TECHNOLOGY

Confidentiality Preserving Rank-Ordered Search

OVERVIEW

In the current information era, efficient and effective search capabilities for digital collections have become essential for information management and knowledge discovery. Meanwhile, a growing number of collections are professionally maintained in data centers and stored in encrypted form to limit their access to only authorized users in order to protect confidentiality and privacy. Examples include medical records, corporate proprietary communications, and sensitive government documents. An emerging critical issue that must be addressed is how to protect data collections and indices through encryption, while providing efficient and effective search capabilities to authorized users.

Consider the case when an authorized user remotely accesses the data collection to search and retrieve desired documents, the large size of the collections often makes it unfeasible to ship all encrypted data to the user's side, and then perform decryption and search on the user's trusted computers. Therefore, new techniques are needed to encrypt and organize the data collections in such a way as to allow the data center to perform efficient search in encrypted domain. Other scenarios exist where the content owner may want to grant a user limited access to search a confidential collection. For example, the searcher could be a scholar or a low-level analyst who wants to identify relevant documents from a private/classified collection, and may need clearance only for the top-ranked documents; the searcher could also be the opposing side during document discovery phase of a litigation, who would request relevant documents from the content owner's digital collection (say, emails) be turned over.

Researchers at the University of Maryland have developed techniques to securely rank-order the documents to retrieve the most relevant document(s) from an encrypted database based on search queries. The proposed methods are highly secure (relying on well-studied cryptographic encryption and hashing primitives), accurate (comparable to conventional searching systems designed for unencrypted data), and efficient (in terms of computational complexity and communication bandwidth). The proposed method maintains the confidentiality of the query as well as the content of retrieved documents.

The invention has a wide range of applications in numerous fields that would require or benefit from privacy-preserving search and multiple clearance/access levels. This includes, but is not limited to, out-sourcing data storage to third party service provider, data management in government and intelligent operations, enabling scholarly study of sensitive data, facilitating document discovery process in litigation. Overall, the invention forms a primary building block to facilitate various knowledge discovery tasks for which security and privacy is a major requirement.

This technology is patent pending. If you would like to review additional information or further discuss the technology with the inventors please contact the Office of Technology Commercialization at 301-405-3947 or otc@umd.edu.

CONTACT INFO

UM Ventures
0134 Lee Building
7809 Regents Drive
College Park, MD 20742
Email: umdtechtransfer@umd.edu
Phone: (301) 405-3947 | Fax: (301) 314-9502

Additional Information

INSTITUTION

University of Maryland, College Park

PATENT STATUS

Patent(s) pending

LICENSE STATUS

Contact OTC for licensing information

CATEGORIES

- Information Technology

EXTERNAL RESOURCES

IS-2007-098