INTERNATIONAL JOURNAL · ENGINEERING RESEARCH ·

International Journal for Innovative Engineering and Management Research

A Peer Revieved Open Access International Journal

www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 21st Sept2017. Link :http://www.ijiemr.org/downloads.php?vol=Volume-6&issue=ISSUE-8

Title: **PRIVACY SECURITY FOR USER UPLOADED IMAGES ON CONTENT SHARING SITES**

Volume 06, Issue 08, Pages: 274– 279.

Paper Authors

B SHANMUK KUMAR,M VENKATESH NAIK,DR.G.PRAKASH BABU

St Mark Educational institution society group of institution, AP..

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

PRIVACY SECURITY FOR USER UPLOADED IMAGES ON CONTENT SHARING SITES

B SHANMUK KUMAR[1], M VENKATESH NAIK[2] , DR.G.PRAKASH BABU[3]

[1]PG Scholar, CSE, St Mark Educational institution society group of institution, AP.

[2]Assistant Professor, CSE, St Mark Educational institution society group of institution, AP.

[3]Professor, CSE, St Mark Educational institution society group of institution, AP.

ABSTRACT:Social Network is an emerging E-service for content sharing sites (CSS). It is emerging service which provides a reliable communication, through this communication a new attack ground for data hackers; they can easily misuses the data through these media. Some users over CSS affects users privacy on their personal contents, where some users keep on sending unwanted comments and messages by taking advantage of the users' inherent trust in their relationship network. By this privacy of the user data may be loss for this issue this paper handles the most prevalent issues and threats targeting different CSS recently. This proposes a privacy policy prediction and access restrictions along with blocking scheme for social sites using data mining techniques. To perform this, the system utilizes APP (Access Policy Prediction) and Access control mechanism by applying BIC algorithm (Bayesian Information Criterion).

Keywords: Adaptive Privacy Policy Prediction (A3P), A3P- Core, A3P- Social, Polar Fourier Transform (PFT)

I. INTRODUCTION

Social media is the two way communication in Web 2.0 and it means to communicate, share, and interact with an individual or with a large audience. Social networking websites are the most famous websites on the Internet and millions of people use them every day to engage and connect with other people. Twitter, Facebook, LinkedIn and Google Plus seems to be the most popular Social networking websites on the Internet. Today, for every single piece of content shared on sites like Facebook—every wall post, photo, status update, and video—the up loader must decide which of his friends, group members, and other Facebook users should be able to access the content. As a result, the issue of privacy on sites like Facebook has received significant attention in both the research community [1] and the mainstream media [2]. Our goal is to improve the set of privacy controls and defaults, but we are limited by the fact that there has been no in-depth study of users' privacy settings on sites like Facebook. While significant privacy violations and mismatched user expectations are likely to exist, the extent to which such privacy violations occur has yet to be quantified. Images are now one of the key enablers of users' connectivity. Sharing takes place both among previously established groups of known people or social circles (e.g. Google+, Flickr or Picasa), and also increasingly with people outside the users social circles, for purposes of social discovery to help them identify new peers and learn about peers interests and social surroundings. With the increasing volume of images users share through social sites, maintaining privacy has become a major problem, as demonstrated by a recent wave of publicized incidents where users inadvertently shared personal information. In light of these incidents, the need of tools to help users control access to their shared content is apparent. An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning,keywords or descriptions to the image retrieval can be performed over the annotation words. Manual image annotation is time consuming, laborious and expensive to address this, there has been a large amount of research done on automatic image annotation. Additionally, the increase social web applications and the semantic web have inspired the development of several web-based image annotation tools. Automatic image annotation [6] is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image. This application of computer vision techniques is used in image retrieval systems to organize and locate images of interest from a database. This method can be regarded as a type of multi-image classification with a very large number of classes large as the vocabulary size. Typically, image analysis in the form of extracted feature vectors and training annotation words are used by machine learning techniques to attempt to automatically apply annotations to new images

II. LITERATURE SURVEY

Privacy Suites [1] is proposed by Jonathan Anderson which allows users to easily choose —suites" of privacy settings. Using privacy programming a privacy suite can be created by an expert. Privacy Suites could also be created directly through existingconfiguration UIs or exporting them to the abstract format. To the members of the social sites the privacy suite is distributed through existing distribution channels. Transparency is the main goal, which is essential for convincing influential users that it is safe to use. The disadvantage of a rich programming language is less understandability for end users. To verify a Privacy Suite sufficiently high-level language and good coding practice, motivated users are able.Privacy-Aware Image Classification and Search [2] is a technique to automatically detect private images, and to enable privacy-oriented image search introduced by Sergej Zerr. To provide security policies technique combines textual meta data images with variety of visual features. It uses various classification models trained on a large scale dataset with privacy assignments obtained through a social annotation game. In this the selected image features (edges, faces, color histograms) which can help discriminate between natural and man-made objects/scenes (the EDCV feature) that can indicate the presence or absence of particular objects (SIFT).

A tag based access control of data [3] is developed by Peter F. Klemperer. It is a system that creates access-control policies from photo management tags. Every photo is incorporated with an access grid for mapping the photo with the participant's friends. A suitable preference can be selected by participants and access the information. Based on the user needs photo tags can be categorized as organizational or communicative. There are several important limitations .First, our results are limited by the participants recruited and the photos provided by them. Machine generated access-control rules are the second limitation.Algorithm used here has no access to the context and meaning of tags and no insight into the policy the participant intended when tagging for access control. Hence, some rules appeared strange to the participants who makes them to tag explicitly like —private‖ and —public

A decentralised authentication protocol [4], is a access control system proposed by Ching-man Au Yeung based on a descriptive tags and linked data of social networks in the Semantic websites. Here users can specify access control rules based on open linked data provided by other parties and it allows users to create expressive policies for their photos stored in one or more photo sharing.Adaptive Privacy Policy Prediction (A3P) [5] system is introduced by Anna Cinzia Squicciarini. Personalized policies can be automatically generated by this system. It makes use of the uploaded images by users and a hierarchical image classification is done. Images content and metadata is handled by the A3P system .It consists of two components: A3P Core and A3P Social. The image will be first sent to the A3P-core, when the user uploads the image. The A3P-core classifies the image and determines whether there is a need to invoke the A3P-social. When meta data information is unavailable it is difficult to generate accurate privacy policy. This is the disadvantage of this system. Privacy violation as well as inaccurate classification will be the after effect of manual creation of meta data log information.

III. PROBLEM STATEMENT

Consider social context such as one's friend list. While interesting, they may not be sufficient to address challenges brought by image files for which privacy may vary substantially not just because of social context but also due to the actual image content. As far as images, authors in have presented an expressive language for images uploaded in social sites. This work is complementary to ours as we do not deal with policy expressiveness, but rely on common forms policy specification for our predictive algorithm. In addition, there is a large body of work on image content analysis, for classification and interpretation, retrieval, and photo ranking, also in the context of online photo sharing sites. Of these works, probably the closest to ours. explores privacy-aware image classification using a mixed set of features, both content and meta-data. This is however a binary classification (private versus public), so the classification task is very different than ours. Also, the authors do not deal with the issue of cold-start problem.

IV. EXISTING SYSTEM

Most content sharing websites allow users to enter their privacy preferences. Unfortunately, recent studies have shown that users struggle to set up and maintain such privacy settings. One of the main reasons provided is that given the amount of shared information this process can be tedious and error-prone. Therefore, many have acknowledged the need of policy recommendation systems which can assist users to easily and properly configure privacy settings. However, existing proposals for automating privacy settings appear to be inadequate to address the unique privacy needs of images, due to the amount of information implicitly carried within images, and their relationship with the online environment wherein they are exposed.

V. PROPOSED SCHEME

Our work is related to the concept on privacy setting configuration in social sites, recommendation systems, and privacy analysis of online images. We propose a novel personalized image search framework by simultaneously considering user and query information. The user's preferences over images under certain query are estimated by how probable he/she assigns the query-related tags to the images.

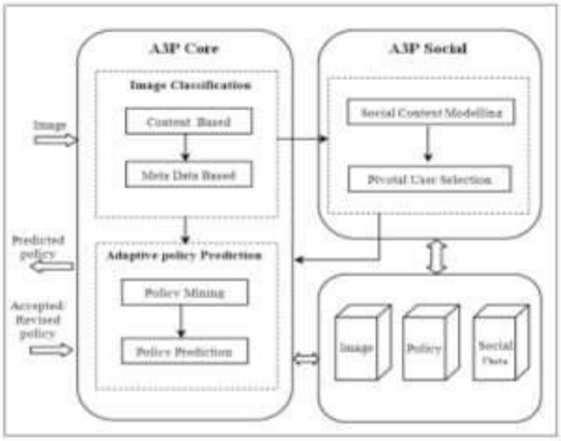

Fig.1. SYSTEM ARCHITECTURE

A ranking based tensor factorization model named RMTF is proposed to predict users' annotations to the images. To better represent the query-tag relationship, we build user-specific topics and map the queries as well as the users' preferences onto the learned topic spaces.

A. User-Specific Topic Modelling

Users may have different intentions for the same query, e.g., searching for "jaguar" by a car fan has a completely different meaning from searching by an animal specialist. One solution to address these problems is personalized search, where user-specific information is considered to distinguish the exact intentions of the user queries and rerank the list results. Given the large and growing importance of search engines, personalized search has the potential to significantly improve searching experience.

B. Personalized Image Search

In the research community of personalized search, evaluation is not an easy task since relevance judgment can only be evaluated by the searchers themselves. The most widely accepted approach is user study, where participants areMaintain both efficiency and high prediction accuracy of a system. asked to judge the search results. Obviously this approach is very costly. In addition, a common problem for user study is that the results are likely to be biased as the participants know that they are being tested. Another extensively used approach is by user query logs or click through history. However, this needs a large-scale real search logs, which is not available for most of the researchers.

C. Ranking – Multi Correlation based

Photo sharing websites differentiate from other social tagging systems by its characteristic of self-tagging: most images are only tagged by their owners. The tagger statistics for Flickr and the webpage tagging system delicious. We can see that in Flickr, 90% images have no more than 4 taggers and the average number of tagger for each image is about 1.9. However, the average tagger for each webpage in delicious is value 6.1. The severe sparsity problem calls for external resources to enable information propagation. In addition to the ternary interrelations, we also collect multiple intra-relations among users, images and tags. We assume that two items with high affinities should be mapped close to each other in the learnt factor subspaces. In the following, we first introduce how to construct the tag affinity graph, and then incorporate them into the tensor factorization framework. To serve the ranking based optimization scheme, we build the tag affinity graph based on the tag semantic relevance and context relevance. The context relevance of tag is simply encoded by their weighted co-occurrence in the image collection

System Overview

The A3P system consists of two main components: A3Pcore and A3P-social. The overall data flow is the following.

When a user uploads an image, the image will be first sent to the A3P-core. The A3P-core classifies the image and determines whether there is a need to invoke the A3P-social. In most cases, the A3P-core predicts policies for the users directly based on their historical behavior. If one of the following two cases is verified true, A3P-core will invoke A3P-social: (i) The user does not have enough data for the type of the uploaded image to conduct policy prediction; (ii) The A3Pcore detects the recent major changes among the user's community about their privacy practices along with user's increase of social networking activities like addition of new friends, new posts on one's profile etc. In above cases, it would be beneficial to report to the user the latest privacy practice of social communities that have similar background as the user. The A3P-social groups users into social communities with similar social context and privacy preferences, and continuously monitors the social groups. When the A3P-social is invoked, it automatically identifies the social group for the user and sends back the information about the group to the A3P-core for policy prediction. At the end, the predicted policy will be displayed to the user. If the user is fully satisfied by the predicted policy, he or she can just accept it. Otherwise, the user can choose to revise the policy.

A3P framework

Users can express their privacy preferences about their content disclosure preferences with their socially connected users via privacy policies. Our policies are inspired by popular content sharing sites i.e. Facebook, Picasa, Flickr, although the actual implementation depends on the specific content management site structure and implementation. In the definition, users in S can be represented by their identities, roles e.g., family, friend, co-workers, or organizations e.g., non-profit organization, profit organization. ID will be the set of images in the user's profile. Each image has a unique ID along with some associated metadata like tags "vacation", "birthday". Images can be further grouped into albums. As for A, we consider four common types of actions: {view, comment, and tag, download}. Last, the condition component C specifies when the granted action is effective. C is a Boolean expression on the grantees' attributes like time, location, and age. For better understanding, an example policy is given by an example. Alice would like to allow her friends and co-workers to comment and tag images in the album named "vacation album" and the image named "summer.jpg" before year 2012. Her privacy preferences can be expressed by the following policy: P: [{friend, coworker}, {vacation album, summer.jpg}, {comment, tag}, (date< 2012)].The policy prediction algorithm provides a predicted policyof a newly uploaded image to the user for his/her reference.More importantly, the predicted policy will reflect the possible changes of a user's privacy concerns. The prediction process consists of three main phases: policy normalization; policy mining; and Policy prediction.

1 Policy Normalization

The policy normalization is a simple decomposition process to convert a user policy into a set of atomic rules in which the data (D) component is a single-element set. An example of policy normalization is shown below. Example 2: Consider policy P in Example 1. Suppose that the album "vacation album" contains k images, namely img1 .jpg, img2 .jpg, imgk.jpg. P is normalized into the following set of atomic rules. 2 Policy mining We propose a hierarchical mining approach for policy mining. Our approach leverages association rule mining techniques to discover popular patterns in policies. Policy mining is carried out within the same category of the new image because images in the same category are more likely under the similar level of privacy protection. The basic idea of the hierarchical mining is to follow a natural order in which a user defines a policy. Given an image, a user usually first decides who can access the image, then thinks about what specific access rights e.g., view only or download should be given, and finally refine the access conditions such as setting the expiration date.

Correspondingly, the hierarchical mining first look for popular subjects defined by the user, then look for popular actions in the policies containing the popular subjects, and finally for popular conditions in the policies containing both popular subjects and conditions.

3 Policy Prediction

We propose a hierarchical mining approach for policy mining. Our approach leverages association rule mining techniques to discover popular patterns in policies. Policy mining is carried out within the same category of the new image because images in the same category are more likely under the similar level of privacy protection. The basic idea of the hierarchical mining is to follow a natural order in which a user defines a policy. The policy mining phase may generate several candidate policies while the goal of our system is to return the most promising one to the user. Thus, we present an approach to choose the best candidate policy that follows the user's privacy tendency. To model the user's privacy tendency, we define a notion of strictness level. The strictness level is a quantitative metric that describes how "strict" a policy is. In particular, a strictness level L is an integer with minimum value in zero, wherein the lower the value, the higher the strictness level. It is generated by two metrics: major level denoted as l and coverage rate (α), where l is determined by the combination of subject and action in a policy, and α is determined by the system using the condition component. All combinations of common subject and common actions are enumerated and assigned an integer value according to the strictness of the corresponding subjects and actions. For example, "view" action is considered more restricted than "tag" action. Given a policy, its l value can be looked up from the table by matching its subject and action. If the policy has multiple subjects or actions and results in multiple l values, we will consider the lowest one. It is worth noting that the table is automatically generated by the system but can be modified by users according to their needs. Then, we introduce the computation of the coverage rate α which is designed to provide fine-grained strictness level. A is a value ranging from 0 to 1 and it will just adjust but not dominate the previously obtained major level. In particular, we define α as the percentage of people in the specified subject category who satisfy the condition in the policy. For example, a user has 5 family members documented in the system and two of them are kids. When he specifies a policy with the condition age > 18, only three family members will satisfy this condition. The corresponding α is then 3/5=0.6. The larger the value of α, the more people are allowed to access the image and hence the policy is less restricted. Therefore, we subtract (1-α) from 1 to obtain the final strictness level. Policies, we now need to determine which strictness level fits best to the user's privacy trend. For this purpose, we propose the following approach.

VI CONCLUSION

We have proposed an Adaptive Privacy Policy Prediction (A3P) system that helps users automate the privacy policy settings for their uploaded images. The A3P system provides a comprehensive framework to infer privacy preferences based on the information available for a given user. We also effectively tackled the issue of cold-start, leveraging social context information. Our experimental study proves that our A3P is a practical tool that offers significant improvements over current approaches to privacy.

REFERENCES

[1] Anna Cinzia Squicciarini, Member, IEEE, Dan Lin, Smitha Sundareswaran, and Joshua Wede, "Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites", IEEE Transactions on Knowledge and Data Engineering, Vol. 27, NO. 1, January 2015.

[2] J. Bonneau, J. Anderson, and L. Church, "Privacy suites: Shared privacy for social networks," in Proc. Symp. Usable Privacy Security, 2009.

[3] A. Kapadia, F. Adu-Oppong, C. K. Gardiner, and P. P. Tsang, "Social circles: Tackling privacy in social networks," in Proc. Symp. Sable Privacy Security, 2008.

[4] Kambiz Ghazinour, Stan Matwin and Marina Sokolova, "Your privacy protector: A Recommender System For Privacy Settings In Social Networks", International Journal of Security, Privacy and Trust Management (IJSPTM) Vol 2, No 4, August 2013.

[5] Alessandra Mazzia Kristen LeFevre and Eytan Adar, The PViz Comprehension Tool for Social Network Privacy Settings, Tech. rep., University of Michigan, 2011.

[6] Peter F. Klemperer, Yuan Liang, Michelle L. Mazurek, “Tag, You Can See It! Using Tags for Access Control in Photo Sharing”, Conference on Human Factors in Computing Systems, May 2012.

[7] C. A. Yeung, L. Kagal, N. Gibbins, and N. Shadbolt, “Providing access control to online photo albums based on tags and linked data,” in Proc. Soc. Semantic Web: Where Web 2.0 Meets Web 3.0 at the AAAI Symp., 2009, pp. 9–14.

[8] Sergej Zerr, Stefan Siersdorfer, Jonathon Hare, Elena Demidova , I Know What You Did Last Summer!:Privacy-Aware Image Classification and Search , Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval, 2012.

[9] Anna Cinzia Squicciarini, “Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites”, IEEE Transactions On Knowledge And Data Engineering, vol. 27, no. 1, January 2015.