Accepted Manuscript

Image search scheme over encrypted database

Jun Ye, Zheng Xu, Yong Ding

Please cite this article as: J. Ye, Z. Xu, Y. Ding, Image search scheme over encrypted database, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.02.045

Image Search Scheme over Encrypted Database

Jun Ye*[a,b], Zheng Xu[c], Yong Ding[d]

[a]*School of Mathematics and Statistics, Sichuan University of Science & Engineering, Sichuan, China*
[b]*Guangxi Key Laboratory of Cryptography and Information Security, Guangxi, China*
[c]*Tsinghua University, Beijing, China*
The Third Research Institute of the Ministry of Public Security, Shanghai, China
[d]*School of Computer Science and Information Security,*
Guangxi Key Laboratory of Cryptography and Information Security,
Guilin University of Electronic Technology, Guilin, Guangxi, China

Abstract

In big data era, too much ordinary information leakage may lead to the leakage of private information. Image search is widely used in many fields. Though there are many studies on image search, most of them are search in plaintext databases. In this paper, we study the image retrieval techniques over encrypted databases. A content-based retrieval scheme for encrypted images is proposed. In the scheme, a blind technique based on discrete logarithm problem is introduced to keep the privacy of feature vectors, and a novel retrieval way is used. This is a flexible scheme, which support fuzzy search. The client can control the search range. And in the search process, the contents of the original image will not be revealed.

Keywords: Cloud Computing, Image Retrieval, Feature Vector, Privacy

Highlights

- A novel request based image search model is proposed.

- A secure search pattern is generated, and a novel matching strategy is used.

- The search accuracy can be controlled by users.

- Low computation cost on the user side.

1. Introduction

Cloud computing, in which a large amount of computing resources and storage space are collected, can provide various computing and storage services for people. Due to the great power of cloud computing, it brings much convenience to the clients. The complex computation tasks can be easily solved by the cloud server. For purpose of saving the local storage space, the client can stores the files in the cloud servers, and retrieve the files when they are needed.

Cloud storage, in which a lot of storage devices are gathered together, is a kind of service-oriented distributed storage system, can provide the storage and management services of massive data. Such as space leasing, data storage, backup, sharing and so on. The data storage services are achieved through the application of cluster computing and distributed computation. In the system, a large number of different network resources will be used. In order to deal with the files anytime and anywhere, the clients usually store the files in the cloud.

Images are the common files people often use. With the help of the Internet, people can easily share the images for each other. The clients can retrieval the images that they needed from the cloud server conveniently. There are two main techniques for image search, text-based image retrieval (TBIR) and content-based image retrieval (CBIR).

TBIR is the traditional image retrieval method. The traditional text retrieval technology was tried to be used in the image retrieval. Thus, the image retrieval is the same as keyword search, which is a kind of matching search process based on keywords. The indices of a image are generated according to the understanding of the image, such as, name, number, size, contents description, image resource, author, time, storage location and so on. The image retrieval is actually becoming the corresponding text retrieval.

CBIR is the search technology based on feature vectors. The image is analyzed by software, and the content information is extracted. The information of color, shape and texture is combined together as the feature vectors, and stored in the feature database. For a given image, in the retrieval process, the feature vectors are extracted, and the similarity between the extracted vectors and the feature vectors stored in the database is computed by using the similarity matching algorithm. The retrieval result is output according to the value of similarity.

The main retrieval principle are the three points. The first one is, form a database retrieval model according to the requirement of the clients. The second one is, collect and process image resources, such as, feature extraction, analysis and indexing. The last one is, use similarity algorithm to retrieval the images

*Corresponding author
E-mail address: yejun@suse.edu.cn (J. Ye), Xuzheng@shu.edu.cn (Z. Xu), stone_dingy@126.com (Y. Ding)