keys)

Figure 1: System Architecture

This mechanism should take full advantage of the use of computing powers thus decreasing environmental damage as well, since less power, air conditioning and so on, is necessary for the same functions. The expression "moving to cloud" also explains to an organization moving away from a traditional CAPEX model i.e buy the devoted hardware and decrease in value it over a period of time to the OPEX model i.e use a shared cloud infrastructure and pay as you use it. Proponents maintain that cloud computing Permit Corporation to avoid direct infrastructure costs, and focus on projects that distinguish their businesses as an alternative of infrastructure. Proponents also maintain s that cloud computing permit scheme s to get their applications should run faster, with better manageability and less maintenance, and enable IT to more quickly adjust resources to meet random and changeable business demand.

The secure kNN algorithm is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encrypted index and query vectors. To resist different attacks in different threat models, we construct two secure search schemes: the basic dynamic multi-keyword ranked search (BDMRS) scheme in the known ciphertext model, and the enhanced dynamic multi-keyword ranked search

(EDMRS) scheme in the known background model. Our contributions are summarized as follows: 1) We design a searchable encryption scheme that supports both the accurate multi-keyword ranked search and flexible dynamic operation on document collection. 2) Due to the special structure of our tree-based index, the search complexity of the proposed scheme is fundamentally kept to logarithmic. And in practice, the proposed scheme can achieve higher search efficiency by executing our “Greedy Depth-first Search” algorithm. Moreover, parallel search can be flexibly performed to further reduce the time cost of search process.

II. RELATED WORK

1.Enabling Public Verifiability and Data Dynamic For Storage Security In Cloud Computing [1]-

Cloud computing represents today's most exciting computing paradigm shift in information technology [1]. However, security and privacy are perceived as primary obstacles to its wide adoption. Here, the author's outline several critical securities challenges and motivate further investigation of security solutions for a trustworthy public cloud environment.

2. Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing [2].From This Paper I Referred-

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. Author [2][3] describes, 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 an architecture would provide to both customers and service providers and give an overview of recent advances in cryptography motivated specifically by cloud storage.

3. Ensuring Data Storage Security in Cloud Computing [3]From This Paper I Referred-

Software protection is one of the most important issues concerning computer practice [3]. There exist many heuristics and ad-hoc methods for protection, but the problem as a whole has not received the theoretical treatment it deserves. In this paper author provide theoretical treatment of software protection. author reduce the problem of software protection to the problem of efficient simulation on oblivious RAM. A machine is oblivious if the sequence in which it accesses memory locations is equivalent for any two inputs with the same running time. For example, an oblivious Turing Machine is one for which the movement of the heads on the tapes is identical for each computation.

4.Dynamic Audit Services for Outsourced Storages in Clouds [4] From This Paper I Referred-

We study the problem of searching on data that is encrypted using a public key system [4]. 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. author define and 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. We refer to this mechanism as Public Key Encryption with keyword Search. As another example, consider a mail server that stores various messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some specific keyword, but learn nothing else. Authordefine the concept of public key encryption with keyword search and give several constructions.

5. Public key encryption that allows pir queries.[5]From This Paper I Referred

Consider the following problem: Alice wishes to maintain her email using a storage provider Bob (such as a Yahoo! or hotmail e-mail account) [5]. This storage-provider should provide for Alice the ability to collect, retrieve, search and delete emails but, at the same time, should learn neither the content of messages sent from the senders to Alice (with Bob as an intermediary), nor the search criteria used by Alice. A trivial solution is that messages will be sent to Bob in encrypted form and Alice, whenever she wants to search for some message, will ask Bob to send her a copy of the entire database of encrypted emails. This however is highly inefficient. We will be interested in solutions that are communication-efficient and, at the same time, respect the privacy of Alice. In this paper, we show how to create a publickey encryption scheme for Alice that allows PIR searching over encrypted documents. Our solution provides a theoretical solution to an open problem posed by Boneh, DiCrescenzo, Ostrovsky and Persiano on “Public-key Encryption with Keyword Search”, providing the first scheme that does not reveal any partial information regarding user’s search (including the access pattern) in the public-key setting and with non-trivially small communication complexity

III. PROPOSED ALGORITHM

A. *KEYWORD-NNE:*

In previous work, BKC algorithm drops its performance when the number of query keywords is increases. To solve this problem, here developed a more efficient keyword nearest neighbour expansion (keyword-NNE) which uses the different strategy. In this algorithm, one query is considered as a principal query keyword. Those objects are associated with principal query keyword are considered as principal objects. Keyword-NNE computes local best solution for each principal object. BKC algorithm returns the lbkc with having highest evaluation. For each of the principal object, its lbkc can be simply selects few closest and highly rated objects by the viewer/customer. Compared with the k-meansclustering, the keyword covers significantly reduced. These keyword covers a further processe inkeyword-NNE-algorithm that will be optimal, and each keyword candidate covers processed generates very less new candidate keyword are covers.

Algorithm to provide efficient multi-keyword ranked search.

The secure kNN algorithm is utilized to encrypt the index and query vectors.

Propose a “Greedy Depth-first Search” algorithm based on this index tree.

Algorithm achieves better-than-linear search efficiency but results in precision loss.

The LSH algorithm is suitable for similar search but cannot provide exact ranking.

I′s ; ci} ← GenUpdateInfo (SK; Ts; i; up type)) This algorithm generates the update information {I′s ; ci} which will be sent to the cloud server.

a. DESIGN GOALS:

To enable secure, efficient, accurate and dynamic multi data under the above models, our system has the following Dynamic: The proposed scheme is designed to provide not only multi - keyword query and accurate result ranking, but also dynamic update on document collections. Search Efficiency: The scheme aims to achieve sublinear search efficiency by exploring a special tree - based index and an efficient search algorithm.

A.Privacy - preserving: The scheme is designed to prevent the cloud server from le arning additional information about the document collection, the index tree, and the query. The specific privacy requirements are summarized as follows,

B. Index Confidentiality and Query Confidentiality: The underlying plaintext information, including keywords in the index and query, TF values of keywords stored in the index, and IDF values of query keywords, should be protected from cloud server;

C. Trapdoor Unlinkability: The cloud server should not be able to determine whether two encrypted queries (trapdoors) are generated from the same search request;

Keyword Privacy: The cloud server could not identify the specific keyword in query, index or document collection by analyzing the statistical information like term frequency. Note that our proposed scheme is not designed to protect access pattern, i.e., the sequence of returned documents

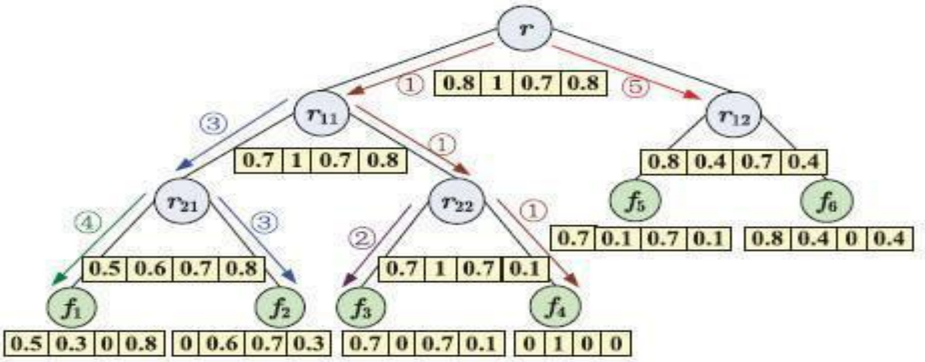

Figure 2. An example of the tree-based index with the document collection

We construct a special keyword balanced binary tree as the index, and propose a “Greedy Depth-first Search” algorithm to obtain better efficiency than linear search.

IV. SYSTEM ARCHITECTURE

A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data We construct a special tree-based index structure and propose a “Greedy Depth-first Search” algorithm to provide efficient multi-keyword ranked search. The proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly.

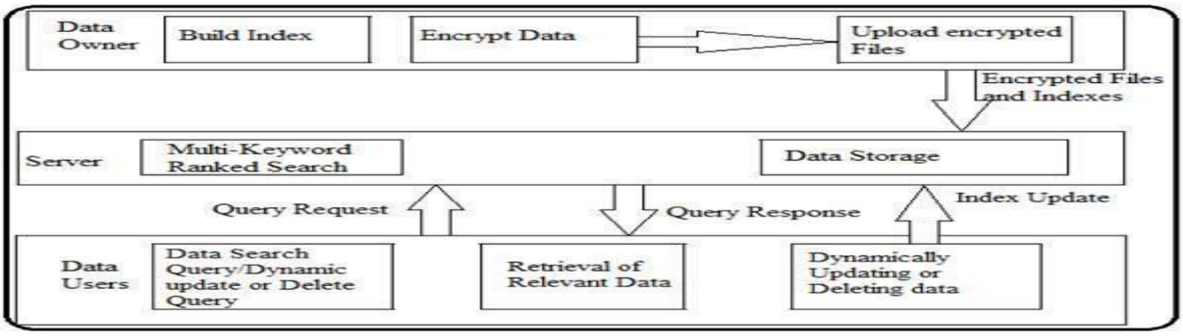
Data Owner
Build Index
Encrypt Data
Upload encrypted Files
Encrypted Files and Indexes
Server
Multi-Keyword Ranked Search
Data Storage
Query Request
Query Response
Index Update
Data Users
Data Search Query/Dynamic update or Delete Query
Retrieval of Relevant Data
Dynamically Updating or Deleting data

Figure 3: Block Diagram of proposed system

Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

1. Abundant works have been proposed under different threat models to achieve various search functionality,

2. Recently, some dynamic schemes have been proposed to support inserting and deleting operations on document collection.

3. This paper proposes a secure tree-based search scheme over the encrypted cloud data, which supports multi keyword ranked search and dynamic operation on the document collection.

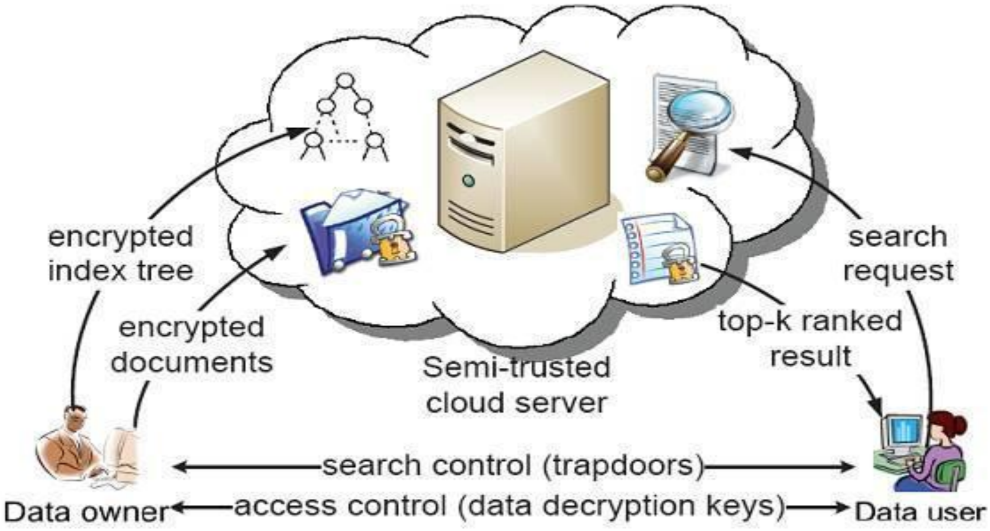

Figure 3. The architecture of ranked search over encrypted cloud data

Despite of the various advantages of cloud services, outsourcing sensitive information such as e-mails, personal health records, company finance data, government documents, etc.

V. RESULT

In this secation we present comparision result of Single Key word Search Ranked search and Multi Keyword Ranked Search Over A Encrypted Data On Cloud as shown in following figers .In this Result Exch Ranked search and Multi Keyword Ranked Search Over A Encrypted Data On Cloud.In this Result Existing System is Single Keyword Search System and proposed System is nothing but MRES System

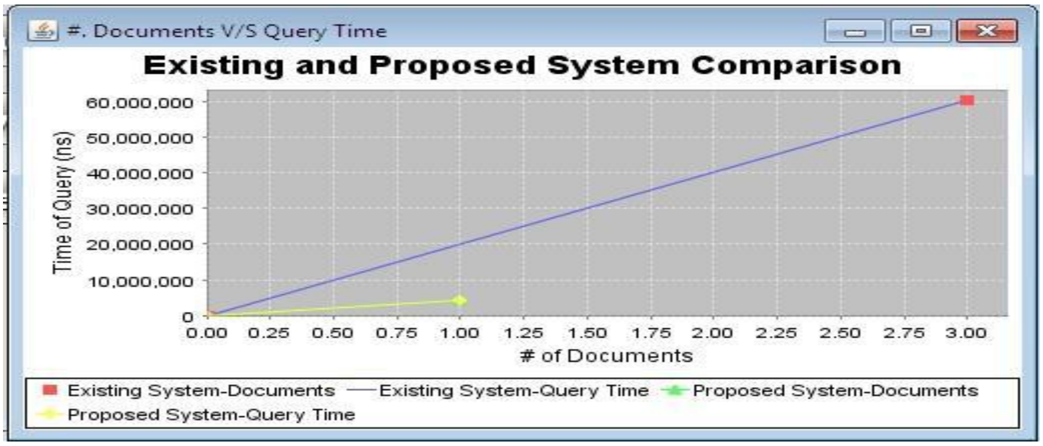

Fig 1: Comparison Graph- No. Of Documents V/S Query Time

Fig 1 is a Comparison graph of Existing System and Our System. The graph is plotted Number of Documents that the respective system's search result returned and Time required to return the documents; in respective System. As shown in the graph Our system requires less time which is less than around 5 ns with most specific result of number of document equal to one which is less than the documents returned by existing system which are three and time required is around 6ns

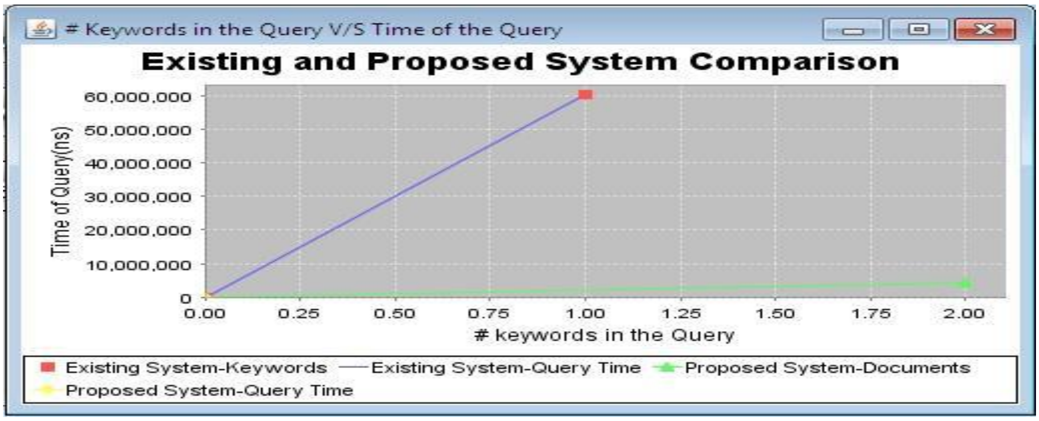

Fig 2: Comparison Graph- No. Of Keywords V/S Query Time

Fig 2 is a Comparison graph of Existing System and Our implemented System. The graph is plotted against Number of Keywords fired in the respective system's search and Time required in respective System. As shown in the graph Our system requires less time which is less than around 5 ns with multiple Keyword Query and existing system requires around 6ns even though a single Keyword query is fired. So Our System Works Better in each and every aspect then existing System.

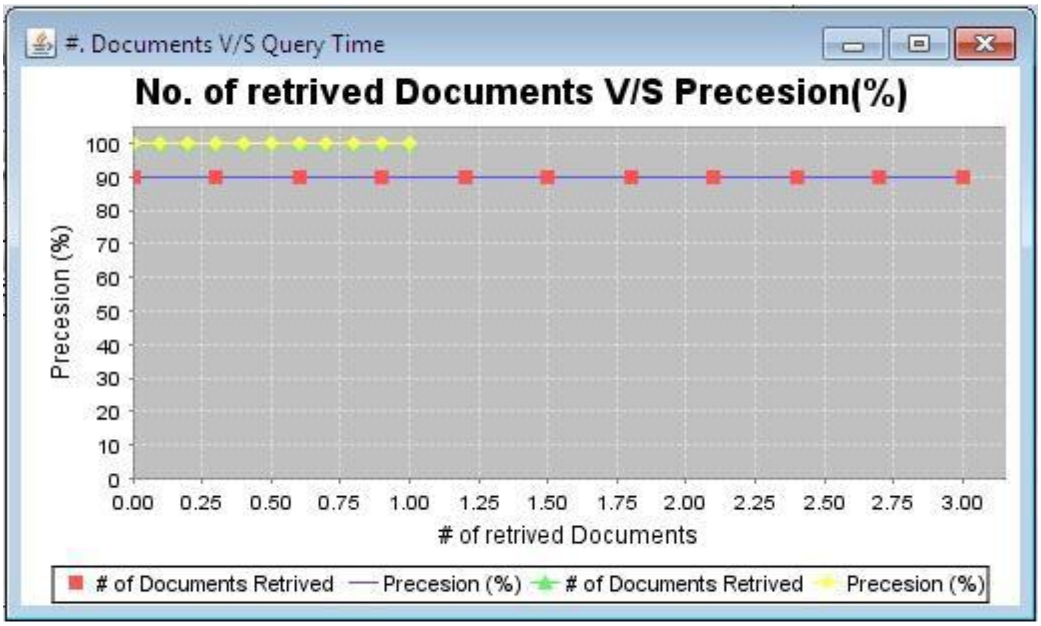

Fig 3: No. Of Retrieved Document V/S Precision.

Fig 3 shows The Comparison graph of Existing System and Our implemented System plotted against No of Documents returned by respective system V/S percentage Precision(Perfectness). Our system gives much better precession than existing system as shown in graph.

In this method the major merits are: (1) data security (2) privacy shield (3) Auditing details to the data owner (4) Audit aptitude aware data scheduling at this time we are going to evaluate the performance of our projected scheme in terms of the computation overhead introduce by each operation. Request and resources are taken as the computing parameter. When the number of requests increase at the same time, it is to check whether they are served within a particular time. The waiting time is measured for each request.

V. CONCLUSION

Thus we proposedthe problem of multiple-keyword ranked search over encrypted cloud data, and construct a variety of security.requirements.From various multikeywordconcepts, we choose the efficient principle of coordinatematching. We first proposesecureinner data computation.Also we achieve effective ranking result using knearestneighbour technique. This system is currently work on single cloud, Infuture is will extended up to sky computing & Provide better security in multi-user systems.

REFERENCES

1. Ren, C. Wang, and Q. Wang, “Security challenges for the public cloud,” IEEE Internet Comput., vol. 16, no. 1, pp. 69–73, Jan-Feb. 2012.
2. S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Proc. Financ. Cryptography Data Secur., 2010, pp. 136–149.
3. Gentry, “A fully homomorphic encryption scheme,” Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 2009.
4. O. Goldreich and R. Ostrovsky, “Software protection and simulation on oblivious rams,” J. ACM, vol. 43, no. 3, pp. 431–473, 1996.
5. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Proc. Adv. Cryptol.-Eurocrypt, 2004, pp. 506–522.
6. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, “Public key encryption that allows pir queries,” in Proc. Adv. Cryptol.,
7. 2007, pp. 5067.
8. X. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. IEEE Symp. Secur. Privacy, 2000, pp. 44–55.
9. E.-J. Goh, “Secure indexes,” IACR Cryptol. ePrint Archive, vol. 2003, p. 216, 2003.
10. Y.-C. Chang and M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data,” in Proc. 3rd Int. Conf. Appl. Cryptography Netw.Secur., 2005, pp. 442–455.
11. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, “Searchable symmetric encryption: Improved definitions and efficient constructions,” in Proc. 13th ACM Conf. Comput. Commun. Secur., 2006, pp. 79–88.
12. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, “Fuzzy keyword search over encrypted data in cloud computing,” in IEEE Proc. INFOCOM, 2010, pp. 1–5.
13. Kuzu, M. S. Islam, and M. Kantarcioglu, “Efficient similarity search over encrypted data,” in Proc. IEEE 28th Int. Conf. Data Eng., 2012, pp. 1156–1167.
14. C. Wang, K. Ren, S. Yu, and K. M. R. Urs, “Achieving usable and privacy-assured similarity search over outsourced cloud data,” in Proc. IEEE INFOCOM, 2012, pp. 451–459.