Cybersecurity

RESEARCH Open Access

A privacy-preserving image retrieval scheme with access control based on searchable encryption in media cloud

Miao Tian[1,2], Yushu Zhang[1,2*], Yongming Zhang[1,2], Xiangli Xiao[1,2] and Wenying Wen[3]

Abstract

With the popularity of the media cloud computing industry, individuals and organizations outsource image computation and storage to the media cloud server to reduce the storage burden. Media images usually contain a large amount of private information. To prevent disclosure of privacy of the image owners, media images are encrypted before uploading to the server. However, this operation will greatly limit the utilization of the image for the user, such as content-based image retrieval. We propose an efficient similarity query algorithm with access control based on Bkd-tree in this paper, in which a searchable encryption scheme is designed for similarity image retrieval, and the encrypted image is used to extract image features by a pre-trained CNN model. The Bkd-tree is utilized to generate an index tree for the image features to speed up retrieval and make it faster than linear indexing. Finally, the security performances of the proposed scheme is analyzed and the performance of this scheme is evaluated by experiments. The results show that the security of the image content and image features can be ensured, and it has a shorter retrieval time and higher retrieval efficiency.

Keywords Privacy-preserving, Searchable encryption, Content-based image retrieval, Access control, Bkd-tree

Introduction

With the growing ubiquity of electronic products, such as smartphones, tablets, and digital cameras, millions of media images are generated every moment, and these images occupy a large amount of memory when storing them on devices. Due to the limited storage of electronic devices, a great many image owners choose to store numerous images in the media clouds, e.g., Ali Cloud, Baidu Cloud, Tencent Cloud, and Apple iCloud. If unencrypted images uploaded to the cloud are leaked, there are many privacy issues for the image owner. To protect the privacy of the image owner, many image encryption schemes (Wang et al. 2023; Yang and Wu 2022; Zheng and Zeng 2022; Paul et al. 2022) have been proposed by researchers. For big encrypted image datasets, effective image retrieval becomes a universal demand. Image retrieval schemes are divided into content-based image retrieval (CBIR) (Huang et al. 2020; Xw et al. 2021; Bella and Vasuki 2019) and text-based image retrieval (TBIR) (Ahamd and Jang 2003; Li et al. 2011; Samet et al. 2016; Zaidi et al. 2019). TBIR schemes need a text label to describe an image, and it is not easy to find the appropriate text label. Besides, generating image labels will take a large amount of time. Therefore, TBIR schemes have low retrieval accuracy and slow retrieval speed. Researchers have constructed some CBIR schemes (Raveendra and Vinothkanna 2019; Magdy et al. 2019; Shamna et al. 2018; Majhi and Mallick 2022; Shamna et al. 2022), which can

*Correspondence:
Yushu Zhang
yushu@nuaa.edu.cn
[1] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, China
[2] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
[3] School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330032, China

reduce the retrieval time and improve the retrieval accuracy. CBIR schemes can be separated into two types on the basis of image features, i.e., encryption-image-based (Yang et al. 2022; Zhu et al. 2022) and encryption-feature-based (Li et al. 2022; Xia et al. 2022). Because features of plaintext images contain private information, the extracted features are generally encrypted for retrieval. In addition, feature encryption methods are required to support retrieval operations. Homomorphic encryption is an encryption function that satisfies the plaintext operation is equivalent to the ciphertext operation. However, it requires a large amount of computing power, and image owners must perform heavy computing tasks.

Searchable encryption (Song et al. 2000) can be introduced to realize fast image search over cipher-text domain without any loss of data confidentiality (He et al. 2021; Li et al. 2020). It can be seen that this great method is more effective and feasible in practical application. Images are encrypted using searchable encryption and then the image owner uploads encrypted images to the server. The extraction of image features and similarity calculation are completed by the server, which reduces the computational burden of the image owner and the user. For some special images, such as medical images, police case images, and military project images, special people are usually authorized access to specific partial data. Therefore, similarity query needs to be implemented with access control.

Contribution. This paper proposes a novel CBIR scheme based on searchable encryption and Bkd-tree. The main contributions can be summarized as follows:

- An image retrieval scheme based on searchable encryption is proposed to realize CBIR with access control based on Bkd-tree, which effectively solves the challenge of encrypted image retrieval. The proposed scheme not only retrieves users' desired images efficiently, quickly, and accurately, but also realizes access controlled image retrieval.
- The security of this scheme is analyzed in the field of image content and features, query results and query requests, and searchability with access control. In view of the characteristics of Bkd-tree and the adopted cryptology mechanism of searchable encryption, the proposed scheme effectively protects the privacy of image owners and users, while enabling efficient image retrieval.
- The performance of this scheme is evaluated using the Corel image dataset, and the results confirm its high search precision and efficiency. Additionally, the experimental results show the feasibility of this scheme.

The rest of this paper is organized as follows. We reviewed the related work in "Related work" section, and the system and threat models are described in "Models and design goal" section. In "Preliminary" section, we give a brief introduction of preliminary background. The proposed scheme and corresponding security analysis are described in "The image retrieval scheme" section and "Security analysis" section, respectively. The actual performance of this scheme is evaluated and the novel CBIR scheme is concluded in the final two sections.

Related work

Since the 1970s, people began to investigate image retrieval techniques, which mainly involved text-based image retrieval methods. The image owner usually generates a label to describe an image, and the user utilizes the label to retrieve the image. However, the process of text-based image retrieval can be cumbersome, as it requires a time-consuming labeling process, and the retrieval accuracy is entirely reliant on the accuracy of these labels. To improve the retrieval accuracy and speed up the retrieval process, some content-based image retrieval methods have been proposed.

Encrypted feature retrieval

In this class of methods, the image owner first extracts the image features, and then carries out the retrieval operation. However, the features extracted by this method are plaintext features, which will expose the search intention of the query user. To protect the privacy of the query user, Lu et al. (2009) constructed the first CBIR scheme based on encrypted feature, which utilizes the min-hash algorithm to compare the Jaccard similarity between the visual word representations of two images. To protect image features, Lu et al. (2009) encrypted the features by three methods, including random projections, random unary encoding, and bitplane randomizations in their next work. However, this scheme results in low retrieval accuracy. Xia et al. (2017) investigated a privacy-protected CBIR model that encrypts features using the KNN method and adopts K-means and LSH to improve efficiency. Weng et al. (2016) designed a privacy-protected CBIR scheme using partial encryption. However, features are extracted from the unencrypted images by the attacker, which can cause serious privacy leakage. Xia et al. (2018) used Scale-Invariant Feature Transformation and the BOW model to describe images and calculated the Earth Mover's Distance to measure similarity between two corresponding images. Nevertheless, the aforementioned scheme involves extraction of image features and encryption of both images and features, which brings a huge computational burden on the image owner. Researchers proposed

CBIR schemes based on encrypted images to solve this problem, in which the image owner only encrypts the images and uploads the ciphertext images to the servers, and feature extraction and similarity calculation are outsourced to the server.

Encrypted image retrieval

Images usually contain sensitive private information, and CBIR scheme (Shashank et al. 2008) that directly exposes unencrypted image database to the cloud server can suffer from information leakage. Therefore, it is extremely necessary to encrypt images before storing them on the cloud server. Researchers have proposed various image retrieval schemes to protect the privacy of image owners. Ferreira et al. (2019) designed a subcontracted image storage and retrieval framework using the BOVW model, in which the Hamming distance of the histogram is adopted to evaluate the similarity among images. However, this framework may reveal some additional information. Xia et al. (2019) proposed an outsourcing scheme for privacy protection in the field of industrial internet of things based on the Local Binary Pattern, which firstly uses order-preserving encryption technology to encrypt images. In this scheme, the cloud server extracts the LBP value directly from the encrypted image, and calculates the histogram of the LBP value, and then normalizes it with the total number of LBP values. Based on color histograms and ac-coefficients, Xia et al. (2019) put forward a Manhattan-distance-based image retrieval scheme, in which the DC coefficient obtained by discrete cosine transform of Y component is encrypted by stream cipher technology, and the AC coefficient and the other two color components are encrypted by value permutation and position scrambling. However, the retrieval precision of this scheme is relatively low. Xia et al. (2021) devised a novel image search scheme using secure LBP to improve search accuracy, in which the Manhattan distance is utilized to measure the similarity. This scheme achieves high security and retrieval accuracy. However, none of the above schemes can realize data access control and access mode privacy protection.

Models and design goal

In this section, we first introduce the design goal, as well as the security and the system model.

System model

In this scheme, a privacy-preserving image retrieval scenario is considered. As shown in Fig. 1, the system model consists of three entities: the cloud server, as well as the query user and the image owner. Their respective roles are elaborated as follows.

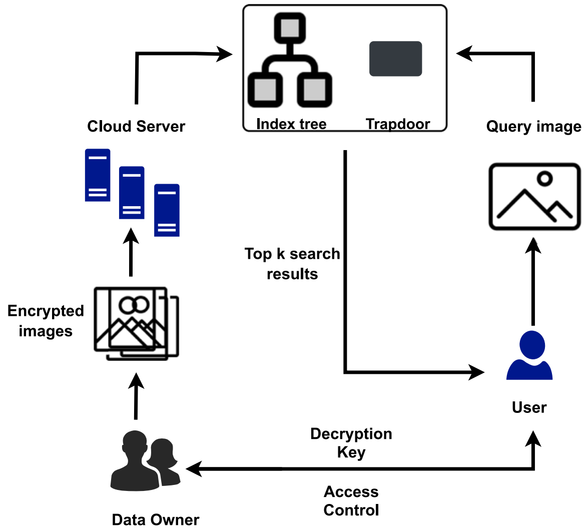

Fig. 1 System model

- *Image owner*. The image owner generates the image dataset and encrypts the images. Namely, the image owner runs specific encryption algorithm to generate the encrypted images and uploads the encrypted results to the cloud to reduce the storage burden.
- *Query user*. The query user retrieves similarity images according to the query image. In addition, the query user decrypts the retrieval result to obtain the plaintext image. During the image search process, the query image is encrypted by the query user using an encryption algorithm. Afterwards, the query user uploads it to the server for a similar image search. When the query results are sent to the query user by the server, the query user decrypts the images to get the plaintext images.
- *Cloud server*. In addition to storage, cloud servers can perform search operations. Specifically, the cloud server first extracts the features of each image with a pre-trained convolutional neural network. Afterwards, the cloud server constructs an index tree with the Bkd-tree and calculates the distance between the feature of images. Finally, the top-k similarity images are returned to the query user.

Security model

The cloud server is considered as honest-but-curious in this model, i.e., the cloud server honestly stores images for data owners and provide retrieval services with access control for the users. On the other hand, the server may be curious about the query intent of the users and the image itself. Moreover, the users are assumed to be honest, indicating that they honestly follow the proposed scheme to launch similarity retrieval requests to the cloud.

Design goal

In this scheme, a scheme is designed to support users in efficiently and accurately retrieving similar images, but also takes into account the access pattern of similarity retrieval. The design goals of the proposed scheme can be specified in three aspects.

- *Privacy preservation*. Both the image feature and image content should be preserved in a CBIR scheme. In addition, the privacy of query requests and query results should be properly protected.
- *Retrieval accuracy*. This scheme adopts searchable encryption technique to ensure the accuracy of retrieval, which supports similarity retrieval on encrypted data and outsources as much work as possible to the server.
- *Efficiency*. Image features are extracted and an index tree is built according to the Bkd-tree by the server in this scheme. We make full use of the index tree to improve the query efficiency.

Preliminary

In this section, we briefly introduce the techniques, namely the Bkd-tree, hidden vector based on access structure, and searchable encryption technique.

Bkd-tree

The Bkd-tree is a data structure based on the kd-tree structure for searching multi-dimensional data (Procopiuc et al. 2003). We give a simple example to illustrate the concrete process of the Bkd-tree. Let $D = \{(x_i, y_i)\}_{i=1}^{8}$ be a dataset containing eight records, where (1, 2); (7, 9); (5, 7); (3, 8); (6, 4); (7, 1); (2, 3); (8, 2).

Step 1. Obviously, the difference between the minimum value and maximum value of X dimension is 7, while the difference in the Y dimension is 8. Therefore, the splicing dimension of the current node is Y dimension.

Step 2. Sort the eight point data based on the the value of the Y dimension to generate left and right subtrees, and the order from smallest to largest is as follows: (7, 1); (1, 2); (8, 2); (2, 3); (6, 4); (5, 7); (3, 8); (7, 9).

Step 3. Divide the dataset into left and right subtree datasets. The current number of nodes is 8, and the sorted point data is split such that the first half belongs to the left subtree, while the remaining data belongs to the right subtree.

Left subtree: (7, 1), (1, 2), (8, 2), (2, 3);

Right subtree: (6, 4), (5, 7), (3, 8), (7, 9).

Follow the above steps repeatedly until both left and right subtrees only contain a single data pair. Finally, an index tree of image feature vectors is constructed as shown in Fig. 2.

Hidden vector based on access structure

The framework is an access control based on attribute, which is defined as $\kappa = (\kappa_1, \kappa_2, \cdots, \kappa_m) \in (\Upsilon \cup *)^m$, where Υ represents all attributes and $*$ represents the don't care attribute. In addition, $(\Upsilon \cup *)^m$ denotes the domain of m-dimensional vectors over $(\Upsilon \cup *)$. An attribute vector is defined as $\tau = (\tau_1, \tau_2, \cdots, \tau_m) \in \Upsilon^m$, where Υ^m is the domain of m-dimensional attribute vectors. Let $\mathbf{M}(\kappa, \tau)$ represent the relationship of matching between κ and τ, where

$$\mathbf{M}(\kappa, \tau) = \begin{cases} 1 & \textbf{if} \quad \kappa_i = \tau_i \quad \textbf{or} \quad \kappa_i = *, \quad 1 \leq i \leq m; \\ 0 & \textbf{otherwise}. \end{cases} \tag{1}$$

Searchable encryption

Searchable encryption (SE) is a method for searching encrypted data that supports controlled and hidden search, as well as query isolation, achieving simple and fast retrieval without increasing space and communication overhead. The SSE algorithm defined on a dictionary $\Delta = \{W_1, W_2, \cdots, W_\iota\}$ can be described as a tuple of several algorithms, namely SSE = (**KeyGen**, **Enc**, **Trapdoor**, **Search**, **Dec**).

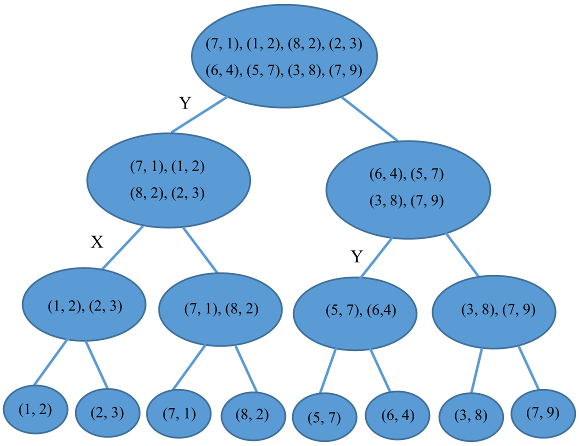

Fig. 2 An example of Bkd-tree

- $K = \textbf{KeyGen}(\lambda)$. It is an algorithm for key generation, in which the secret key K is obtained based on the security parameters λ.
- $(I, C) = \textbf{Enc}(K, D)$. It is an encryption algorithm that takes as input the secret key set K and the image collection D, and outputs the encrypted image collection C and the index I, where I is defined as ϕ if there is no need to construct an index.
- $T_W = \textbf{Trapdoor}(K, W)$. It is a trapdoor generation algorithm, in which the trapdoor T_W is generated by the secret key set K and the keyword W.
- $D_W = \textbf{Search}(I, T_W)$. It is a search algorithm that takes as input the index I and the query trapdoor T_W, and outputs the query result D_W containing the keyword W.
- $D = \textbf{Dec}(K, C)$. It is an image decryption algorithm, in which the set of decrypted files D is generated by the secret key K and the encrypted files C.

Content-based image retrieval

The framework of CBIR technology mainly includes image feature extraction and similarity measurement. Just as the name implies, feature extraction is to extract distinguishing feature information from images, and similarity measure is to calculate the distance between feature of query image and features in the feature database, and sort them according to their similarity to return the first k retrieval results.

From the perspective of visual perception, image features mainly include texture, shape, and color features. Texture features reflect the gray distribution of pixels in the image and their neighboring pixels (Saikia et al. 2021). Shape features describe the shape of the object, and represent the boundary outline and internal shape of the object. Color features are generally extracted from pixels in the image for transformation. Methods for extracting texture features mainly include four categories: structure method, statistical method, frequency method, and model method. Statistical method mainly analyzes the distribution law of gray scale in the image, such as Scale-Invariant Feature Transform (SIFT), autocorrelation function, and Local Binary Pattern (LBP). Frequency methods include Gabor transform, wavelet transform, and Fourier transform (Ashraf et al. 2018).

Shape features generally describe the contour of the target object in the image and its surrounding area, and have nothing to do with brightness, color, and other information. However, color feature extraction includes the global method and the spatial method. The global method mainly includes color histogram, dominant color of image and color moment, and color distribution entropy. The existing research methods include color correlation vector, color correlation graph, color ring, angle, and mixture histogram (Karakasis et al. 2014).

The image retrieval scheme

Overview

First of all, the image owner generates secret keys K and secret keys are stored locally. In the second place, a set of plaintext images $\varsigma = \{\varsigma_i\}_{i=1}^{n}$ are encrypted to generate the set $\vartheta = \{\vartheta_i\}_{i=1}^{n}$, and encrypted images ϑ is stored in the cloud. After the encrypted images are received, the server stores them and extracts feature vectors $F_i = \{F_{ij}\}_{j=1}^{m}$ by a pre-trained CNN model from $\vartheta = \{\vartheta_i\}_{i=1}^{n}$. In addition, the index generation algorithm is executed by the cloud server to generate a retrieval index tree. When a user sends a query request, the query image is encrypted and submitted to the server. Finally, the server calculates a query trapdoor based on the trapdoor generation algorithm and retrieves similar images by querying the index tree. The top-k most similar images are sent to the user. The protocol for image retrieval is shown in Fig. 3.

Proposed image retrieval scheme with access control

Based on Bkd-tree and searchable encryption, we construct a CBIR scheme with access control. Let $\Lambda = \{(I_i, \iota_i)_{i=1,2,\cdots,n}\}$ be an image set, where ι_i is an access-policy-based hidden vector with l-dimensional. Let (q, e) be a query, where q is the feature of the encrypted query image C_q, and e is the attribute vector. In the end, the top-k most similar images are retrieved by the similarity query algorithm.

- *Image encryption.* The disclosure of unencrypted images brings some privacy concerns. Therefore, encrypting images can protect the privacy of the image owners. Each image I_i in the image database I is encrypted by block permutation, intra-block permutation, stream cipher and polyalphabetic cipher. The permutation cipher is a cryptographic algorithm that changes the order of characters in a string according to certain rules. Stream encryption is a symmetric encryption algorithm, in which the encipherer and decrypter use the same pseudo-randomstream as the key and the plaintext data and the key data flow are encrypted sequentially each time to obtain the ciphertext data flow. The polyalphabetic cipher is a substitution-based cipher that uses multiple replacement alphabets.
- *Index generation.* For a feature vector $F_i = (F_{i1}, F_{i2}, \cdots, F_{in})$ extracted by the pre-trained CNN from the encrypted dataset C_i, the server

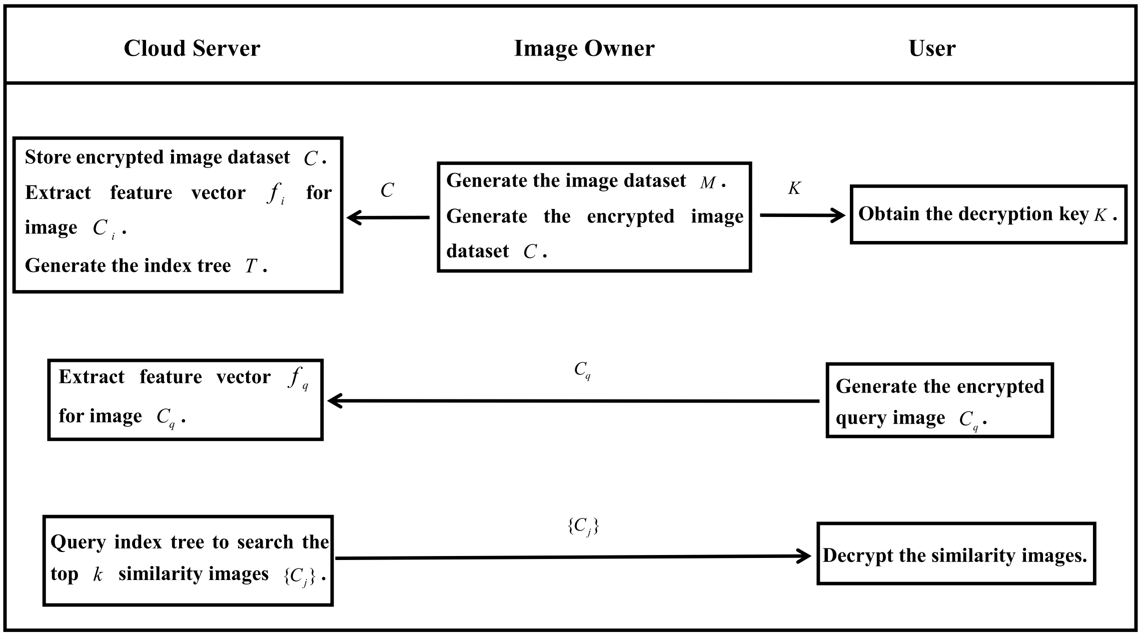
Cloud Server
Image Owner
User
Store encrypted image dataset C.
Extract feature vector f_i for image C_i.
Generate the index tree T.
C
Generate the image dataset M.
Generate the encrypted image dataset C.
K
Obtain the decryption key K.
Extract feature vector f_q for image C_q.
C_q
Generate the encrypted query image C_q.
Query index tree to search the top k similarity images $\{C_j\}$.
$\{C_j\}$
Decrypt the similarity images.

Fig. 3 The protocol of image retrieval

builds an index tree using the Bkd-tree T. Here are the detailed steps in this process. Firstly, the server selects the splicing dimension by comparing the difference between the minimum and maximum values of X and Y, and selects the large one as the splicing dimension of the current node. Secondly, the server sorts the data according to the value of the larger one, and the first half of the sorted data is divided into a left subtree, and the rest of the data is divided into right subtree. The above steps are repeated until both left and right subtrees contain one data. Finally, an index tree of image feature vectors is constructed for similarity image retrieval.

- *Trapdoor generation.* In the process of searchable encryption, if the user submits the keyword directly to the server when submitting the search request, it defeats the purpose of this technology. Therefore, searchable encryption requires no information about plaintext being disclosed. Therefore, the searchable encryption does not directly submit keywords to the server. Instead, the keywords are encrypted before submission, i.e., the encrypted keywords are called trapdoors. The process of trapdoor generation is as follows. Firstly, the query image I_q is encrypted to generate the encrypted image C_q, and it is uploaded to the server. Secondly, the server extracts the feature vector q of the encrypted query image C_q by the pretrained CNN model, namely, the feature vector q is the trapdoor.
- *Image search.* Firstly, the trapdoor q is uploaded to the server. Afterwards, the server verifies that the equation $\mathbf{M}(\iota_i, e) = 1$, and if not, No access permission is outputed. Otherwise, the cloud server queries the index tree and the top-k most similar images $R = \{C_j\}_{j \in (1,\cdots,k)}$ are returned to the query user. The overall process of image search is summarized in Algorithm 1.

Algorithm 1 Image Search

```
Input: Trapdoor $q = (q_x, q_y)$, index tree $T$, attribute vector $e$,
       hidden vector $\iota_i$.
Output: The top-$k$ most similar images $E$.
1:  if $\mathbf{M}(\iota_i, e) == 1$ then
2:      if The splicing dimension of the root node is Y then
3:          Compute the median $Y_m$ of Y for all data;
4:          if $q_y < Y_m$ then
5:              Query the left child node;
6:          else
7:              Query the right child node.
8:      else
9:          Compute the median $X_m$ of X for all data.
10:     Repeat steps 2 to 9 until top $k$ similarity images are retrieved.
11:     return top $k$ similarity images to the user.
12: else
13:     return "no access permission".
```

- *Image decryption.* After the retrieval results $R = \{C_j\}_{j \in (1,\cdots,k)}$ are returned, the query user decrypts images $\{C_j\}_{j\in(1,\cdots,k)}$ using the secret key K to obtain similarity images $\{I_j\}_{j\in(1,\cdots,k)}$.

Security analysis

Security of image content and image features

Images typically convey a wealth of information, including time, location, and people, and so on. To prevent privacy breaches, the images are encrypted using block permutation, intra-block permutation, stream cipher and polyalphabetic cipher in the proposed scheme, which is secure enough to ensure that the adversary cannot obtain any information from the images without the keys. Therefore, the confidentiality of the content in the proposed system is well-protected.

For image encryption, the proposed scheme is compatible with existing encryption algorithms that support search, and the security of the image content in this scheme is inherited from the security of the encryption algorithm used. The security of existing encryption algorithms that support search has been proven and widely recognized. Therefore, the privacy protection of images in this paper is secure enough. In addition, the key space setting in the proposed scheme is intended to be quite large and not feasible for brute force attacks.

Theorem 1 *In the encryption of image content in this paper, the key space is established to be large enough to withstand exhaustive attacks from probabilistic polynomial-time adversaries.*

Proof In the image encryption phase, each image is encrypted by block permutation and intra-block permutation. The key length of block permutation is equal to the number of block l_1 and the key length of intra-block permutation is equal to the size of intra-block l_2, and then each image is encrypted by stream cipher, the key length is the length of the key stream in stream cipher encryption, which is the same as the length of the message stream l_3. Finally, each image is encrypted by polyalphabetic cipher whose key length is l_4. For simplicity, the size of the image is expressed as $n \times m$, the number of block is expressed as s, the size of intra-block is expressed as $n \times m/s$, and the length of the key stream in stream cipher encryption is expressed as L. Therefore, the total security strength is the combination of four steps. When the image size is 64×64, l_1, l_2, l_3, and l_4 equals 64, 64, 256, and 256 bits, respectively. Then, we get a 640 bits key space which can be computationally secure under current computation power.

Image features extracted from images can be used for CBIR, and the query user's image features contain the user's search intention. When it comes to research on privacy-preserving image retrieval schemes, the earliest scheme was searchable encryption, which aims to achieve image retrieval tasks while encrypting data. Some existing searchable encryption schemes (Li et al. 2021; Liang et al. 2019) protect the content utilizing partial encryption, in which features are extracted from the unencrypted images. These schemes cannot ensure security against searching attacks. However, in the proposed scheme, the features are extracted from encrypted images, not the plaintext image. Therefore, the image features will not reveal any information about the image. For the security of image content, we carry out security analysis above. In brief, the proposed scheme ensures the security of image content and image features and protects the user's search intention.

Security of query requests and query results

Query requests are generally associated with user identity information. For example, users who query images of lung cancer are most likely to be doctors or family members of patients, and users who query images of stars are most likely to be fans, users who query images of cars are likely to be car salespeople or buyers. In our proposed scheme, the server calculates the feature of the encrypted query using the CNN model to ensure the security of the user's query requests. Namely, the user's search intention is not exposed to anyone. Additionally, the similarity query results returned by the cloud server are encrypted, which does not reveal any additional privacy. In short, the scheme ensures the security of both query requests and query results.

Searchability with access control

Access control allows users to set permissions on their own files and data, and only authorized users can access data. The universal access control mechanism is to set access control policies based on user roles. Some images contain sensitive information, so it is necessary to have image retrieval with access control. In this paper, it achieves the goal of accuracy of retrieval results and access control, which meets the requirement of safe CBIR. This is also a characteristic of the basic searchable encryption scheme, which can enable the user to search encrypted images uploaded in the server. In addition, the user is able to generate the required trapdoor for any given query image, i.e., the proposed scheme still maintains search capability.

Comparative analysis

To more comprehensively demonstrate the relevance and performance of the proposed scheme in the field of security data management, we compare and analyze the proposed scheme with other oblivious RAM schemes and TEE-based solutions, and review the impact of the proposed scheme on image privacy preservation in this section.

Comparative analysis with other oblivious RAM schemes

With the rapid development of cloud computing, more and more enterprises and individuals outsource their data to the public cloud. Images contain a wealth of sensitive information, therefore, data security and privacy protection are extremely important. That is, protecting data confidentiality in the cloud environment has become a necessary research topic. The typical method of protection is to encrypt data before uploading it to the cloud server. Nevertheless, attackers can still infer sensitive information from the data access pattern, even if the data is encrypted. This refers to the leakage of information resulting from a series of program accesses to the memory, including command addresses and data. To address this issue, researchers proposed the Oblivious Random Access Machine (ORAM).

Goldreich (1987) proposed a theory of software protection and simulation by oblivious RAMs, which has significantly contributed to the theoretical treatment of software protection. However, this scheme dose not directly address the problem of fingerprint software. To address this problem, Ostrovsky (1990) proposed a scheme for software protection based on a generic one-processor RAM model of computation. Subsequently, Goldreich and Ostrovsky (1996) refined the definition of ORAM and gave two classical constructs of ORAM, square root-ORAM and hierarchical ORAM. Both implementations experience significant periodic delays due to excessive bandwidth blowup in the worst-case scenario.

Later studies have shifted their focus towards the practical performance of ORAM instead of the asymptotic cost, as a small increase in client storage, server computation, or server storage can significantly reduce the bandwidth blowup. In any case, the goal is to optimize ORAM performance to mitigate bandwidth blowup. While some scenarios at the extreme end of the spectrum achieve a bandwidth blowup of O(1), other costs related to ORAM remain exorbitantly high. For instance, some methods rely on homomorphic encryption to optimize ORAM, but the computational cost is so high that it increases the ORAM request time significantly, making it impractical. Additionally, adding client storage is not suitable for mainstream smart devices like mobile phones and iPads. Furthermore, utilizing multiple servers results in the more larger total server storage and more higher cost for the user, making it unfeasible as well. In short, this direction focuses on the construction of purely theoretical schemes, it is difficult to break through basic theories. Therefore, researchers need to devote more energy to practical applications.

Comparative analysis with TEE-based solutions

A Privacy-protecting image retrieval scheme generally considers a semi-honest cloud server, which will correctly provide the storage and similar image searching services but may be tempted by the content of images and their features. Therefore, unencrypted images stored in the cloud are not secure, with the cloud being regarded as the adversary in such a scheme. A Trusted execution environment (TEE) (Dai et al. 2010) is a secure computing environment that protects the code and data running inside it from external attacks, including attacks from the operating system, hardware, and other applications. This technology has extensive application value in cloud computing, identity authentication, the Internet of things, and other scenarios, while it circumvents additional communication procedures and substantial computational overhead in public key cryptography (Deng et al. 2022; Jang et al. 2018). Despite the immense potential of TEE technology, there are still some challenges:

Cross-platform compatibility. Currently, TEE implementations from different processor vendors are not fully compatible, which can result in additional costs and learning curve for developers and users.

Resource constraints. TEE typically runs inside the processor and may be limited by processing power and memory resources, which can lead to performance degradation and resource competition.

Side-channel attacks. Its security is largely dependent on hardware implementation, and it is difficult to give a specific definition of security boundaries. Therefore, it is more vulnerable to side-channel attacks from different attack surfaces.

Standardization. TEE technology has not been fully standardized, which may affect its wide application and development. Currently, TEE security standards are mainly formulated by GlobalPlatform, and few products have passed GlobalPlatform security certification. How to further develop clear TEE security standards is also a difficult problem.

Therefore, it is necessary to use cryptographic techniques for further processing to ensure the security of the search.

The impact of the proposed scheme on image privacy preservation

With the development of information technology and new forms of digital economy, a large amount of image data continues to be generated and it increases the demand for image data storage and computation. Cloud computing is a convenient resource sharing model, and users can outsource heavy storage and computing tasks to cloud servers at low costs (Zhu et al. 2011; Ren et al. 2012; Troncoso-pastoriza and Perez-Gonzalez 2013). However, the outsourcing of storage and computing has led to a loss of direct physical control of images by users, which greatly increases the security risk of images in the cloud platform. To prevent the disclosure of personal privacy information, image privacy protection is essential. Therefore, ensuring images security and achieving privacy protection have become a research hotspot in recent years. In addition, the leakage of personal data have also attracted great attention from the relevant government departments in various countries, and relevant policies have been introduced one after another. For example, the Cybersecurity Law of the People's Republic of China was formally implemented in 2017, and the General Data Protection Regulation of the European Union came into effect in 2018. At the technical level, image encryption can prevent the privacy disclosure of data owners. However, this approach destroys the availability of data such as image retrieval.

To solve this problem, encrypted image retrieval schemes (Shashank et al. 2008; Ferreira et al. 2019; Xia et al. 2019, 2019, 2021) have been proposed to protect data privacy while maintaining the retrieval functionality. In the case of specific images, the image owner does not want anyone to be able to retrieve images, but only authorized users should be entitled to access them. Therefore, image retrieval methods with access control are necessary. To meet the requirements of both image owners and retrieval users, we propose an image retrieval scheme with access control based on searchable encryption in media cloud. This scheme enables efficient and controllable image retrieval while protecting the privacy of the image owner and the retrieval user.

Performance analysis

Experiments are conducted to evaluate the proposed content-based image retrieval scheme on the Corel image dataset in this section. This image database is a benchmark dataset for the image retrieval, which includes 100 categories of images and each category contains 100 similar images (Wang et al. 2001). The experimental scheme is implemented on an Intel (R) Core (TM) i7-6500 CPU. The ciphertext image search results are sorted based on the Manhattan distance of feature vectors. Therefore, the accuracy of image search is affected by the encryption algorithm, and depends on the algorithm of image feature extraction. The performance parameters of simulation platform are presented in Table 1.

Table 1 Simulation platform

Operating system	CentOS 7.8. x86 64
CPU	Intel(R) Core (TM) i7-6500
Memory	8GB RAM
Program language	Java

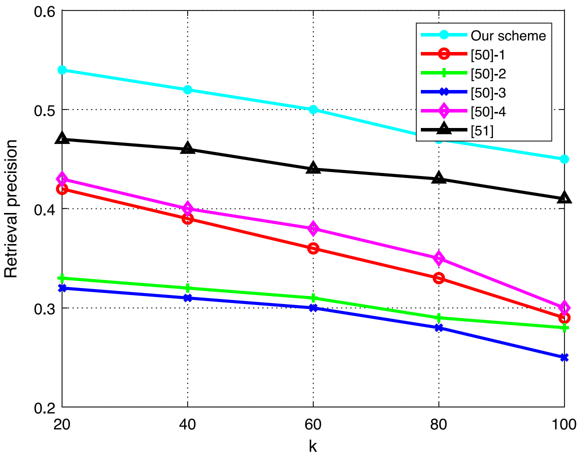

Fig. 4 Retrieval precision

Retrieval precision

The retrieval precision is represented by $P = p_1/p_2$ in the proposed scheme, i.e., the ratio of the number of queried similarity images to the total number of similar images. Effectively measuring the similarity between image features can also improve the precision of image search. As can be seen in Fig. 4, the retrieval precision is mainly dependent on the feature descriptors. There are many methods to extract image features, e.g., HOG, LBP, SURF, and SIFT. Traditional image feature descriptors, such as the local feature descriptors based on local area, have good stability and can maintain invariance to image brightness, translation, rotation, and other distortions. However, this type of image feature has limitations. Compared with traditional feature descriptors, convolutional neural network has a strong ability to abstractly represent the content contained in images and can comprehensively and effectively organize image features. In this scheme, the feature descriptors are extracted by the CNN model, which makes the precision of retrieval higher than other schemes. As can see from Fig. 4, the retrieval precision decreases as the number of retrieved images k grows, and the proposed scheme achieves higher retrieval precision than the methods proposed in Xia et al. (2022), Li et al. (2021).

Time consumption

We analyze the time consumption of the proposed scheme from three aspects, i.e., the stage of image encryption, the stage of index generation, and the stage of search operation.

Image encryption time

It is widely acknowledged that images contain a wealth of information, especially private information, e.g., identity, occupation, location. Therefore, image owners typically encrypt images before storing them in the server to avoid private information leakage. In this scheme, we encrypt images using block permutation, intra-block permutation, stream cipher and polyalphabetic cipher. The permutation cipher is a cryptographic algorithm that changes the order of characters in a string according to certain rules. Stream encryption is a symmetric encryption algorithm in which the encoder and decoder uses the same pseudo-randomstream, and the plaintext data and the key data flow are encrypted sequentially each time to obtain the ciphertext data flow. The polyalphabetic cipher is a substitution-based cipher that uses multiple replacement alphabet. Fig. 5 presents the time consumption of cryptographic operations in the Corel database under diverse block sizes, which clearly shows that the encryption time decreases as the block size increases.

Index generation time

In this scheme, the cloud server firstly extracts feature vectors using a pre-trained CNN model for encrypted images. Afterwards, the server builds an index tree using a Bkd-tree. Here are the detailed steps in this process. Firstly, the server selects the splicing dimension by comparing the difference between the minimum and maximum values of X and Y, and selects the large one as the splicing dimension of the current node. Secondly, the server sorts the data according to the value of the larger one, and the first half of the sorted data is divided into a left subtree, and the rest of the data are divided into a right subtree. The above steps are repeated until both left and right subtrees contain one piece of data. Finally, an index tree of image feature vectors is constructed for similarity image retrieval. To evaluate the performance of this scheme, we show the time cost of index generation in Fig. 6, which depends on two parameters: the attribute values l and the query record d. l is set to 4 and the main assessment is that the computational cost of generating query tokens varies with d. Figure 6 shows that the time consumption of index generation increases linearly with the increase of the dimension of query record.

By comparing the experimental results, it is clear that the index generation time of the proposed scheme is less than that of the scheme (Xia et al. 2022; Li et al. 2021).

Search time

In the process of image search, this paper adopts Manhattan distance to realize similarity measurement between ciphertext index and query trapdoor. In addition, the scheme adopts a BKD-tree to assist retrieval

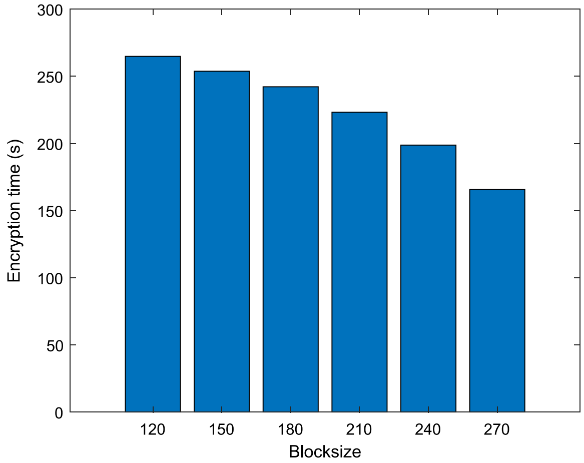

Fig. 5 The time consumption of image encryption

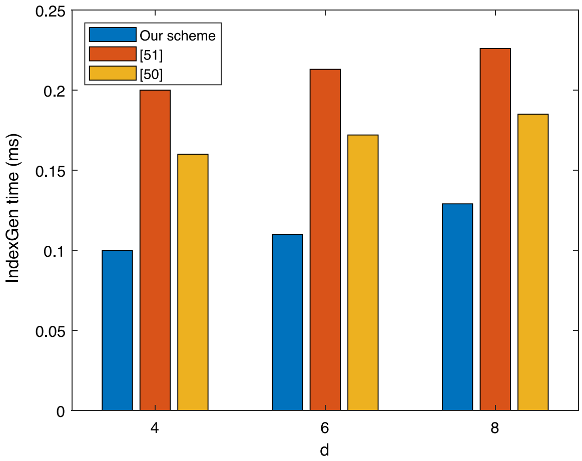

Fig. 6 The time consumption of index generation

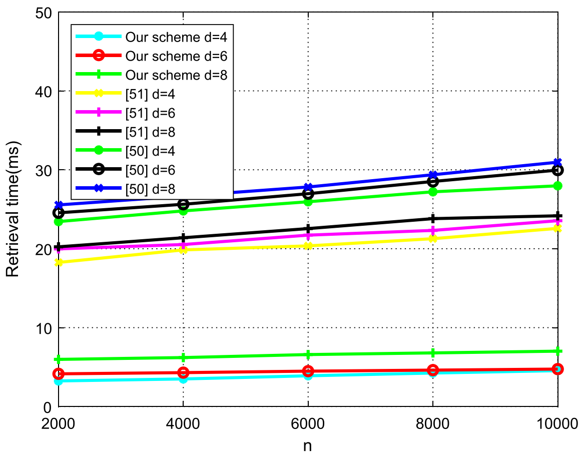

Fig. 7 The time consumption of search operation

and returns the top-k images to the query user after similarity calculation. In the proposed scheme, an index tree is constructed to expedite the retrieval process and achieve preferable retrieval efficiency compared to other schemes using linear index. The calculation cost of image query is affected by the dimension of attribute values l, the dimension of data records d, and the size of the dataset n. In our experiment, l is set to 4 and the main assessment is that the query time varies with n and d. We set $n \in \{2000, 4000, 6000, 8000, 10000\}$ and $d \in \{4, 6, 8\}$. From Fig. 7, it is obvious that the retrieval time grows as the dataset grows. Obviously, the time consumption of retrieval in the proposed scheme is far less than that of the schemes (Xia et al. 2022; Li et al. 2021).

Conclusion

In this paper, we have proposed a novel CBIR scheme with access control. Specifically, the images are encrypted and uploaded to the server. Afterwards, the feature descriptors are extracted by the CNN model. Afterwards, the index tree of image features is constructed using Bkd-tree by the cloud server, and the distances between features are calculated, which greatly expedites the process of retrieval. Finally, the top-k most similar images are sent to the user, who decrypts these images to obtain plaintext images. Compared with traditional feature extraction algorithms, using a CNN model for feature extraction simplifies the process of image feature calculation. Additionally, in the retrieval process, this method reduces computational overhead and enhances the user's search experience. The experimental analysis demonstrates that the proposed scheme is feasible and accuracy in secure cloud computing.

Acknowledgements

This work was supported by the National Key R &D Program of China (Grant No. 2021YFB3100400) and the State Key Laboratory of Information Security (Grant No. 2022-MS-02).

Author contributions

Miao Tian: Conceptualization, Methodology, Formal analysis, Investigation, Writing-original draft. Yushu Zhang: Resources, Data Curation, Funding acquisition. Yongming Zhang: Project administration.. Xiangli Xiao: Writing - review & editing, Software, Validation, Visualization. Wenying: Supervision.

Funding

This work was supported by the National Key R &D Program of China (Grant No. 2021YFB3100400) and the State Key Laboratory of Information Security (Grant No. 2022-MS-02).

Availability of data and materials

The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Competing interest

The authors declare that they have no competing interest.

Received: 15 May 2023 Accepted: 22 January 2024
Published online: 03 July 2024

References

Ahamd I, Jang T-S (2003) Old fashion text-based image retrieval using FCA. Proc Int Conf Image Process 3:33

Ashraf R, Ahmed M, Jabbar S, Khalid S, Ahmad A, Din S, Jeon G (2018) Content based image retrieval by using color descriptor and discrete wavelet transform. J Med Syst 42(3):44

Bella MT, Vasuki A (2019) An efficient image retrieval framework using fused information feature. Comput Electr Eng 75:46–60

Dai W, Jin H, Zou D, Xu S, Zheng W, Shi L, Yang LT (2010) TEE: A virtual drtm based execution environment for secure cloud-end computing. In: Proceedings of 17th ACM conference on computer communications and security, pp 663–665

Deng Y, Wang C, Yu S, Liu S, Ning Z, Leach K, Li J, Yan S, He Z, Cao J, Zhang F (2022) Strong box: a GPU TEE on arm endpoints. In: Proceedings of 29th ACM SIGSAC conference in computing communications and security (CCS'22), pp 769–783

Ferreira B, Rodrigues J, Leitao J, Domingos H (2019) Practical privacy-preserving content-based retrieval in cloud image repositories. IEEE Trans Cloud Comput 7(3):784–798

Goldreich O (1987) Towards a theory of software protection and simulation by oblivious rams. In: Proceedings of 19th annual ACM symposium theory and computing, pp 182–194

Goldreich O, Ostrovsky R (1996) Software protection and simulation on oblivious rams. J ACM 43(3):431–473

He K, Chen J, Zhou Q, Du R, Xiang Y (2021) Secure dynamic searchable symmetric encryption with constant client storage cost. IEEE Trans Inf Forensics Secur 16:1538–1549

Huang Y, Zhang J, Pan L, Xiang Y (2020) Privacy protection in interactive content based image retrieval. IEEE Trans Depend Secure Comput 17(3):595–607

Jang J, Choi C, Lee J, Kwak N, Lee S, Choi Y, Kang BB (2018) Privatezone: providing a private execution environment using arm trustzone. IEEE Trans Depend Secure Comput 15(5):797–810

Karakasis EG, Papakostas GA, Koulouriotis DE, Tourassis VD (2014) A unified methodology for computing accurate quaternion color moments and moment invariants. IEEE Trans Image Process 23(2):596–611

Li W, Duan L, Xu D, Tsang IW-H (2011) Text-based image retrieval using progressive multi-instance learning. Int Conf Comput Vis 8:2049–2055
Li W, Laurent A, Teddy F (2016) Privacy-preserving outsourced media search. IEEE Trans Knowl Data Eng 28:2738–2751
Li H, Yang Y, Dai Y, Bai J, Yu S, Xiang Y (2020) Achieving secure and efficient dynamic searchable symmetric encryption over medical cloud data. IEEE Trans Cloud Comput 8(2):484–494
Li B, Ding S, Yang X (2021) A privacy-preserving scheme for jpeg image retrieval based on deep learning. J Phys Conf Ser 1856(1):012007
Li B, Ding S, Yang X (2021) A privacy-preserving scheme for jpeg image retrieval based on deep learning. J Phys 1856:1
Li Y, Ma J, Miao Y, Wang Y, Liu X, Choo KKR (2022) Similarity search for encrypted images in secure cloud computing. IEEE Trans Cloud Comput 10(2):1142–1155
Liang H, Zhang X, Cheng H (2019) Huffman-code based retrieval for encrypted jpeg images. J Vis Commun Image Represent
Lu W, Swaminathan A, Varna AL, Wu M (2009) Enabling search over encrypted multimedia databases. In: Media forensics security, vol 7254 . International Society for Optics and Photonics
Lu W, Varna AL, Swaminathan A, Min W (2009) Secure image retrieval through feature protection. In: IEEE international conference on acoustics
Magdy S, Abouelseoud Y, Mikhail M (2019) Effect of chosen features on performance of privacy preserving image retrieval systems. Comput Electr Eng 76:411–424
Majhi M, Mallick AK (2022) Random projection and hashing based privacy preserving for image retrieval paradigm using invariant and clustered feature. J King Saud Univ Comput Inf Sci
Ostrovsky R (1990) Efficient computation on oblivious rams. In: Proeedings of 22th annual ACM symposium and theory computing, pp 514–523
Paul A, Kandar S, Dhara BC (2022) Image encryption using permutation generated by modified Regula–Falsi method. Appl Intell 52(10):10979–10998
Procopiuc O, Agarwal PK, Arge L, Vitter JS (2003) Bkd-tree: a dynamic scalable kd-tree, vol 2750, pp 46–65
Raveendra K, Vinothkanna R (2019) Hybrid ant colony optimization model for image retrieval using scale-invariant feature transform local descriptor. Comput Electr Eng 74:281–291
Ren K, Wang C, Wang Q (2012) Security challenges for the public cloud. IEEE Int Comput 16(1):69–73
Saikia S, Fernáindez-Robles L, Alegre E (2021) Image retrieval based on texture using latent space representation of discrete Fourier transformed maps. Neural Comput Appl 33(20):13301–13316
Samet N, Hiçsönmez S, Sener F (2016) Creating image tags for text based image retrieval using additional corpora. In: 24th Signal processing and communication application conference (SIU), pp 1321–1324
Shamna P, Govindan VK, Abdul Nazeer KA (2022) Content-based medical image retrieval by spatial matching of visual words. J King Saud Univ Comput Inf Sci. 34(2):58–71
Shamna P, Govindan VK, Nazeer KAA (2018) Content-based medical image retrieval by spatial matching of visual words. J King Saud Univ Comput Inf Sci
Shashank J, Kowshik P, Srinathan K, Jawahar CV (2008) Private content based image retrieval. In: IEEE conference on computer vision and pattern recognition, pp 1–8
Song DX, Wagner D, Perrig A (2000) Practical techniques for searches on encrypted data. In: Proceedings under IEEE symposium on security and privacy(SP), pp 44–55
Troncoso-pastoriza JR, Perez-Gonzalez F (2013) Secure signal processing in the cloud: enabling technologies for privacy-preserving multimedia cloud processing. IEEE Signal Proc Mag 30(2):29–41
Wang JZ, Li J, Wiederhold G (2001) Simplicity: semantics-sensitive integrated matching for picture libraries. IEEE Trans Pattern Anal Mach Intell 23(9):947–963
Wang X, Lan R, Wang H, Liu Z, Luo X (2021) Fine-grained correlation analysis for medical image retrieval. Comput Electric Eng 90:106992
Wang Y, Chai X, Gan Z, Zhang Y, Chen X, He X (2023) TPE-ISE: approximate thumbnail preserving encryption based on multilevel dwt information self-embedding. Appl Intell 53(4):4027–4046
Xia Z, Wang X, Zhang L, Qin Z, Sun X, Ren K (2017) A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE Trans Inf Forensics Secur 11(11):2594–2608
Xia Z, Zhu Y, Sun X, Qin Z, Ren K (2018) Towards privacy-preserving content-based image retrieval in cloud computing. IEEE Trans Cloud Comput 6(1):276–286
Xia Z, Jiang L, Ma X, Yang W, Ji P, Xiong N (2019) A privacy-preserving outsourcing scheme for image local binary pattern in secure industrial internet of things. IEEE Trans Ind Inf 16:629–638
Xia Z, Lu L, Qiu T, Shim HJ, Chen X, Jeon B (2019) A privacy-preserving image retrieval based on ac-coefficients and color histograms in cloud environment. Comput Mater Contin 58(1):27–43
Xia Z, Wang L, Tang J, Xiong N, Weng J (2021) A privacy-preserving image retrieval scheme using secure local binary pattern in cloud computing. IEEE Trans Network Sci Eng 8(1):318–330
Xia Z, Jiang L, Liu D, Lu L, Jeon B (2022) BOEW: a content-based image retrieval scheme using bag-of-encrypted-words in cloud computing. IEEE Trans Services Comput 15(1):202–214
Xia Z, Ji Q, Gu Q, Yuan C, Xiao F (2022) A format-compatible searchable encryption scheme for jpeg images using bag-of-words. ACM Trans Mult Comput Commun Appl 18(3):1–18
Yang J, Wu Y (2022) An approach of bursty event detection in social networks based on topological features. Appl Intell 52:6503–6521
Yang T, Ma J, Miao Y, Wang Y, Liu X, Choo K-KR, Xiao B (2022) Mu-teir: traceable encrypted image retrieval in the multi-user setting. IEEE Trans Services Comput 8:9. https://doi.org/10.1109/TSC.2022.3149962
Zaidi SAJ, Buriro A, Riaz M, Mahboob A, Riaz MN (2019) Implementation and comparison of text-based image retrieval schemes. Int J Adv Comput Sci Appl 10(1):611–618
Zheng J, Zeng Q (2022) An image encryption algorithm using a dynamic s-box and chaotic maps. Appl Intell 52(13):15703–15717
Zhu W, Luo C, Wang J, Li S (2011) Multimedia cloud computing. IEEE Sig Process Mag 28(3):59–69
Zhu D, Zhu H, Wang X, Lu R, Feng D (2022) An accurate and privacy-preserving retrieval scheme over outsourced medical images. IEEE Trans Services Comput. https://doi.org/10.1109/TSC.2022.3149847

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Yushu Zhang Yushu Zhang received the Ph.D. degree in computer science and technology from the College of Computer Science, Chongqing University, Chongqing, China, in 2014. He held various research positions with Southwest University, Chongqing, China, City University of Hong Kong, Hong Kong, China, University of Macau, China, and Deakin University, Victoria, Australia. He is currently a Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. He is currently an Associate Editor for the Information Sciences and Signal Processing. He has authored or coauthored more than 100 refereed journal articles and conference papers in these areas. His current research interests include multimedia security, blockchain, and artificial intelligence.