Recognition of Encrypted Data on Cloud

Dayanand G Savakar[1], **Ravi Hosur**[2]

[1]Rani Chennamma University, Belagavi

[2]BLDEA's Vachana Pitamaha Dr. P.G.Halaktti College of Engg. & Tech, Bijapur

Abstract: ***The work introduces to classify the encrypted data which are outsourced to the cloud. Nowadays cloud computing is growing in popularity with most of the companies are moving towards it because of its interesting characteristics. It is a network which stores large data, manages and processes, and provides the security to it by storing in an encrypted form; so as to protect sensitive information of the users. But classifying the data which is in encrypted form is a complex procedure. So, we have proposed privacy preserving k-nearest neighbor algorithm to maintain the privacy of the data even after classifying it. This algorithm will give the accurate data mining results. Thus, the system tends to provide efficient functionalities and promises to have security guarantees to the data incorporated into cloud.***

Keywords: Network, Cloud, Security, k-NN algorithm, classification

1. Introduction

Cloud computing is the trending technology which is rapidly growing day by day in a blink of an eye in today's IT world. It provides a rich set of potential computing features like storage space, server maintenance in automated way as required without involvement of any external entity (human). Thus, the organization's way of handling their data is becoming easier. To get relaxed from troublesome activities of maintaining a set of resources, organizations are moving towards cloud computing technology.

Nowadays insurance companies are using this technology so as to preserve their client's data which are confidential. In present system most of the insurance companies are facing problem of managing data stores, hence affects data privacy. To overcome these issues, the insurance companies tend to utilize the advantages of cloud. Here data should be encrypted earlier for outsourcing to the cloud if it is highly sensitive. But categorization of data without decrypting it is the challenging task for organizations, whereas the classification is the commonly used task in data mining process. So we introduce a novel method called "Privacy Preserving k-Nearest Neighbor (PPkNN)" algorithm for encrypted data.

2. Review of Literature

Literature review is the document which contains the ground study of the existing system's drawbacks to introduce new proposed system. It includes the place, person or publications which contains the relevant topics to get a better planning for what to do and how to go about it. In existing-system companies are fed up with maintenance of databases and heavy amount is invested on them. The data is also stored in database in a human readable form. When unauthorized person tries to access the database he can easily get the sensitive form of data. Thus there is a problem of handling the data to be private. In some systems even though the companies are using cloud technology, they are unable to classify the encrypted data properly.

The method [1], witness's huge notice and a wealth of assure in content-based image recovery as a rising technology. It also a horizontal way for a huge number of new techniques and systems, get various new citizens include. In this piece, we survey almost 300 new hypothetical and experimental charity in the existing decade related to image recovery and regular image clarification. We also discuss significant challenges involved in the difference of existing image recovery techniques to build systems that can be useful in the genuine world. In retrospect of what has been achieved so far, we also work out what the prospect may hold for image recovery study.

Predictable methods [2] of image revival require that metadata is connected with the image, usually known as keywords. Though some content based image retrieval systems utilize together semantic and prehistoric attributes to relation search principle, history has proven that it is tricky to remove linguistic in sequence from a 2D picture. In this observe, activity theory is used as a foundation to express how semantic in sequence can be retrieved from objects recognized in a picture. Via an picture segmentation method. By The Berkeley Digital Library Project, and merge it with, a high-level accepting of he picture can be established Content-Based Image Retrieval [3] has become one of the popular most research areas. Many diagram attribute representations contain been explored and many systems build. While, these research information found the foundation of satisfied based image recovery, the kindness of the future approaches is incomplete. Specially, these efforts have comparatively overlooked two different characteristics of systems the space between towering level concepts and low level skin texture bias of human compassion of visual content. Which electively takes into account the above two uniqueness in CBIR. During the recovery process, the user's high level query and insight partisanship are captured by dynamically updated weights based on the user's advice. The provisional results over more than 70,000 images show that the future approach greatly reduces the user's effort of composing a doubt and capture the user's in sequence.

Application feedback [4] scheme based on support vector equipment have been generally used in content-based image retrieval. However, the arrangement of based application criticism is frequently abridged when the figure of labeled positive advice sample is little. This is mostly due to three reasons a classifier is disturbed on a little sized teaching locate, and over suitable happens since the number of characteristic dimensions is much senior than the size of the preparation set. In this document, we expand a device to overcome these troubles. To speak to the first two troubles, we propose an asymmetric container based. For the third problem, we combine the random subspace method and SVM for application feedback, which is named random subspace SVM (RS-SVM). Finally, by AB-SVM and RSSVM, an asymmetric bag and accidental subspace SVM (ABRS-SVM) is build to solve these three problems and further improve the application feedback performance. Some researchers used Image processing techniques for security[5][6] and for agriculture and horticulture produce[7][8].

3. Proposed System

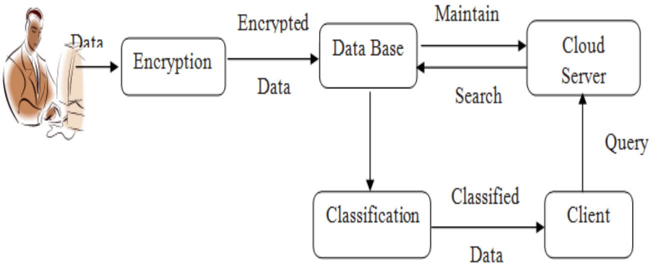
Implementation of k-nearest neighbor algorithm to preserve the privacy of encrypted data in the cloud is a major thing in proposed system. This algorithm is used to overcome through the disadvantages of existing system. The data mining over encrypted problem will be solved using K-nearest neighbor classifier.

In this project the owner will collect the information from the customers who are interested in taking the policy from his/her company. Then he/she will encrypt the data using advanced encryption standard (AES) algorithm and stores in the data base. He/she can view and update the details of the customers too. After storing the encrypted data in the cloud, customer will get the email which contains policy number and password. Using these two inputs customer can login and see his/her details and risk value. The agent of the company is responsible to request details for classification. Agent will give the

K-value as classification limits so as to find the risk value of the customer and then he/she will be classified with HIGH, LOW, MID as class labels.

To classify the data k-nearest neighbor algorithm is used because it is the best solution to overcome through the disadvantages of existing system. Here the major task is to classify the encrypted data which is achieved by it, that meets the following privacy requisites:

- Classifying the encrypted data with taking care of the privacy of the information in the outsourced data
- Once the data is outsourced by the owner he/she will not participate in any other computations of the customer's details. He/she can only update the customer details
- Only the class label of the customer risk value known to the Agent. Other than this owner is not permitted to go through the customer details for any computations



This scheme is introduced to consider the security of data in the cloud while classifying the data. So using k-nearest neighbor algorithm is a good way to handle this situation where other classifying algorithms are not so good to consider. Initially, data owner will collect and store the customer details in the cloud with encoded form. When agent sends the query to cloud requesting for customer details, cloud will search the database for agents' request and classifies data.

4. Functional Requirements

This will specify the important functionalities of software that must be present in the system. In brief they suggest that what a system should be capable to do. Functional requirements will provide the method of tracking the system progress concern the completion of project on time.

Owner registration page

Owner must register with his valid details such as owner name, password, email id, address, gender, and phone number. After successful registration he can login to the system with valid username and password.

Owner Login page

After successful login owner must collect the information from customer to store his/her data in the cloud. Owner can upload, update, view the customer details and as well as of agent details he can see.

Upload customer details

Initially all the personal information, contact information and then some policy details are uploaded to the database.

Customer login page

After successful registration of customer he will get the policy number and username to his/her mail id through which the successful login will be done to see risk value and respective agent details.

Agent login page

After successful registration agent can login with valid username and password to view the customer details and to know the risk value particular customer.

View risk value of customer

To know the risk label of customer initially agent must send records of particular customer and then he should enter the k value as classification limit. Then the cloud will send top nearest neighbor class labels.

5. Non-Functional Requirements

These are fundamentals of a system that a system must have. These requirements will not change the behavior of the system. The software should capture the purpose of the applicable functionality. Such as:

Performance requirements
Performance of the system is good as it responds to any users of the system very fast. Number of the simultaneous users of the system shall be more to give the performance of system good

Accuracy
System will give the accurate results in finding out the risk level of users by calculating it with the help of customers given detailed and accurate policy information

Security
We are storing all information in cloud so there is no doubt that system will leak the secrecy of data because it will be stored in encrypted format.

Modifiability
While updating the customer details owner will ask required information to make changes for example to change the marital status from single to married customer should submit the marriage certificate

Portability
The system can be portable to any platform as we have designed in java technology it is platform independent

Usability
Usability of the system is increased by advertising the system and by the interest of agent towards the system

Reliability
System is reliable towards finding out the errors occur in some situation such as in giving valid email id, contact number etc.

6. Conclusion

To protect user data there are different classification methods are introduced long ago. But those aren't suitable to outsourced data where information will be in encoded form. So we have proposed the special k-nearest neighbor algorithm to assure the secrecy of the data in the cloud. In this project we have implemented the work using k-nearest neighbor algorithm on encrypted data to classify all the customers taking k-value as classification limit. This system saves the privacy of the data, agents' request query, and customers' risk value over the cloud with remarkable efficiency.

7. Future Enhancement

Improving the effectiveness of the cloud security is a valuable first step for improving the accomplishment of our Privacy Preserving k-Nearest Neighbor protocol, we can think of new ideas to search different and more effective solutions to the task of classification of encrypted data in our upcoming work. Also we will examine and increase our search to other classification algorithms.

References

[1] R. Datta, D. Joshi, and J.Z. Wang (2007), "Image Retrieval: Ideas, Influences, and Trends" ACM Computing Surveys, vol. 40, article 5

[2] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain (2000),"Content-Based Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380

[3] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra(1998), "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655

[4] X.S. Zhou and T.S. Huang (2003), "Relevance Feedback in Image Retrieval: A Comprehensive Review," Multimedia Systems, vol. 8,pp. 536-544D.G.Savakar, Anand Ghuli (2015), "Digital Watermarking A Combined Approach by DWT, Chirp-Z and Fast Walsh-Hadamard Transform", IJCTA, Vol. 5 No.6, pp 2006-2010.

[5] D.G.Savakar, Anand Ghuli (2015), "Digital Watermarking as a distributed noise by Discrete Wavelet Transformation, Fast Fourier Transformation and Fast Walsh-Hadamard Transform to study the sensitivity between Robustness and Fidelity", IJCA, Issue 1, Volume 5, pp 102-107

[6] Dayanand G. Savakar (2012), Identification and Classification of Bulk Fruits Images using Artificial Neural Networks. International Journal of Engineering and Innovative Technology (IJEIT), Volume 1, Issue 3, Pages: 35-40

[7] Dayanand G. Savakar (2012), Recognition and Classification of Similar Looking Food Grain Images using ANN, Journal of Applied Computer Science and Mathematics ,Volume 13(6), Pages: 61- 65