Design and analysis of new multi keyword ranked search schema called SSEDU in cloud computing

[1]Pawan Kumar Tanwar, [2]Ajay Khunteta, [3]Vishal Goar

[1]Research Scholar, [2]Professor, [3]Assistant Professor

[1,2]Department of Computer Science, Department of Computer Application

[1,2]Poornima University, Jaipur, [3]Govt. Engineering College, Bikaner

[1]pktbkn@gmail.com, [2]khutetaajay@poornima.org, [3]dr.vishalgoar@gmail.com

Abstract: In these days, required information can be searched through the cloud. Henceforth, this paper deals with the cloud computing, where information seeking and privacy preservation are the main area of emphasis. Therefore, keeping multi-keyword ranked search with dynamic updation as an area of data searching (information seeking), have been chosen here as a research dimension. One more thing is that the in most of the cases searching is bounded to unit keyword only. Hence, we have chosen the other dimension of searching which is multi keyword ranking search. Moreover, efforts have not been made in the area of dynamic updation earlier. Here dynamic updation caters the addition and deletion of documents dynamically. For covering the dynamic updation part the scheme SSEDU (searchable symmetric encryption with dynamic updation) has been designed and deployed. In this paper we have presented the efforts made for the performance evaluation of the said SSEDU scheme.

Keywords: Cloud Computing, Dynamic Updation, Multi-keyword ranked search, SSEDU

INTRODUCTION

We have introduced and designed a new scheme for dynamic updation in multikeyword ranked search called SSEDU (Searchable symmetric encryption with dynamic updation). The scheme is extremely parallel and updates can be tackled in easy way. The scheme supports parallel keyword searching as well as parallel addition and deletion of documents.

In this work, we have focused for the formations of SSEDU (Searchable symmetric encryption with dynamic updation) schema. Following task has been performed to justify our work:

1. We have presented a fundamental definition of security for proposed SSEDU schema. Specifically, the definition obtains a powerful concept of security for searchable symmetric encryption that is resilient against chosen-keyword attacks (ACKA).
2. We have formed the initial SSE schema which is dynamic, ACKA secure and gains ideal searching time. It is noted that, the proposed formation is secure in the random oracle model in comparison to earlier presented schemas.

3. We have demonstrated the initial deployment and analysis of the SSEDU schema founded on inverted index method. The deployment demonstrates that the proposed schema is largely effective.

4. We have done a performance analysis of the proposed schema which shows the progressive cost of increasing confidentiality to the searchable storage system of cloud.

DETAILS OF ALGORITHM DESIGN

Notations

The collection of total strings (binary) of length l is defined as {0,1}l, and the collection of all countable strings (binary) as {0,1}#. The symbol [l] expresses the collection of integers {1,…….....,l }. We write z←Z to express a member z being taken from a group Z, and z < $ Z to express a member x being taken uniformly and randomly from the collection Z. The result x of a probable algorithm PA is expressed from z ←PA and the actual algorithm PB by z: =PB. Provided a sequencing of members w we refer to its jth member either as wi or w[i] or to its all quantity of members as #u. If C is a collection then #C refers to its cardinality. U expresses the universe of words. If d = (u1,…….,um) ϵ Um is a document, then #d expresses its total quantity of words and |d| is its length of bit. Also, d- is the document which outputs from deleting all identical elements from d (i.e., d- includes only the non identical words in d pipelined as per the order in which they comes in d). If t is a string then |t| expresses to its bit length. We have expressed the addition of n strings t1,……., tn by (t1,…… tn).

Different data structures like arrays, linked lists and dictionaries. If J is a list then #J defines the total quantity of links. If Ar is an array then #Ar is its total quantity of boxes, Ar[i] is the value saved at place i ϵ [#Ar] and Ar[i] := w defines the action which saves w at place i in Ar. The associative array (dictionary or key value) is a data structure which saves key-value twins (t, w). If the twin (t, w) is in T, then T[s] is the quantity w linked with t. T[t] := w defines the action which saves the quantity w under searching key s in Ts and #Ts is the quantity of twins in T.

j ϵ N will define the parameter of security and estimate that whole algorithms receipts j essentially as input. A function w: N → N is minute in j if for each absolute polynomial p(.) and big enough j, w(j) < 1/p(j). d (j) = poly (j) shows that there is a polynomial p(.) i.e.

d(j)<= p(j) for whole big enough jϵN alternatively d(j) <= negl(j) shows that there is a minute function w(.) i.e d(j)<=w(j) for whole big enough j. Z and Z' are alike computationally indistinguishable if for whole PPT (probabilistic polynomial-time) differentiators D, |Pr [D(Z) = 1] - Pr [D(Z') = 1] | <= negl(j).

Primitives of Cryptography

SKE = (gen; enc; dec) is the collection of 3 polynomial algorithms of time used for a private-key encryption schema where gen is the probability algorithm which receipts an argument j for security and provides SK, a secret key, enc is also a probability algorithm which receipts the key SK and the message ms provides a cipher text cp, dec is the deterministic algorithm which receipts the key SK and a cipher text cp and provides ms if SK was the key inside that cp was generated. Particularly the encryption (private-key) schema is secure(CPA) if the cipher text whose results does not disclose any information(partial) regarding the plaintext even to the intruder which could adaptively query the oracle of encryption. With the encryption schema, PRF (pseudo-random functions) and PRP (permutations), those are timing computable (polynomial) functions and could not be differentiated from functions (random) by any probable polynomial timing intruder.

Definitions

A searchable encryption gives the permission to the client for encrypting text in such a way that this could further derive search token for sending as queries to the server. Provided a search token, a server could find over the encrypted text and output the required encrypted documents.

The data could be seen as a series of n documents d= (d1,……..,dn), where document di is a series of words (u1,…….,um) from the world U. Suppose that every document has a particular identifier id(d_i). The text is dynamic; hence at any particular time a document can be removed or added. The document may have text or any other type for that there is an algorithm available which maps every file with a document of keywords from U. Provided a keyword u and we express by du the collection of documents in d which includes u. If $c_p = (c_{p1},\ldots\ldots\ldots\ldots,c_{pn})$ is a collection of encryptions of documents in d, then cpu referring to cipher texts those are encryptions of documents in d_u.

The drawback of all known SSE schemes is that the tokens produced by them can be determined, so that the particular token will be produced always for the particular keyword. Hence searching leaks statistical data about searching pattern of user. At present, it could not be said about the construction of effective SSE schemes with probable trapdoors.

As we are considering SSE as dynamic, hence the schema must give permission for the deletion and addition of documents. Tokens are used to apply these operations. For addition of a document, the client produces the addition token Ta and with provided Ta and y, the encrypted index updated by the provider. In the similar way, for deletion of a document d, the delete token Td, is produced by the client, which is used by the provider for updating the y.

Def.-1–SSEDU (Symmetric Searchable Encryption with Dynamic Updation)

The SSEDU scheme is the aggregation of 9 polynomial-time algorithms. SSEDU = (gen, enc, search token, add token, deltoken, srch, add, del, dec):

SK← gen (1) j: It is a probable algorithm which receipts as input j as a parameter of security and provides the SK as secret key.

(y,cp)← enc (SK, d): It is a probable algorithm which receipts as input SK the secret key and a series of documents d. Further it provides a series of cipher texts cp and the encrypted index y.

Ts←searchtoken (SK,u): It is a probable (possibly) algorithm which receipts as input a secret key SK ,a keyword u. It provides a search token Ts.

(Ta, cpd) ← addtoken (SK, d): is a probable (possibly) algorithm which receipts as input a secret key SK and a document d. It provides a cipher text cpd and addtoken Ta.

Td←deltoken (SK, d): It is a probable (possibly) algorithm which receipts as input SK a secret key and a document d. It provides a delete token Td.

i_d := srch(y,cp,Ts): It is a determined algorithm which receipts as input an encrypted index y, a series of cipher texts cp and a search token Ts. It provides a series of identifiers Id.

(Y', cp') := add(y,cp,Ta,cp): It is a determined algorithm which receipts as input an encrypted index y, a series of cipher texts cp, an add token Ta and a cipher text cp. It provides a new encrypted index y' and series of cipher texts cp'.

(y',cp') := del(y,cp,Td): It is a determined algorithm which receipts as input an encrypted index y , a series of cipher texts cp, and a delete token Td. It provides a new encrypted index y'and new sequence of cipher texts cp'.

d := dec(SK, cp): It is a determined algorithm which receipts as input a secret key SK and a cipher text cp and provides a document d.

The SSEDU schema is right if for all j∈N, for all keys SK produced by gen (1)j, for whole d, for all (y, cp) resulted by enc(SK, d), and for whole series of deletion, addition or searching on y, searching produces the right collection of indices always.

Particularly, the guarantee of security, we want from the SSEDU schema is that given an encrypted index and a series of cipher texts cp, any intruder could not know any data partially regarding the documents d and along with it, a series of tokens T = (T1,…, Tn) for an adaptively produced series of queries q = (q1,……..qn) (for the addition, deletion or searching process), no intruder could know any data partially regarding either q or d. The intuition is very tedious to get exactly and the most known non-interactive and effective SSE schemes disclose the searching and accessing patterns. Hence it is required to weak the definition relevantly by providing few information regarding the queries and messages to be disclosed to the third party. For applying this, we should follow some method and partially fulfill the proposed definition with a collection of leaking functions which is capturing particularly what is to be leaked through the tokens and cipher text.

The other issue regarding to security of SSE is that the schema is secure from ACKA (adaptive chosen-keyword attacks) or only from NACKA (non-adaptive chosen keyword attacks). ACKA guarantees about security even at the time queries of client's are based on output of previous queries and encrypted index. The NACKA only guarantees about security if the queries of client's are non dependent of the previous output and index.

We will expand the relevancy of ACKA security with dynamic operation in the following definition.

Def.-2-(ACKA security with dynamic operation) Suppose SSE = (gen, enc, searchtoken, addtoken, deltoken, srch, add, del, dec) is a index-based SSE (dynamic) schema and considering the probable experiments in the next step, here A is the intruder, S is the simulator and L1, L2, L3 and L4 are the algorithms of leakage:

RealA (j): challenger is running gen (1) j to produce a key SK. A results d and gets (y,c) such that (y, cp)← encSK(d) through the challenger. The third party forms a number (polynomial) of queries (adaptive) {u, d1, d2} and, for every query q, gets through the challenger is either a Ts (search token) such that Ts←srchtokenK (w), a token of addition and pair of cipher text (Ta,cd1) such that (Ta,cd1)←addtokenK(d1) or a token of deletion Td such that Td←deltokenK(d2). At the end, A outputs the bit b which is resulted from the experiment.

IdealA, S (j): A results d. Provided L1 (d), S produces and forwards a twin(y, cp) to A. The intruder forms a polynomial value of queries (adaptive) q∈{u,d1,d2}and, for every query q, the simulator is provided either L2(d, u), L3(d ,d1) or L4(d,d2). The simulator outputs the relevant token T and, in case of an operation of addition, a cipher text cp. At the end, A outputs the bit b which is result from the experiment.

SSEDU is (L1, L2, L3, L4) safe from ACDCK (adaptive dynamic chosen-keyword attacks) if for whole PPT intruders A, there is a PPT simulator (S) such that

| Pr [RealA(j) = 1] - Pr [Ideal A,S(j) = 1] | <= negl(j):

It is to be noted that after the inclusion of dynamic operations the gap between our definitions and the existing definitions are stylistic further we deploy functions of leakage in the style also.

CONCLUSION

It is finally concluded that searchable encryption is a significant cryptography primitive which is highly impacted by the popularization of cloud services such as Microsoft skydrive, Apple icloud, Dropbox, etc. and infrastructures like Amazon S3 and Microsoft Azure public cloud storage. The SSE schema, however, should satisfy some properties like adaptive security, sublinear (optimal) search, compactness as well as the ability to supporting adding and deleting of documents. The initial SSE formation has achieved all the said properties. Along with we have deployed our schema and analyzed its effectiveness. The experiments demonstrated that the formation is very efficient and implementation ready

REFERENCES

[1] Narendra K. and Narsimhareddy Gkv. , "Dynamic Multi-Keyword Ranking Scheme on Encrypted Cloud Data", International Journal of Innovative Technology and Research, Volume No.4, Issue No.4, June – July 2016.

[2] Gomathi M. and Seenivasan D., "Dynamic multi-keyword rank scheme using Top key over encrypted cloud data", International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 04 | April-2016.

[3] Narayankar Ajaykumar, Rathod Gajanan, Londhe Sanket , Wankhade Ashish and Ansari M.A., "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 2, February 2016.

[4] Neeshima P.P., Hegde Pavitra Shankar, P. Poojashree and Pallavi G.B, "A multi keyword ranked search technique with provision for dynamic update of encrypted documents in cloud", International Journal of Computer Engineering and Applications, Volume X, Issue III, March 16.

[5] Karthick K.S. and Deepa P , "A Secure and Dynamic Multi-keyword Ranking Search On Encrypted Cloud Data using GDFS", International Journal on Advanced Computer Theory and Engineering (IJACTE), Volume -5, Issue -2, 2016

[6] HARIKA HAMPI K. S., LAKSHMI K. and PREM KUMAR S., "A Secure and Dynamic Multi Keyword Ranked Search Scheme Over Encrypted Cloud Data", International Journal of Innovative Technologies, Volume.04, Issue No.08, July-2016, Pages: 1406-1411

[7] Saravanan K.S.and Karthika S., "A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 2, February 2016

[8] Metkari Siddheshwar S. and Sonkamble S.B., "Multi-keyword Ranked Search Over Encrypted Cloud Data Supporting Synonym Query", International Journal of Science and Research (IJSR), Volume 5 Issue 6, June 2016

[9] Jain Purva and Banubakode Abhijit, "A Review Paper on Multi keyword Ranked Search on Encrypted Cloud Data", IOSR Journal of Computer Engineering (IOSR-JCE), PP 28-32, 2015

[10] Xia Zhihua, Wang Xinhua, Sun Xingming and Wang Qian , "A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE Transactions on Parallel and Distributed Systems, Vol. 1, p.p.1-13, 2015

[11] Strizhov Mikhail, "Towards a Practical and Efficient Search over Encrypted Data in the Cloud", IEEE International Conference on Cloud Engineering, Vol. 15, p.p. 496-498, 2015

[12] Chen Chi, Zhu Xiaojie, Shen Peisong, Hu J., Guo S., Tari Z.and Zomaya Albert Y. "An Efficient Privacy Preserving Ranked Keyword Search Method", IEEE Transactions on Parallel and Distributed Systems, Vol. 1, p.p. 1-14,2015

[13] Chen Chi, Zhu Xiaojie, Shen Peisong, Hu J., Guo S., Tari Z and Zomaya Albert Y. "An Efficient Privacy Preserving Ranked Keyword Search Method", IEEE Transactions on Parallel and Distributed Systems, Vol. 1, p.p. 1-14, 2015

[14] Cash D., Jarecki S., Julta C., Krawczyk H., Rosu M. –C. and Steiner M. "Dynamic Searchable encryption in very large databases: Data structures and implementation", in Proc of NDSS, vol. 14, 2014

[15] Wang B., Yu S., Lou W. and Hou Y.T. "Privacy preserving multi-keyword fuzzy search over encrypted data in the cloud", in IEEE INFOCOM, 2014.

[16] Zhao Ruihui, Li Hongwei, Yang Yi and Liang Yu, "Privacy-preserving Personalized Search over Encrypted Cloud Data Supporting Multi-keyword Ranking", Sixth International Conference on Wireless Communications and Signal Processing (WCSP), Vol. 14, p.p. 1-6,2014

[17] Ren Yanzhi, Chen Yingying, Yang Jie, Xie Bin, "Privacy-preserving Ranked Multi-Keyword Search Leveraging Polynomial Function in Cloud Computing", Vol. 15,p.p. 594-600, 2014

[18] Sun Wenhai, Wang Bing, Cao Ning, Li Ming, Lou Wenjing, Hou Y. Thomas, Fellow and Hui Li, "Verifiable Privacy Preserving Multi Keyword Text Search in the Cloud Supporting Similarity Based Ranking", IEEE Transactions On Parallel And Distributed Systems, Vol. 25, No. 11, p.p. 3025-3035, November 2014

[19] Ajai Ajni.K. and Rajesh R.S., "Hierarchical Multi-keyword Ranked Search for Secured Document Retrieval in Public Clouds", International Conference on Communication and Network Technologies (ICCNT), Vol. 14, p.p. 33-37, 2014

[20] Zhang Wei, Xiao Sheng, Lin Yaping, Zhou Ting and Zhou Siwang, “Secure Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing”, 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Vol. 14, p.p. 276-286, 2014

[21] Wang Qinqin, Zhu Yanqin and Luo Xizhao,” Multi-user Searchable Encryption with Fine-Grained Access Control without Key Sharing”, 3rd International Conference on Advanced Computer Science Applications and Technologies, Vol. 15,p.p.145-150,2014

[22] Cao Ning , Wang Cong, Li Ming, Ren Kui and Lou Wenjing, “Privacy - Preserving Multi Keyword Ranked Search over Encrypted Cloud Data”, IEEE Transactions On Parallel And Distributed Systems, Vol. 25, No. 1, p.p. 222-333, January 2014.