http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (5): 824-831, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.824.831

Enabling Similarity Search over Encrypted Images in Cloud

[1,2]Yi Zhu, [1,2]Xingming Sun, [1,2]Zhihua Xia, [1,2]Li Chen, [1,2]Tao Li and [3]Daxing Zhang
[1]Jiangsu Engineering Center of Network Monitoring,
[2]School of Computer and Software, Nanjing University of Information Science and Technology,
Nanjing, 210044, China
[3]Hangzhou Dianzi University, Hangzhou, Zhejiang, China

Abstract: With the growing popularity of cloud computing, more and more users are willing to outsource their private data to the cloud. To ensure the security of data, data owners usually encrypted their private data before outsourcing them to the cloud server. Though data encryption improves the security of data, it increases the difficulty of data operating. This study focuses on the search of encrypted images in the cloud and proposes an efficient similarity retrieval scheme over encrypted images. The proposed scheme enables data owners to outsource their personal images and the content-based image retrieval service to the cloud without revealing the actual content of the image database to the cloud. The proposed scheme in this study supports the global feature based image retrieval methods under the Euclidean distance metric. Besides, rigorous security analysis and extensive experiments show that the proposed scheme is secure and efficient.

Key words: Similarity content-based image retrieval, locality sensitive hashing, secure data outsourcing

INTRODUCTION

Recent years, owning to strong data storage and management ability of the cloud, more and more data owners are likely to outsource their private data (such as emails, hospital records, personal medical images) to the cloud. Outsourcing data to the cloud saves software and hardware infrastructure costs for enterprises or individuals and also reduces the costs of managing software and hardware. Besides, data owners can easily share their personal data with authorized users by outsourcing data to the cloud.

However, cloud storage may bring security problems to the private data as data owners don't know where their data are stored and therefore lose control of them. Thus, the security of sensitive data has become the most concerned problem. To ensure the security of sensitive data, data owners usually encrypt their data before outsourcing to the cloud server. Data encryption improves the security of sensitive data to some extent, however, makes the data operation on sensitive data extremely complicated. It has a great need that the cloud server provides efficient data search service for the data owners to manage their sensitive data conveniently. Currently, many schemes about encrypted data search have been proposed However, these works mainly concentrated on encrypted text search and did not suit for similarity search over encrypted images. Therefore, an efficient similarity search scheme over encrypted images is desired.

This study proposes a content-based image retrieval scheme supporting similarity search over encrypted image in the cloud. We utilize p-stable Locality Sensitive Hashing (Datar *et al*., 2004) algorithm in our scheme for the construction of secure search index. Firstly, we can easily filter images with the search index and therefore, get similar images with the query image. Then, the cloud server compares the distance between the query image and the remaining images after filtering. Finally, top k images with smallest distance with the query image are returned to data users as results. The proposed scheme saves time by constructing an efficient search index based on LSH algorithm.

Contributions: For the first time, we propose a practical content-based image retrieval solution over encrypted images in cloud computing. Besides, we exploit p-stable LSH in content-based image retrieval. Our scheme is based on the global feature of images. By leveraging the computation power of the cloud, the proposed scheme costs low local computation while achieving high retrieval accuracy. In this work, we give rigorous security analysis and conclude that under certain security model our similarity image search scheme is secure. At last, extensive experiments are conducted and acquire satisfying experiment results and search efficiency.

Corresponding Author: Xingming Sun, No. 219, Ningliu Road, Pukou, Nanjing City, Jiangsu Province, 210044, China
Tel/Fax: 025-58731575

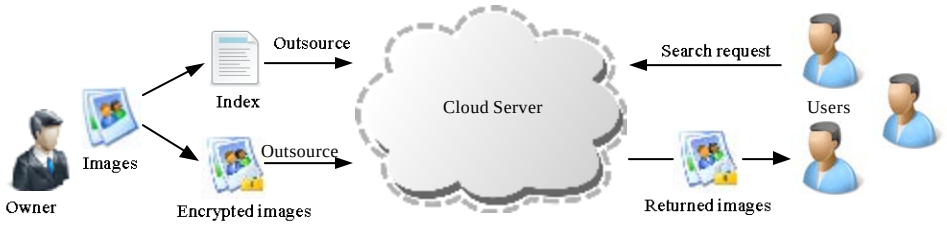

Fig. 1: System Architecture

PROBLEM FORMULATION

System model: A similarity search problem involves a collection of objects (documents, images, etc.) that are characterized by a collection of relevant features and represented as points in a high-dimensional attribute space. Given queries in the form of points in this space, we are required to find the nearest (most similar) object to the query. Our scheme is designed to not only support similarity search, but also prevent the leaking of information about the image database. In this study, we consider a cloud data system involving three different entities as illustrated in Fig. 1: Image owner, image user and cloud server.

Image owner has a collection of n images $M = (m_1, m_2,\ldots, m_n)$ that he wants to outsource to the cloud server in encrypted form while still keeping the capability to search through them for effective image utilization reasons. Image owners will first build a secure searchable index I from a set of n features $F = (f_1, f_2,\ldots, f_n)$ extracted from the image database M before outsourcing data to cloud server. Then, image owner encrypts all features with an invertible matrix R. Finally, image owner stores the secure index I, encrypted features and the encrypted image database M on the cloud server.

Image users are the authorized ones to use the images. We assume the authorization between the image owner and image users is appropriately done. For a given query image, an authorized user generates and submits a search request in a secret form a trapdoor $T(f_q)$ of the feature f_q and encrypted query feature f'_q to the cloud server. Upon receiving the search request $T(_q)$ and f'_q, the cloud server is responsible to search the index I and return the corresponding set of images to the user.

Cloud server stores the encrypted images and the index I for image owner and processes the query of image users. After receiving the query trapdoor $TD(f_q)$ and f'_q, cloud server compares the trapdoor $T(f_q)$ with the items in index I to return k most similar images. In the proposed scheme, cloud server is considered to be "honest but curious". The cloud server tries to learn more information during the search. So, it is necessary to design a secure similarity search scheme over encrypted images.

Design goals: To enable secure and efficient similarity search over encrypted images in cloud under the aforementioned model, our scheme should achieve the following design goals:

- **Accuracy:** The proposed scheme in this study should have high retrieval accuracy, achieving the same results as standard CBIR system
- **Efficiency:** Our scheme should reduce the computational complexity and communication overhead as far as possible
- **Security:** In the procedure of search, we should ensure the security of sensitive data without leaking information such as image databases and search index

BACKGROUND AND DEFINITIONS

The basic building block of our solution is p-stable locality sensitive hashing. In this section, we present an overview of Locality sensitive hashing and LSH based on p-stable distribution.

Locality sensitive hashing: Locality sensitive hashing is an approximate nearest neighbor algorithm used in high-dimension space (Indyk and Motwani, 1998; Gionis *et al*., 1999).

The definition of LSH is as follows.

Definition 1:A family $H = \{h : S \to U\}$ is called (r, cr, p_1, p_2)-sensitive if:

$$\begin{cases} \Pr\{h(x) = h(y)\} \geq p_1 & \text{for } D(x,y) \leq r \\ \Pr\{h(x) = h(y)\} \leq p_2 & \text{for } D(x,y) \geq cr \end{cases} \quad (1)$$

where, D(x, y) is the distance of feature vector x and y, the constant c>1 and probabilities $p_1>p_2$. We can learn from the definition that if the distance between two elements is less than r, then hash values of them have the possibility of p_1 to equal with each other; if the distance is larger than cr, then the possibility that their hash values are equal is less than p_2. As described in (Rajaraman and Ullman, 2011), we can easily construct a hash family satisfy the above conditions. Given a LSH family H as in Definition 1, in order to receive more accurate results, we can flexibly amplify the gap between the "high" probability p_1 and "low" probability p_2 by concatenating λ hash functions randomly selected from the hash family. In particular, for a specified λ, a new LSH family $G = \{g: S \rightarrow U^{\lambda}\}$ can be defined, where $g(x) = (h_1(x),...h_\lambda(x))$, $h_i \in \mathcal{H}$ which maps a d-dimensional feature vector to a λ-dimensional vector.

In the proposed scheme, we store each feature vector f in the buckets $g_i(f)$, i = 1, 2,..., s. To process a query f_q, we search all the buckets $g_1(f_q),..., g_L(f_q)$. All features that collision with f_q is gathered. Finally, top k images are returned as the query results.

Locality sensitive hashing based on p-stable distribution: In this study, we use a locality-sensitive hashing family based on p-stable distribution which works directly in Euclidean space and is quite easy to implement (Datar *et al*., 2004). The p-stable distribution can be defined as follows.

Definition 2: A distribution D is called p-stable, if exists $p \geq 0$ such that for any n real numbers $v_1, v_2,... v_n$ and variables $X_1, X_2,... X_n$ with distribution D, the random variable $\sum_i v_i X_i$ and variable $(\sum_i |v_i|^p)^{1/p} X$ has the same distribution, where X is a random variable with distribution D.

It is proved (Zolotarev, 1986) that p-stable distributions exist for any $p \in (0,2]$, a Gaussian distribution, defined by the density function:

$$G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

is 2-stable. The final LSH $h: \mathcal{R}^d \rightarrow \mathcal{N}$ maps a d dimensional vector v onto the set of integers and is defined as follows:

$$h_{a,b}(v) = \left\lfloor \frac{a \cdot v + b}{r} \right\rfloor \quad (2)$$

where, a is a d dimensional vector whose elements are chosen independently from Gaussian distribution and b is a real number chosen uniformly from the range [0,r]. Then, the LSH h is used to generate LSH family $G = \{g = (h_1,... h_\lambda)\}$.

PROPOSED SCHEME

In this section, we describe our encrypted image search scheme in two phase. To search similar images outsourced to the cloud, data owner should construct a secure index and outsource it along with the encrypted images to the cloud server. Then, server performs search on the index according to the requests submitted by the data users. The cloud server learns nothing about the query request or the image databases itself.

Setup phase: In the setup phase, image owner needs to build a secure index and encrypt the images. Then, the index and the encrypted images are uploaded to the cloud.

Key generation: Firstly, the image owner generates private keys k_j (j = 0, 1, 2,..., L) and k_m. In particular, k_j is the secret key for the encryption of keywords $B_{i,j}$ and k_m is the secret key for the encryption of image database. Let k_j, $k_m \in \{0,1\}^{\psi}$ be the secret keys of size ψ. Besides, a $(l+1)\times(l+1)$ invertible matrix R is generated to encrypt the feature vectors.

Feature extraction: For each image f_i in the image database, we extract a global feature m_i from it. So, every image m_i in the image database has one corresponding global feature f_i.

Secure index construction: Features of the images are high-dimensional vectors and p-stable LSH scheme can work directly on these vectors in the Euclidean space without embedding them. Learning from the p-stable LSH algorithm, we have a locality sensitive function:

$$h_{a,b}(v) = \left\lfloor \frac{a \cdot v + b}{r} \right\rfloor$$

The data owner constructs the hash function $g(f_i) = (h_1(f_i), h_2(f_i),..., h_\lambda(f_i))$ where h_i is randomly chosen from hash family H. Then, the data owner applies $g(f_i) = (h_1(f_i), h_2(f_i),..., h_\lambda(f_i))$ to all global features $\{f_i\}_{i=1}^n$ so as to build the search index. Let $\{B_i\}_{i=1}^N$ denote the derived set of LSH hash values, where, N refers to the total number of cluster. In fact, the $g(f_i)$ function maps a d-dimensional feature vector to a λ-dimensional vector. So, each $B_{i,j}$ is a λ-dimensional vector. However, in general, such LSH function is not necessary to have one-way property. Therefore, it is necessary to encrypt the keywords $B_{i,j}$ in

Table 1: Secure search index

Variables	Description
$Enc(k_j, B_{1,j})$	$(ID(m_1),, f'_1), (ID(m_5),, f'_5), (ID(m_9),, f'_9), (ID(m_{13}),, f'_{13})\ldots$
$Enc(k_j, B_{2,j})$	$(ID(m_2),, f'_2), (ID(m_6),, f'_6), (ID(m_{10}),, f'_{10}), (ID(m_{14}),, f'_{14})\ldots$
$Enc(k_j, B_{3,j})$	
…	…
$Enc(k_j, B_{mi,j})$	$(ID(m_2),, f'_3), (ID(m_7),, f'_7), (ID(m_{11}),, f'_{11}), (ID(m_{15}),, f'_{15})\ldots$

the hash table for the security of them before outsourcing to the cloud. Here, we utilize k_j as the secret key, let $Enc: \{0,1\}^{\ell_h} \times \kappa \to \{0,1\}^{\ell_h}$ be a pseudorandom permutation.

Global features are used to compare the distance between images in the search phase. So, we must encrypt global features in such way that the cloud server is able to compute the distance of them. Firstly, each feature vector $f_i = (f_{1,i}, \ldots f_{l,i})^T$ is modified as:

$$\hat{f}_i = (f_{1,i}, \cdots, f_{\ell,i}, \|f_i\|_2^2)^T$$

Then, the data owner randomly picks $a(\ell+\ell)\times(\ell+\ell)$ invertible matrix R to encrypt the modified feature vector as $f'_i = R^T \cdot \hat{f}_i$.

We assume every image m_i has an unique identifier $ID(m_i)$, $(ID(m_i), f'_i)$, $(ID(m_j), f'_j)\ldots$, $(ID(m_k), f'_k)$ be a list of image identifiers and their corresponding modified feature vector. Finally, one search index I is as shown in Table 1. Furthermore, to increase the clustering accuracy, we repeat the index construction process L times by generating L hash tables as Table 1.

Upload: After constructing the index, image owner encrypts the image database with the key k_m and sends all encrypted images, encrypted features along with the search index (Table 1) to the cloud server for search purpose. Though the encrypted images outsourced, authenticated users should be able to retrieval images from the cloud server. To reach this goal, image owner should share some information with authenticated users:

k_j(j = 0, 1, 2,…, L): Secret key for keyword $B_{i,j}$ encryption
k_m: Secret key for image database encryption
R: Secret matrix
g: p-stable LSH function used in the index construction

Search phase: In search phase, image user wants to retrieve images that are similar to one query image from the cloud server. In order to avoid the information leaking, image user generates a secure trapdoor with the query image. Then, the trapdoor is submitted to the cloud server. With the trapdoor, the cloud server returns k most similar images by searching on the Index.

Trapdoor generation: The authorized user first needs to extract a global feature f_q from the query image when he wants to search images similar to the query image. Then, we apply L p-stable LSH functions $g(f_q) = (h_q(f_i), h_2(f_q),\ldots, h_\lambda(f_q))$ on the feature f_q and therefore, generate L hash values for the feature f_q. Finally, to protect the security of the trapdoor, the data user applies pseudorandom permutation Enc on each hash value such that $T(f_q) = Enc(k_1, g_1(f_q)), Enc(k_2, g_2(f_q)),\ldots, Enc(k_L, g_L(fq)))$. Besides, for a query feature vector $f_q = (f_{1,q},\ldots, f_{l,q})$, the data user first constructs a modified vector as:

$$\hat{f}_q = (-2f_{1,q}, -2f_{2,q}, \cdots, -2f_{\ell,q}, 1)^T$$

He next chooses a random positive value $r \in \mathbb{R}$ and uses it with the secret matrix R to encrypt the modified query vector as $f'_q = rR^{-1} \cdot \hat{f}_q$. At last, the data user sends $T(f_q)$ along with the f'_q to the cloud server.

Search: Once receiving a search request, the cloud server performs search on the index for each component of the $T(f_q)$. Then, the cloud server gets an identifier list of images corresponding to every component of $T(f_q)$. The global feature extracted from the image is mapped to L hash tables during the index period. So, during search period, the global feature extracted from the query image is also mapped to L hash tables with L hash functions. If the query image and one specific image in the collection satisfies the condition that $g(f_q) = g(f_i)$, then they are considered to be similar. Finally, cloud server gathers all these image identifiers that have the same hash value with the query image. However, in order to get more accurate results, the cloud server need to compute the distance between the query image and all these similar images. The cloud server conducts the scalar product:

$$f_q'^T \cdot f'_i = (rR^{-1} \cdot \hat{f}_q)^T \cdot R^T \hat{f}_i = r(\hat{f}_q)^T \cdot \hat{f}_i = r(\|f_i\|_2^2 - 2\sum_{j=1}^{\ell} f_{j,i} f_{j,q}) = r(\|f_q - f_i\|_2^2 - \|f_q\|_2^2) \quad (3)$$

The distance $\|f_q\text{-}f_i\|^2_2$ is hidden by the secret scalar r and the unknown $\|f_q\|^2_2$. When $x \geq 0$, the function $f(x) = x^3$ is order preserving, i.e., $f(x_1)>f(x_2)$ implies $x_1>x_2$. Also, because for each query f_q the values of $\|f_q\|^2_2$ and r is fixed, the cloud server can directly find closest feature vectors by simply sorting out the set of scalar products $r(\|f_q\text{-}f_i\|^2_2\text{-}\|f_q\|^2_2)$, without knowing the sensitive information from the feature vectors.

After computing the distances, the cloud server ranks these images according to their distance with the query image. Finally, the cloud server sends the top k most similar encrypted images back to the data user as search results.

Once receiving the encrypted images returned by the cloud server, data user decrypts these images with the secret key k_m shared by the data owner and obtains the plain images similar to the query image. Now, the round of search is completed.

SECURITY ANALYSIS

If we want to perfect protect the privacy of sensitive data, the computational complexity of cloud server is at least O(n). To make the search more efficient, we inevitably reveal some information to the cloud server. In the process of our query, we firstly filter images that are not similar to the query image using search index. While constructing index, a pseudo-random permutation is used to encrypt hash values in the index. However, the cloud server knows that images in the same bucket are similar to each other. In addition, though the server can't learn additional information about the query image itself or returned results, he/she learns that all these returned images are similar to query image and to each other. The above information leakage is a compromise for efficiency.

In our scheme, image database, search index, feature vector and search request are encrypted before sent to the cloud server. The image database is encrypted with AES algorithm, AES algorithm is CPA secure. So, the image database is secure under AES encryption algorithm. The keywords in search index are encrypted with pseudo-random permutation. And keywords in different index are encrypted with different key, so keywords are indistinguishable and secure. In addition, image features are encrypted with invertible matrix before sent to the cloud server, therefore, feature vectors are also secure. Above all, except for some inevitably information leakage, the proposed scheme in this study is secure.

EXPERIMENT RESULTS

In this section, we describe the details of our experiments and analysis the experimental results of our scheme. And an image database consisting of 130000 images is used to test the performance of the proposed scheme.

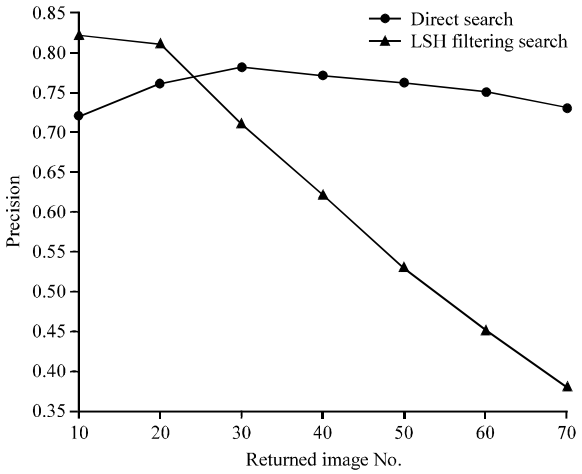

Fig. 2: Comparison of search precision

Result accuracy: This criterion is used to evaluate the correction of the returned results, is evaluated by precision defined as follows:

$$\text{Precision} = \frac{\text{num}_{\text{hit}}}{\text{num}_{\text{return}}} \quad (4)$$

where, $\text{num}_{\text{return}}$ is the number of the returned images, num_{hit} is the number of correct similar images. The accuracy of the scheme is mainly decided by the feature extraction method in common image retrieval systems. The accuracy of a query in our scheme is shown in the Fig. 2. In our experiment, we compare the retrieval performance of our proposed scheme with the direct distance computation scheme. The direct distance computation scheme compares the distance between different feature vectors as shown in Eq. 3. The direct search scheme linearly scans all images in the database and finds the similar images with the query image.

Time complexity: In theory, the LSH filtering search scheme should search faster than the direct search scheme since the direct search time is linear to the total number of image in the database. The experiment also shows the prospective results as the theoretical analysis as shown in Fig. 3. The experiment results illustrate that the LSH filtering scheme efficiently improves search efficiency and saves search time. We find that search time of different returned image number is same. That is because we always compute the distance of all remaining images with the query image and finally return specific number of images to the user.

Though our LSH scheme does not improve the precision, it reduces the computational costs of similarity

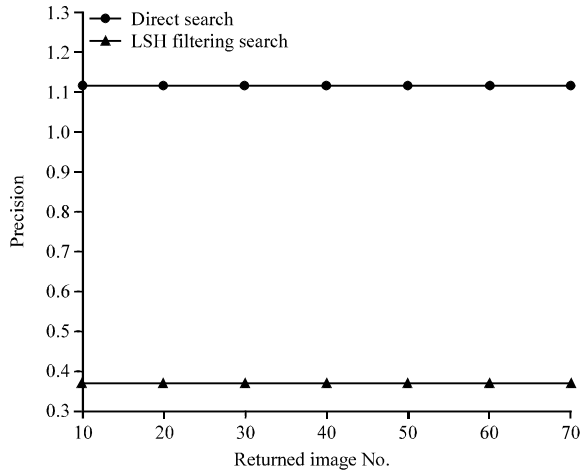

Fig. 3: Comparison of search time

image search by filtering images using search index. The direct search scheme finds the similar images by linearly scanning all images in the database. So the time complexity of the direct search scheme is O(n) which n is the total number of image in the database. However, our scheme saves search time by exploiting LSH algorithm. The experiment result shows that LSH algorithm makes sense for similarity image search. Our work aims to prove that LSH is suit for similarity search over encrypted image and eventually gets the prospective results. To further improve the retrieval accuracy, we will exploit local image features in our future work.

RELATED WORK

Searchable Symmetric Encryption (SSE) on text domain has been widely studied in the literature. The practical searchable encryption was first proposed by Song *et al.* (2000). Song *et al.* (2000) described their cryptographic schemes for searching on encrypted data and proved the security of the proposed scheme. Goh (2003) defines the secure index which possesses the characteristics of semantically security against adaptive chosen keyword attack (IND-CKA). They use the Bloom filters and pseudo-random functions to construct the secure index, for which the searching complexity is proportional to the number of files containing this keyword in the collection. Curtmola *et al.* (2006) propose a keyword based scheme, in which an encrypted hash table index is built for each single file. This scheme is very efficient and there is constant search complexity for each returned file. It also can achieve adaptive SSE security.

However, these proposed keyword based schemes do not enable similarity search. In order to enhance the search flexibility and usability, some researchers proposed works to support similar keyword search which could tolerate typing errors (Li *et al.*, 2010; Chuah and Hu, 2011; Wang *et al.*, 2012a). Li *et al.* (2010) for the first proposed fuzzy keyword search over encrypted cloud data in cloud computing. They exploit edit distance to quantify keyword similarity. Fuzzy keywords tolerates errors to some extent, it is only applicable to strings under edit distance. If we have long words, then the fuzzy keywords set is very big which is inefficient for search. On the other hand, some of the works focused on multi-keyword searches which could return more accurate results ranked according to some predefined criterions (Golle *et al.*, 2004; Boneh and Waters, 2007; Katz *et al.*, 2008; Cao *et al.*, 2011; Wang *et al.*, 2012b; Xu *et al.*, 2012a, b; Sun *et al.*, 2013).

Kuzu *et al.* (2012) proposed an scheme for similarity search over encrypted data. They utilize locality sensitive hashing for fast near neighbor search. Their index is constructed based on LSH algorithm. Their search scheme supports similarity search over encrypted data. Their work is more applicable to keyword search and is not suit for similar image retrieval. So far, all these keyword based schemes are not suit for image retrieval because of image's high dimensionality.

Lu *et al.* (2009) proposed a search scheme over encrypted multimedia databases. Visual words are extracted from images and build indexes based on them. They considered two indexing schemes, inverted index and min-Hash and perform secure image retrieval over encrypted image databases. This work is not suit for other image features except visual words and their index makes the search result less accurate. In short, secure, efficient and accurate search over encrypted images is still an open problem.

CONCLUSION

In this study, we proposed an efficient similarity search scheme over encrypted images in the cloud. We exploit the widely used p-stable locality sensitive hashing for the construction of index in our scheme. P-stable LSH algorithm is often used in high dimension spaces and is used to construct secure search index in this study. Our scheme enables efficient similarity image search and less time consuming. In the cloud computing, it is critical to prove the security of outsourced images. This study gives a rigorous security definition and proved the security of the proposed scheme under the provided security model. To verify the proposed scheme, we conduct experiments on a image library with 130000 images. Experiment results show that our scheme is suit

for similarity search over encrypted images. In short, our works have a great significance to the further development of similarity CBIR in the cloud.

ACKNOWLEDGMENTS

This study is supported by the NSFC (61232016, 61173141, 61173142, 61173136, 61103215, 61373132, 61373133), National Basic Research Program 973 (2011CB311808), 2011GK2009, GYHY201206033, 201301030, 2013DFG12860, SBC201310569 and PAPD fund.

REFERENCES

Boneh, D. and B. Waters, 2007. Conjunctive, subset and range queries on encrypted data. Proceedings of the 4th Theory of Cryptography Conference, February 21-24, 2007, Amsterdam, The Netherlands, pp: 535-554.

Cao, N., C. Wang, M. Li, K. Ren and W. Lou, 2011. Privacy-preserving multi-keyword ranked search over encrypted cloud data. Proceedings of the 30th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, April 10-15, 2011, Shanghai, China, pp: 829-837.

Chuah, M. and W. Hu, 2011. Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data. Proceedings of the 31st International Conference on Distributed Computing Systems Workshops, June 20-24, 2011, Minneapolis, MN., pp: 273-281.

Curtmola, R., J.A. Garay, S. Kamara and R. Ostrovsky, 2006. Searchable symmetric encryption: Improved definitions and efficient constructions. Proceedings of the 13th ACM Conference on Computer and Communications Security, October 30-November 3, 2006, Alexandria, USA., pp: 79 -88.

Datar, M., N. Immorlica, P. Indyk and Y.S. Mirrokni, 2004. Locality-sensitive hashing scheme based on p-stable distributions. Proceedings of the 20th Annual Symposium on Computational Geometry, June 8-11, 2004, Brooklyn, NY., USA., pp: 253-262.

Gionis, A., P. Indyk and R. Motwani, 1999. Similarity search in high dimensions via hashing. Proceedings of the 25th International Conference on Very Large Data Bases, September 7-10, 1999, Edinburgh, Scotland, UK., pp: 518-529.

Goh, E.J., 2003. Cryptology eprinta archive: Report 2003/216. Secure indexes. October 2003. http://eprint.iacr.org/2003/216.

Golle, P., J. Staddon and B. Waters, 2004. Secure conjunctive keyword search over encrypted data. Proceedings of the 2nd International Conference on Applied Cryptography and Network Security, June 8-11, 2004, Yellow Mountain, China, pp: 31-45.

Indyk, P. and R. Motwani, 1998. Approximate nearest neighbors: Towards removing the curse of dimensionality. Proceedings of the 30th Annual ACM Symposium on Theory of Computing, May 24-26, 1998, Dallas, TX., USA., pp: 604-613.

Katz, J., A. Sahai and B. Waters, 2008. Predicate encryption supporting disjunctions, polynomial equations and inner products. Adv. Cryptol., 4965: 146-162.

Kuzu, M., M.S. Islam and M. Kantarcioglu, 2012. Efficient similarity search over encrypted data. Proceedings of the IEEE 28th International Conference on Data Engineering, April 1-5, 2012, Washington, DC., pp: 1156-1167.

Li, J., Q. Wang, C. Wang, N. Cao, K. Ren and W. Lou, 2010. Fuzzy keyword search over encrypted data in cloud computing. Proceedings of the 9th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, March 15-19, 2010, San Diego, CA., USA., pp: 1-5.

Lu, W., A. Swaminathan, A.L. Varna and M. Wu, 2009. Enabling search over encrypted multimedia databases. Proceedings of the IS and T/SPIE Electronic Imaging, Volume 7254, January 18, 2009, International Society for Optics and Photonics.

Rajaraman, A. and J.D. Ullman, 2011. Mining of Massive Datasets. Cambridge University Press, UK., ISBN-13: 978-1107015357, Pages: 326.

Song, D., D. Wagner and A. Perrig, 2000. Practical techniques for searches on encrypted data. Proceeding of the IEEE Symposium on Security and Privacy, May 14-17, 2000, Berkeley, CA., USA., pp: 44-55.

Sun, W., B. Wang, N. Cao, M. Li, W. Lou, Y.T. Hou and H. Li, 2013. Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking. Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, May 8-10, 2013, Hangzhou, China, pp: 71-82.

Wang, C., K. Ren, S. Yu and K.M.R. Urs, 2012a. Achieving usable and privacy-assured similarity search over outsourced cloud data. Proceedings of the IEEE INFOCOM, March 25-30, 2012, Orlando, FL., pp: 451-459.

Wang, C., N. Cao, K. Ren and W. Lou, 2012b. Enabling secure and efficient ranked keyword search over outsourced cloud data. IEEE Trans. Parallel Distib. Syst., 23: 1467-1479.

Xu, J., W. Zhang, C. Yang, J. Xu and N. Yu, 2012a. Two-step-ranking secure multi-keyword search over encrypted cloud data. Proceedings of the International Conference on Cloud and Service Computing, November 22-24, 2012, Shanghai, pp: 124-130.

Xu, Z., W. Kang, R. Li, K. Yow and C.Z. Xu, 2012b. Efficient multi-keyword ranked query on encrypted data in the cloud. Proceedings of the IEEE 18th International Conference on Parallel and Distributed Systems, December 17-19, 2012, Singapore, pp: 244-251.

Zolotarev, V.M., 1986. One-Dimensional Stable Distributions. Vol. 65, American Mathematical Socity, USA., ISBN: 9780821898154, Pages: 284.