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 4, April 2017

Data Security Analysis Using Probabilistic OPE on Cloud Search

Ashlesha Khatke[1], Neha Sharma[1], Sachin Kade[1], Sofia Sayyad[1], Prof.Sonali Patil[2]

B.E. Students, Department of Computer Engineering, JSPM's BSIOTR, Wagholi, Pune, India[1]

Assistant Professor, Department of Computer Engineering, JSPM's BSIOTR, Wagholi, Pune, India[2]

ABSTRACT: The data on cloud computing is encrypted because to security concern or the factor of malicious user. As the consequence to this, the search over encrypted data becomes a challenging task. The conventional method like searching in plain text cannot be apply over encrypted data. Due to this searchable encryption techniques are being used. In searchable encryption techniques the order of relevance must be consider as the concern as when it is large amount of data it becomes complicated as relevant documents are more in number. Probabilistic OPE technique is also known as one-to-many OPE. The outcome is that cloud server cannot alter in trusted user data and produce the search on encrypted data and results will appear in order of relevance score. When having some background information on the outsourced documents, the cloud server can correctly interpret the encrypted keywords using the predicated distributions. Hence, sensitive data have to be encrypted before being outsourced to a commercial public cloud. Thus, through the frequency score we can fetch the file from cloud.

KEYWORDS: Searchable Encryption, Order Preserving Encryption, Cloud Computing, Security

I. INTRODUCTION

Now a day around everything is moving to cloud. Cloud has been the most expandable and cost efficient way to store our data .Uploading the data on the cloud is very easy process. As cloud preferable because the data is increasing so fast and there is problem of storage on personal system and organization. Thus if sensitive data such as chemical formulae, combinations, patents, health records of patient, bank statements, password etc. are store in cloud and so it may be hacked by the malicious user so it is very necessary to deal with the security of the system. Downloading all the data from the cloud and decrypt it is not feasible, since the information mending is a complex task through encrypted domain as quantity of the outsourced file can be huge and so the traditional search method cannot decrypt the cipher text directly. To avoid this Searchable Encryption (SE)[1] was proposed to do query in the encrypted domain through the trusted user. As there are many limitation of searchable encryption.

To deal with the limitation of searchable encryption, we are using order preserving encryption which raise the ability of the ranked search. This algorithm was proposed in 2004 to determine the encrypted query problem in the database system. The order-preserving property means that if the plaintexts $x1 < x2$, then the equivalent ciphertexts $E(x1)$ and $E(x2)$ satisfy $E(x1) < E(x2)$. The security definition and the according to [2] it state that OPE is deterministic encryption schema state that plaintext will be encrypted as same ciphertext. Hence the dissemination of the plaintext will be same as of ciphertext thus the security of the data is not achieved properly. For example the in privacy-preserving keyword search ,OPE is used to encrypt relevance scores in the inverted index [3].As noted by Wang et al.[2] when using deterministic OPE ,the resulting ciphertext gives the accurately the same dissemination as the relevance score ,by which the server can indicate the keywords[4].Wang et al.[2] revised the OPE in [5] and suggested a "Probabilistic OPE " in this secure keyword search schema were eliminated and they overcome the distribution of the plaintexts.

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 4, April 2017

II. RELATED WORK

In [10] author used secure multi-keyword ranked search on the encrypted data which concurrently support dynamic update procedure. The vector space model and the mostly used TF x IDF model are united in index and query generation. The "Greedy Depth-first Search" algorithm is used to provide multi-keyword ranked search. The use of special tree-based index structure, the described theory can achieve sub-linear search time and deal with dynamic operation efficiently. Author meets challenging conditions where the outsourced dataset can be shared by many authorized users and searched by many users. Attribute-based Encryption (ABE) here first attribute-based keyword search scheme with user revocation that enable file-level search authorization. This theory allows multiple owners to encrypt and outsource the data to cloud server individually. In [2] author defined and solves the problem of Secure ranked search on encrypted cloud data. The author proposed Order Preserving Encryption (OPE) technique to support search process ranked manner [1]. The relevance score and inverted index are secured with the Order Preserving Encryption (OPE). The distribution and index differences are utilized to estimate the search keyword in differential attacks. Someone introduces new schemes for confidentiality preserving rank-ordered search and regeneration over large document collections [4]. The described scheme not only assures document/query confidentiality against an unauthorized user, but also prevents an untrusted data center from learning information about the query and the document collection. In [5] we propose a security notion in the spirit of pseudorandom functions (PRFs) and related primitives asking that an OPE scheme look "as-random-as-possible" subject to the order-preserving constraint. We then design an efficient OPE scheme and prove its security under our notion based on pseudo randomness of an underlying block cipher. Here we aim at constructing efficient and programmable OPE framework for outsourced database and implementation detail including how to use our OPE framework in database application [8]. Author reports challenging open issues by stating and enforcing access policies based on data attributes and gives the data owner alternate task involved in fine-grained data access control to untrusted cloud server without exposing the data content [9]. This is a survey of the different security risks that pose threat to the cloud ispresented [11]. In this system, we explain and solve the interesting problem of privacy preserving multi keywords ranked search over encrypted cloud data, and create a set of strict privacy necessities for such a safe cloud data application system to be effected in real. We analyze the security of the OP E encryption scheme SE, m, n and give the upper bound on the probability for the adversary to recover the plain text encrypted by SE, m, n under chosen plain text attacks [12].

III. PROPOSED ALGORITHM

A. *OPE Algorithm:*

1. The Binary Search based OPE algorithm take input as secret key K which is generated by using randomized key generation algorithm, plane textspace as $D = \{0,1,2..M\}$and cipher-text space as $R = \{1,2,...N\}$and m as plain text.
2. The algorithm at first maps the middle range gap "y" to a domain gap. y is the gap between the middle two range point.
3. Random coins are generated by using random coin generator i.e. *TapeGen*() algorithm that takes the input as K, D, R and y.
4. Hyper geometric probability distribution function is used to generate x.
5. If the input domain point m that represents the plain text is less than or equal to domain gap x that represents the cipher text that is generated from above steps then the algorithm recurses on the lower i.e. respective upper half of the range and the lower i.e. respective upper part of the domain.

B. *Description of the Proposed Algorithm :*

The objective of both OPE and One-to-Many OPE is to avoid information leakage to the cloud server. If a deterministic OPE is used to encrypt relevance scores, the cipher texts will provide the same dissemination as its plaintext, by which the server can identify the keywords. To overcome this limitation of deterministic OPE, the above

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 4, April 2017

OPE deterministic algorithm is modified as One-to-Many OPE. For m i.e. plaintext relevance score, the "One-to-Many-OPE" first executes above deterministic OPE algorithm to select a bucket for "m", and then randomly determine a value in the bucket as the cipher text. The randomly determination operation in the bucket is succeeded by the unique file IDs together with the plaintext m, and thus the same relevance score in the inverted Index will be encrypted as different cipher texts. Below In [6,7], comparison between plaintext distribution and cipher text distribution obtained by two kindsofOPEona particular keyword.

C. *Mathematical Model:*

System = S:

S = {*I, P, O*}

Success condition:

User will get the file which he/she has requested. And file user get decryption key to decrypt the encrypted file.

Failure Condition:

User will not get the decryption key we can say user can not access the file.

Input :

I=Input

O = Output

P = Processing

I = {F, K}

F = Plain text Files

K = K is the set of keywords / trapdoors, which will be used to find the Files.

O = { F}

F = Searched Files from encrypted data.

P = {*O,F,C,W,T*}

O: the data owner collection, denoted as a set of *m* data owners

$O = (O_1, O_2, \ldots, O_m)$.

F : the plaintext file collection of *O*, denoted as a set of *n* data file

$F=(F_1, F_2, \ldots, F_n)$.

C: the ciphertext file collection of *F*, denoted as

$C=(C_1, C_2, \ldots, C_n)$.

W: the keyword collection, denoted as a set of *u* keywords

$W = (w_1, w_2, \ldots, w_u)$.

Wi: *Oi*'s encrypted keyword collection of *W*, denoted as

$W_i = (b_{wi1}, b_{wi2}, \ldots, b_{wiu})$.

$_fW$: the subset of *W* which represents queried keywords, denoted as

$_fW = (w_1, w_2, \ldots, w_q)$.

- T_fW: the trapdoor for $_fW$, denoted as

$T_fW = (Tw_1, Tw_2, \ldots, Tw^q)$.

- *Si;j;t*: the relevance score of t^{th} keyword to j^{th} file of i^{th}data owner.

The data Owner "*O* "Uploads the plaintext file F.

Where ;

Data Owner is:$O= (O1, O2, \ldots, Om)$.

Files are : $F= F=(F_1, F_2, \ldots, F_n)$.

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 4, April 2017

This Plain text files are getting Encrypted and give the encrypted files are as;

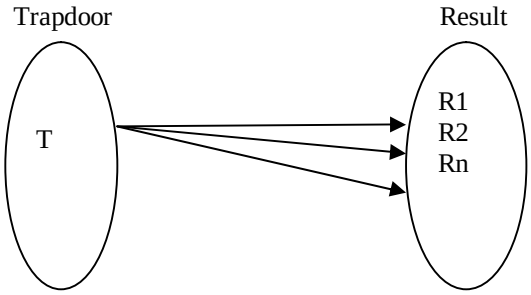
Cipher text files $C=(C_1, C_2, \ldots, C_n)$.

At the same time from these files the Index file is created and uploaded to the application server.

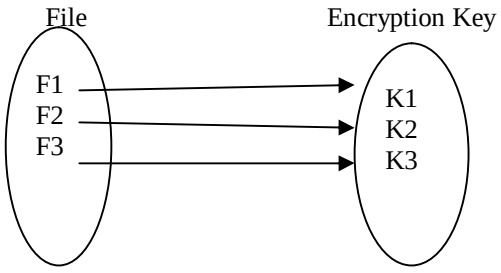
$Ind = \{Ind_F \mid Ind_{F1}, Ind_{F2} \ldots Ind_{Fn}\}$

After the file has been uploaded, the Data owner exchange the key with the data user, he wants to access these file.The Data user enters the trapdoor to the system, The trapdoors are nothing but the keywords set that can be searched in the encrypted file on cloud.The trapdoor is first submitted to the Application Server then application server will find the required file with the help of index file an Application server and the Semi trusted Cloud Server.If the trapdoor matched with the index file then, the files are returned to the data user. Data user with keys exchanged from the Data owner has to decrypt and use those files.

Relation between trapdoor and result is one to many some time one to one (if match)



Each file f1, f2, f3, only one secret key so relation between file and encrypted key is one to one



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

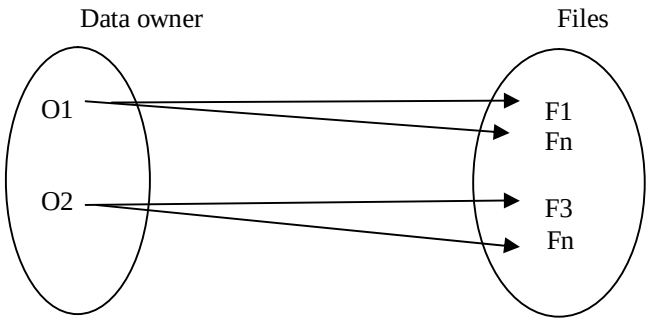
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Each data owner can upload many file so relation between data owner and file is one to many



IV.SIMULATION RESULT

In the existing system the searching time required for any file search was more than the proposed system. Previous they were used to Fuzzy Search, Index-based Search in which we have mention file content, and if it was incorrect the user was not used to get the required data. But by OPE Search user can just mention the file name and can get required data easily. Fig 1.represents the result of the cloud data search file searching takes less than 10seconds.System provide multiple security using cipher text key for avoiding hacking. It provide multiple file uploading

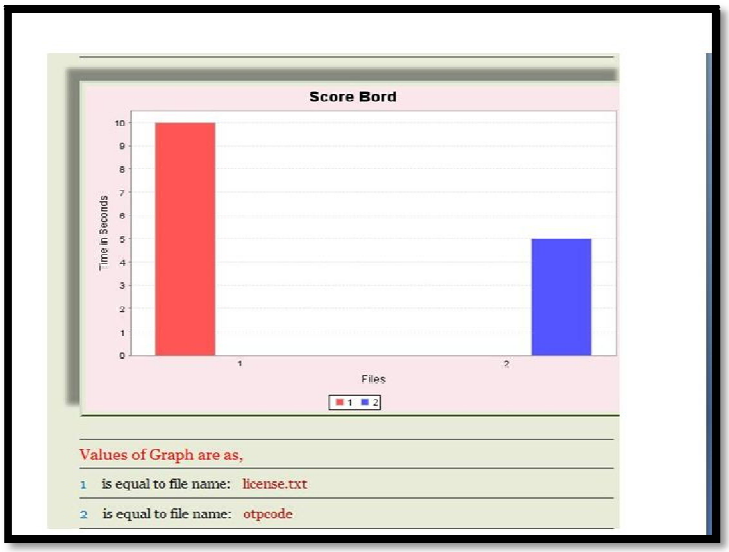


Fig .1. Result of Cloud Data Search

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 4, April 2017

V. CONCLUSION AND FUTURE WORK

One-to-Many OPE is constructed for ranked search of encrypted data over the cloud and to maintain the order of relevance scores and mask their dissemination. As contributed in [6][7] it is seen that cloud server can evaluate the dissemination of relevance scores by change point analysis on the differences of cipher texts of One-to-Many OPE. In future work the author has described to improve one-to- Many OPE in two ways. One way is to divide the plaintext into many sets and divide the analogous bucket into many sub-buckets by which some new change points will appear in the differential attack, which will cover up the original distribution of plaintexts. Another way is to add noise in the inverted index by adding some dummy documents IDs and keywords.

REFERENCES

1. Protecting Data and Query against Differential Attacks in Outsourced Cloud Search: International Journal On Engineering Technology andSciences – IJETS™ ISSN(P): 2349-3968, ISSN (O): 2349-3976 Volume III, Issue VII, July- 2016
2. C.Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data,"IEEETrans.ParallelDistrib. Syst., vol.23, no.8, pp. 1467-1479, Aug. 2012.
3. A. Boldyreva, N. Chenette and A. ONeill, "Order-preserving encryption revisited: improved security analysis and alternative solutions," Advances in CryptologyCCRYPTO, 2011. Springer Berlin Heidelberg, pp. 578-595, 2011
4. Swaminathan, Y. Mao and G.-M Su, "Confidentiality-preserving rank-ordered search," Proceedings of the 2007 ACM workshop on Storage security and survivability. ACM, pp. 7-12, 2007.
5. A. Boldyreva, N. Chenette and Y. Lee , "Order-preserving symmetric encryption," Advances in Cryptology-EUROCRYPT, 2009. Springer Berlin Heidelberg, pp. 224-241, 2009.
6. Ke Li, Weiming Zhang, Ce Yang, and Nenghai Yu "Security Analysis on One-to-Many Order Preserving Encryption- Based Cloud Data Search" In IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL.10 NO. 9, SEPTEMBER 2015
7. AshleshaKhatke,NehaSharma,SachinKhade,SayyadSofia,SonaliPatil,``Security Analysis Using Probabilistic OPE Based Cloud Searching" IJIRCCE Vol.4,Issues 9,September 2016
8. HarshaliAnantAgutale,`` A Survey and Security Analysis on One-To-Many Order Preserving Technique on Cloud Data",International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 11
9. S. Yu, C. Wang and K. Ren, "Achieving secure, scalable, and fine grained data access control in cloud computing," *INFOCOM, 2010Proceedings IEEE*. IEEE, pp. 1-9, 2010.
10. A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data :Zhihua Xia, Member, IEEE, XinhuiWang,Xingming Sun, Senior Member, IEEE, and Qian Wang, Member, IEEE, 2015
11. N Cao, C. Wang, M. Li,K.Ren, and W. Lou, "Privacy preserving keyword ranked search over encrypted cloud data," IEEE Trans. Parallel Distrib. Syst., vol, 25, no. 1, pp. 222-223, Jan. 2014.
12. L. Xiao, I.-L Yen, "Security analysis for order preserving encryption schemes," *Proc. of 46th Annual Conference on Information Sciences and System*, pp. 1-6, 2012.
13. DhamaleSwapnali , BagulSonali , DhadgeMadhuri , Garad Priyanka, Prof.Sonali A. Patil,``A Survey on Efficient Data Integrity Checking with Group User Revocation in Cloud "IJIRCCE Vol.4,Issues 9,September 2016