

A Modified Approach on Security Analysis Using Probabilistic Order Preserving Encryption Based on Cloud Data Serach

Ingale Ashwini Nagnath¹, Kanherkar Harshali Balaso², Kanse Sonali Dilip³, Naykinde Pallavi Babasaheb⁴, Prof. Belsare P.P.⁵

¹Pune University, Department of Computer Engineering,
 S.B.Patil Collage of Engg, Vangali, Indapur 413106, INDIA
 ashwini10ingale@gmail.com

²Pune University, Department of Computer Engineering,
 S.B.Patil Collage of Engg, Vangali, Indapur 413106, INDIA
 kanherkarh@gmail.com

³Pune University, Department of Computer Engineering,
 S.B.Patil Collage of Engg, Vangali, Indapur 413106, INDIA
 kansesonali1995@gmail.com

⁴Pune University, Department of Computer Engineering,
 S.B.Patil Collage of Engg, Vangali, Indapur 413106, INDIA
 naykindepallavi28@gmail.com

⁵Pune University, Department of Computer Engineering,
 S.B.Patil Collage of Engg, Vangali, Indapur 413106, INDIA
 pritambelsare@gmail.com

Abstract: Order preserving encryption (OPE) is an efficient tool to encrypt relevance score of the inverted index for ranked search in encrypted cloud data. From last few years cloud computing has a efficiently growing. So security is major issue for cloud computing .Existing system is based on searchable encryption (SE) with the popularity in cloud computing security and data. Considering the large number of data users and documents in cloud it is necessary for the search service to allow multi keyword query and provide results based on similarity ranking to meet effective data retrieval. In this proposed system to design secure methods of probability order preserving encryption and schemes for search in encrypted data.

Keywords: Binary Search, Cloud Computing, Multikeyword Search, Probabilistic Order Preserving Encryption.

1. Introduction

Now a days Users linked to the internet may store their personal data such as photos,vedio,e-mail,documents,web browse cache etc and let the server monitor or process their data .They can enjoy convenient and most effective service without paying too much money and time ,as one of the most gorgeouse feature of cloud computing is the However, no matter how advantageouse cloud computing may sound , lot of people still worry about Defence theof this technology .if cloud server get direct access to all these user data ,it may try to scrutinize the document s to get the private information .the initial aim of this action may kind .The server want to provide better service by digging into these data and then displaying customer-oriented Advertisement ,which could be appropriate but also aggravating .Beside ,when we consider the sensitive data such as personal data health information, photos, financial documents the situation become more grave theoretically, the server is not supposed to have access to sensitive data at all; therefore we should ensure the server has no given access to leaking this sensitive data information to an untrusted third party. Thus, sensitive data have to be encrypted before being outsourced to a commercial public cloud.

2. Searching Model

- 2.1 Plaintext Searching Model
 - 2.1.1 Cipher text Searching Model
 - 2.1.2 Privacy Threat Model

2.1 Plaintext Seraching Model

To realize fast serach the keyword , ID of file and the relevance score are usually organized as an index structure named “Inverted Index”.an example on posting list of the inverted index as shown in TABLE 1.with a complete inverted index ,a server can complete retrival task by simply comparing relevance score in the index which represent the important level of each file for a certain keyword.

Table 1: Example Of Posting List Of The Inverted Index

keyword	W			
file ID	F1	F2	Ffw
Relevance Score	8.6	6.1	7.3

2.1.1 Cipher text Searching Model

Due to distinct background of cloud computing ,unlike traditional plaintext information retrieval ,there are there objects in cloud data retrieval as shown in fig 1:

Three objects are data vendor, cloud server and users.a data vendor can be an individual or a co-orporation I.e ., it is the entity that own the collection of documents $D_c=\{D_1,D_2,\dots,D_{Nd}\}$ that is want to share with a trusted user s . The keyword set is marked as $W=\{w_1,w_2,\dots,w_{Nw}\}$ for security and privacy apprehension ,document have to be encrypted into $I=\{E(D_1),E(D_2),\dots,E(D_{Nd})\}$ before being uploaded to the cloud server .Additionally the plaintext index has to be encrypted into I to Avoid information outflow . The encrypted form of example of the posting list of the inverted index as shown in TABLE 2 in which the keyword w_i is protected by a hash function $hash()$,and the relevance score are encrypted by a encryption scheme $E'()$.

Table 2: Example Of Encrypted Posting List Of The Inverted Index

keyword	hash(W)			
file ID	F1	F2	Ffw
Relevance Score	$E'(8.6)$	$E'(6.1)$	$E'(7.3)$

We can use TABLE 2 as an example to understand how a cloud server conduct a secure search based on an encrypted list.In this search procedure ,a user first generate a serach request i.e.,user query in a secret form –a trapdoor $T(w)$.In this case Trapdoor is just the hash value of concern. Once cloud server reiceve the the trapdoor $T(w)$,it compare it with the hash values of all keyword in the list I ,then the desired document which are corresponding to keyword w are found .Next,the server returns the matched file IDs:F1,F2,.....,Ffw to the user Finally ,the user can download all the encrypted documents based on the given IDs and decrypt them . A required system is supposed to return the document in a ranked order by their relevance with the queried keyword,but using traditional encryption schemes will disorder relevance score s . Therefore in Order preserving encryption (OPE) Is applied to encrypt the relevance score ,which enable the server to quickly perform ranked serach without knowing the plain text relevance scores.

2.1.2 Privacy threat model s

The purpose of both OPE and one-to –many OPE is to avoid information outflow to the cloud server . the cloud server is deliberated as “semi-honest” also called “honest but nosy”.Specifically,the cloud server will not attempt to remove encrypted data file or list from the storage ,and it will also correctly follow the designed protocol specification and excute the procedure faithfully .However it is nosy to handle the stored data and tries to scrutinize the data to learn additional information.When talking about the “honest but nosy”model,usually there are two attack models””Known Ciphertext Model” and “Known background Model”.”Known Ciphertext Model ”assumes that the cloud server can only get access to get access to the encrypted files and the encrypted list.In this model the server can only dig into the ciphertext s without any other background information , and thus security means that the keyword and document

information are strictly and protected there is no indirect way to venture these information. “Known Background Model” is closer to the real-world situation in the cloud application.The cloud server is soupposed to possess more knowledge then what can be accessed in the known Ciphertext Model.it may intentionally collect related statistical information about outsourced document,and with this information the server can infer more sensitive information.

3. Figure

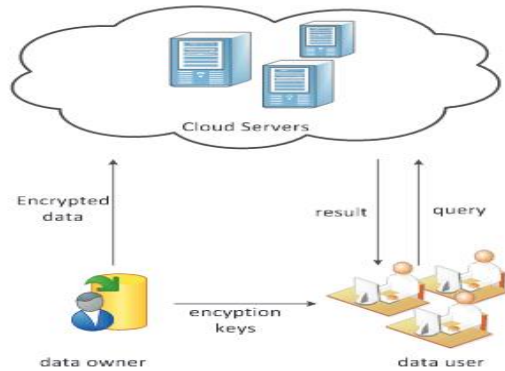


Figure 1: System Architecture

In this paper our There are three main component data owner, remote cloud server and users. A data owner can be an individual or a corporation, i.e., it is the entity that owns a collection of documents $D_c=\{D_1,D_2,\dots,D_{Nd}\}$ that it wants to share with trusted users. The keyword set marked as $W=\{w_1,w_2,\dots,w_{Nw}\}$. For security and privacy concerns, documents have to be encrypted into $E=\{E(D_1),E(D_2),\dots,E(D_{Nd})\}$. before being uploaded to thecloud server. Additionally, the plaintext index has to be en-crypted into I to prevent information leakage.

4. Equations

Input: User query as input.
 Output: Return resulted data depend on query.
 Functions: Encrypt data, validating user, indexing of input query, protecting, searching and returning result data
 Let D_c is the set of document.
 $D_c = \{D_1,D_2,\dots,D_{Nd}\}$
 $W =$ Set of Keywords
 $W = \{fW_1,W_2,\dots,W_{Nw}\}$
 $I =$ Encrypted Information
 $I = \{fE(D_1),E(D_2),\dots,E(D_{Nd})\}$
 secret form $T(W)$ request compare TI returns the match document _le F1,F2.....Ffw
 Success Conditions : when valid user query input.
 Failure Conditions : when invalid user query input

References

[1] N.Cao,C.wang and M.Li,Privacy-preserving multi-Keyword serach over encrypted cloud data,INFOCOM,2011 proceeding IEEE.IEEE,PP,829-837,2011.

- [2] S.Subashini and .Kavita,"A survey on security issues in service delivery,model of cloud computing",Journal of network and computer Application,34(1):1-11,2011.
- [3] Security Analysis on one to many order preserving encryption basd on cloud data serach,Ke Li Weiming zhang,Ce Yang and Nenghai Yu.
- [4] Analysis of Symmetric Serachable Encrypation and Data Retrival in Cloud Computing Vasudha Arora S.S.Tyagi.
- [5] A Secured and high Octane Rank Based Analysis in Cloud Computing Enviroment Poojitha Koneru,Dr.S.Prabakaran Dept of CSE,SRM University,Chennai.
- [6] C.Wang,N.Cao and K.Ren,Enabling secure and efficient ranked keyword serach over out souced cloud data Parallel and Distributed System,IEEE Transaction 23(8),pp.1467-1479,2012.
- [7] 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.