FAST PHRASE SEARCH FOR ENCRYPTED CLOUD STORAGE

Ravada Phani Pratyusha , Sri.G.Ramesh Kumar, Sri.V.Bhaskara Murthy,
MCA Student, Assistant Professor, Associate Professor
Dept Of MCA
B.V.Raju College, Bhimavaram

ABSTRACT

Cloud computing has generated much interest in the research community in recent years for its many advantages, but has also raise security and privacy concerns. The storage and access of confidential documents have been identified as one of the central problems in the area. In particular, many researchers investigated solutions to search over encrypted documents stored on remote cloud servers. While many schemes have been proposed to perform conjunctive keyword search, less attention has been noted on more specialized searching techniques. In this paper, we present a phrase search technique based on Bloom filters that is significantly faster than existing solutions, with similar or better storage and communication cost. Our technique uses a series of n-gram filters to support the functionality. The scheme exhibits a trade-off between storage and false positive rate, and is adaptable to defend against inclusion-relation attacks. A design approach based on an application's target false positive rate is also described.

I. INTRODUCTION

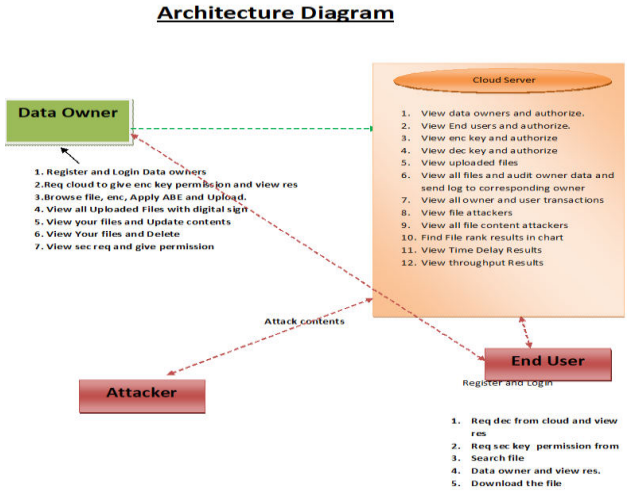
AS organizations and individuals adopt cloud technologies, many have become aware of the serious concerns regarding security and privacy of accessing personal and confidential information over the Internet. In particular, the recent and continuing data breaches highlight the need for more secure cloud storage systems. While it is generally agreed that encryption is necessary, cloud providers often perform the encryption and maintain the private keys instead of the data owners. That is, the cloud can read any data it desired, providing no privacy to its users. The storage of private keys and encrypted data by the cloud provider is also problematic in case of data breach. Hence, researchers have actively been exploring solutions for secure storage on private and public clouds where private keys remain in the hands of data owners.

Boneh et al. [1] proposed one of the earliest works on keyword searching. Their scheme uses public key encryption to allow keywords to be searchable without revealing data content. Waters et al. [2] investigated the problem for searching over encrypted audit logs. Many of the early works focused on single keyword searches. Recently, researchers have proposed solutions on conjunctive keyword search, which involves multiple keywords [3], [4]. Other interesting problems, such as the ranking of search results [5], [6], [7] and searching with keywords that might contain errors [8], [9] termed fuzzy keyword search, have also been considered. The ability to search for phrases was also recently investigated [10], [11], [12], [13]. Some [14] have examined the security of the proposed solutions and,

where flaws were found, solutions were proposed [15].

In this paper, we present a phrase search scheme which achieves a much faster response time than existing solutions. The scheme is also scalable, where documents can easily be removed and added to the corpus. We also describe modifications to the scheme to lower storage cost at a small cost in response time and to defend against cloud providers with statistical knowledge on stored data. We begin by presenting the communication framework in section 2 and various backgrounds including related works in section 3. Although phrase searches are processed independently using our technique, they are typically a specialized function in a keyword search scheme, where the primary function is to provide conjunctive keyword searches. Therefore, we describe both the basic conjunctive keyword search algorithm and the basic phrase search algorithm in section 4 along with design techniques in section 4.3. Performance analysis and experimental results are included in section 5 and 6.

II. SYSTEM ARCHITECTURE



III. EXISTING SYSTEM

- In the existing system, Ding et al. [3] extended Boneh et al.'s scheme using bilinear mapping to perform multiple keyword search and described a solution that did not include expensive pairing operations in the encryption and trapdoor generation phase.
- Kerschbaum et al. [4] considered the search of unstructured text, where positions of keywords are unknown. The use of encrypted index for keyword search was examined in [22] and a scheme secure against chosen keyword attack was proposed. The ranking of search results was looked at by Wang et al. in [17]. The authors described a solution based on the commonly used TFIDF (Term Frequency x Inverse Document Frequency) rule

and the use of order preserving symmetric encryption.

- Liu et al. [23] considered the search for potentially erroneous keywords termed fuzzy keyword search. The index-based solution makes use of fuzzy dictionaries containing various misspelling of keywords including wildcards.

Disadvantages

- There is less security on outsourced data due to lack of Modified phrase search scheme against IR attacks.
- There is no Data integrity technique to audit outsourced data.

IV. PROPOSED SYSTEM

- In the proposed system, the system presents a phrase search scheme which achieves a much faster response time than existing solutions. The scheme is also scalable, where documents can easily be removed and added to the corpus. The system also describes modifications to the scheme to lower storage cost at a small cost in response time and to defend against cloud providers with statistical knowledge on stored data.
- In the proposed system, the system also presents a phrase search technique based on Bloom filters that is significantly faster than existing solutions, with similar or better storage and communication cost. The proposed system technique uses a series of n-gram filters to support the functionality.
- The scheme exhibits a trade-off between storage and false positive rate, and is adaptable to defend against inclusion-relation attacks. A design approach based on an application's target false positive rate is also described.

Advantages

- The Data retrieval is fast due to Conjunctive keyword search scheme.
- The security is more on outsourced data due to Modified phrase search scheme against IR attacks.

V. IMPLEMENTATION

- **Data Owner**

In this module, the data owner uploads their data in the cloud server. For the security purpose the data owner encrypts the file and the index name and then store in the cloud. The data encryptor can have capable deleting of a specific file. And also he can view the transactions based on the files he uploaded to cloud and will do the following operations like Register and Login Data owners,Req cloud to give enc key permission and view res,Browse file, enc, Apply ABE and Upload, View all Uploaded Files with digital sign, View your files and Update contents, View Your files and Delete , View sec req and give permission.

- **Data User**

In this module, user logs in by using his/her user name and password. After Login user requests search control to cloud and will Search for files based on the index keyword with the Score of the searched file and downloads the file. User can view the search of the files and also do some operations like Req dec from cloud and view res,Req sec

key permission from, Search file, Data owner and view res,Download the file.

- **Cloud Server**

The cloud server manages a cloud to provide data storage service. Data owners encrypt their data files and store them in the cloud for sharing with Remote User. To access the shared data files, data consumers download encrypted data files of their interest from the cloud and then decrypt them.

The cloud server authorizes the data owner and the data user and provides the search requests sent from the users. Also in this module it shows personalized search model and the interest search model. Can view all the file attackers and doing following operations View data owners and authorize, View End users and authorize, View enc key and authorize, View dec key and authorize, View uploaded files, View all files and audit owner data and send log to corresponding owner, View all owner and user transactions, View file attackers, View all file content attackers, Find File rank results in chart, View Time Delay Results, View throughput Results.

VI. CONCLUSION

In this paper, we presented a phrase search scheme based on Bloom filter that is significantly faster than existing approaches, requiring only a single round of communication and Bloom filter verifications. The solution addresses the high computational cost noted in [13] by reformulating phrase search as n-gram verification rather than a location search or a sequential chain verification. Unlike [10], [12],[13], our schemes consider only the existence of a phrase, omitting any information of its location. Unlike [11], our schemes do not require sequential verification, is parallelizable and has a practical storage requirement. Our approach is also the first to effectively allow phrase search to run independently without first performing a conjunctive keyword search to identify candidate documents. The technique of constructing a Bloom filter index introduced in section 4.2 enables fast verification of Bloom filters in the same manner as indexing. According to our experiment, it also achieves a lower storage cost than all existing solutions except [13], where a higher computational cost was exchanged in favor of lower storage. While exhibiting similar communication cost to leading existing solutions, the proposed solution can also be adjusted to achieve maximum speed or high speed with a reasonable storage cost depending on the application. An approach is also described to adapt the scheme to defend against inclusion-relation attacks. Various issues on security and efficiency, such as the effect of long phrases and precision rate, were also discussed to support our design choices.

REFERENCES

[1] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in In proceedings of Eurocrypt, 2004, pp. 506–522.

[2] B. Waters, D. Balfanz, G. Durfee, and D. K. Smetters, “Building an encrypted and searchable audit log,” in Network and Distributed System Security Symposium, 2004.

[3] M. Ding, F. Gao, Z. Jin, and H. Zhang, "An efficient public key encryption with conjunctive keyword search scheme based on pairings," in IEEE International Conference on Network Infrastructure and Digital Content, 2012, pp. 526–530.
[4] F. Kerschbaum, "Secure conjunctive keyword searches for unstructured text," in International Conference on Network and System Security, 2011, pp. 285–289.
[5] C. Hu and P. Liu, "Public key encryption with ranked multi keyword search," in International Conference on Intelligent Networking and Collaborative Systems, 2013, pp. 109–113.
[6] Z. Fu, X. Sun, N. Linge, and L. Zhou, "Achieving effective cloud search services: multi-keyword ranked search over encrypted cloud data supporting synonym query," IEEE Transactions on Consumer Electronics, vol. 60, pp. 164–172, 2014.
[7] C. L. A. Clarke, G. V. Cormack, and E. A. Tudhope, "Relevance ranking for one to three term queries," Information Processing and Management: an International Journal, vol. 36, no. 2, pp. 291–311, Jan. 2000.
[8] H. Tuo and M. Wenping, "An effective fuzzy keyword search scheme in cloud computing," in International Conference on Intelligent Networking and Collaborative Systems, 2013, pp. 786–789.
[9] M. Zheng and H. Zhou, "An efficient attack on a fuzzy keyword search scheme over encrypted data," in International Conference on High Performance Computing and Communications and Embedded and Ubiquitous Computing, 2013, pp. 1647–1651.
[10] S. Zittrower and C. C. Zou, "Encrypted phrase searching in the cloud," in IEEE Global Communications Conference, 2012, pp. 764–770.
[11] Y. Tang, D. Gu, N. Ding, and H. Lu, "Phrase search over encrypted data with symmetric encryption scheme," in International Conference on Distributed Computing SystemsWorkshops, 2012, pp. 471–480.
[12] H. Poon and A. Miri, "An efficient conjunctive keyword and phrase search scheme for encrypted cloud storage systems," in IEEE International Conference on Cloud Computing, 2015.
[13] ——, "A low storage phrase search scheme based on bloom filters for encrypted cloud services," to appear in IEEE International Conference on Cyber Security and Cloud Computing, 2015.
[14] H. S. Rhee, I. R. Jeong, J. W. Byun, and D. H. Lee, "Difference set attacks on conjunctive keyword search schemes," in Proceedings of the Third VLDB International Conference on Secure Data Management, 2006, pp. 64–74.