International Journal of Advances in Applied Sciences (IJAAS)
Vol. 9, No. 1, March 2020, pp. 67~69
ISSN: 2252-8814, DOI: 10.11591/ijaas.v9.i1.pp67-69

Top-K search scheme on encrypted data in cloud

K. Pushpa Rani[1], L. Lakshmi[2], Ch. Sabitha[3], B. Dhana Lakshmi[4], S. Sreeja[5]
[1,2]Department of Computer Science and Engineering, MLR Institute of Technology, India
[3]Department of Computer Science and Engineering, Vardhaman College of Engineering, India
[4]Department of Computer Science and Engineering, IARE College of Engineering, India
[5]Department of Computer Science and Engineering, St. Martins College of Engineering, India

Article Info

Article history:

Received May 8, 2019
Revised Sep 20, 2019
Accepted Dec 10, 2019

Keywords:

Cloud
Data proprietors
Multi keyword ranked seek over encrypted cloud facts
OTP
Product resemblance

ABSTRACT

A Secure and Effective Multi-keyword Ranked Search Scheme on Encrypted Cloud Data. Cloud computing is providing people a very good knowledge on all the popular and relevant domains which they need in their daily life. For this, all the people who act as Data Owners must possess some knowledge on Cloud should be provided with more information so that it will help them to make the cloud maintenance and administration easy. And most important concern these days is privacy. Some sensitive data exposed in the cloud these days have security issues. So, sensitive information ought to be encrypted earlier before making the data externalized for confidentiality, which makes some keyword-based information retrieval methods outdated. But this has some other problems like the usage of this information becomes difficult and also all the ancient algorithms developed for performing search on these data are not so efficient now because of the encryption done to help data from breaches. In this project, we try to investigate the multi- keyword top-k search problem for encryption against privacy breaks and to establish an economical and secure resolution to the present drawback. we have a tendency to construct a special tree-based index structure and style a random traversal formula, which makes even identical question to supply totally different visiting ways on the index, and may additionally maintain the accuracy of queries unchanged below stronger privacy. For this purpose, we take the help of vector area models and TFIDF. The KNN set of rules are used to develop this approach.

Corresponding Author:

K. Pushpa Rani,
Department of Computer Science and Engineering,
MLR Institute of Technology,
Hyderabad-500043, India.
Email: rani536@gmail.com

1. INTRODUCTION

These days, cloud computing [1] has emerged as an essential mechanism for plenty utilities, where cloud customers can keep their statistics into the cloud that allows them to take advantage from on-demand extremely good request and offerings from a shared pool of configurable computing assets. Cloud computing is now a days a trend in most of the IT industries because of its extra ordinary features like Pay-as-you-go basis. This will help people in achieving their necessities with basic cost on all the resources. People need not spend so much in the starting stage itself. As in the beginning, any company needs only basic services and based on its growing demand, it can set up all the resources and necessities. Hence cloud computing has become a good trend as it is eliminating most of the unnecessary costs to a basic startup company. Nowadays, additional and additional corporations and people from an outsized variety of huge information

applications have source their information and maintain their data in cloud servers for simple information management, economical process and question processing tasks. In such cases there is a high risk of security [2, 3] issues as there are many sensitive data like e-mail, health records of individuals etc. The owner of the data [4] is concerned about privacy apart from enjoying the benefits of the cloud. Their outsourced information must be in a way that is more secure so that it is not possible for illegitimate users [5-8] to access their data. For this purpose, the outsourced data must be converted to a format that is not readable easily but is able to accessible only to those who is a legal or valid user of the data. In our project, this can be done by encrypting the data before it is put in the cloud. And as a result of this, the ancient methods for data retrieval are not so efficient on such data. So, there is another method proposed in the project that makes the Search on this type of data fast and efficient. It is known as Top-k approach [9, 10]. This approach will help construct tree-based indexes which are nearer to search criteria. Different tree-based paths are obtained which when traversed gives unique search results.

2. RESEARCH METHOD

The framework consists of explorable encoded method that helps the accurate multiple key word ranked seek and bendy vigorous running on document group. This framework is a relaxed tree shaped based totally exploring model on the enciphered cloud statistics, which helps multi key word ranked seek. The so-called vector space model and the extensively casted off "time period frequency (TF) × inverse record frequency (IDF)" groups are binded to the index construction and also for generation of question of search.

Algorithm: Term Frequency-Inverse Document Frequency
Input: *Data d.*
Output: *result r.*

```
Let data d,
        Collection c;
        c=getWords(d); //Using Split("\\s+")

Term Frequency  tf1;
α= the count of terms t appearing in a document;
        tf1=( α);.
Inverse Document Frequency idf1;

α= The count of the terms t that are present in a document;
β= Total number of terms in the document;
        IDF1(t) = ( α)/( β);.
End;
```

3. RESULTS AND ANALYSIS

The suggested one, data users/people can acquire specific necessities on search correctness of privateness with the aid of the standard deviation of adjustment that can be dealt with as a compensation parameter. The assessment of structures with a recent painting prove that it gives a high seek performance. PMRSE scheme calls the hunt results with the aid of specific reckoning of two types of vectors i.e document and query. Thus, the seek accuracy of PMRSE scheme is 100%. But based totally and similarity Multi- key-word rectangular seek pattern, the basic scheme is affected by lack of precision because of various factors like accumulation of sub-vectors along with the index creation. The validation is iterated 16 number of times. Average accuracy of 91%. During the quest, whilst the relevance of the node is higher in Rlist, examines the server of the cloud. Because it is a balanced binary tree, its height of the index n should be taken care. The convolution of the calculation is ranked relevance order of m. We carried out an experimental assessment of the existing System of RSSE and the proposed one PMRSE. The Comparison graph is drawn. The graph coordinates are based on number of documents the corresponding system's search end result given back along with the time required to go back the documents. The complete experiment machine is carried out with the aid of Java on a Windows Server with Intel i5 2.93GHz. Figure 1 – axis will show the Time (milliseconds), Y- axis represents no. of documents retrieved. By this comparision PMRSE got best results compare with RSSE. The result values is presented in Table 1.

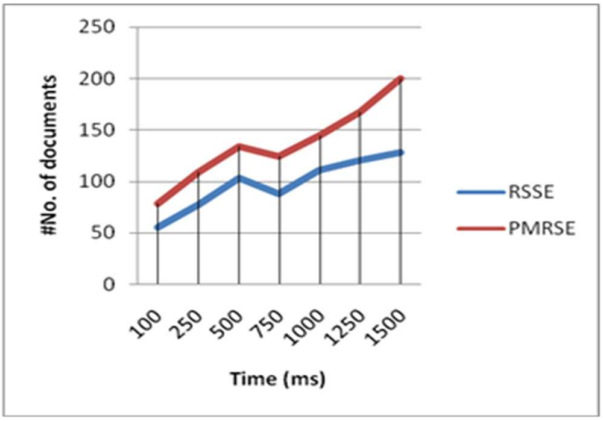

Figure 1. Comparison PMRSE with RSSE

Table1. Results table

Time (ms)	RSSE	PMRSE
100	55	78
250	77	109
500	103	134
750	88	124
1000	111	144
1250	120	167
1500	128	200

4. CONCLUSION AND FUTURE WORK

The proposed model tries to improve the coherence of the top- k multiple keyword search over encrypted data. For this purpose, we try two same questions with unalike keys, for which the server traverse through two unassociated byways to give the user most accurate search results. Then, we also tried to divide the entire dictionary into multiple groups top-ck documents while building index. Traversal algorithm used is RGTMS. Finally, the experimental upshots will teach that our methods are extra added efficient along with a safer than the state-of-the-art methods.

REFERENCES

[1] K. Ren, C. Wang, Q. Wang, *et al.*, "Security challenges for the public cloud," *IEEE Internet Computing*, vol. 16, no. 1, pp. 69–73, 2012.
[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in *Financial Cryptography and Data Security*, Springer, pp. 136–149, 2010.
[3] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.
[4] O. Goldreich and R. Ostrovsky, "Software protection and simulation on oblivious rams," *Journal of the ACM (JACM)*, vol. 43, no. 3, pp. 431–473, 1996.
[5] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Advances in Cryptology-Eurocrypt 2004*, Springer, pp. 506–522, 2004.
[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir queries," in *Advances in Cryptology-CRYPTO 2007*. Springer, pp. 50–67, 2007.
[7] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Security and Privacy, 2000. S&P 2000.Proceedings*, 2000 IEEE Symposium, pp. 44–55, 2000.
[8] E.-J. Goh *et al.*, "Secure indexes," IACR Cryptology ePrint Archive, vol. 2003, p. 216, 2003.
[9] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Proceedings of the Third international conference on Applied Cryptography and Network Security*. Springer-Verlag, pp. 442–455, 2005.
[10] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in *Proceedings of the 13th ACM conference on Computer and communications security*. ACM, pp. 79–88, 2006.