Privacy Protection In Interactive Content Based Image Retrieval(CBIR)

Mercy A[1], Nandhini M[2], Sangeetha K[3], RajaRamanan V[4]
[1, 2, 3] Dept of Computer Science and Engineering
[4]Assistant Professor, Dept of Computer Science and Engineering
[1, 2, 3, 4] Adhiyamaan College of Engineering, Hosur, India.

Abstract- *Privacy protection in Content Based Image Retrieval (CBIR) is a new research topic in cyber security and privacy. The state-of-art CBIR systems usually adopt interactive mechanism, namely relevance feedback, to enhance the retrieval precision. How to protect the user's privacy in such Relevance Feedback based CBIR(RF-CBIR) is a challenge problem. In this paper, we investigate this problem and propose a new Private Relevance Feedback CBIR(PRF-CBIR) scheme.PRF-CBIR can leverage the performance gain of relevance feedback and preserve the user's search intention at the same time. The new PRF-CBIR consists of three stages: 1) private query; 2) private feedback; 3) local retrieval. Private query performs the initial query with a privacy controllable feature vector; private feedback constructs the feedback image set and local retrieval finally re-ranks the images in the user side. Privacy analysis shows that PRF-CBIR fulfils the privacy requirements. The experiments carried out on the real -world collection confirm the effectiveness of the proposed PRF-CBIR scheme.*

Keywords *Cyber security, Relevance feedback, Private query, Private feedback.*

I. INTRODUCTION

Due to the popularity of various digital cameras and the rapid growth of social media tools for internet-based photo sharing, recent years have witnessed an explosion of the number of digital photos captured and stored by consumers. A large portion of photos shared by users on the Internet are human facial images. Some of these facial images are tagged with names, but many of them are not tagged properly. This has motivated the study of auto face annotation, an important technique that aims to annotate facial images automatically. Auto face annotation can be beneficial to many real world applications. For example, with auto face annotation techniques, online photo-sharing sites (e.g., Facebook) can automatically annotate user's uploaded photos to facilitate online photo search and management.

One challenge faced by such SBFA paradigm is how to effectively exploit the short list of candidate facial images and their weak labels for the face name annotation task. To tackle the above problem, we investigate and develop a search-based face annotation scheme. In particular, we propose a novel unsupervised label refinement (URL) scheme by exploring machine learning techniques to enhance the labels purely from the weakly labeled data without human manual efforts. We also propose a clustering- based approximation (CBA) algorithm to improve the efficiency and scalability. We investigate and implement a promising searchbased face annotation scheme by mining large amount of weakly labeled facial images freely available on the WWW. We note that a short version of this work had appeared in SIGIR2011. This journal article has been significantly extended by including a substantial amount of new content. We conducted an extensive set of experiments, in which encouraging results were obtained.

II. LITERATURE REVIEW

For content level access, very often database needs the query as a sample image. However, the image may contain private information and hence the user does not wish to reveal the image to the database. Private content based image retrieval (PCBIR) deals with retrieving similar images from an image database without revealing the content of the query image - not even to the database server. We propose algorithms for PCBIR, when the database is indexed using hierarchical index structure or hash based indexing scheme. Experiments are conducted on real datasets with popular features and state of the art data structures. It is observed that specialty and subjectivity of image retrieval (unlike SQL queries to a relational database) enables in computationally efficient yet private solutions. Our solution enables both encrypted storage and searching using CBIR queries while preserving privacy. Hence our proposed work achieves better performance compared to existing methods.

III. EXISTING SYSTEM

Most of existing research efforts focused on privacy protection for CBIR in the cloud computing environment. In this scenario, the privacy issue arises because of untrustworthy cloud. The main challenge is to leverage the tremendous computing power of the cloud for CBIR, whilst preventing the

cloud from learning anything useful about the image dataset and query. While the images can be protected separately using image encryption technologies, the main concern is how to perform the similarity computation among the image features in a privacy preserving way. Feature/index randomization and encryption are the commonly adopted methods. Very few works have been devoted to address the privacy issue in the viewpoint of the untrustworthy CBIR service provider. The major privacy concern in this situation is that the user's search intention could be learned by the service provider. Based on the search intention, service provider can infer the user's profile, such as user's interest, living place, health condition and even commercial secret. For example, in the medical field, query with tumor images could leak the user's health condition to service provider. In the viewpoint of privacy protection for big data, user's search intention on CBIR is a primary source to profile user for potential malicious activities.

IV. PROPOSED SYSTEM

The issues faced during image based content search now days are user privacy issue and relevant images with respect to user query is not accurate. Our proposed method initially focuses on input image and extract its histogram value and given it as input for search in cloud. Keeping this as a query cloud could not identify or predict user's interest hence privacy issue of user is reduced. The other most important issue faced by clients are most of the search result does not meet user expectation which means current data retrieval method incorporates some irrelevant results which are may be related to that particular query. To overcome this issue user feedback approach has been taken as input once a user receives certain image as output for his/her query he/she may give feedback to the results generated. That is according to particular user query the received results which are relevant to that particular query will be marked as most relevant and this information is collected as feedback from clients. With this information our database will be retrained and stored. Hence our proposed work achieves accurate result and secure image based content search in cloud.

V. MODULES

REGISTRATION MODULE:

Both user and admin should initially register in our system to upload and search data from our database. Admin module plays an important role in our project. Admin module is responsible for uploading and deleting file in a database. Similarly receiving feedback from users and acting respective to it are the responsibility of admin. Once admin login to our system he/she can upload an image with its respective keyword for efficient search result.

HISTOGRAM VALUE EXTRACTION:

In this module, users will login to our system and select image for search. Now the selected image is considered as input and its histogram value is calculated. The first step is to divide the range of values into a series of intervals (bin) of equal sizes. These intervals are consecutive, non-overlapping and adjacent. In the second step frequency of each bin is generated which is nothing but the number of values in each bin. A gray scale image's histogram generation process typically plots the frequency at which each grey-level occurs from 0 to 255. Histogram represents the frequency of occurrence of all grey-levels in the image, that means it tell us how the values of individual pixel in an image are distributed. The color histogram generation approach uses a method relatively different from the gray scale even though values in all the three (3) channels such as what is found in the RGB (Red, Green, and Blue) images are used as well.

EDGE DETECTION AND CLASSIFIER:

Edge detection methods, transforming original images into edge images, benefit from the changes of grey tones in the image. In image processing especially in computer vision, the edge detection treats the localization of important variations of a grey level image and the detection of the physical and geometrical properties of objects of the scene. It is a fundamental process which detects and outlines of an object and boundaries among objects and the background in the image and it is the most familiar approach for detecting significant discontinuities in intensity values. Finding the edge direction is trivial once the gradient in the x and y directions are known. However, you will generate an error whenever sum X is equal to zero. So in the code there has to be a restriction set whenever this takes place. Whenever the gradient in the x direction is equal to zero, the edge direction has to be equal to 90 degrees or 0 degrees, depending on what the value of the gradient in the y-direction is equal to. If GY has a value of zero, the edge direction will equal 0 degrees. Otherwise the edge direction will equal 90 degrees.

FEEDBACK MODULE:

In this module, the result obtained from the database according to user query (input histogram and edge detection) is shown. User will select most relevant images with respective to his/her query and submit it as feedback to our system. This feedback is used to avoid irrelevant data retrieval from database with respect to user query. Our system

calculates edge detection, text classification and majority voting for reorder the results for improving performance of our system.

VI. CONCLUSION

This problem addressed a new problem of privacy protection in RF-CBIR. A new PRF-CBIR scheme is proposed to protect the user's search intention and leverage the performance gain of relevance feedback. PRF-CBIR consists of three stage: private query, private feedback and local retrieval. PRF-CBIR can deal with query attack, result attack and feedback attack existing in RF-CBIR. We provided a theoretical analysis on privacy protection of the new scheme. It shows the new scheme can effectively control privacy leakage and significantly reduce the attack success probability. Moreover, we carried out a large number of experiments on a real world image collection. The results demonstrate that the privacy preserving performance of PRF-CBIR is significantly improved compared to RF-CBIR, while the retrieval performance sacrifice is acceptable.

VII. RESULTS

Fig.1. Login Page

Fig.2. Main Page

Fig.3. Advanced Search

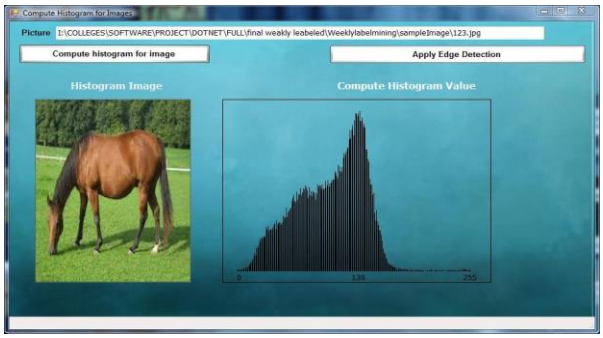

Fig.4. Computing Histogram

Fig.5. Edge Detection

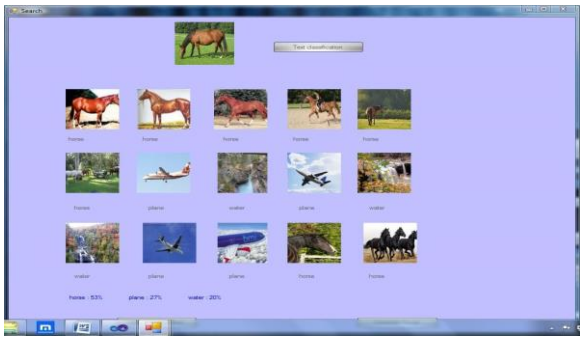

Fig.6.Text Classification

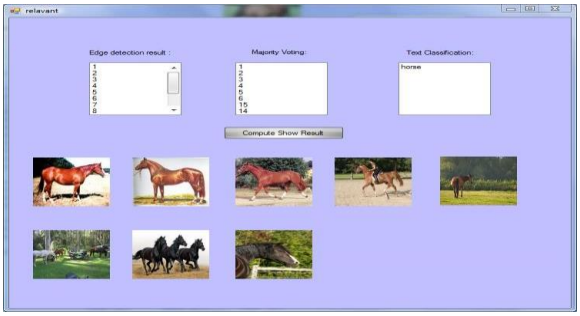

Fig.7. Relevant Images

REFERENCES

[1] W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, 2000.

[2] Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," Journal of Visual Communication and Image Representation, vol. 10, no. 1, pp. 39-62, 1999.

[3] R. Datta, J. Li, and J. Z. Wang, "Content-based image retrieval: Approaches and trends of the new age," in The 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, Hilton, Singapore, 2005, pp. 253-262.

[4] W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, "Enabling search over encrypted multimedia databases," in Proc. SPIE, vol. 7254, San Jose, CA, 2009.

[5] W. Lu, A. L. Varna, A. Swaminathan, and M. Wu, "Secure image retrieval through feature protection," in The 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, 2009, pp. 1533–1536.

[6] W. Lu, A. L. Varna, and M. Wu, "Confidentiality-preserving image search: A comparative study between homomorphic encryption and distance-preserving randomization," IEEE Access, vol. 2, pp. 125-141, 2014.

[7] J. Yuan, S. Yu, and L. Guo, "SEISA: Secure and efficient encrypted image search with access control," in 2015 IEEE Conference on Computer Communications (INFOCOM), Kowloon, 2015, pp.2083-2091.

[8] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing," IEEE Transactions on Cloud Computing, no. 99, 2015.

[9] Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Privacy-preserving content-based image retrieval in the cloud," in The 2015 IEEE 34th Symposium on Reliable Distributed Systems (SRDS),, Montreal, QC, 2015, pp. 11-20.

[10] L. Zhang, T. Jung, P. Feng, K. Liu, X. Y. Li, and Y. Liu, "PIC: Enable large-scale privacy preserving content-based image search on cloud," in The 44th International Conference on Parallel Processing (ICPP), Beijing, 2015, pp.949-958.

[11] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G.Lee, D. Patterson, A. Rabkin, I. Stoica, and M.Zaharia, "A view of cloud computing," Communications of the ACM, vol. 53, no. 4, pp. 50–58, 2010.

[12] J. Shashank, P. Kowshik, K. Srinathan, and C. V. Jawahar, "Private content based image retrieval," in IEEE Conference on Computer Vision and PatternRecognition (CVPR), Anchorage, AK, 2008, pp. 1–8.

[13] L. Weng, L. Amsaleg, A. Morton, and S. Marchand-Maillet, "A privacypreserving framework for large-scale contentbased information retrieval," IEEE Transactions on Information Forensics and Security, vol. 10, no. 1, pp. 152– 167, 2015.

[14] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval," IEEE Transactions on Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644–655, 1998.