Towards Security Protecting Substance Based Picture Recovery in Cloud

[1]Dr. R. Jegadeesan [2]N. Amulya, , [3]K. Lavanya, [4]Dr.S.Prabakaran. [5]G.Saiteja, [6]CH. Mehersai,

[1,4]Associate Professor-CSE [2,3,5,6]Final year Student Computer science and Engineering,

[1,2,3,4,5,6]Jyothishmathi Institute of Technology and Science, Karimnagar, India

***ABSTRACT*:** Now a days Cloud storage is one of the huge repository to store and usage of services for the remote purpose. In our regular days we used to share data in many ways. While sharing the private data, privacy preserving is becoming an increasingly significant issue.in such case images also exposed to public when the related term or keyword is searched.in these note we use AIES(Advanced Image Encryption Schemes)-SCBIR(Specific Content Based Image Retrieval)-According to the image content based on color,shapes etc. ,the images can be retrieved.so the people who know the content can access the images.it helps to secure the images which are in private from the public.this method helps to secure images by using encryption schemes and privacy preserving in cloud environment in which it can include the benefit of encryption on the server side also..

***Index terms*:** Content Based Image Retrieval,Cloud computing, searchable encryption, image retrieval

1. INTRODUCTION

Thanks to affordable storage and straight forward internet hosting, the world has witnessed an incredible growth within the quantity, handiness and importance of pictures in our daily life. Pictures begin to play a vital role in numerous fields like medication, journalism, advertising, design, education and amusement, etc. The need for economical storage and retrieval of pictures is strengthened by the rise of large scale image databases among all types of areas. Meanwhile, as a rising technology, Specific Content Based Image Retrieval(SCBIR) shows enough promise and maturity to be useful in several real-world image retrieval applications. For example, clinicians might use SCBIR to retrieve the similar cases of the patients to facilitate the clinical decision-making process[1]. As another example, enforcement agencies usually compare the proof from the crime scene with the records in their archives[2].However, such type of SCBIR service is intensive in each computation and storage intensive. An oversized image information sometimes consists of countless images. Sometimes, one digital image would possibly contain more than twenty million dimensions and its size may be higher than forty megabytes, like diagnostic procedure pictures[3]. Moreover, SCBIR usually has high machine complexness because of the high spatial property of image information. Cloud computing offers an excellent chance to supply on-demand access to ample computation and storage resource, that makes it a primary alternative for image storage and SCBIR outsourcing. By deploying such image retrieval outsourcing, the data owner is no longer required to keep up the image information domestically. An authorized data user can query the cloud for CBIR service while not interacting with the data owner. In spite of the enormous advantages, privacy becomes the most important concern regarding SCBIR outsourcing. For example, the patients won't wish to disclose their medical images. In fact, the Health Insurance Portability and Accountability Act (HIPAA) sets legal necessities to shield patients' privacy.

Contribution. In this paper, we have a tendency to study the privacy-preserving SCBIR outsourcing drawback and present a practical solution. We tend to utilize the techniques from security, image processing and data retrieval domains to attain secure and economical looking over encrypted pictures. The proposed theme supports local-feature based SCBIR with the Earth Mover's Distance (EMD) as a resemblance metric. In particular, a secure transformation is intended in order that the cloud server will solve the EMD drawback with the privacy preserved. local sensitive hash in utilized to realize constant search efficiency.

2. RELATED WORK

The data confidentiality is explained in a way that if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query, without loss of data confidentiality [16].Encrypted storage protects the data against illegal access, but it complicates some basic, yet important functionality such as the search on the data. To achieve search over encrypted data without compromising the privacy, considerable amount of searchable encryption schemes have been proposed in the literature [17].Secure index is a system that permits a query, with a "trapdoor" for a word x to check in O(1) time given that the index contains x; The index reveals no info concerning its contents without valid trapdoors, and trapdoors will solely be generated with a secret key. Secure indexes are a natural extension of problem of constructing data structures with privacy guarantees such as those provided by history free data structures[18].Image content based retrieval is rising as a very important research space with application to digital libraries and multimedia system databases. the main focus of this paper is on the image process aspects and specially using texture data for browsing and retrieval of large image information [19].Searchable Symmetric Encryption (SSE) permits a part to outsource the storage of his data to another party in an private way, whereas maintaining the flexibility to search over it [20].The comparison the strategies of shape-based feature extraction and illustration. About forty techniques for extraction of shape features are shortly represented and compared.not like the standard classification, the approaches of shape-based feature extraction and illustration were classified by their process approaches. These process approaches enclosed form signatures, two-dimensional figure approximation strategies, special inter- relation feature, moments approaches, scale-space strategies and form remodel domains: in such method, one will simply select the acceptable process approach[21].One of the major drawbacks of cloud computing, however, is the lack of robust mechanisms for the users to control the privacy of the data they farm out to the clouds.so, to develop an image encoding scheme that enhances the privacy of image data that is outsourced to the clouds for processing. Unlike previously proposed image encryption schemes [22].The process of figuring out the worth, amount, or quality of something of Depot, a cloud storage system that makes something as small as possible something important as unimportant trust ideas you think are true. Depot tolerates buggy or evil and cruel behaviour by any number of clients or servers, yet it provides safety and liveness promises that something will definitely happen or that something will definitely work as described to correct clients [23].The image can be approximately reconstructed based on the output of a black box local description software such as those classically used for image indexing. this approach consists first in using an off-the-shelf image database to find patches which are visually similar to each region of interest of the unknown input image, according to associated local descriptors. These patches are then warped into input image domain according to interest region geometry and seamlessly stitched together [24].

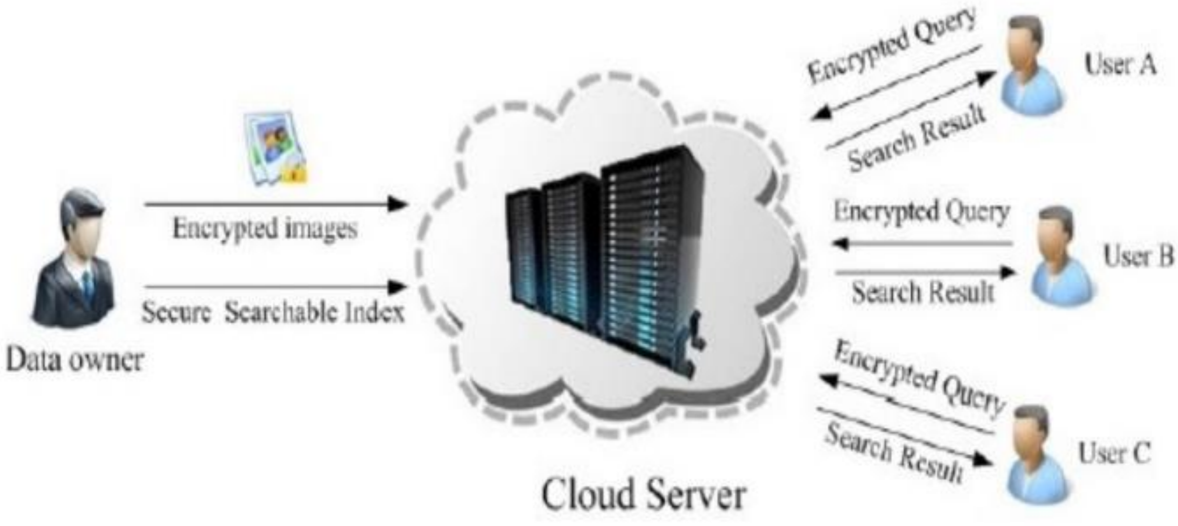


Fig.1: Architecture of proposed scheme

3.2 Security model

The honest-but-curious cloud servers are considered here. This cloud servers can properly follow the selected protocol specification, but keep and analyze the communication history, making an attempt to derive sensitive information. The data owner and approved users are perpetually sure. This concept aims at preventing the cloud server from knowing the content of image information and users queries. Just like searchable coding theme, we will not take in account the data leakage because of access pattern. For example, if pictures m_i and $m_j$are obtained because the search results of the identical query, its simple to deduce that pictures m_i and m_j are just like one another. In fact, this issue maybe effectively solved by applying associate existing ORAM scheme [4].

3.3 Bag-of-words model

CBIR for the most part includes extraction of visual highlights and seek in the visual element space for comparative pictures. So it At that point, the picture $m_t \in M$ can be indicated as a sack of words: scientifically portray a picture, which is alluded to as the component extraction step. An element is characterized to catch a specific visual property of a picture, either all inclusive for the whole picture or locally for a little gathering of pixels. The usually utilized worldwide features contain color histograms [5], texture features[6], shape features[7], and so forth. The promotion vantage of global component is its rapid for both separating highlights and registering similitude. Be that as it may, as an exchange off among exactness and calculation, the worldwide component based CBIR is frequently too unbending to even think about representing a picture. Neighborhood highlights in view of neighborhood invariants, for example, corner focuses or intrigue focuses, are commonly increasingly powerful for spatial change what's more, more often than not recover increasingly precise outcomes. The most popular neighborhood highlights contain Scale-invariant feature transform (SIFT) features [8] and picture patches [9]. For nearby component based CBIR, a lot of neighborhood highlights are extricated from nearby districts of a picture. One well known CBIR approach utilizing nearby highlights is called sack of-words display. In this model, nearby highlights are removed from all pictures in the database and after that mutually bunched. The bunch focuses are utilized as 'words' to shape the vocabulary. Note that every neighborhood highlight is a component vector comprising of different components. Subsequent to being bunched, the vectors of bunch focuses will be kept. For every one of other neighborhood highlights, just the identifier of the most close group focus is kept. At that point, the picture $m_t \in M$ can be indicated as a sack of words:

$S_t = \{(c(t)1, w(t)1), \ldots ,(c(t)i, w(t)i), \ldots ,(c(t)k_t, w(t)k_t)\}$,

Where c_i denotes a cluster center, w_i is the number of local features that are clustered to the class centered by c_i in m_t, and k_t is the total number of relevant cluster centers to image m_t. In this paper, s_t is regarded as the signature of the image m_t. Then, the similarity between a query image m_q and an image $m_t \in M$ can be calculated as the earth mover's distance between their signatures s_q and s_t.

3.4 Earth Movers Distance

The Earth Mover's Distance (EMD) is a method to evaluate dissimilarity between two multi-dimensional distributions[10],[11] in some feature space where a distance measure between single features, which we call the ground distance is given. The EMD ``lifts" this distance from individual features to full distributions.

Intuitively, given two distributions, one can be seen as a mass of earth properly spread in space, the other as a collection of holes in that same space. Then, the EMD measures the least amount of work needed to fill the holes with earth. Here, a unit of work corresponds to transporting a unit of earth by a unit of ground distance.

A distribution can be represented by a set of clusters where each cluster is represented by its mean (or mode), and by the fraction of the distribution that belongs to that cluster. We call such a representation the signature of the distribution. The two signatures can have different sizes, for example, simple distributions have shorter signatures than complex ones.

Computing the EMD is based on a solution to the well-known transportation problem. Suppose that several suppliers, each with a given amount of goods, are required to supply several consumers, each with a given limited capacity. For each supplier-consumer pair, the cost of transporting a single unit of goods is given. The transportation problem is then to find a least-expensive flow of goods from the suppliers to the consumers that satisfies the consumers' demand. Matching signatures can be naturally cast as a transportation problem by defining one signature as the supplier and the other as the consumer, and by setting the cost for a supplier-consumer pair to equal the ground distance between an element in the first signature and an element in the second. Intuitively, the solution is then the minimum amount of ``work" required to transform one signature into the other.

This can be formalized as the following linear programming problem: Let

$$P = \{(p_1, w_{p_1}), \ldots, (p_m, w_{p_m})\}$$

be the first signature with m clusters, where p_i is the cluster representative and w_{pi} is the weight of the cluster;

$$Q = \{(q_1, w_{q_1}), \ldots, (q_n, w_{q_n})\}$$

the second signature with n clusters;

$$\mathbf{D} = [d_{ij}]$$

the ground distance matrix where d_{ij} is the ground distance between clusters p_i and q_j. We want to find a flow

$$\mathbf{F} = [f_{ij}]$$

with f_{ij} the flow between p_i and q_j, that minimizes the overall cost

$$\mathrm{WORK}(P, Q, \mathbf{F}) = \sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} d_{ij} \;,$$

$$\begin{aligned}
f_{ij} &\geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq n \\
\sum_{j=1}^{n} f_{ij} &\leq w_{p_i} \qquad 1 \leq i \leq m \\
\sum_{i=1}^{m} f_{ij} &\leq w_{q_j} \qquad 1 \leq j \leq n \\
\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} &= \min\left(\sum_{i=1}^{m} w_{p_i}, \sum_{j=1}^{n} w_{q_j}\right) \;,
\end{aligned}$$

The first constraint allows moving ``supplies" from P to Q and not vice versa. The next two constraints limits the amount of supplies that can be sent by the clusters in P to their weights, and the clusters in Q to receive no more supplies than their weights; and the last constraint forces to move the maximum amount of supplies possible. We call this amount the total flow. Once the transportation problem is solved, and we have found the optimal flow $\mathbf{F}$, the earth mover's distance is defined as the work normalize

$$\mathrm{EMD}(P, Q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} d_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}} \;.$$

The normalization factor is introduced in order to avoid favoring smaller signatures in the case of partial matching.

The EMD has the following advantages

- Naturally extends the notion of a distance between single elements to that of a distance between sets, or distributions, of elements.
- Can be applied to the more general variable-size signatures, which subsume histograms. Signatures are more compact, and the cost of moving ``earth" reflects the notion of nearness properly, without the quantization problems of most other measures.
- Allows for partial matches in a very natural way. This is important, for instance, for image retrieval and in order to deal with occlusions and clutter.
- Is a true metric if the ground distance is metric and if the total weights of two signatures are equal. This allows endowing image spaces with a metric structure.
- Is bounded from below by the distance between the centers of mass of the two signatures when the ground distance is induced by a norm. Using this lower bound in retrieval systems significantly reduced the number of EMD computations.
- Matches perceptual similarity better than other measures, when the ground distance is perceptually meaningful. This was shown by for color- and texture-based image retrieval.

4. The Proposed Scheme Design

Firstly Security-Protecting SCBIR scheme is designed. Initially, the proposed scheme's outline is introduced. For construction of index we use two technologies. The whole scheme can be detail demonstrated as below

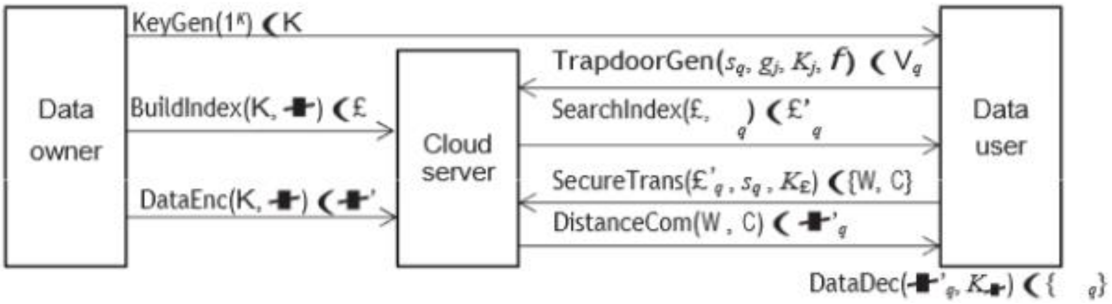

Data owner
KeyGen(1^κ) ❨K
BuildIndex(K, ■) ❨£
DataEnc(K, ■) ❨■'
Cloud server
TrapdoorGen(s_q, g_j, K_j, f) ❨ V_q
SearchIndex(£, $_q$) ❨£'$_q$
SecureTrans(£'$_q$, s_q, $K_£$) ❨{W, C}
DistanceCom(W , C) ❨ ■'$_q$
Data user
DataDec(■'$_q$, $K_■$) ❨{ $_q$}

Fig 2: Flowchart of the Proposed Scheme

4.1 The Framework Of The Scheme

In proposed scheme a tuple of probabilistic polynomial time algorithm ج={KeyGen, BuildIndex, DataEnc,TdGen,SearchIndex,SecureTrans,DisCom,DataDec}.Figure2 Demonstrates the proposed scheme's flowchart.

At the very first phase ie., Initiation phase, when an image database M, is given, the owner of the data runs KeyGen(1^K), BIndex(K,M) and DataEnc(K,M) and generate K, I and M. Then the cloud server receives I and M from data owners then the data users who are authorized receives K.

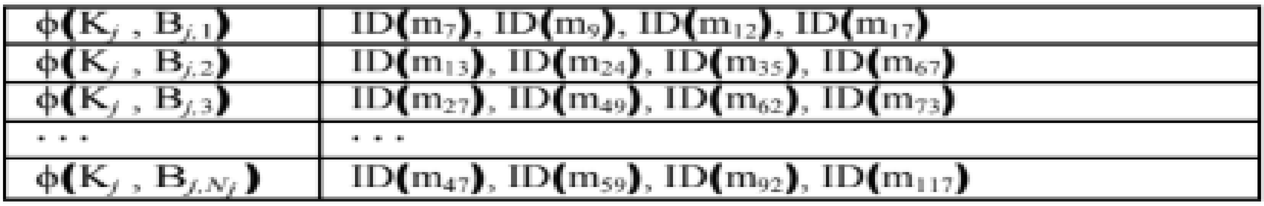
In second phase ie., image retrieval phase, A trapdoor $T_q$is generated by an authenticated data userby running TdGen(S_q, $g_{j\ ,}$ K ,ϕ,).This trapdoor T_q is to be submitted to the cloud server. here a set of encrypted signature $S_q$is obtained by running SearchIndex(I,I_q)this process is done by the cloud server. Hence it proved that set images M is subset of query images M_q. The data user then receives S_q from the cloud server. By using S_q the data users starts SecureTrans($s^1{}_q$, s_q, K_j) to develop the set of encrypted EMD problems Ω, which is to be delivered to the cloud server. When cloud server receives this set of encrypted EMD problems Ω ,it runs DisCom(Ω) and evaluate the most related images to the query image, to send this image set M_q to the data users. After receiving this the data user runs DDec($M_{q,}$ K_m) and obtain the set of similar images $M_{q.}$

4.2 Local Sensitive Hash On Signature Centroid

There will be a time complexity while calculating the EMD problems in-between the images in the database and query image. This time complexity will be linear to the number of images I the image set. So this will be difficult to be applied for a large number of images in the real world. In this case, we require an other way to eliminate the unrelated images within short span. After filtering out those images it will be easy to calculate the EMD problem with the remaining images. This process of filtering can be done by employing local sensitive hash (LSH) which has a signature centroid.

4.3 Centroid of the signature

Euclidean distance of the centroids can be found by the lower bound of the EMD in between two signatures. A point in multiple dimensional space is known as centroid of the signature. The lower bound of the EMD in between signatures of s_t and $s_q$is given by



$\phi(K_j\ ,\ B_{j,1})$	ID(m_7), ID(m_9), ID(m_{12}), ID(m_{17})
$\phi(K_j\ ,\ B_{j,2})$	ID(m_{13}), ID(m_{24}), ID(m_{35}), ID(m_{67})
$\phi(K_j\ ,\ B_{j,3})$	ID(m_{27}), ID(m_{49}), ID(m_{62}), ID(m_{73})
. . .	. . .
$\phi(K_j\ ,\ B_{j,N_j})$	ID(m_{47}), ID(m_{59}), ID(m_{92}), ID(m_{117})

Table: An example of j-th hash table

4.4 Local Sensitive Hash

According to LSH the probability of collision is high when compared with close items than that with distant ones. This property can be used in approximating queries of vector space hash functionality 'H' can be given by H={h:U→R} is known as (p,αp,p_1,p_2) sensitive for any X,Y €U if

Pr{h(x)=h(y)}>=p1 for d(X,Y)<=ṗ

Pr{h(x)=h(y)},=p2 for d(X,Y).=αṗ

Where, α>1 and P1>P2..

To increase the gap betweenP1 and P2, we use multiple hash functions and build an other function family G=$\{g:U\rightarrow R\}^{\lambda}$. Here g(v)=($h_1$(v),$h_2$(v),…..$h_\lambda$(v)). This function is the concatenation function of 'λ' LSH function h_i€H.

Algorithm 1:Build Index

Input: the centroid database Ξ, the set of hash functions {g_i}L , the set of keys for hash value {K_i}L , the one way hash function ϕ.

Output: secure index . 1. For each j = 1, , L, data owner builds the j-th hash table by applying function gj over all the elements in centroid database Ξ. One example For each j = 1, , L, data owner picks a random key K_j and replaces each LSH hash digest $B_{j,i,}$ i = 1, , N_j in the j-th hash table with $\phi(K_j, B_{j,i})$. 3. For each j = 1, · · ·, L, data owner further fills the j-th hash tables with identifiers of corresponding images $ID(m_i)$.

Return: Index I consists of L secure hash tables.

5. Security Analysis

In order to increase the search efficiency the proposed system uses profiler table to group similar images together .By this the cloud server will understand that the images in the bucket are similar to each other.The images which are retrieved by the same query are similar too.

Algorithm 2. Secure Transformation

Input: Original problem $\Psi = (c, U, \tau, V, E)$ and secure key $K_T = (G, \Lambda, r, \gamma)$

Output: Transformed problem $\Omega = (c', U', \tau', V', E')$

1: Pick a non-singular k_tk_q*$k_tk_q$matrix Λ and k_tk_q *1 vector r to perform the transformation as

Formula(6);

2: Pick an $(k_t+ k_q)$* $(k_t + k_q)$ generalized permutation matrix G and multiple it to the constraints

as Formula (7);

Return: The transformed problem Ω as shown in Formula (8).

```
K ← KeyGen($1^\kappa$)
  1. Generate secret keys $K_M$, $K_S$ , $K_\Xi$, $\{K_j\}^L$   ;
  2. Output the secret key K = {$K_M$, $K_S$ , $K_\Xi$, $\{K_j\}^L$   }.
I ← BuildIndex(Ξ, {$g_j$, $K_j\}^L$ , ϕ)
     See Algorithm 1. {M′, S′} ← DataEnc(M, S, $K_M$, $K_S$ )
  1. Encrypt all of image $m_t$∈ M with secure key $K_M$, generating encrypted image set M′;
  2. Encrypt all of image $s_t$∈ S with secure key $K_S$ , generating encrypted image set S′;
  3. Output {M′, S′}
$T_q$ ← TrapdoorGen($s_q$, {$g_j$, $K_j\}^L$   , ϕ)
  1. For query signature sq, compute its centroid ξq and initialize the trapdoor q = T;
  2. For each LSH function $g_j$ and key $K_{j,\,j}$ = 1, ..., L,
     1) Generate subtrapdoor$T_{q,j}$= ϕ($K_j$, $g_j(\xi_q)$),
     2) s$T_q$← $T_q$∪$T_{q,j}$;
  3. Output $T_q$.
$S_q'$ ←SearchIndex(I, $T_q$)
  1.  Initialize the candidate encrypted signature set $S_q$ ′ = ∅, image identity set of candidate
  imageID$_q$ = ∅;
  2. For each $T_{q,j,}$ j = 1, ..., L:
    1) In the j-th hash table of index I, retrieve the set of $ID_{Tq,j}$from bucket $B_{Tq,j}$where the
```

```
encrypted bucket value equals to $T_{q,j}$;
        2) $ID_q \leftarrow ID_q \cup ID_{Tq,j}$;
  3. Retrieve the corresponding encrypted signatures $Sq' = \emptyset$ according to identity set $ID_q$;
  4. Output retrieval encrypted signature set $S_q'$ .
$\{\Omega, C\} \leftarrow$ SecureTrans($S_q'$ , $s_q$, $K_S$ )
  1. For each encrypted $s'_t \in S_q'$ ;
      1) Calculate $s_t = Dec_{KS}(s'_t)$;
      2) Compute the distance vector $c^T$ for signature $s_t$and query signature $s_q$;
      3) Formulate the LP problem $\Psi_{q,t}= (c_{q,t}, U_{q,t}, \tau_{q,t}, V_{q,t}, E_{q,t})$ by using the weights information ins$_t$;
      4) Generate secure transformation key $K_{Tq,t}= (G_{q,t}, \Lambda_{q,t}, r_{q,t,}\gamma_q)$;
      5) Compute the transformed problem $\Omega_{q,t}= (c'_{q,t}, U_{' q,t}, \tau_q ' ,t, V'_{q,t}, E_{' q,t})$ using $K_{Tq,t}$ according to Algorithm 2.
      6) Calculate the offset values $C_{q,t} = \{\gamma_q c^T r_{q,t}, \sum k_q k_t r_{q,t,i}\}$;
   2. Output the transformed problem set $\Omega = \{\Omega_{q,t|} s_t \in S_q\}$ and the corresponding offset set $C = \{C_{q,t} | s_t \in S_q\}$.
$M'_q \leftarrow$ DistanceCom($\Omega$, C)
    1.  Solve the transformed problems in $\Omega$, and use the corresponding offset values to compute the EMDs in an order preserving way;
  2. Sort the results according to EMD;
  3. Output the set of top-k ranked encrypted image $M'_q$ .
$M_q \leftarrow$ DataDec($M'_q$, $K_M$)
  1. Decrypt all of the encrypted image $m ' t \in M'_q$ , generating the unencrypted image set $M_q$;
  2. Output $M_q$.
```

5. Implementation and Performance

We have three modules. They are Data Owner, Data Use, Cloud Server

Data Owner

The data owner holds a large-scale image database M={m1,...,m n } to be outsourced, where n is the image number of the database. The data owner generates a searchable index for the image database M. For privacy preserving, the data owner needs to encrypt the image database and the search index, and then outsources the encrypted image database and index to the cloud.

Data User

In the CBIR query phase, the authorized user submits an encrypted query trapdoor to the cloud server. Then, the cloud server compares the similarities between the query image and the images in the database, and returns the encrypted similar images to the data user. Finally, the authorized user decrypts the received images.

Cloud Server

We consider the semi-honest (also known as honest-but-curious) cloud server, who will correctly follow the designated protocol specification, but keep and analyze the communication history, trying to derive sensitive information. The scheme is designed to prevent the cloud server from knowing the content of image database and users' queries. The cloud can provide the CBIR service without interacting with the data owner once the database is outsourced; we need to construct a special searchable encryption scheme that supports CBIR over encrypted data framework and formally prove its security.

5.1 AIES-SCBIR Design and Implementation

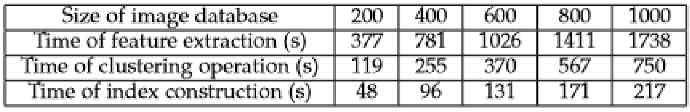
The main element on the users' aspect leverages a unique cryptographic scheme specifically designed for pictures and privacy preserving CBIR, dubbed AIES-SCBIR. Before describing AIES-SCBIR well, we tend to give a definition of image privacy underlines our work. Informally, we tend to define image privacy because the ability to keep the contents of a picture secret to public or merely unauthorized revelation. Typically speaking, image contents are characterized by the mixture of its color and texture information. These two parts kind what one will promptly identify in associate image: objects, people, etc. As such, to safeguard image privacy entails preventing unauthorized entities from recognizing objects in those pictures. We have a tendency to any remark that image color and texture information will be separated from each other. In fact, color data is given from pixel color values within the completely different channels of some color models; whereas texture data is given by the relative position of pixels and strong color changes across neighbouring pixels. We also remark that texture data is sometimes additional relevant in images for visual perception. Finally, we have a tendency to conclude that no sub-component alone (i.e. color or texture information) will be used to infer the precise contents of a picture, as color information on itself is sometimes ambiguous (e.g. robust blue will translate into sky, ocean, etc.) and texture data depends not solely on pixel positions but additionally on their color values. These conclusions are any supported by the foremost recent works in image reconstruction, that not solely rely upon native features extracted from sub-parts of the pictures in this work we concentrate on global options extracted from every image as a whole, however conjointly on those native options not being encrypted. Leveraging the previous definition and observations, we design IES-CBIR, a picture cryptography theme that separates color from texture data, applying completely different cryptography techniques for safeguarding each: action that texture is usually additional relevant than color for visual perception, we design AIES-SCBIR to guard image texture with probabilistic encryption and color data with deterministic cryptography. This way, content-based image indexing and retrieval, based on color data, will be performed on the cloud servers in an exceedingly privacy-preserving method and while not intervention of users, whereas texture data remains protected with the highest level of security we give a close and formal security analysis.

5.2 EFFICIENCY

5.2.1 Index construction

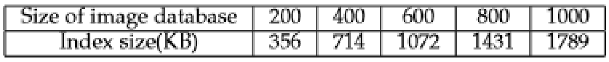
the extraction of SIFT features are used to represent the images.we consider the two preprocesses such as feature extraction and the clustering for the index construction.the process of index construction mainly includes centroid calculation,hash calculation and signature generation.

The time consumption in feature extraction, clustering and index construction is listed in Table



Size of image database	200	400	600	800	1000
Time of feature extraction (s)	377	781	1026	1411	1738
Time of clustering operation (s)	119	255	370	567	750
Time of index construction (s)	48	96	131	171	217

TABLE: Time consumption of feature extraction, clustering operation, and index construction



Size of image database	200	400	600	800	1000
Index size(KB)	356	714	1072	1431	1789

TABLE: Storage consumption of index

5.2.2 Trapdoor Generation

For a query request, the trapdoor generation includes centroid calculation, hash calculation, feature extraction, and signature generation. Similar to the index construction, the feature extraction consumes the most of the time. The average time consumption of trapdoor generation is 2.084 seconds in our 40 times of experiment.

5.2.3 Time of search operation

After receiving the query trapdoor, the cloud server searches on the index (i.e. the LSH tables) to get a identity set of candidate pictures. Then, the corresponding signature set is created and sent to the question user. The query user decrypts the signatures, constructs the secure remodeled EMD problems and sends the transformed problems back to the cloud server. the transformed problems solved by the cloud server are to obtain the top-k ranked images. Finally, the ranked images are sent to the query user for decryption.

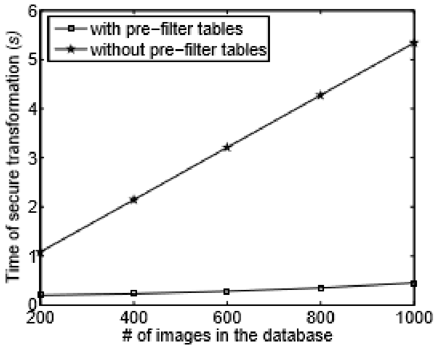


Fig 3: Time of Search Index

the cloud in our scheme consumes less time than that during a query process, while the query user in our scheme consumes more time than that . The previous works, that are mainly devote to secure outsourcing of global-feature based CBIR, but our work solves the problems in local-feature based CBIR. Thus, the comparison of performance may not be quite fair. The local feature-based schemes generally consume more time than the global ones. Please note that, most of the time consumed on the query user side in our scheme is spent on the extraction of local features.

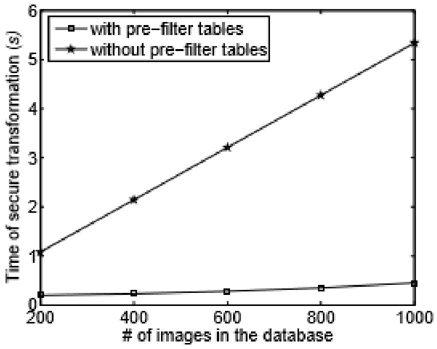


Fig 4:Time of secure transformation

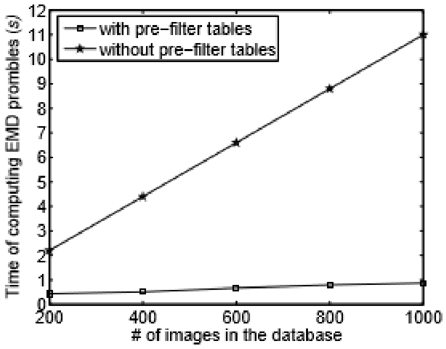


Fig 5: Time of calculation of EMD problems

6. Conclusion and Future work

A privacy-preserving content based image retrieval scheme, which allows the data owner to outsource image database and the CBIR service to the cloud without revealing the actual content of the database. Local features are utilized to represent the images, and earth mover's distance (EMD) is employed to evaluate the similarity of images. We transform the EMD problem so that the cloud server can solve the problem without learning the sensitive information. In order to improve the search efficiency, we design a two-stage structure with LSH. In the first stage, dissimilar images are filtered out by pre-filter tables to shrink the search scope. In the second stage, the remaining images are compared under EMD metric one by one for refined search results. The security analysis and experiments show the security and efficiency of the proposed scheme. In the future, we will study how to outsource the feature extraction to the cloud server so as to further relieve the burden of data owner and data user.

REFERENCES

[1] C. Pavlopoulou, A. C. Kak, and C. E. Brodley, "Content-based image retrieval for medical imagery," in Medical Imaging 2003. International Society for Optics and Photonics, 2003, pp. 85–96.

[2] A. K. Jain, J.-E. Lee, R. Jin, and N. Gregg, "Content-based image retrieval: An application to tattoo images," in Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, 2009, pp. 2745–2748.

[3] R.Jegadeesan, Dr.N.Sankar Ram October -2013 "ENROUTING TECHNICS USING DYNAMIC WIRELESS NETWORKS" International Journal of Asia Pacific Journal of Research Ph.D Research Scholar 1, Supervisor2, VOL -3 Page No: Print-ISSN-2320-5504 impact factor 0.433

[4] J. M. Lewin, R. E. Hendrick, C. J. DOrsi, P. K. Isaacs, L. J. Moss, A. Karellas, G. A. Sisney, C. C. Kuni, and G. R. Cutter, "Comparison of full-field digital mammography with screen-film mammography for cancer detection: Results of 4,945 paired examinations 1," Radiology, vol. 218, no. 3, pp. 873–880, 2001.

[5] B.Pinkas and T.Reinman, "Oblivious RAM revisited" in Advances in Cryptology-CRYPTO 2010.Springer, 2010 ,pp. 502-519.

[6] J. R. and S.-F. Chang. "Tools and techniques for color image retrieval".in Storage and Retrieval for image and video Databases (SPIE), vol. 2670, 1996, pp. 2-7.

[7] R.Jegadeesan , Dr.N.Sankar Ram, M.S.Tharani (September-October, 2013)"Enhancing File Security by Integrating Steganography Technique in Linux Kernel" Global journal of Engineering,Design & Technology G.J. E.D.T., Vol. 2(5): Page No:9-14 ISSN: 2319 – 7293

[8] B. S. Manjunath and W.-Y. Ma, "Texture features for browsing and retrieval of image data," Pattern Analysis and Machine Intelligence,IEEE Transactions on, vol. 18, no. 8, pp. 837–842, 1996.

[9] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," Pattern recognition, pp. 43–90, 2008.

[10] D. G. Lowe, "Distinctive image features from scale-invariant key-points," International journal of computer vision, vol. 60, no. 2, pp.91–110, 2004.

[11] T. Deselaers, D. Keysers, and H. Ney, "Discriminative training for object recognition using image patches," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2. IEEE, 2005, pp. 157–162.

[12] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," International Journal of Computer Vision, vol. 40, no. 2, pp. 99–121, 2000.

[13] H. Ling and K. Okada, "An efficient earth mover's distance algorithm for robust histogram comparison," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, no. 5, pp. 840–853, 2007.

[14] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi- keyword ranked search scheme over encrypted cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. PP, no. 99, p. 1, 2015.

[15] H. Mu¨ller, W. Mu¨ller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: Overview and proposals," Pattern Recognition Letters, vol. 22, no. 5, pp. 593–601, 2001.

[16] J. Z. Wang, J. Li, and G. Wiederhold, "Simplicity: Semantics-sensitive integrated matching for picture libraries," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, no. 9, pp.947–963, 2001.

[17] E. Levina and P. Bickel, “The earth mover’s distance is the Mallows distance: Some insights from statistics,” in Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, vol. 2. IEEE, 2001, pp. 251–256.

[18] Dawn Xiaodong Song David Wagne "Practical Techniques for Searches on Encrypted Data": concept of searchable cryptography, 2000.

[19] M. Kuzu, M. S. Islam, and M. Kantarcioglu "Efficient similarity search over encrypted data":search over encrypted data,2012.

[20] R.Jegadeesan1, Dr.N.SankarRam,T.Karpagam March-2014 “Defending wireless network using Randomized Routing process” International Journal of Emerging Research in management and Technology

[21] Eu-Jin Goh "Secure Indexes": The index reveals no info concerning its contents without valid trapdoors, 2012.

[22] B.S. Manjunath, W.Y. Ma "Texture features for browsing and retrieval of image data": Content based retrieval, IEEE Vol. 18, pp. 837-842, 1996.

[23] R. Curtmola, Juan Garay, Seny Kamara and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions" in CCS'06, 2006, pp. 79-88.

[24] A. Nourian and M. Maheshwaran, "Privacy aware image template matching in cloud using ambient data", J.Supercomput, vol. 66, no. 2, pp. 1049-1070, 2013.

[25] Mingqiang Yang, KidiyoKpalma,and Joseph Ronsin "A Survey of Shape Feature Extraction Techniques",Pattern Recognition, IN-TECH, pp. 43-90, 2008.

[26] R. Chow,P. Golle,M. Jakobsson,E. Shi,J. Staddon "Controlling data in the cloud: outsourcing computation without outsourcing control", in CCSW'09, 2009.

[27] James Z. Wang, Jia Li, and GioWiederhold, "SIMPLIcity:Semantics-sensitive Integrated Matching for Picture LIbraries", IEEE Trans.Pattern Anal. Mach. Intell., vol.23, no. 9, pp. 947-963, 2001.

[28] Jegadeesan,R., Sankar Ram , and J.Abirmi “Implementing Online Driving License Renewal by Integration of Web Orchestration and Web Choreogrphy“ International journal of Advanced Research trends in Engineering and Technology (IJARTET) ISSN:2394-3785 (Volume-5, Issue-1, January 2018

[29].Jegadeesan,R.,Sankar Ram M.Naveen Kumar JAN 2013 “Less Cost Any Routing With Energy Cost Optimization” International Journal of Advanced Research in Computer Networking,Wireless and Mobile Communications.Volume-No.1: Page no: Issue-No.1 Impact Factor = 1.5