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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

Survey on Phrase Search Scheme over Encrypted Data on Cloud Systems

Komal K. Sharma, Prof. Mrunalinee Patole

M.E. Student, Department of Computer Engineering, RMDSSOE, Pune, Maharashtra, India

Department of Computer Engineering, RMDSSOE, Pune, Maharashtra, India

ABSTRACT: Cloud computing provides computing, storage, services, and applications over the internet. Cloud computing is convenient for on demand access to a shared pool of resources. Before outsourcing the data on cloud, the data should be encrypted to protect privacy of sensitive information. But it is difficult to search over encrypted data. So many researchers have proposed efficient search schemes over encrypted cloud data. All existing schemes uses keywords and semantic words as the document feature. The idea of system, is to search over encrypted data using phrase i.e. with collection of words. Multiple keyword search problems cannot be used to perform phrase search on encrypted documents, because they are unable to determine the positional relationship of the keywords composing a phrase in the encrypted environment. System uses P3, an efficient privacy-preserving phrase search scheme. To determine positional relationship of multiple queried keywords over encrypted data, scheme uses homomorphic encryption and bilinear map. And to protect user search pattern it utilizes a probabilistic trapdoor generation algorithm. This P3 scheme improves search accuracy.

KEYWORDS: Phrase search, encrypted data, artificial intelligence, IoT, cloud

I. INTRODUCTION

The IoT enables articles to be detected or controlled remotely crosswise over existing system foundation, creating opportunities for more direct integration of the physical world into computer-based systems, and resulting in improved efficiency, accuracy and economic benefit in addition to reduced human intervention. The growth of IoT leads to the generation of large amounts of data, which possess massive computing resources, storage space and communication bandwidth. Cloud serves as the brain to effectively transform data to insight and drive productive, cost effective action resulting to improve accuracy of decision making and optimize internet-based interactions. Despite the benefits of the integration of cloud computing and IoT are attractive, cloud computing is not a panacea that can address all the problems in IoT.

Storing and retrieving such a large amount of data consumes lot of time as data in the cloud needs to be always stored in encrypted format while storing and needs to be decrypted while searching. There are a number of propositions for executing queries over encrypted data. This implement the client to encrypt data before outsourcing it to the cloud in a database scheme. To avoid this massive consumption of time, data searching speed can be increased by directly searching over encrypted data in the cloud. There are many methods used for searching the encrypted data over cloud. In keyword-based search schemes ignore the semantic representation information of users retrieval, and cannot completely meet with users search intention. The semantic and multi-keyword searching schemes cannot be used for phrase search. The main challenge is to enable cloud servers to make judgement on whether the keywords occurring in an encrypted document are consecutive or not, without leaking sensitive information.

II. LITERATURE SURVEY

In recent years, many researchers have proposed a series of efficient search schemes over encrypted cloud data. Paper, ‘Practical Techniques for Searches on Encrypted Data’ published by, Dawn Xiaodong Song [10] in this, describe our cryptographic schemes for the problem of searching on encrypted data and provide proofs of security for the resulting crypto systems. Our techniques have a number of crucial advantages. They are provably secure: they provide provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the ciphertext; they provide query isolation for searches, meaning that the untrusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching, so that the untrusted server cannot search for an arbitrary word without the user's authorization; they also support hidden queries, so that the user may ask the untrusted server to search for a secret word without revealing the word to the server.

Paper, 'Evaluating 2-DNF Formulas on Ciphertexts' published by, Dan Boneh [9] in this, homomorphic encryption scheme that supports addition and one multiplication. We require that the values being encrypted lie in a small range as is the case when encrypting bits. These homomorphic properties enable us to evaluate multi-variate polynomials of total degree 2 given the encrypted inputs.

Paper, **'**Secure kNN Computation on Encrypted Databases' published by, W. K. Wong [7] in this, author discuss the general problem of secure computation on an encrypted database and propose a SCONEDB (Secure Computation ON an Encrypted DataBase) model, which captures the execution and security requirements. Author focus on the problem of k-nearest neighbor (kNN) on encrypted datasets.

Paper, 'Fuzzy Keyword Search over Encrypted Data in Cloud Computing' published by, Jin Li [6] in this, author exploit edit distance to quantify keywords similarity and develop leading technique while constructing fuzzy keyword sets, which reduces the storage.

Paper, **'**Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data' published by, Ning Cao [5] in this, author define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted data in cloud computing (MRSE).

Paper, 'Phrase Search over Encrypted Data with Symmetric Encryption Scheme' published by, Yinqi Tang[4] in this, we have further studied the problem of searchable encryption, which solves the dilemma of maintaining the confidentiality of data and the ability for a client to search. We first introduce the model of phrase search with symmetric encryption and its security definition, then propose a construction and its security proof. Lastly, we analyze our scheme and evaluate how it performs when updating.

Paper, 'Semantic-aware Searching over Encrypted Data for Cloud Computing' published by Zhangjie Fu, [2] ,in this to address the problem of semantic retrieval, author propose effective schemes based on concept hierarchy. To improve accuracy, author extend the concept hierarchy to expand the search conditions.

Research paper, 'Secure Phrase Search for Intelligent Processing of Encrypted Data in Cloud-Based IoT' published by Meng Shen, [1] in this paper author proposes P3 scheme to perform phrase search over encrypted data. Scheme exploits homomorphic encryption and bilinear map to determine the pairewise location relationship of queried keywords on the cloud server side. It eliminates need of trusted third party.

III. SYSTEM MODEL

The P3 scheme i.e. privacy-preserving phrase search scheme over encrypted data involves three entities, namely data owner, cloud server and one or multiple user, as shown in fig 1.

Data owner: The data owner generates a secure searchable index for the document set and outsources the secure index along with the encrypted document set to the cloud server.

Data users: The authorized data user makes phrase search over the encrypted documents, user first acquires the corresponding trapdoor from the data owner through the search control mechanism and then submits the trapdoor to the cloud server.

Cloud Server: Using receiving user's trapdoor, the cloud server executes the predesigned search algorithms and relies to the user with the corresponding set of encrypted documents as the search results. And after that, user decrypts the received documents with the help of the data owner.

We resort to the homomorphic encryption and bilinear map, which enables the client to obtain exact search results from a single interaction with the cloud server.

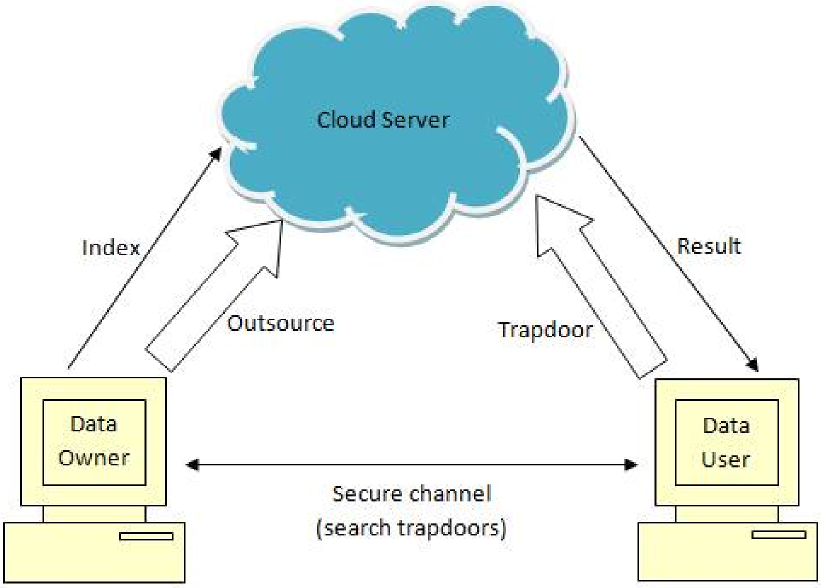

Fig 1. System Model

- Bilinear map:

Let G1, G2, and Gt be cyclic groups of the same order.
Definition : A bilinear map from G1 × G2 to Gt is a function
$e : G1 \times G2 \rightarrow Gt$
such that for all $u \in G1$, $v \in G2$, $a, b \in Z$,
$e(u^a , v^b) = e(u, v)^{ab}$.
Bilinear maps are called pairings because they associate pairs of elements from G1 and G2 with elements in Gt . Note that this definition admits degenerate maps which map everything to the identity of Gt .
Groups with a bilinear map allow us to build public key encryption schemes with new properties that are otherwise difficult to obtain using groups without a bilinear map.

- Homomorphic encryption:

Homomorphic encryption enables “computing with encrypted data” and is hence a useful tool for secure protocols. Homomorphic encryption schemes have many applications, such as protocols for electronic voting schemes, computational private information retrieval (PIR) schemes, and private matching. It is a cryptographic primitive that allows us to perform operations over encrypted data without knowing the secret key or decrypting the data.

IV. SECURE PHRASE SEARCH SCHEME

System overview:
The phrase search procedure can be defined as follows: When the cloud server receives the trapdoor for a specific phrase query from a user, it first locates the inverted lists for the queried keywords, and then finds the documents that contain all of the queried keywords. After that, the cloud server identifies whether the locations of the keywords are consecutive and returns only the relevant documents that contain the exact phrase. The overall workflow of the scheme is as shown in fig 2.

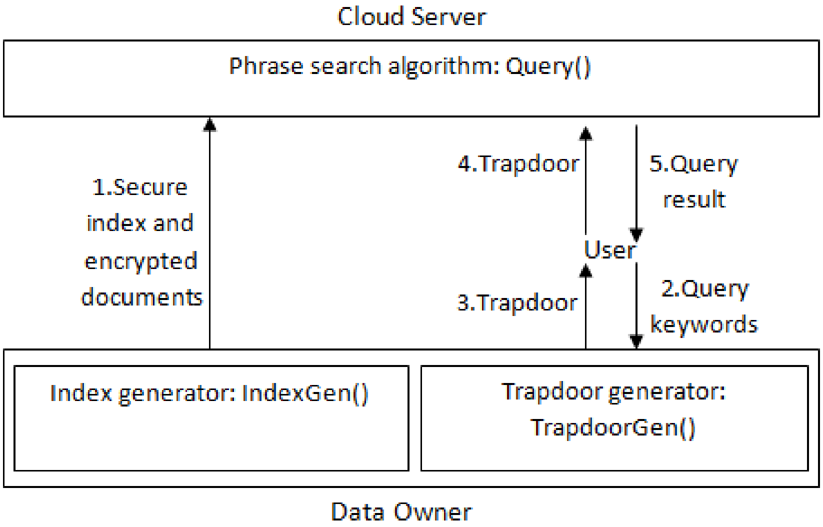
Cloud Server
Phrase search algorithm: Query()
1.Secure index and encrypted documents
4.Trapdoor
5.Query result
User
3.Trapdoor
2.Query keywords
Index generator: IndexGen()
Trapdoor generator: TrapdoorGen()
Data Owner

Fig 2. Workflow of the scheme P3

Fig 2. Shows three modules mainly which are used in scheme are as follows:

- Index Generator: It executed on the data owner side. It takes the documents as the input and outputs the corresponding secure index, as well as the encrypted documents.
- Trapdoor Generator: It also executed on the data owner side. Given a user's queried phrase, it generates the corresponding secure trapdoor and replies to the user.
- Phrase Search Algorithm: It executed on the cloud server side. Upon receiving a trapdoor from a user, it performs a phrase search procedure over the secure index and returns the search results.

V. COMPARATIVE ANALYSIS

Here we analyses many schemes used for searching over encrypted data. But they are not useful for phrase search. Fig 3 shows comparative analysis between various search schemes over encrypted data.

According to fig.3 we see different search schemes used to search over encrypted data in detail as follows:

A. Single keyword search scheme:

This is the method used to search over encrypted data. In this scheme only single word can be search. So this is less useful in future.

B. Multi keyword search scheme:

In this scheme, multiple keywords search can be occur by single query, but the drawback of this scheme is, it is unable to find positional relevance of the words in the encrypted documents.

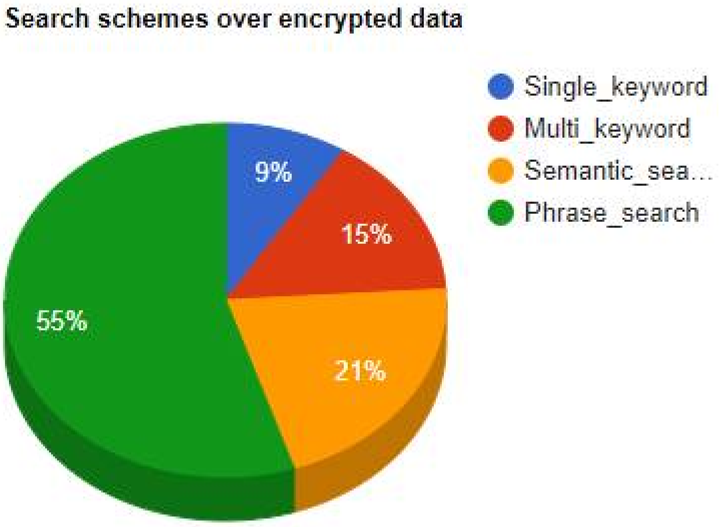

Fig 3. Comparative analysis between various search schemes over encrypted data

C. Semantic search scheme:

Semantic search scheme searches semantic relationship between words. If the words are semantically correct then this scheme gives result. This is fast as compared to overall, but the drawback of this scheme is, it only searches word not phrase.

D. Phrase search scheme:

This is very efficient scheme as compared to others. Phrase search scheme over encrypted data uses P3 scheme which greatly improves the search accuracy.

VI. CONCLUSION

In this survey, firstly we see different schemes used in search over encrypted data on cloud. In phrase search scheme, P3 i.e. privacy-preserving phrase search method is used. This scheme enables search without relying on trusted third-party. In this scheme, to determine pairwise positional relationship of queried keywords on the cloud server side the homomorphic encryption and bilinear map are used. Then we perform comparative analysis on many schemes used for search and after analysis we conclude that the phrase search scheme over encrypted data is the efficient scheme.

REFERENCES

[1] Meng Shen, Baoli Ma, Liehuang Zhu, Xiaojiang Du, Ke Xu, “Secure phrase search for intelligent processing of encrypted data in Cloud-Based IoT” , IEEE Internet of Things Journal (Early Access) , 2018.
[2] Zhangjie Fu, Lili Xia, Xingming Sun, Alex X. Liu, Guowu Xie, “Semantic-aware Searching over Encrypted Data for Cloud Computing”, IEEE Transactions on Information Forensics and Security, 2018.
[3] Komal K. Sharma, Prof. Mrunalinee Patole, “Survey on semantic-aware searching over encrypted data on cloud systems”, International Journal of General Science and Engineering Research, 2018.
[4] Y. Tang, D. Gu, N. Ding, and H. Lu. Phrase search over encrypted data with symmetric encryption scheme. In Workshops of IEEE ICDCS, pages 471–480, June 2012.
[5] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou. Privacy-preserving multikeyword ranked search over encrypted cloud data. In IEEE INFOCOM, pages 829–837, April 2011.
[6]] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, “Fuzzy keyword search over encrypted data in cloud computing,” in Proc. of IEEE INFOCOM’10 Mini-Conference, San Diego, CA, USA, March 2010, pp. 1–5

[7] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure knn computation on encrypted databases," in Proc. of SIGMOD, 2009, pp. 139–152.
[8] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. Searchable symmetric encryption: Improved definitions and efficient constructions. In Proc. of ACM CCS, pages 79–88, New York, NY, USA, 2006. ACM.
[9] D. Boneh, E.-J. Goh, and K. Nissim. Evaluating 2-dnf formulas on ciphertexts. In TCC, pages 325–341. Springer, 2005.
[10] D. X. Song, D. Wagner, and A. Perrig. Practical techniques for searches on encrypted data. In IEEE S&P, pages 44–55, 2000.