An overview of: A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data

David E. Craciunescu

Universidad de Alcala

Abstract. The usage of cloud computing is growing at a rate never experienced before, and the number of data owners that outsource their data to cloud servers is growing with it. This outsourcing process has obvious advantages and conveniences, such as the reduced cost in data management, making this trend unsurprising to the careful observer. Data, however (especially private data) must always be encrypted for obvious privacy reasons, action that would destroy any kind of performance and speed when indexing it for latter retrieval. This paper is an overview of the original study, whose objective was to solve such inconvenience by creating a multi-keyword ranked search scheme over encrypted cloud data, which supports dynamic update operations of documents. The referenced paper also proposes the use of a special "*tree-based* index structure and (. . .) a 'Greedy Depth-first Search' algorithm". This proposed scheme can achieve sub-linear search time thanks to its careful design and crafting.

Introduction

The "new" technology which cloud computing represents has existed for many years, but it hasn't been until very recently when its use has seen an increase of unimaginable proportions. This may very well be because of the advancements in other fields such as networking or connectivity, but it is generally considered to be a joint effort. It is slowly starting to be considered as the new way of enterprise information technology infrastructure. Currently it provides users with the usage of gigantic computing resources, on-demand access to a shared group of configurable resources that come with remarkable efficiency, and little to no economic overhead. Individuals as well as companies, being extremely drawn to these attractive advantages and features, have the motivation to outsource their information to the cloud, instead of storing it themselves and having to buy and maintain software and hardware on-premises. The ability to avoid the need of purchasing licenses and hardware to store one's own information is also a big factor to be considered.

Motivation

The cloud system raises some flags regarding privacy. Even though there are clear advantages to cloud usage over on-site storage, outsourcing sensitive and private information to remote servers isn't trusted by many. In fact, there are many that directly don't trust the cloud service providers themselves. Since they keep the data for users, they could technically access private information without the proper authorization. It is generally believed that encrypting the data before outsourcing is the best approach when dealing with security and privacy concerns. This is, however, a big hit to usability. One's information should ideally be secure and be usable at the same time for indexing or searching, for example, and the most used modern encryption techniques are not compatible with current information retrieval methods. Another possibility would be downloading all the data and decrypting it downloaded, but it is obvious why this approach is impractical.

Work and Development

Researchers realized these problems and tried to solve them by creating methods such as *fully-homomorphic* encryption[1]. Unfortunately, effective as they may be, these methods come with an enormous computational overhead, and are not suited to the necessities of the cloud. The processing and information retrieval had to be done evaluating the encrypted information without any need of decryption computation. With this goal in mind, *Searchable Encryption* methods were created. These schemes enable the client to store the information encrypted on the cloud and execute keyword search over the encrypted data. Many different kind of implementations and methods have been proposed, but the most promising seems to be the *multi-keyword* ranked search, given its practicality and usability. In addition to this, some very specific schemes have been proposed that enable the insertion and deletion operations on encrypted documents. As the avid reader might have already realized, the overviewed paper mixes both methods, for there are but a few dynamic schemes that support efficient multi-keyword ranked search.

The authors of the aforementioned paper have designed a *"secure tree-based search scheme over the encrypted cloud data, which supports multi-key-word ranked search and dynamic operations on the document collection"* [1]. To put it in layman's terms, the paper proposes the implementation of a scheme that will enable the retrieval of information in the encrypted data without the necessity of decryption. This method is able to rank results according to their degree of similarity to the multiple words used as search keys. Furthermore, the basic structure on which the whole system is based upon is a somewhat standard

[1] This type of encryption is special because any operation done to the encrypted data will have been done to the data itself unencrypted without the need to decrypt it for the operation.

general tree data structure, which uses a custom definition of the Depth-first Search algorithm as a lookup mechanism.

Summary

After explaining all this, it is considered that the reader understands the paper at a high level. As as short summary, what the authors of the original paper aimed for and achieved were the following two points: (1) design a *"searchable encryption scheme that supports both the accurate multi-keyword ranked search and flexible dynamic operation on document collection"*, and (2) achieve a logarithmic search complexity, given the specific tree structure chosen by the implementers[2].

[2] Quantitatively, the search can be even faster, given the use of the Greedy Depth-first Search algorithm

References

[1] Z. Xia et al. "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data". In: *IEEE Transactions on Parallel and Distributed Systems* 27.2 (Feb. 2016), pp. 340–352. ISSN: 1045-9219. DOI: 10.1109/TPDS.2015.2401003.