International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

# Secure and Efficient Item Information Retrieval in Cloud Computing

**Dr. Shameem Akther[1], Tahseen Begum[2]**

[1]Associate Professor, Department of Computer Science and Engineering, KBN College of Engineering, Kalaburagi, India
*Shameemakther150[at]gmail.com*

[2] M.Tech Student, Department of Computer Science and Engineering, KBN College of Engineering, Kalaburagi, India
*bibifatima236[at]gmail.com*

**Abstract:** ***Cloud computing is a promising information technique (IT) that can organize a large amount of IT resources in an efficient and flexible manner. Increasingly numerous companies plan to move their local data management systems to the cloud and store and manage their item information on cloud servers. For protecting data privacy, sensitive data has to be encrypted before outsourcing. An accompanying challenge is how to protect the security of the commercially confidential data, while maintaining the ability to search the data. Privacy is the protection for the truthful use of personal information of cloud user. In this paper, privacy-preserving data search scheme is proposed, that can support both the identifier-based and feature-based item searches. This scheme is designed to ensure that only legitimate users based on identifiers or keywords, and have the ability to search the data.***

**Keywords:** Item information retrieval, cloud computing, information security

## 1. Introduction

As increasingly numerous data files are being stored locally in enterprises, the pressure on local data storage systems greatly increases. Local hardware failures lead to great damage or loss of data, which greatly affects the daily operations of the enterprises. Fortunately, cloud storage techniques came into being under such circumstances. As Cloud Computing becomes prevalent, sensitive information are being increasingly centralized into the cloud. For the protection of data privacy, sensitive data has to be encrypted before outsourcing. Privacy-preserving data search scheme is designed to ensure that only legitimate users based on identifiers or keywords, and have the ability to search the data. In the proposed system an encrypted item information retrieval system is designed. This system includes two index structures: a hash value index tree, known as an ID-AVL tree, and a height-balanced index tree, known as a product retrieval feature (PRF) tree. Based on the two index trees, two data search methods are supported, i.e., the data users can search the desired item by the identifier or feature vector.

## 2. Related Work

In secure multi-keyword ranked search scheme over encrypted cloud data [1] authors proposed a multi-keyword ranked search scheme, which supports both multi-keyword ranked search and dynamic update. In ranked keyword search, Relevance score is employed to make a secure searchable index and order-preserving mapping function is used. The proposed scheme is designed to provide not only multi-keyword query and accurate result ranking, but also dynamic update on document collections.

In [2] authors defined and solved the problem of effective yet secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria. The authors give a straightforward yet ideal construction of ranked keyword search under the state-of-the-art searchable symmetric encryption (SSE) security definition, and demonstrate its inefficiency. To achieve more practical performance, they proposed a definition for ranked searchable symmetric encryption, and give an efficient design by properly utilizing the existing cryptographic primitive, order-preserving symmetric encryption (OPSE).

In [3] authors defined and solved the challenging problem of privacy preserving multi-keyword ranked search over encrypted cloud data (MRSE).They established a set of strict privacy requirements for a secure cloud data utilization system and further used “inner product similarity” to quantitatively evaluate similarity measure. They proposed a basic idea for the MRSE based on secure inner product computation, and then give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models.

In [4] authors proposed a multi-keyword ranked search over encrypted data based on hierarchical clustering index (MRSE-HCI). The author investigated the problem of maintaining the close relationship between different plain documents over an encrypted domain and proposed a clustering method to solve the problem. The author proposed the MRSE-HCI architecture to speed up server-side searching phase and designed a search strategy to improve the rank privacy.

## 3. Proposed System

An encrypted item information retrieval system is designed in proposed system. This system includes two index structures: a hash value index tree, known as an ID-AVL

tree, and a height-balanced index tree, known as a product retrieval feature (PRF) tree. Based on the two index trees, two data search methods are supported, i.e., the data users can search the desired product by the identifier or feature vector. The elements in the ID-AVL tree are the hash values of the product identifiers and the tree can be directly outsourced to the cloud. Meanwhile, the elements in the PRF tree are plaintext data, and they are encrypted by the secure kNN algorithm before being outsourced. In addition, a detailed depth-first product search algorithm is designed for the PRF tree. Privacy-preserving data search scheme is designed to ensure that only legitimate users based on identifiers or keywords, and have the ability to search the data.

Advantages of proposed system

- An item information outsourcing and searching system model including the data owner, cloud server and data users is designed.
- Two index structures supporting efficient item retrieval are constructed.
- To improve the security of the file, the file is encrypted by a secret key, and the keys of different files are independent.
- To improve the search efficiency, an index structure is constructed for the outsourced data.

The data users can retrieve the interested item in two ways, i.e., retrieving the items by their identifiers or the product feature vector. When a data user wants to search an item based on its identifier, she first needs to encrypt the identifier based on hash function. Next, the hash value of the identifier is sent to the cloud server. The cloud server is responsible for searching for the hash value in the ID-AVL tree, and once the hash value is found, the corresponding encrypted item information is sent to the data user. Finally, the data user can decrypt the item information based on the secret key, and the data retrieval process is completed.

Moreover, in certain cases, the data user may want to search the item based on the features. Initially, the data user needs to construct the feature vector of the item. Then, a depth-first search algorithm is design for the PRF tree. In the initial phase, we need to first locate the most relevant leaf node with the query vector in the tree to initialize RList. Then, the result list is continuously updated by searching the necessary paths in the tree until the final search result is obtained.

**Algorithm 1: DepthFirstSearch**(a PRF Tree With Root $r$,a Query Vector $VQ$)

```
1:  u ← r;
2:  while u is not a leaf node
3:  Calculate all the relevance scores between the child nodes
      of u with VQ;
4:  u ← the most relevant child node;
5:  end while
6:  Select the most relevant k document vectors in u by
    RScore(Vi, VQ) and construct RList;
7:  Stack:push(r);
8:  while Stack is not empty
9:  u ← Stack:pop();
10: if the node u is not a leaf node
11:  if RScore (Vu, max, VQ)>kthScore
12:  Sort the child nodes of u in ascending order based on
       the relevant scores with VQ;
13:  Push the children of u into Stack in order, i.e., the most
       relevant child is latest inserted into Stack;
14:  else
15:    break;
16:  end if
17: else
18:  Calculate the relevance scores between the document
       vectors in the leaf node with VQ and update RList;
19:  end if
20: end while
21: return RList;
```

## 4. System Design

**System Architecture**

As shown in below figure the entire item retrieval system model is composed primarily of three entities: the data manager, the cloud server and the data user.

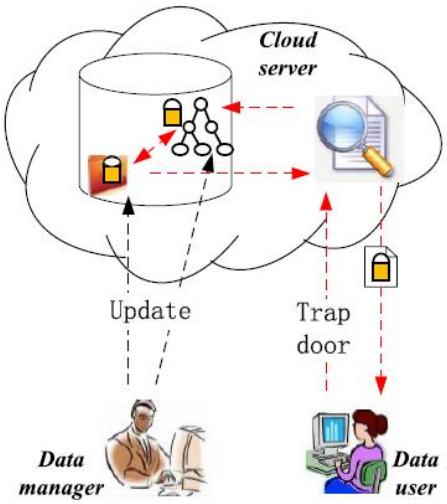



**Figure 1:** System Architecture

The role of data manager is managing the item and collecting the item information. In addition, the data manager needs to encrypt the item information file by a symmetric encryption technique before outsourcing the data to the cloud server. The cloud server stores all the data uploaded by the data manager .When a data user wants to search an item, she needs to generate a trapdoor, which is sent to the cloud server. The cloud server returns the encrypted file. The data user needs to decrypt it by secret key, which is provided by the data manager.

## 5. Modules

**Data Manager**

Data manager should register himself before collecting and managing the item information. After registration is successful, he has to Sign In by using authorized user name and password. The role of data manager is managing and collecting the item information. In addition, the data manager needs to encrypt the item information file by a symmetric encryption technique before outsourcing the data to the cloud server. To improve the search efficiency, an index structure is constructed for the outsourced data. To improve the security of the file, the file is encrypted by a secret key, and the keys of different files are independent.

**Data User**

User should register before searching any item information. After registration is successful, he has to Sign In by using authorized user name and password. When a data user wants to search an item, she needs to generate a trapdoor, which is sent to the cloud server. The cloud server returns the encrypted file. The data user needs to decrypt it by symmetric secret key. This secret key is provided by the data manager.

**Cloud Server**

The cloud server stores all the data uploaded by the data manager. When a data user needs to search the data in the cloud, she first generates a trapdoor, which is sent to the cloud server. A search engineer is employed by the cloud server to act as a bridge between the data users and the encrypted data. Though the cloud server cannot get the plaintexts of the data, it should be capable of sending the accurate search result of the trapdoor to the data users. Of course, the returned data are encoded, and the data user needs to decrypt it by the symmetric secret key which is provided by the data manager.

## 6. Results

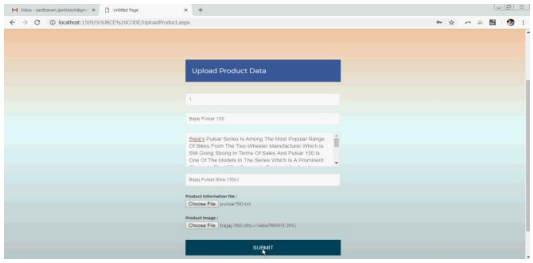



**Figure 2:** Upload item data

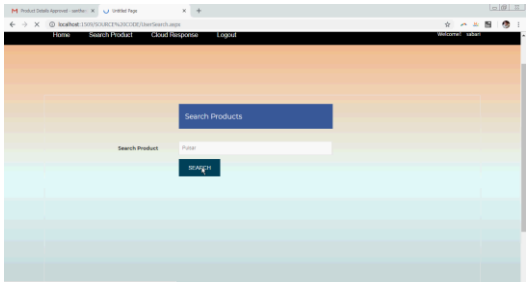



**Figure 3:** Search item

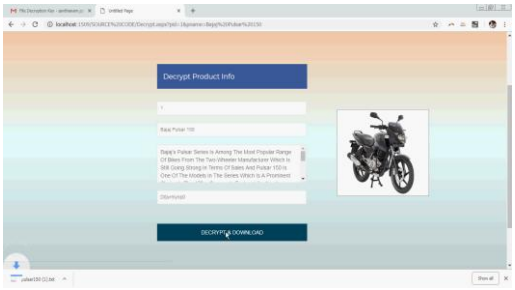
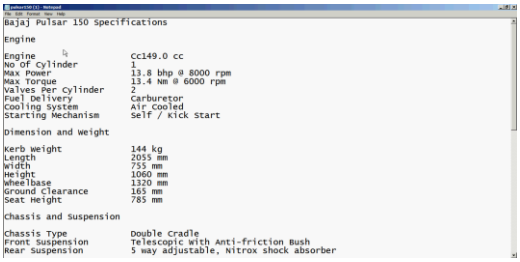



**Figure 4:** Decrypt and download



```
Bajaj Pulsar 150 Specifications

Engine

Engine                  cc149.0 cc
No Of Cylinder          1
Max Power               13.8 bhp @ 8000 rpm
Max Torque              13.4 Nm @ 6000 rpm
Valves Per Cylinder     2
Fuel Delivery           Carburetor
Cooling System          Air Cooled
Starting Mechanism      Self / Kick Start

Dimension and Weight

Kerb Weight             144 kg
Length                  2055 mm
Width                   755 mm
Height                  1060 mm
Wheelbase               1320 mm
Ground Clearance        165 mm
Seat Height             785 mm

Chassis and Suspension

Chassis Type            Double Cradle
Front Suspension        Telescopic With Anti-friction Bush
Rear Suspension         5 way adjustable, Nitrox shock absorber
```

**Figure 5:** Plaintext file

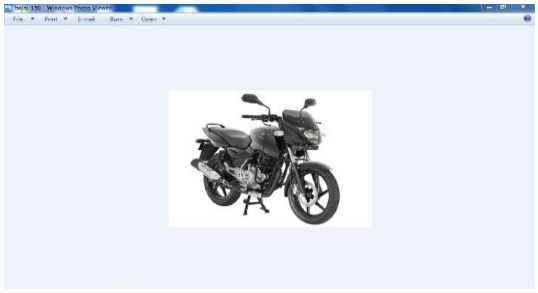


**Figure 6:** Image file

## 7. Conclusion and Future Scope

In this paper, a secure and efficient item information retrieval scheme based on cloud computing is designed. Two index structures supporting efficient item retrieval are constructed, and they support both identifier based and feature based item search. To protect the item information privacy, all the outsourced data are encrypted. The item information is symmetrically encrypted based on a secret key. The future work can be carried out to seamlessly integrate more index structures into the scheme to support more search patterns. The difficult and promising challenge is further improving the search efficiency.

## References

[1] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 8, pp. 1467-1479, Aug. 2012.

[2] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in *Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2010, pp. 253-262.

[3] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi keyword ranked search over encrypted cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 222-233, Jan. 2014.

[4] C. Chen *et al.*, "An efficient privacy-preserving ranked keyword search method," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 4, pp. 951-963, Apr. 2016.
[5] Z. Fu, K. Ren, J. Shu, X. Sun, and F. Huang, "Enabling personalized search over encrypted outsourced data with efficiency improvement," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 9, pp. 2546-2559, Sep. 2016.