Secure and Efficient Product Information Retrieval in Cloud Computing

[1]Kodam.Anupama [2]Mr.N.Naveen Kumar

[1]M.Tech Student, School of Information Technology JNTUH, Kukatpally, Medchal-Malkajigiri, Hyderabad.

[2]Assistant Professor, School of Information Technology JNTUH, Kukatpally, Medchal-Malkajigiri, Hyderabad.

Abstract—Cloud computing is a promising IT technique that can organize a large amount of IT resources in an efficient and flexible manner. Increasingly numerous companies plan to move their local data management systems to the cloud and store and manage their product information on cloud servers. An accompanying challenge is how to protect the security of the commercially confidential data while maintaining the ability to search the data. In this paper, a privacy-preserving data search scheme is proposed that can support both the identifier-based and feature-based product searches. Specifically, two novel index trees are constructed and encrypted that can be searched without knowing the plaintext data. Analysis and simulation results demonstrate the security and efficiency of our scheme.

1. INTRODUCTION

Driven by the revolution of information technology in recent years and with the slowdown in the economic growth, there is an urgent need to transform China's entire industrial chain. To promote an all-around industrial upgrading, China has proposed the strategy of "Internet +", and the integration of China's ecommerce with its traditional economy has been significantly improved. Ecommerce has accelerated its expansion from consumption to various industries and infiltrated all aspects of social and economic activities, thereby driving the development of enterprise-level ecommerce, both in scope and in depth, and facilitating the transformation and upgrading of enterprises. The Monitoring Report on the Data of China's Ecommerce Market [1] shows that in 2016, the volume of ecommerce transactions in China reached approximately 3.5 trillion dollars, a year-on-year growth rate of approximately 25.5%. The rapidly rising number of cyber-transactions has

spawned ecommerce big data. As increasingly numerous data files are being stored locally in enterprises, the pressure on local data storage systems greatly increases. Local hardware failures lead to great damage or loss of data, which greatly affects the daily operations of the enterprises. Fortunately, cloud storage techniques came into being under such circumstances. Cloud computing can collect and organize a large number of different types of storage devices by means of various functions, such as cluster applications, network technology and distributed file systems. There have already been a number of typical cloud service products at home and abroad, such as Amazon Web Services [2], Microsoft Azure [3], i Cloud [4], and App Engine [5].

2. RELATED WORK

[11] 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. In this paper, 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 we present are simple, fast (for a document of length , 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.

We have described new techniques for remote searching on encrypted data using an

untrusted server and provided proofs of security for the resulting crypto systems. Our techniques have a number of crucial advantages: they are provably secure; they support controlled and hidden search and query isolation; they are simple and fast (More specifically, for a document of length , the encryption and search algorithms only need O (n) stream cipher and block cipher operations); and they introduce almost no space and communication overhead. Our scheme is also very flexible, and it can easily be extended to support more advanced search queries. We conclude that this provides a powerful new building block for the construction of secure services in the untrusted infrastructure.

[13] Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions.

Searchable symmetric encryption (SSE) allows a party to outsource the storage of his data to another party in a private manner, while maintaining the ability to selectively search over it. This problem has been the focus of active research and several security definitions and constructions have been proposed. In this paper we begin by reviewing existing notions of security and propose new and stronger security definitions. We then present two constructions that we show secure under our new definitions. Interestingly, in addition to satisfying stronger security guarantees, our constructions are more efficient than all previous constructions. Further, prior work on SSE only considered 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 this multi-user setting, and present an efficient construction.

In this article, we have revisited the problem of searchable symmetric encryption, which allows a client to store its data on a remote server in such a way that it can search over it in a private manner. We make several contributions including new security definitions and new constructions. Motivated by subtle problems in all previous security definitions for SSE, we propose new definitions and point out that the existing notions have significant practical drawbacks: contrary to the natural use of searchable encryption, they only guarantee security for users that perform all their searches at once. We address this limitation by introducing stronger definitions that guarantee security even when users perform

more realistic searches. We also propose two new SSE constructions. Surprisingly, despite being provably secure under our stronger security definitions, these are the most efficient schemes to date and are (asymptotically) optimal (i.e., the work performed by the server per returned document is constant in the size of the data). Finally, we also consider multi-user SSE, which extends the searching ability to parties other than the owner.

3. FRAMEWORK

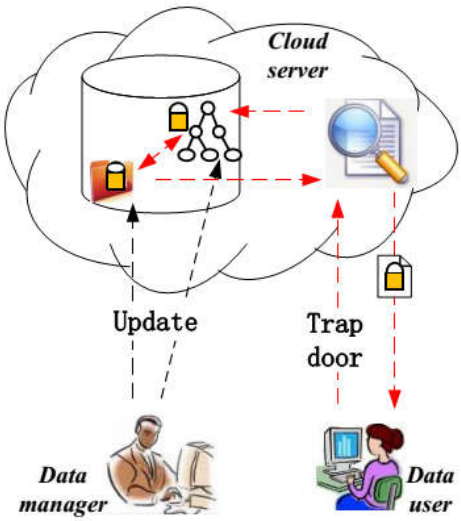

Fig.1. Encrypted product information retrieval system model

As shown in Fig. 1, the entire product retrieval system model is composed primarily of three entities: the data manager, the cloud server and the data user. The primary responsibilities of these three entities are presented in the following.

The data manager is responsible for managing the product and collecting the product information. In addition, the data manager needs to encrypt the product information file by a symmetric encryption technique before outsourcing the data to the cloud server. To improve the security of the files, each file is encrypted by a single secret key, and the keys of different files are independent. Furthermore, to improve the search efficiency, an index structure is constructed for the outsourced data. At first, an identifier index structure is constructed based on the hash function and height-balanced binary search tree. Then, a feature vector tree is built for all the feature vectors of the product, and it is encrypted by the secure kNN algorithm.

When a data user wants to search a set of chosen products, she needs to generate a trapdoor to describe her interest. Two types of the trapdoor can be provided, i.e., a set of hash values of the desired product information files or a set of feature vectors. For the first type of trapdoor, a set of encrypted files with the same hash identifiers are returned, and for the second type trapdoor, the most relevant encrypted

files are returned. The data user can obtain the plaintext files by decrypting the returned files with the help of the symmetric secret keys. These secret keys are provided by the data manager.

The cloud server stores all the data uploaded by the data manager. When a data user needs to search the data in the cloud, she first generates a trapdoor, which is sent to the cloud server. A search engineer is employed by the cloud server to act as a bridge between the data users and the encrypted data. Though the cloud server cannot get the plaintexts of the data, it should be capable of sending the accurate search result of the trapdoor to the data users. Of course, the returned data are ciphertext, and the data user needs to decrypt them by the symmetric secret keys which are provided by the data manager.

4. EXPERIMENTAL RESULTS

We evaluate the search efficiency of our scheme. First, we evaluate the construction time of the index structures of the product information. Specifically, we compare our scheme with the MRSE scheme [14]. To decrease the bias of the data manager who is responsible for generating the vectors and the hash values, in this paper we employ the Enron Email Data Set [15] to test our scheme.

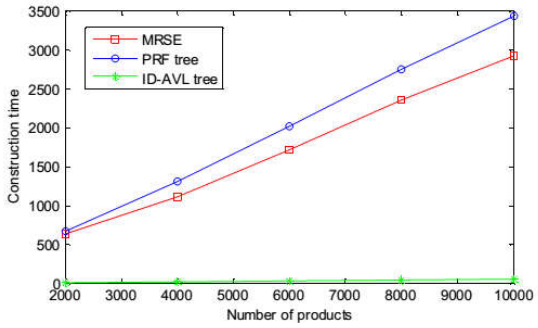

Fig.2. Construction time of the index structures

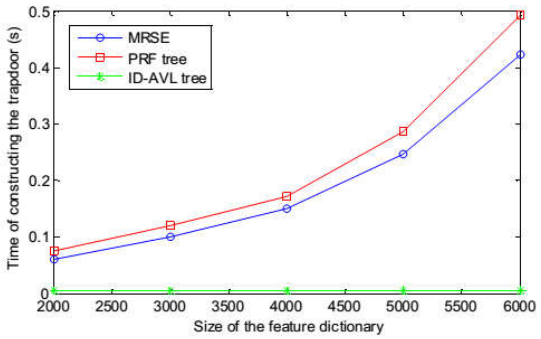

Fig.3. Time of constructing the trapdoors

5. CONCLUSION

In this paper, we designed a secure and efficient product information retrieval scheme based on cloud computing. Specifically, two index structures, including a hash value AVL tree and a product vector retrieval tree, are constructed, and they support an identifier-based product search and feature-vector-based product search, respectively. Correspondingly, two search

algorithms are designed to search the two trees. To protect the product information privacy, all the outsourced data are encrypted. The product information is symmetrically encrypted based on a set of independent secret keys, and the product vectors are encrypted based on the secure kNN algorithm. Security analysis and simulation results illustrate the security and efficiency of the proposed scheme.

REFERENCES

[1] www.100EC.cn. 2016 Monitoring Report on the Data of China's Ecommerce Market [EB/OL]. http://www.100ec.cn/zt/16jcbg/,2017- 05-24

[2] Amazon. Amazon S3. http://aws.amazon.com/s3/

[3] Windows azure. http://www.microsoft.com/windowsazure/

[4] Apple i Cloud. http://www.icloud.com/

[5] Google App Engine. http://appengine.google.com/

[6] Golle P.Staddon J,Waters B. Secure Conjunctive Keyword Search over Data[C]. Springer, 2004.

[7] Song D X,Wanger D.Perrig A. Practical Techniques for Searched on Encrypted Data[C].IEEE,2000.

[8] Boneh D,Di Crescenzo G,Ostrovsky R. et al. Public Key Encryption with Keyword Search: EUROCRYPT[C].Springer,2004.

[9] Rhee H S.Park J K,Susilo W. et al. Trapdoor Security in A Searchable Public-Key Encryption Scheme with A Designated Tester[J].Journal of Systems and Software,2010,83(5):763-771

[10] Ren, Kui, Cong Wang, and Qian Wang. "Security challenges for the public cloud." IEEE Internet Computing 16.1 (2012): 69-73.

[11] Song, Dawn Xiaoding, David Wagner, and Adrian Perrig. "Practical techniques for searches on encrypted data." Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000.

[12] Goh, Eu-Jin. "Secure indexes." IACR Cryptology ePrint Archive 2003 (2003): 216.

[13] Curtmola, Reza, et al. "Searchable symmetric encryption: improved definitions and efficient constructions." Journal of Computer Security 19.5 (2011): 895-934.

[14] Cao, Ning, et al. "Privacy-preserving multi-keyword ranked search over encrypted cloud data." IEEE Transactions on parallel and distributed systems 25.1 (2014): 222-233.

[15] W.W. Cohen, "Enron Email Data Set," https://www.cs.cmu.edu/~./enron/, 2015.