ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)
Vol. 3, Issue 2 (Apr. - Jun. 2015)
International Journal of Advanced Research in Computer Science & Technology (IJARCST 2015)

Concealment of Data in Multi-keyword Ranked Search over Encrypted Cloud

[I]S.Nagapadhmapriyaa, [II]S.Aravinth Kumar
[I,II]Dept. of Computer Science, Galgotias University, Delhi, India

Abstract
Globally, data outsourcing has been experiencing vast growth in commercial cloud – especially in the current economic climate where cost-minimization and data privacy is a big question. Many enterprise owners are turning to public cloud mainly for the reason capital expenses are shifted to operational expenses – pay as you use, but when considering data privacy, it is a big challenging issue.This paper focus on privacy for multi-keyword ranked search over encrypted cloud data (MKRS).To protect the privacy of the data, data owners encrypt the data and then outsource it to the public cloud. And for efficient retrieval of data, the search query has multiple keywords. The earlier paper works are done only for single keyword search over encrypted cloud. Two different attacks are considered on the data and strict privacy is established through improved MKRS schemes.

Keywords
Cloud privacy,Multi-keyword Ranked Search,Verified objects.

I. Introduction

The new era of industry advances in Cloud computing by cutting costs in the areas of capital hardware, data center management, software development and enables a distributed workforce, but learning how the data functions, is essential to ensure secure services for cloud customers. In today's world, we are bombarded with huge data in cloud sector which forces multi-keyword ranked search(MKRS) for relevant data retrieval [2].For the challenging security issue in cloud it is essential to encrypt the data before outsourcing the sensitive data to the untrusted cloud server.The cloud may contain any sensitive data like credit card data, health insurance, passport data, student information etc.Hence the privacy of theses data are to be strictly maintained.The cloud server has the tendency to do cryptanalysis.This paper gives improved MKRS SCHEME based on two different threats.Any leakage of the keyword will lead to drastic result such that even a single query with known keyword may drop the whole database.

For authentication the public key cryptography technique is used in RSA and AES algorithm along with root signature from SHA1. In earlier Data mining paper – Authenticating skyline queries, the data owner builds an authenticated data structure (ADS), tree like index structure where the root is signed by the data owner.Data owner sends data,ADS and the root signature to the cloud server. When search query is shot out, the cloud server returns results along with VO(Verified objects) [4] for result verification.This technique is used in privacy preserving cloud paper along with top-k queries also for ranked result to avoid unnecessary traffic and storage space.

In early paper works single keyword search over encrypted cloud [1] has been done.And the most popular one,Google search does not use encrypted data.Hence this new era demands privacy on multi-keyword ranked search(MKRS) over encrypted cloud.The multi-keyword search query will retrieve the relevant data and the ranking done will avoid network traffic by returning only the most relevant data.So this paper provides stringent security mechanisms for multi-keyword ranked search(MKRS) over encrypted cloud. The overall summary of the paper is as follows:

1. This paper focus on privacy for multi-keyword ranked search over encrypted cloud data (MKRS) coupled with VO (Verified objects) and top-k queries.
2. Literature provides only single keyword search over encrypted cloud and coordinate matching is done only for plain text information retrieval [2].
3. In Fig.1 Data owner first builds index from the data document and then encrypts both data and index using RSA and AES algorithm.The public key is used during encryption and the secret key is used during decryption as in public key cryptography.
4. Encryped data and index is outsourced to the cloud server along with root signature from SHA1,after registration of data owner with the cloud server.
5. The cloud server has the tendency to do cryptanalysis.This can be of two different threats.One if it tries to analyze the encrypted data and index.Another if it knows additional information from the keyword frequency.Identified keywords may result in drastic result.
6. The user has to register with the data owner to get the decryption keys.When search query is initiated from registered users the results are ranked with top-k scheme to avoid unnecessary traffic and storage space.
7. Finally after decryption user generates Verification object using md5 and then compares it with the received Verification object.

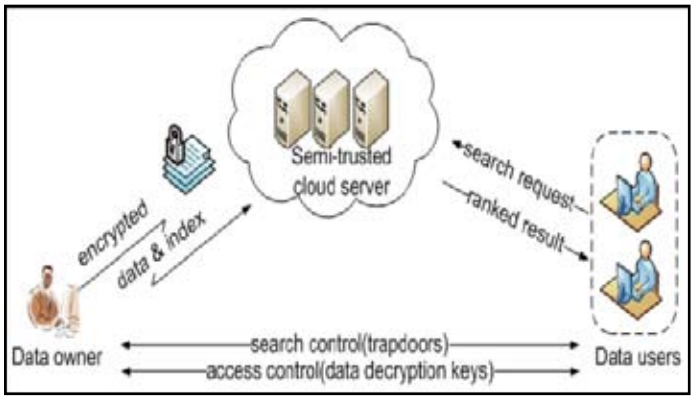

Fig. 1: Multi-keyword Ranked Search Architecture [2]

II. Problem Definition

In this section we present the overall framework and the threats faced.The notations are given in Table1.

A. Framework

From Fig.2 Data owner does authentication request (generate Sk

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

and Pk) and builds index I from the data document C and then encrypts both data and index .The private key is used during encryption and the public key is used during decryption . During index construction , both the data document and search query are considered as binary vectors(bits) for quantitatively measuring the matches.The trapdoor T is generated from set of keywords and then encrypted and sent to cloud server which does score calculation and top-k ranking for retrieving the files.

Table 1

Notation	Meaning
Sk	Secret key
Pk	Public Key
F	File
C	Encrypted File
I	Index
I0	Encrypted Index

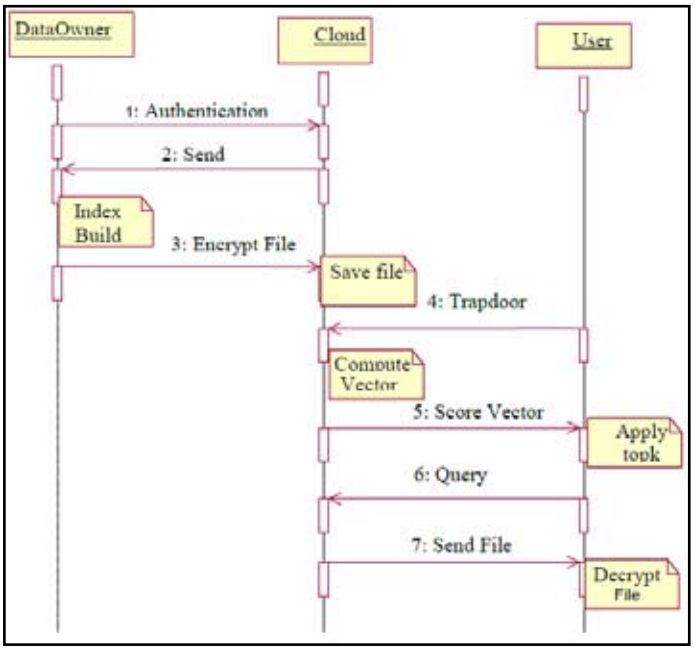

Fig. 2: Processing of Multi-keyword Ranked Search

B. Cryptanalysis

Classic Crypt Attack:
In this the cloud server tries to analyze the encrypted data and encrypted index without keys.
Frequency Analysis Attack:
In this the cloud will have more knowledge other than the encrypted information.It learns more information from keyword frequency and with known trapdoor it breaks all.

C. Cons from existing system:

- Single-keyword search without ranking.[1]
- Boolean- keyword search without ranking.[2]
- Single-keyword search with ranking.

D. Pros of proposed system:

- Concealment of data maintained on enriched multi-keyword search through md5.
- Root signature of SHA1 used for additional integrity.
- RSA with AES used for double security.

III. List of Modules

A. Data owner Initialization

In this module data owner first registers with cloud server and gets the secret key Sk and public key Pk.The public key is used for encryption and secret key is used for decryption.

B. Building Index

- Now the data owner is authenticated to outsource the file C.
- So first build index I from the file and the encrypt both file and index.The extract keyword extracts all keywords from the file and also gives the count of it.
- The data owner can select their own index also.

C. Encryption and File Outsourcing

The file F and index I are encrypted using public key Pk. The RSA key main value passed to AES algorithm.The root signature is also generated using SHA1 algorithm.Then the encrypted index I0 ,encrypted file C and the root signature are sent to cloud server.

D. User registration and Verification

- The user registers with data owner and then does the search process.
- The trapdoor T is generated from the set of keywords and the encrypted using public key.
- Cloud server receives the user encrypted query and sends back the result vector file.
- The user applies top-k and sends it back to cloud.The cloud generates Verification object VO from md5 and send the encrypted file C to the user.
- The user decrypts the file using secret key Sk.
- The user generates Verification object and the compares it with the received Verification object.
- When both are same then the file is authenticated and confidential.

IV. Conclusion

This paper gives security mechanisms for multi-keyword ranked search(MKRS).Mainly for authentication and security purpose the data owner registers with cloud server and similarly the user registers with data owner.The search process includes verification done on user side.This verification technique compares the received Verification object VO and the user-side generated Verification object.If the values are same then the file received is authenticated. Thus privacy of data in encrypted cloud is maintained by coupling with data mining Verification Object.

References

[1] *C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. of ICDCS'10, 2010.*
[2] *Ning Cao†, Cong Wang‡, Ming Li†, Kui Ren‡, and Wenjing Lou†, "Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data" in IEEE transactions on parallel and distributed systems vol:25 no:1 year 2014.*
[3] *C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in Proc. Of INFOCOM, 2010.*

[4] *Xin Lin, Jianliang Xu, Senior Member, IEEE, Haibo Hu, and Wang-Chien Lee, "Authenticating Location-Based Skyline Queries in Arbitrary Subspaces" in IEEE transactions on knowledge and data engineering, vol. 26, no. 6, june 2014.*

[5] *H. Hu, Q. Chen, and J. Xu. "VERDICT: Privacy-preserving authentication of range queries in location-based services," in ICDE, Brisbane, QLD, Australia, 2013*

[6] *Q. Chen, H. Hu, and J. Xu, "Authenticating Top-k queries in location-based services with confidentiality," in PVLDB,Hangzhou, China, 2014.*

[7] *S. Kamara and K. Lauter, "Cryptographic cloud storage," in RLCPS,January 2010, LNCS. Springer, Heidelberg.*