ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Back Synonymous With Advanced Multi-Word Search Queries Cloud System Rated Encrypted

Y.Ganesh
M.Tech (CSE),
Prasad College of Engineering.

E.Madhu
Assistant Professor,
Prasad College of Engineering.

ABSTRACT:

As an enhancement we enhance the existing system and in this paper we propose an effective approach to solve the problem of multi-keyword ranked search over encrypted cloud data supporting synonym queries. The main contribution of this paper is summarized in two aspects: multi-keyword ranked search to achieve more accurate search results and synonym-based search to support synonym queries. Meanwhile, existing search approaches over encrypted cloud data support only exact or fuzzy keyword search, but not semantics-based multi-keyword ranked search. Therefore, how to enable an effective searchable system with support of ranked search remains a very challenging problem.

INTRODUCTION:

A distributed system is a software system in which components located on networked computers communicate and coordinate their actions by passing messages. The components interact with each other in order to achieve a common goal. There are many alternatives for the message passing mechanism, including RPC-like connectors and message queues. Three significant characteristics of distributed systems are: concurrency of components, lack of a global clock, and independent failure of components.

An important goal and challenge of distributed systems is location transparency. Examples of distributed systems vary from SOA-based systems to massively multiplayer online games to peer-to-peer applications. A computer program that runs in a distributed system is called a distributed program, and distributed programming is the process of writing such programs. Distributed computing also refers to the use of distributed systems to solve computational problems. In distributed computing, a problem is divided into many tasks, each of which is solved by one or more computers, which communicate with each other by message passing. The word *distributed* in terms such as "distributed system", "distributed programming", and "distributed algorithm" originally referred to computer networks where individual computers were physically distributed within some geographical area. The terms are nowadays used in a much wider sense, even referring to autonomous processes that run on the same physical computer and interact with each other by message passing. Distributed systems are groups of networked computers, which have the same goal for their work.

The terms "concurrent computing", "parallel computing", and "distributed computing" have a lot of overlap, and no clear distinction exists between them. The same system may be characterised both as "parallel" and "distributed"; the processors in a typical distributed system run concurrently in parallel. Parallel computing may be seen as a particular tightly coupled form of distributed computing, and distributed computing may be seen as a loosely coupled form of parallel computing. The situation is further complicated by the traditional uses of the terms parallel and distributed algorithm that do not quite match the above definitions of parallel and distributed systems; see the section Theoretical foundations below for more detailed discussion. Nevertheless, as a rule of thumb, high-performance parallel computation in a shared-memory multiprocessor uses parallel algorithms while the coordination of a large-scale distributed system uses distributed algorithms.

RELATED WORK:

The cloud server is considered as "honest-but-curious" in our model, which is consistent with related works on cloud security. Specifically, the cloud server acts in an "honest" fashion and correctly follows the designated protocol specification. However, it is "curious" to infer and analyze data (including index) in its storage and message flows received during the protocol so as to learn additional information. Based on what information the cloud server knows, we consider two threat models with different attack capabilities as follows.

Known ciphertext model. In this model, the cloud server is supposed to only know encrypted data set C and searchable index I, both of which are outsourced from the data owner. Known background model. In this stronger model, the cloud server is supposed to possess more knowledge than what can be accessed in the known ciphertext model. Such information may include the correlation relationship of given search requests (trapdoors), as well as the data set related statistical information. As an instance of possible attacks in this case, the cloud server could use the known trapdoor information combined with document/keyword frequency to deduce/identify certain keywords in the query.

As a hybrid of conjunctive search and disjunctive search, "coordinate matching" is an intermediate similarity measure which uses the number of query keywords appearing in the document to quantify the relevance of that document to the query. When users know the exact subset of the data set to be retrieved, Boolean queries perform well with the precise search requirement specified by the user. In cloud computing, however, this is not the practical case, given the huge amount of outsourced data. Therefore, it is more flexible for users to specify a list of keywords indicating their interest and retrieve the most relevant documents with a rank order. To efficiently achieve multi-keyword ranked search, we propose to employ "inner product similarity" to quantitatively evaluate the efficient similarity measure "coordinate matching."

Specifically, Di is a binary data vector for document Fi where each bit Di½j_ 2 f0; 1g represents the existence of the corresponding keyword Wj in that document, and Q is a binary query vector indicating the keywords of interest where each bit Q½j_ 2 f0; 1g represents the existence of the corresponding keyword Wj in the query fW.

The similarity score of document Fi to query fW is therefore expressed as the inner product of their binary column vectors, i.e., Di _ Q. For the purpose of ranking, the cloud server must be given the capability to compare the similarity of different documents to the query. But, to preserve strict systemwise privacy, data vector Di, query vector Q and their inner product Di _ Q should not be exposed to the cloud server. In this section, we first propose a basic idea for the MRSE using secure inner product computation, which is adapted from a secure kNN technique, and then show how to significantly improve it to be privacy-preserving against different threat models in the MRSE framework in a step-by-step manner. We further discuss supporting more search semantics and dynamic operation.

SYSTEM PREMELERIES

Data Owner Module

This module helps the owner to register those details and also include login details. This module helps the owner to upload his file with encryption using RSA algorithm. This ensures the files to be protected from unauthorized user.

Data User Module

This module includes the user registration login details. This module is used to help the client to search the file using the multiple key words concept and get the accurate result list based on the user query. The user is going to select the required file and register the user details and get activation code in mail email before enter the activation code. After user can download the Zip file and extract that file.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Encryption Module:

This module is used to help the server to encrypt the document using RSA Algorithm and to convert the encrypted document to the Zip file with activation code and then activation code send to the user for download.

Rank Search Module

These modules ensure the user to search the files that are searched frequently using rank search. This module allows the user to download the file using his secret key to decrypt the downloaded data. This module allows the Owner to view the uploaded files and downloaded files

CONCLUSION

In this paper, for the first time we define and solve the problem of multi-keyword ranked search over encrypted cloud data, and establish a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure.

For meeting the challenge of supporting multi-keyword semantic without privacy breaches, we propose a basic idea of MRSE using secure inner product computation. Then, we give two improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. We also investigate some further enhancements of our ranked search mechanism, including supporting more search semantics, i.e., TF _ IDF, and dynamic data operations. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world data set show our proposed schemes introduce low overhead on both computation and communication.

REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829-837, Apr, 2011.

[2] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50-55, 2009.

[3] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.

[4] S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.

[5] A. Singhal, "Modern Information Retrieval: A Brief Overview,"IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, Mar. 2001.

[6] I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Publishing, May 1999.

[7] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.

[8] E.-J. Goh, "Secure Indexes," Cryptology ePrint Archive, http://eprint.iacr.org/2003/216. 2003.

[9] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.

[10] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. 13th ACM Conf. Computer and Comm. Security (CCS '06), 2006.

[11] D. Boneh, G.D. Crescenzo, R. Ostrovsky, and G. Persiano, “Public Key Encryption with Keyword Search,” Proc. Int’l Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2004.

[12] M. Bellare, A. Boldyreva, and A. ONeill, “Deterministic and Efficiently Searchable Encryption,” Proc. 27th Ann. Int’l Cryptology Conf. Advances in Cryptology (CRYPTO ’07), 2007.

[13] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, “Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous Ibe, and Extensions,” J. Cryptology, vol. 21, no. 3, pp. 350-391, 2008.

[14] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, “Fuzzy Keyword Search Over Encrypted Data in Cloud Computing,” Proc. IEEE INFOCOM, Mar. 2010.

[15] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W.E.S. III, “Public Key Encryption That Allows PIR Queries,” Proc. 27th Ann. Int’l Cryptology Conf. Advances in Cryptology (CRYPTO ’07), 2007.

[16] P. Golle, J. Staddon, and B. Waters, “Secure Conjunctive Keyword Search over Encrypted Data,” Proc. Applied Cryptography and Network Security, pp. 31-45, 2004.

[17] L. Ballard, S. Kamara, and F. Monrose, “Achieving Efficient Conjunctive Keyword Searches over Encrypted Data,” Proc. Seventh Int’l Conf. Information and Comm. Security (ICICS ’05), 2005.

[18] D. Boneh and B. Waters, “Conjunctive, Subset, and Range Queries on Encrypted Data,” Proc. Fourth Conf. Theory Cryptography (TCC), pp. 535-554, 2007.

[19] R. Brinkman, “Searching in Encrypted Data,” PhD thesis, Univ. of Twente, 2007.

[20] Y. Hwang and P. Lee, “Public Key Encryption with Conjunctive Keyword Search and Its Extension to a Multi-User System,” Pairing, vol. 4575, pp. 2-22, 2007.