Privacy Preserving in image annotation by using Multi label Learning

Sanjaykumar J Khandare, Prof Pravin P Nimbalkar

sanjaykhandare@gmail.com, ppnimbalkar1@gmail.com

JSPM's Imperial College of Engineering and Research, Wagholi, Pune

***Abstract*— For privacy concerns, secure searches over encrypted cloud data motivated several researches under the single owner model. However, most cloud servers in practice do not just serve one owner, instead, they support multiple owners to share the benefits brought by cloud servers. To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, we systematically construct a novel secure search protocol. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.**

Keywords— Image Annotation, multilable, cloud computing.

1. Introduction

Cloud storage system, is set of storage servers, and provides long-term storage services over the Internet. Storing data in a third party's cloud system causes grave to connect to over data secret. Normal hidden schemes defend data secret but have some limitation to functionality of the storage system because a few operations are supported over hidden information. Building a grave storage system that compatible several functions is endurance when system is distributed. Service providers of cloud would pledge to owners data security using phenomenon like virtualization and firewalls. These phenomenon do not protect owner's data privacy from the CSP itself, since the CSP control whole of cloud hardware, software, and owners' data. Hiding the sensitive data before send outside can stored data confidentiality against CSP. Data hidden makes the conventional data utilization service based on plain text keyword search a very challenging problem. A solution to this problem is to download all the hidden data and create the original data using the hidden key, but this is not practical cause it create extra overhead.

In this project, we suggest when search multiple owner multiple keywords that time provide the privacy and show the result in ranking form to make easy cloud servers to perform safe search excluding knowing the real value of both keywords and trapdoors, we properly build a novel safe search rule. So that various data owners use distinct keys to hide their files and keywords. Genuine data users can get a query excluding knowing confidential keys of these various data owners. At the same time, the accuracy of image retrieval will be low as it this can retrieve many relevant images. There are three reasons for low accuracy. To protect from disclosing the result we propose a novel dynamic secret key generation protocol and a new data user authentication rule.

2. Motivation

The main motto for this system is cloud does not support only one or two user instead they supports millions of users and hence privacy issues of data is incurred. We are analyzing the schemes to deal with Privacy preserving Ranked multilable Search in a Multi-owner model. According to our analysis this scheme perform secure search without knowing the actual data of both keywords and trapdoors.

3. Objective

Objective of this system is to keep alive the privacy of related scores between keywords and files and rank the search result, we suggest a novel Additive Order and Privacy Preserving Function family and dynamic hidden key creation rule and a new data user to establish as genuine rule.

4. Literature Survey

ISSN: 2395 - 1621

Wei Zhang, Yaping Lin, Sheng Xiao, Jie Wu, and Siwang Zhou, explore the problem of secure multi-keyword search in multi-keyword search. PRMSM model in this system searches a keywords without knowing actual data of Vol-2 Issue-1 2016

IJARIIE-ISSN(O)-2395-4396 1612

www.ijariie.com 473 trapdoors as well as keywords. This system preserves the keywords and files systematically. In this system sum of the relevance scores is used to search result in metric

M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D.

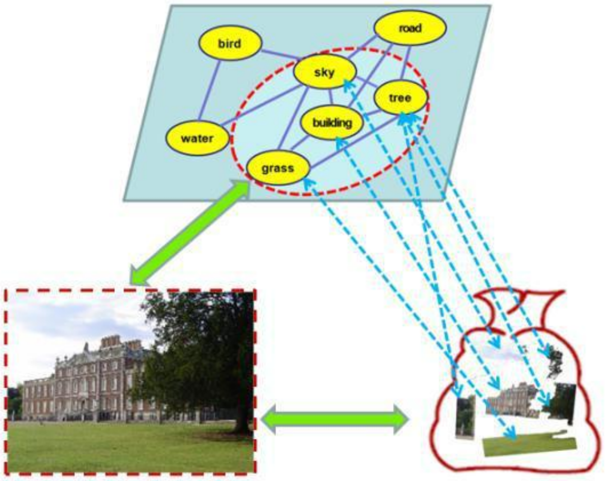
Patterson, A. Rabkin, I. Stoica, and M. Zaharia, provides the simple figure to evaluate the comparison between cloud computing and conventional computing. It also identifies functional and non-functional opportunities of cloud storage.

P. Golle, J. Staddon, and B. Waters, proposed protocols that allow for conjunctive keyword queries on encrypted data. It solves the problem of secure Boolean search.

J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, formalizes and solves the problem of effective fuzzy keyword search over encrypted cloud data and maintains keyword privacy.

Q. Zheng, S. Xu, and G. Ateniese, uses ABE to construct a new primitive called attribute -based keyword search (ABKS). This scheme prevents a data owner from knowing the keywords a data user is searching owners/trusted. In this system authors extends the access tree to privilege trees. In this data files having several operations are executable itself. This system having compromises in terms of authority to tolerate

5. Architectural Diagram



6. Technologies to be used

JAVA:

Java has been tested, refined, extended, and proven by a dedicated community of Java developers, architects and enthusiasts. Java is designed to enable development of portable, high-performance applications for the widest range of computing platforms possible. By making applications available across heterogeneous environments, businesses can provide more services and boost end-user productivity, communication, and collaboration— and dramatically reduce the cost of ownership of both enterprise and consumer applications.

The original and reference implementation Java compilers, virtual machines, and class libraries were originally released by Sun under proprietary licenses. As of May 2007, in compliance with the Specifications of the Java Community Process, Sun re-licensed most of its Java technologies under the GNU General Public License. Others have also developed alternative implementations of these Sun technologies, such as the GNU Compiler for Java (byte code compiler), GNU Class path (standard libraries), and Iced Tea-Web (browser plug in for applets).

Eclipse is an integrated development environment which is mostly written in java. Developing java application is its primary purpose although it may also be used to develop applications in other programming languages like C, C++ etc. Customer module (Android application) is designed using eclipse in this project. Both front end user interface and back end coding of the android application is done using eclipse. The operating systems that support eclipse are Linux, Mac Operating System, Solaris, and Windows. It works on both 32 and 64-bit variant Windows.

MySQL:

MySQL is the most popular Open Source Relational SQL Database Management System. MySQL is one of the best RDBMS being used for developing various web-based software applications. MySQL is developed, marketed and supported by MySQL AB, which is a Swedish company. *MySQL* is the most popular Open Source Relational SQL Database Management System.

MYSQL Enterprise edition includes the most comprehensive set of advanced features & management tools for MYSQL.

MYSQL is the world's most popular open source database. Whether you are a fast-growing web property, technology ISV or large enterprise, MYSQL can cost-effectively help you deliver high performance, scalable database applications

MYSQL is popular choice of database for used in web application & is a central component of widely used LAMP open source web application software stack.

MYSQL Query Analyzer: To optimize performance by visualizing query activity and fixing problem SQL code.

AES:

Each cipher encrypts and decrypts data in blocks of 128 bits using cryptographic keys of 128-, 192- and 256-bits, respectively. The cipher was designed to accept additional block sizes and key lengths, but for AES, those functions were not adopted. It is a symmetric encryption algorithm. The algorithm was developed by two Belgian cryptographer Joan Daemen and Vincent Rijmen. AES was designed to be efficient in both hardware and software, and supports a block length of 128 bits and key lengths of 128, 192, and 256 bits. The first transformation in the AES encryption cipher is substitution of data using a substitution table; the second transformation shifts data rows, the third mixes columns. The last transformation is a simple exclusive or (XOR) operation performed on each column using a different part of the encryption key longer keys need more rounds to complete.

IMAGE ANNOTATION TECHNIQUES:

Making use of Textual Information:

The huge numbers of images are available on the World Wide Web. In order to categorize and competently retrieve them, background information of the images such as surrounding content and associations can be used for image annotation. Automatically we can obtain semantic knowledge for Web images. Similarly context can be assigned to web images by using page layout analysis method. At the same time, the accuracy of image retrieval will be low as it this can retrieve many relevant images. There are three reasons for low accuracy. Firstly, Web images can be used by anyone in the Web pages and there is no standard exists for the relationships between the texts and inserted images in the Web pages. Secondly, Web images are fairly wide-ranging in meaning, because they are created by different group for different reasons. Thirdly, the qualities of the Web images vary significantly. The users require passing through the whole list of retrieved images to search the preferred ones.

Manual Annotation:

In manual annotation users have to enter some descriptive keywords when the images are loaded/registered/browsed. Manual annotation of image content is considered “best case” in terms of accuracy, since keywords are selected based on human determination of the semantic content of images. But at the same time, it is an effort intensive and monotonous process. Manual annotation can have a problem that at the retrieval, users can forget the annotations they have used after a long period of time.

7. Overall Description

7.1 PRODUCT PERSPECTIVE:

Data owners to outsource their data to public cloud servers while allowing data users to retrieve this data. For privacy concerns, secure searches over encrypted cloud data has motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing.

7.2 REQUIREMENTS:

SOFTWARE REQUIREMENTS:

1) Eclipse Mars 1
2) Apache Tomcat Server 7
3) JDK 1.7

HARDWARE REQUIREMENTS:

1. 8 GB RAM
2. 500 GB HDD
3. RSA for encryption

7.3 PRODUCT FUNCTION:

User:

Will have to apply for login credentials.

Admin:

1. Admin will approve or disapprove the User request for login credentials.
2. If admin approves the request, then Admin will send Email to the user Username, Password, Secret Code to download and delete file
3. If admin disapproves the request then user record will be deleted and disapprove Email will send to the user.

User can upload the file.

1. At the time of file uploading index will be generated
2. User can search a File with entering keyword and according to result he/she will request for owner of file to download

Owner of file will approve or disapprove the request

1. If approved secret pin will be shared with requested user
2. User can delete his own file by using secret pin.
3. On Delete following operation will happen
   1. User provides secret pin on rise of delete request.
   2. If file has any pointer then only database entry will be deleted
4. On Download following operation will happen:
   1. User provides secret pin on rise of download request.
   2. If secret pin matched, then only file will be downloaded.

8. Mathematical Model

$S=\{s,e,X,Y,T,F_{main},NDD,DD,Success,Failure\}$

S(System) = Is our proposed system which includes following tuple.

s (initial state at time T) = GUI of search engine. The GUI provides space to enter a query/input for user.

X (input to system):- Input Query. The user has to first enter the query. The query may be ambiguous or not. The query also represents what user wants to search.

Y (output of system):- List of URLs with Snippets. User has to enter a query into search engine then search engine generates a result which contains relevant and irrelevant URL's and their snippets.

T (No. of steps to be performed) :- **2**. These are the total number of steps required to process a query and generates results.

f_{main}(main algorithm):- It contains Process P. Process P contains Input, Output and subordinates functions. It shows how the query will be processed into different modules and how the results are generated.

DD (deterministic data):- It contains Database data. Here we have considered MySQL, which contains number of queries. Such queries are user for showing results. Hence, MySQL is our DD.

NDD (non-deterministic data):- No. of input queries. In our system, user can enter numbers of queries so that we cannot judge how many queries user enters into single session. Hence, Number of Input queries are our NDD.

Memory shared: - MYSQL. MYSQL will store information like User Authentication, Performing Operations like Login credentials, File that is to be uploaded. Since it is the only memory shared in our system, we have included it in the MYSQL.

CPU_{count}:- 1. In our system, we require 1 CPU for server.

Success = successfully recommended best system as per user's interest

Failure = If application will not send the notification to user it will fail.

Subordinate functions:

$S=\{s,e,X,Y,F_{main},NDD,DD,Success,Failure\}$

Where

s=Start State

e=End State

X={Set Of Inputs}

= {x1}

Where x1= Login credentials, File that is to be uploaded, Secret pin on download and delete request.

Y={Set of Outputs}

= {y1}

Where y1= Encrypted file will be uploaded, On download decrypted file will be downloaded

F_{main} = {Set of procedure}

= {f1,f2}

Where

f1= Take x1 Input

f2= Give y1 Output

State Transition Diagram:

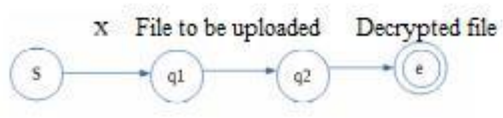


Fig:State Transition Diagram

Where,

s=input state

x=query

q1= Login credentials, File that is to be uploaded, Secret pin on download and delete request.

q2= Encrypted file will be uploaded, On download decrypted file will be downloaded.

Explanation:

The q1 state accept the ambiguous query 'x' from the state 's' which is our initial state.

The q1 state for Login credentials, File that is to be uploaded and sent to state q2.

The q2 state is meant for Encrypted file which will be uploaded, On download decrypted file will be downloaded.

9. CONCLUSION

We conclude that provide the feasible solution for preserving privacy for multi-data owners. In this project, we hide user's identity that is having data on cloud, to level up the security constraint, provide backup facility in which last modified copy of data should preserve. The data backup is in the encrypted format and it is restoring when required.

10. REFERENCES

- http://ijariie.com/AdminUploadPdf/Privacy_Preserving_Ranked_Multi_Keyword_Search_for_Multiple_Data_Owners_in_Cloud_Computing_ijariie1612.pdf
- http://www.ijircce.com/upload/2015/november/96Privacy.pdf
- http://personal.stevens.edu/~ychen6/projects/Privacypreserving%20Ranked%20MultiKeyword20Search%20Leveraging%20Polynomial%20Function%20in%20Cloud%20Computing.pdf