Volume No:2, Issue No:6 (November-2016) ISSN No : 2454-423X (Online)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

Privacy Preserving Ranked Multi-Keyword Serach for Multiple Data Owners in Cloud Computing

Ch.Ramesh Kumar
Associate Professor & HOD,
Department of CSE,
Malla Reddy Engineering College
& Management Sciences, Kistapur,
Medchal, Hyderabad.

Y.Ashwini
Assistant Professor,
Department of CSE,
Malla Reddy Engineering College
& Management Sciences, Kistapur,
Medchal, Hyderabad.

Velamala Venkataramana
M.Tech Student,
Department of CSE,
Malla Reddy Engineering College
& Management Sciences, Kistapur,
Medchal, Hyderabad.

ABSTRACT:

The advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in cloud, it is crucial for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely differentiate the search results. In this paper, for the first time, we define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, we choose the efficient principle of “coordinate matching”, i.e., as many matches as possible, to capture the similarity between search query and data documents, and further use “inner product similarity” to quantitatively formalize such principle for similarity measurement. We first propose a basic MRSE scheme using secure inner product computation, and then significantly improve it to meet different privacy requirements in two levels of threat models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset further show proposed schemes indeed introduce low overhead on computation and communication.

KEYWORDS:

Cloud computing, ranked keyword search, several owners, privacy preserving, dynamic hidden key.

INTRODUCTION:

Cloud storage system, is set of storage servers, and provides long-term storage services over the Internet. Storing data in a third party’s cloud system causes grave to connect to over data secret. Normal hidden schemes defend data secret but have some limitation to functionality of the storage system because a few operations are supported over hidden information. Building a grave storage system that compatible several functions is endurance when system is distributed. Service providers of cloud would pledge to owners data security using phenomenon like virtualization and firewalls. These phenomenon’s do not protect owners data privacy from the CSP itself, since the CSP control whole of cloud hardware, software, and owners’ data. Hiding the sensitive data before send outside can stored data confidentiality against CSP. Data hidden makes the conventional data utilization service based on plaintext keyword search a very challenging problem. A solution to this problem is to download all the hidden data and create the original data using the hidden key, but this is not practical cause it create extra overhead In this paper, we suggest when search multiple owner multiple keywords that time provide the privacy and show the result in ranking form to make easy cloud servers to perform safe search excluding knowing the real value of both keywords and trapdoors, we properly build a novel safe search rule. So that various data owners use distinct keys to hide their files and keywords. Genuine data users can get a query excluding knowing confidential keys of these various data owners.

To rank the search results and preserve the privacy of relevance scores between keywords and files, we suggest a family which preserves privacy, which helps the cloud server return the most relevant search results to data users without revealing any sensitive information. To protect from disclosing the result we propose a novel dynamic secret key generation protocol and a new data user authentication rule[1]. The main contributions of this paper are listed as follows:

•We define search data on clued that data is hidden format and also providing the privacy when search the multiple keywords.

•We suggest an capable data user authentication rule, which stop attackers to disclose hidden key and only genuine data user can do search.

•We suggest a approach that performs multiple key word search and rank them properly. We suggest an Additive Order and Privacy Preserving Function family (AOPPF) which allows the cloud server produces the file that rank properly.

•We supervise experiments on real-world Datasets to verify the effectiveness and capability our suggest schemes.

LITERATURE SURVEY

1) A view of cloud computing:

Cloud computing, the long-held dream of computing as a utility, has the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and purchased. Developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. They need not be concerned about overprovisioning for a service whose popularity does not meet their predictions, thus wasting costly resources, or underprovisioning for one that becomes wildly popular, thus missing potential customers and revenue. Moreover, companies with large batch-oriented tasks can get results as quickly as their programs can scale, since using 1,000 servers for one hour costs no more than using one server for 1,000 hours. This elasticity of resources, without paying a premium for large scale, is unprecedented in the history of IT.

2) Privacypreserving public auditing for secure cloud storage:

Using cloud storage, users can remotely store their data and enjoy the on-demand high-quality applications and services from a shared pool of configurable computing resources, without the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the outsourced data makes the data integrity protection in cloud computing a formidable task, especially for users with constrained computing resources. Moreover, users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. Thus, enabling public auditability for cloud storage is of critical importance so that users can resort to a third-party auditor (TPA) to check the integrity of outsourced data and be worry free. To securely introduce an effective TPA, the auditing process should bring in no new vulnerabilities toward user data privacy, and introduce no additional online burden to user. In this paper, we propose a secure cloud storage system supporting privacy-preserving public auditing. We further extend our result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient. Our preliminary experiment conducted on Amazon EC2 instance further demonstrates the fast performance of the design.

3)Practical techniques for searches on encrypted data:

It is desirable to store data on data storage servers such as mail servers and file servers in encrypted form to reduce security and privacy risks. But this usually implies that one has to sacrifice functionality for security. For example, if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query, without loss of data confidentiality. We 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. The algorithms presented are simple, fast (for a document of length n, the encryption and search algorithms only need O(n) stream cipher and block cipher operations), and introduce almost no space and communication overhead, and hence are practical to use today

4)Searchable symmetric encryption: improved definitions and efficient constructions:

Searchable symmetric encryption (SSE) allows a party to outsource the storage of its data to another party (a server) in a private manner, while maintaining the ability to selectively search over it. This problem has been the focus of active research in recent years. In this paper we show two solutions to SSE that simultaneously enjoy the following properties: Both solutions are more efficient than all previous constant-round schemes. In particular, the work performed by the server per returned document is constant as opposed to linear in the size of the data. Both solutions enjoy stronger security guarantees than previous constant-round schemes. In fact, we point out subtle but serious problems with previous notions of security for SSE, and show how to design constructions which avoid these pitfalls.

Further, our second solution also achieves what we call adaptive SSE security, where queries to the server can be chosen adaptively (by the adversary) during the execution of the search; this notion is both important in practice and has not been previously considered. Surprisingly, despite being more secure and more efficient, our SSE schemes are remarkably simple. We consider the simplicity of both solutions as an important step towards the deployment of SSE technologies.As an additional contribution, we also consider multi-user SSE. All prior work on SSE studied the setting where only the owner of the data is capable of submitting search queries. We consider the natural extension where an arbitrary group of parties other than the owner can submit search queries. We formally define SSE in the multi-user setting, and present an efficient construction that achieves better performance than simply using access control mechanisms.

5) Public key encryption with keyword search secure against keyword guessing attacks without random oracle:

The notion of public key encryption with keyword search (PEKS) was put forth by Boneh et al. to enable a server to search from a collection of encrypted emails given a “trapdoor” (i.e., an encrypted keyword) provided by the receiver. The nice property in this scheme allows the server to search for a keyword, given the trapdoor. Hence, the verifier can merely use an untrusted server, which makes this notion very practical. Following Boneh et al.'s work, there have been subsequent works that have been proposed to enhance this notion. Two important notions include the so-called keyword guessing attack and secure channel free, proposed by Byun et al. and Baek et al., respectively. The former realizes the fact that in practice, the space of the keywords used is very limited, while the latter considers the removal of secure channel between the receiver and the server to make PEKS practical. Unfortunately, the existing construction of PEKS secure against keyword guessing attack is only secure under the random oracle model, which does not reflect its security in the real world.

Furthermore, there is no complete definition that captures secure channel free PEKS schemes that are secure against chosen keyword attack, chosen ciphertext attack, and against keyword guessing attacks, even though these notions seem to be the most practical application of PEKS primitives. In this paper, we make the following contributions. First, we define the strongest model of PEKS which is secure channel free and secure against chosen keyword attack, chosen ciphertext attack, and keyword guessing attack. In particular, we present two important security notions namely IND-SCF-CKCA and IND-KGA. The former is to capture an inside adversary, while the latter is to capture an outside adversary. Intuitively, it should be clear that IND-SCF-CKCA captures a more stringent attack compared to IND-KGA. Second, we present a secure channel free PEKS scheme secure without random oracle under the well known assumptions, namely DLP, DBDH, SXDH and truncated q-ABDHE assumption. Our contributions fill the gap in the literature and hence, making the notion of PEKS very practical. We shall highlight that our scheme is IND-SCF-CKCA secure.

Volume No:2, Issue No:6 (November-2016) ISSN No : 2454-423X (Online)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

EXISTING SYSTEM:

•Secure search over encrypted data has recently attracted the interest of many researchers. Song et al. first define and solve the problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that enables users to perform a keyword-based search on an encrypted dataset, just as on a plaintext dataset. Searchable encryption is further developed.

•Secure search over encrypted cloud data is first defined by Wang et al. and further developed. These researches not only reduce the computation and storage cost for secure keyword search over encrypted cloud data, but also enrich the category of search function, including secure ranked multi-keyword search, fuzzy keyword search, and similarity search.

Limitations:

•Existing schemes are concerned mostly with single or boolean keyword search.

•All the existing schemes are limited to the single-owner model. As a matter of fact, most cloud servers in practice do not just serve one data owner; instead, they often support multiple data owners to share the benefits brought by cloud computing.

PROPOSED SYSTEM:

•In this paper, we propose PRMSM, a privacy preserving ranked multi-keyword search protocol in a multi-owner cloud model.

•We define a multi-owner model for privacy preserving keyword search over encrypted cloud data.

•We propose an efficient data user authentication protocol, which not only prevents attackers from eavesdropping secret keys and pretending to be illegal data users performing searches, but also enables data user authentication and revocation.

•We systematically construct a novel secure search protocol, which not only enables the cloud server to perform secure ranked keyword search without knowing the actual data of both keywords and trapdoors, but also allows data owners to encrypt keywords with self-chosen keys and allows authenticated data users to query without knowing these keys.

•We propose an Additive Order and Privacy Preserving Function family (AOPPF) which allows data owners to protect the privacy of relevance scores using different functions according to their preference, while still permitting the cloud server to rank the data files accurately.

•We conduct extensive experiments on real-world datasets to confirm the efficacy and efficiency of our proposed schemes.

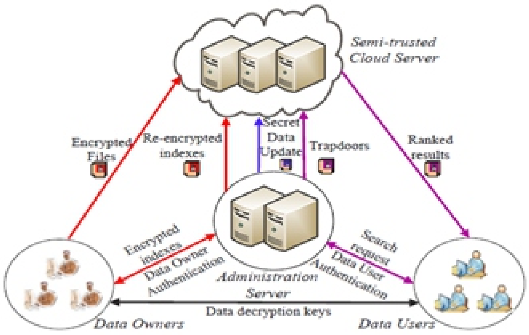


FIG: SYSTEM ARCHITECTURE

Advantages:

•The proposed scheme allows multi-keyword search over encrypted files which would be encrypted with different keys for different data owners.

•The proposed scheme allows new data owners to enter this system without affecting other data owners or data users, i.e., the scheme supports data owner scalability in a plug-and-play model.

•The proposed scheme ensures that only authenticated data users can perform correct searches. Moreover, once a data user is revoked, he can no longer perform correct searches over the encrypted cloud data.

•To enable cloud servers to perform secure search without knowing the actual value of both keywords and trapdoors, we systematically construct a novel secure search protocol. As a result, different data owners use different keys to encrypt their files and keywords. Authenticated data users can issue a query without knowing secret keys of these different data owners.

•To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a new additive order and privacy preserving function family, which helps the cloud server return the most relevant search results to data users without revealing any sensitive information.

•To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.

IMPLEMENTATION

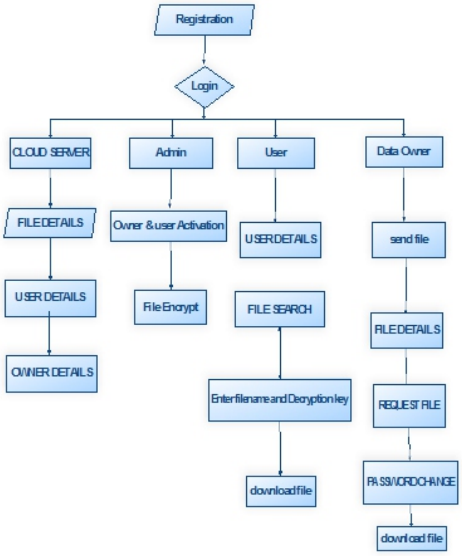
1)System Model:

•In the first module, we develop the System Model to implement our proposed system. Our System model consists of Admin, users, data owners, and Cloud Servers. Admin provides the accessibility to Data-owners. Initially Data-owner needs to register and admin approves the each data owner request. The respective Password and login credentials will be sent to the Email ID of Data owner.

•In Users sub-module, Each user has a global identity in the system. A user may be entitled a set of attributes which may come from multiple attribute authorities. The user will receive a secret key associated with its attributes entitled by the corresponding attribute authorities.

•In data owners sub-module, the proposed scheme should allow new data owners to enter this system without affecting other data owners or data users, i.e., the scheme should support data owner scalability in a plug-and-play model.

•In Cloud Server sub-module of system model, the owner sends the encrypted data to the cloud server through Admin. They do not rely on the server to do data access control. But, the access control happens inside the cryptography. That is only when the user's attributes satisfy the access policy defined in the cipher text; the user is able to decrypt the ciphertext. Thus, users with different attributes can decrypt different number of content keys and thus obtain different granularities of information from the same data.



2) Data User Authentication:

•To prevent attackers from pretending to be legal data users performing searches and launching statistical attacks based on the search result, data users must be authenticated before the administration server re-encrypts trapdoors for data users. Traditional authentication methods often follow three steps. First, data requester and data authenticator share a secret key, say, k0. Second, the requester encrypts his personally identifiable information d0 using k0 and sends the encrypted data (d0)k0 to the authenticator. Third, the authenticator decrypts the received data with k0 and authenticates the decrypted data.

•The key point of a successful authentication is to provide both the dynamically changing secret keys and the historical data of the corresponding data user.

3)Illegal Search Detection:

•In our scheme, the authentication process is protected by the dynamic secret key and the historical information. We assume that an attacker has successfully eavesdropped the secret key. Then he has to construct the authentication data; if the attacker has not successfully eavesdropped the historical data, e.g., the request counter, the last request time, he cannot construct the correct authentication data. Therefore this illegal action will soon be detected by the administration server.

•Further, if the attacker has successfully eavesdropped all data of Uj , the attacker can correctly construct the authentication data and pretend himself to be Uj without being detected by the administration server. However, once the legal data user Uj performs his search, since the secret key on the administration server side has changed, there will be contradictory secret keys between the administration server and the legal data user. Therefore, the data user and administration server will soon detect this illegal action.

4)Search over Multi-owner:

•The proposed scheme should allow multi-keyword search over encrypted files which would be encrypted with different keys for different data owners. It also needs to allow the cloud server to rank the search results among different data owners and return the top-k results. The cloud server stores all encrypted files and keywords of different data owners.

•The administration server will also store a secret data on the cloud server. Upon receiving a query request, the cloud will search over the data of all these data owners.

The cloud processes the search request in two steps. First, the cloud matches the queried keywords from all keywords stored on it, and it gets a candidate file set. Second, the cloud ranks files in the candidate file set and finds the most top-k relevant files. Finally, we apply the proposed scheme to encode the relevance scores and obtain the top-k search results.

CONCLUSION AND FUTURE WORK:

In this paper, we explore the problem of secure multi-keyword search for multiple data owners and multiple data users in the cloud computing environment. Different from prior works, our schemes enable authenticated data users to achieve secure, convenient, and efficient searches over multiple data owners' data. To efficiently authenticate data users and detect attackers who steal the secret key and perform illegal searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. To enable the cloud server to perform secure search among multiple owners' data encrypted with different secret keys, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. Moreover, we show that our approach is computationally efficient, even for large data and keyword sets.

REFERENCES:

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Communication of the ACM, vol. 53, no. 4, pp. 50–58, 2010.

[2] C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacypreserving public auditing for secure cloud storage," Computers, IEEE Transactions on, vol. 62, no. 2, pp. 362–375, 2013.

[3] D.Song, D.Wagner, and A.Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE International Symposium on Security and Privacy (S&P'00), Nagoya, Japan, Jan. 2000, pp. 44–55.

[4] E. Goh. (2003) Secure indexes. [Online]. Available: http://eprint.iacr.org/ [5] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. ACM CCS'06, VA, USA, Oct. 2006, pp. 79–88.

[6] D. B. et al., "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," EUROCRYPT, vol. 43, pp. 506–522, 2004.

[7] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Applied Cryptography and Network Security (ACNS'04), Yellow Mountain, China, Jun. 2004, pp. 31–45.

[8] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proc. Information and Communications Security (ICICS'05), Beijing, China, Dec. 2005, pp. 414–426.

[9] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE Distributed Computing Systems (ICDCS'10), Genoa, Italy, Jun. 2010, pp. 253–262.

[10] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacypreserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM'11, Shanghai, China, Apr. 2011, pp. 829–837.

[11] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacypreserving multi-keyword ranked search over encrypted cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 1, pp. 222–233, 2014.

[12] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 11, pp. 3025–3035, 2014.

[13] Z. Xu, W. Kang, R. Li, K. Yow, and C. Xu, "Efficient multikeyword ranked query on encrypted data in the cloud," in Proc. IEEE Parallel and Distributed Systems (ICPADS'12), Singapore, Dec. 2012, pp. 244–251.

[14] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in Proc. IEEE INFOCOM'10, San Diego, CA, Mar. 2010, pp. 1–5.

[15] M. Chuah and W. Hu, "Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data," in Proc. IEEE 31th International Conference on Distributed Computing Systems (ICDCS'11), Minneapolis, MN, Jun. 2011, pp. 383–392.

[16] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," Computers, IEEE Transactions on, vol. 62, no. 11, pp. 2266–2277, 2013.

Authors Biography

Velamala Venkataramana

Completed his B.Tech(CSE) degree in CMR College of Engineering and Technology in 2014. He is pursuing M.Tech. in Computer Science & Engineering from Department of Computer Science & Engineering in Malla Reddy Engineering College & Management sciences, Kistapur, Medchal, Hyderabad. Affiliated to JNTUH, HYDERABAD, TELANGANA, India.. His research interest include cloud, data mining, Big Data and networking.

Y.Ashwini

working as Assistant Professor, department of computer science and engineering, in Malla Reddy Engineering College & Management Sciences, Kistapur, Medchal ,Hyderabad. Affiliated to JNTUH, HYDERABAD, TELANGANA., India. Her research interests include data mining, computer networks.

Ch. Ramesh Kumar

working as Assoc. Prof & Head of the Department of Computer Science and Engineering in Malla Reddy Engineering College & Management Sciences, Kistapur, Medchal, Hyderabad. Affiliated to JNTUH, HYDERABAD, TELANGANA., India. he has several international publications to his credit. His research interests include Software reuse, Software performance, Software testing ,Data Mining and cloud computing.