Privacy Policy Inference of User-Uploaded Images On Content Sharing Sites

Mr. U Prasad#1,Ms. Kata Priyanka #2Ms. Sk Shabna #3, Ms. Y Susmitha #4, Ms. D Manasa #5

#1 Assistant professor, Dept Of CSE, Qis Institute of Technology, Ongole, Prakasam (Dt)
#2 Student, Dept Of IT, Qis Institute of Technology, Ongole, Prakasam (Dt)
#3 Student, Dept Of IT, Qis Institute of Technology, Ongole, Prakasam (Dt)
#4 Student, Dept Of IT, Qis Institute of Technology, Ongole, Prakasam (Dt)
#5 Student, Dept Of IT, Qis Institute of Technology, Ongole, Prakasam (Dt)

ABSTRACT

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. Toward addressing this need, we propose an Adaptive Privacy Policy Prediction (A3P) system to help users compose privacy settings for their images. We examine the role of social context, image content, and metadata as possible indicators of users' privacy preferences. We propose a two-level framework which according to the user's available history on the site, determines the best available privacy policy for the user's images being uploaded. Our solution relies on an image classification framework for image categories which may be associated with similar policies, and on a policy prediction algorithm to automatically generate a policy for each newly uploaded image, also according to users' social features. Over time, the generated policies will follow the evolution of users' privacy attitude. We provide the results of our extensive evaluation over 5,000 policies, which demonstrate the effectiveness of our system, with prediction accuracies over 90 percent.

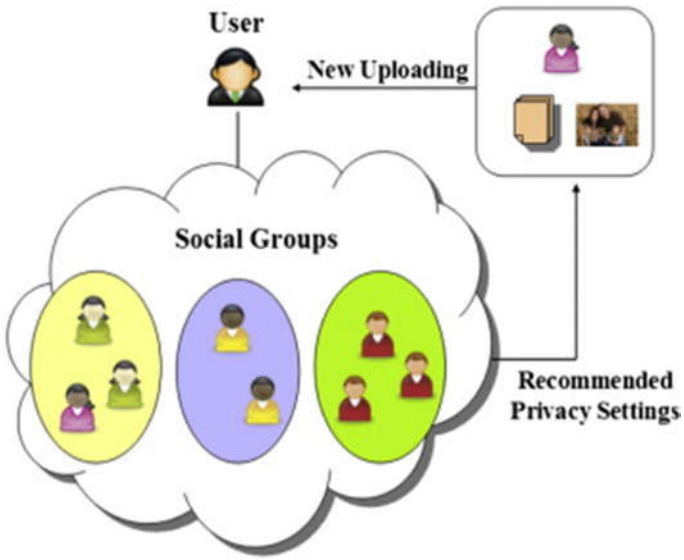
Keywords Online information services, web-based services.

I. INTRODUCTION

Now a days people are sharing their images on content sharing sites with people within their social circles (e. g., Google+, Flickr, Facebook, Picasa etc.,), and also 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. However, semantically rich images may reveal content-sensitive information. Consider a photo of a student's 2012 graduation ceremony, for example. It could be shared within a Google+ circle or Flickr group, but may unnecessarily expose the students'

family members to other friends. Sharing images within online content sharing sites lead to unwanted disclosure and privacy violations. Also, the persistent nature of online media makes it possible for other users to collect rich aggregated information about the owner of the published content and the subjects in the published content. The aggregated information can result in unexpected exposure of one's social environment and lead to abuse or misuse of one's personal information. The content sharing websites allow users to enter their privacy preferences. But the 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.

Social media is very powerful tool to communicate with each other ,user can communicate with social site to exchange idea ,emotion ,information, happiness. Now every user are connect to each other , there are very high volume which are connect with each other using different sites. Social media is the two way communication in Web 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. Number of web site like Facebook, Twitter, etc. users are used to communicate, connected with each other, user can upload, post, tweet, download images video and performing number of action. The aggregated information can result in unexpected exposure of one's social environment and lead to abuse of one's personal information. 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. The Most content sharing websites allows a user 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 reason provided is that the amount of shared information this process can be tedious and error-prone. 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 Face book.



The impact of social environment and personal characteristics. Social context of users, such as their profile information and relationships with others may provide useful information regarding users' privacy preferences. For example, users interested in photography may like to share their photos with other photographers. Users who have

several family members among their social contacts may share with them pictures related to family events. The role of image's content and metadata. In general, similar images often incur similar privacy preferences, especially when people appear in the images. For example, one may upload several photos of his kids and specify that only his family members are allowed to see these photos. He may upload some other photos of landscapes which he took as a hobby and for these photos, he may set privacy preference allowing anyone to view and comment the photos. Analyzing the visual content may not be sufficient to capture users' privacy preferences. Tags and other metadata are indicative of the social context of the image, including where it was taken and why , and also provide a synthetic description of images, complementing the information obtained from visual content analysis.

II. LITERATURE REVIEW

Jonathan Anderson proposed a paradigm called Privacy Suites [4] which allows users to easily choose ―suites" of privacy settings. A privacy suite can be created by an expert using privacy programming. Privacy Suites could also be created directly through existing configuration UIs or exporting them to the abstract format. The privacy suite is distributed through existing distribution channels to the members of the social sites. The disadvantage of a rich programming language is less understandability for end users.

Fabeah Adu-Oppong developed privacy settings based on the concept of social circles [5]. It provides a web based solution to protect personal information. The technique named Social Circles Finder, automatically generates the friend's list. It is a technique that analyses the social circle of a person and identifies the intensity of relationship and therefore social circles provide a meaningful categorization of friends for setting privacy policies. The application will identify the social circles of the subject but not show them to the subject. The subject will then be asked questions about their willingness to share a piece of their personal information. Based on the answers the application finds the visual graph of users.

Kambiz Ghazinour designed a recommender system known as YourPrivacyProtector [5] that understands the social net behavior of their privacy settings and recommending reasonable privacy options. It uses user's personal profile, User's interests and User's privacy settings on photo albums as parameters and with the help of these parameters the system constructs the personal profile of the user. It automatically learned for a given profile of users and assign the privacy options. It allows users to see their current privacy settings on their social network profile, namely Facebook, and monitors and detects the possible privacy risks.

Alessandra Mazzia introduced PViz Comprehension Tool [1], an interface and system that corresponds more directly with how users model groups and privacy policies applied to their networks. PViz allows the user to understand the visibility of her profile according to automatically-constructed, natural sub-groupings of

friends, and at different levels of granularity. Because the user must be able to identify and distinguish automatically-constructed groups, we also address the important sub-problem of producing effective group labels.

Peter F. Klemperer developed a tag based access control of data [8] shared in the social media sites. 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. The participants can select a suitable preference and access the information. Photo tags can be categorized as organizational or communicative based on the user needs. There are several important limitations to our study design. First, our results are limited by the participants we recruited and the photos they provided. A second set of limitations concerns our use of machine generated access-control rules. The algorithm has no access to the context and meaning of tags and no insight into the policy the participant intended when tagging for access control. As a result, some rules appeared strange or arbitrary to the participants, potentially driving them toward explicit policybased tags like ―private‖ and ―public.

Ching-man Au Yeung propose a access control system based on a decentralised authentication protocol [11], descriptive tags and linked data of social networks in the Semantic Web. It allows users to create expressive policies for their photos stored in one or more photo sharing sites, and users can specify access control rules based on open linked data provided by other parties.

Sergej Zerr propose a technique Privacy-Aware Image Classification and Search [10] to automatically detect private images, and to enable privacy-oriented image search. It combines textual meta data images with variety of visual features to provide security policies. 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). It uses various classification models trained on a large scale dataset with privacy assignments obtained through a social annotation game.

Anna Cinzia Squicciarini developed an Adaptive Privacy Policy Prediction (A3P) [3] system, a free privacy settings system by automatically generating personalized policies. The A3P system handles user uploaded images based on the person's personal characteristics and images content and metadata. The A3P system consists of two components: A3P Core and A3P Social. 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. The disadvantage is inaccurate privacy policy generation in case of the absence of metadata information about the images.

III. 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.

DISADVANTAGES OF EXISTING SYSTEM

Sharing images within online content sharing sites, therefore, may quickly lead to unwanted disclosure and privacy violations. Further, the persistent nature of online media makes it possible for other users to collect rich aggregated information about the owner of the published content and the subjects in the published content. The aggregated information can result in unexpected exposure of one's social environment and lead to abuse of one's personal information.

IV. PROPOSED SYSTEM

In this paper, we propose an Adaptive Privacy Policy Prediction (A3P) system which aims to provide users a hassle free privacy settings experience by automatically generating personalized policies.

The A3P system handles user uploaded images, and factors in the following criteria that influence one's privacy settings of images:

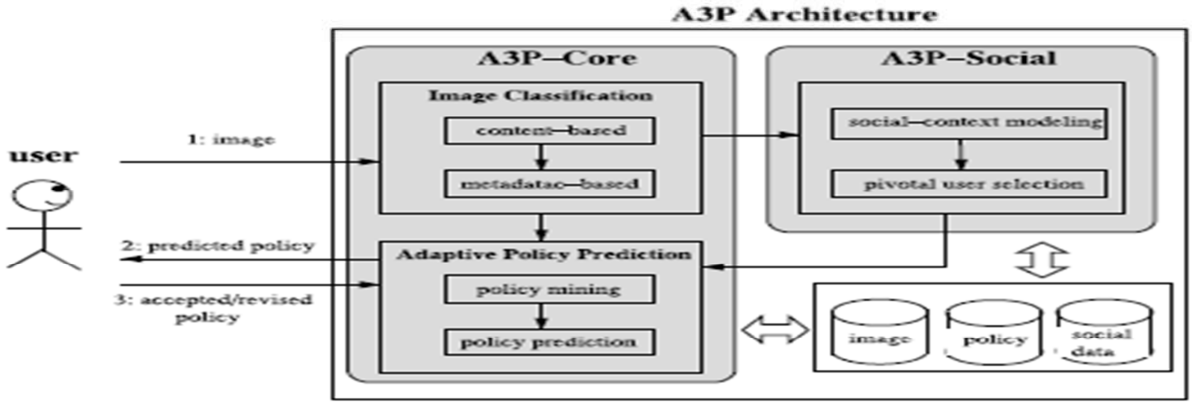
The impact of social environment and personal characteristics. Social context of users, such as their profile information and relationships with others may provide useful information regarding users' privacy preferences.

For example users interested in photography may like to share their photos with other amateur photographers.

The impact of social environment and personal characteristics. Social context of users, such as their profile information and relationships with others may provide useful information regarding users' privacy preferences. For example, users interested in photography may like to share their photos with other amateur photographers. The role of image's content and metadata. In general, similar images often incur similar privacy preferences, especially when people appear in the images. For example, one may upload several photos of his kids and specify that only his family members are allowed to see these photos.

ADVANTAGES OF PROPOSED SYSTEM

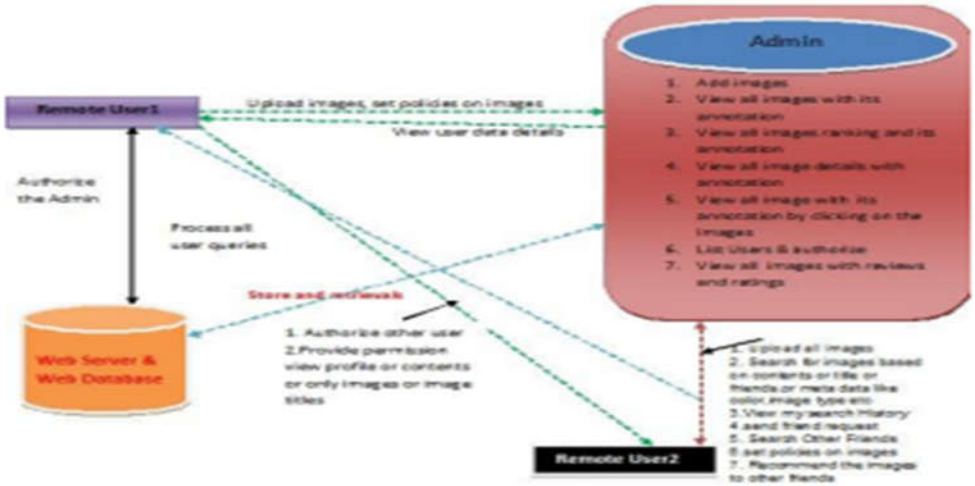
The A3P-core focuses on analyzing each individual user's own images and metadata, while the A3P-Social offers a community perspective of privacy setting recommendations for a user's potential privacy improvement. We design the



interaction flows between the two building blocks to balance the benefits from meeting personal characteristics and obtaining community advice.

V. IMPLEMENTATION

System Construction



The A3P system consists of two main components: A3P-core 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 A3Psocial: (i) The user does not have enough data for the type of the uploaded image to conduct policy prediction; (ii) The A3P-core detects the recent major changes among the user's community about their privacy practices along with user's increase of social networking activities (addition of new friends, new posts on one's profile etc).

Content-Based Classification

To obtain groups of images that may be associated with similar privacy preferences, we propose a hierarchical image classification which classifies images first based on their contents and then refine each category into subcategories based on their metadata. Images that do not have metadata will be grouped only by content. Such a hierarchical classification gives a higher priority to image content and minimizes the influence of missing tags. Note that it is possible that some images are included in multiple categories as long as they contain the typical content features or metadata of those categories. Our approach to content-based classification is based on an efficient and yet accurate image similarity approach. Specifically, our classification algorithm compares image signatures defined based on quantified and sanitized version of Haar wavelet transformation. For each image, the wavelet transform encodes frequency and spatial information related to image color, size, invariant transform, shape, texture, symmetry, etc. Then, a small number of coefficients are selected to form the

signature of the image. The content similarity among images is then determined by the distance among their image signatures.

Metadata-Based Classification

The metadata-based classification groups images into subcategories under aforementioned baseline categories. The process consists of three main steps. The first step is to extract keywords from the metadata associated with an image. The metadata considered in our work are tags, captions, and comments. The second step is to derive a representative hypernym (denoted as h) from each metadata vector. The third step is to find a subcategory that an image belongs to. This is an incremental procedure. At the beginning, the first image forms a subcategory as itself and the representative hypernyms of the image becomes the subcategory's representative hypernyms.

Adaptive Policy Prediction

The policy prediction algorithm provides a predicted policy of 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: (i) policy normalization; (ii) policy mining; and (iii) policy prediction.

IMPLEMENTATION DETAILS

This study involved 88 participants (48 female and 40 males) who were recruited from a large US university community (staff, students, and the community at large). Their average age is 26.3 years old (Range: 18-39). The participants completed at least 90 percent of the ques -tionnaire consisting of two parts. The first part contains questions related to one's background information and online privacy practices and the second part is to collect user-specified policies. In the first part of the questionnaire, the participants were asked to indicate any social networks they were a part of (98 percent indicated Facebook and 37 percent also indicated others like Myspace). In terms of usage frequency, 95 percent of the respondents accessed social network sites at least once a week, with 76 percent of reporting that they were daily users. We also asked participants if they have had concerns about their privacy due to shared images. Over 51 percent of the participants indicated that they had privacy concerns. Users also reported that image content is an important factor when determining privacy settings for an image with 87 per-cent of people agreeing or strongly agreeing with the statement ―When I set privacy settings for a certain image I usually think about the content of the image‖, and over 91 percent of users agreeing or strongly agreeing with the statement ―The content of an image determines whether I upload the image to a social network site.‖ Surprisingly, however, many users indicated that they never changed privacy settings for images (38 percent) or changed their set-tings only 1 or 2 times (36 percent) since joining the social network. There seems to be a clear disconnect between users privacy inclinations and their practice of setting privacy policies. The possible reason could

be ―Changing privacy set-tings for every image uploaded on a social site can be very time consuming‖, as strong agreed or agreed by 70 percent of users. In the second part of the questionnaire, we presented each user 30 images selected by us. For each image, we asked the user to input the privacy settings by assuming these photos as his/her own images. We collected around 3,000 policies. We invited another 41 people to use our A3P system. The goal of this experiment is to assess our system's acceptability, i.e., whether users would consider the predicted policies reasonable, and inline with their over-all preferences. We asked participants to input policies for a few images at first for training purposes. Three images from a given class are sufficient to bootstrap the algorithm. Next, participants enter privacy settings for a set of images that they would upload in their fictitious profile. Upon showing the image, privacy settings for it are suggested to the user. The participant has the option to accept the predicted policy as is, revise some components of it, or disagree with the predicted result and e re-enter preferred settings.

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.

REFERENCES

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

[2] 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.

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

[4] Kambiz Ghazinour, Stan Matwin and Marina Sokolova, ―Yourprivacyprotector: 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] 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.

[6] Karthikram, Mailaivasan, Parthiban and Ganesh, ―Tag Based Image Retrieval (TBIR) Using Automatic Image Annotation‖, IJRET: International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321-7308.

[7] 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.

[8] Sangeetha. J et al, An Improved Privacy Policy Inference over the Socially Shared Images with Automated Annotation Process,

(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 3166-3169.

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

[10] 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.

[11] L. Church, J. Anderson, J. Bonneau, and F. Stajano, "Privacy stories: Confidence on privacy behaviors through end user programming," in Proc. 5th Symp. Usable Privacy Security, 2009.

[12] R. da Silva Torres and A. Falc~ao, "Content-based image retrieval: Theory and applications," Revista de Inform_atica Te_orica e Aplicada, vol. 2, no. 13, pp. 161–185, 2006.

[13] R. Datta, D. Joshi, J. Li, and J. Wang, "Image retrieval: Ideas, influences, and trends of the new age," ACM Comput. Surv., vol. 40, no. 2, p. 5, 2008.

[14] J. Deng, A. C. Berg, K. Li, and L. Fei-Fei, "What does classifying more than 10,000 image categories tell us?" in Proc. 11th Eur. Conf. Comput. Vis.: Part V, 2010, pp.71–84.[Online].

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

[16] L. Geng and H. J. Hamilton, "Interestingness measures for data mining: A survey," ACM Comput. Surv., vol. 38, no. 3, p. 9, 2006.

[17] Image-net data set. [Online]. Available: www.image-net.org, Dec. 2013.

[18] S. Jones and E. O'Neill, "Contextual dynamics of group-based sharing decisions," in Proc. Conf. Human Factors Comput. Syst., 2011, pp. 1777–1786. [Online]. Available:http://doi.acm.org/10.1145/1978942.1979200

[19] P. Klemperer, Y. Liang, M. Mazurek, M. Sleeper, B. Ur, L. Bauer, L. F. Cranor, N. Gupta, and M. Reiter, "Tag, you can see it!: Using tags for access control in photo sharing," in Proc. ACM Annu. 99615.94

Authors Profile

Mr U Prasad, M.Tech., working as an Assistant Professor in the Department of Computer Science & Engineering, Qis Institute of Technology, Ongole, Andhra Pradesh.

Ms Kata Priyanka pursuing B Tech in Information Technology from Qis Institute of Technology, Ongole, Affiliated to Jawaharlal Nehru Technological University, Kakinada in 2016-20 respectively.

Ms Sk Shabna pursuing B Tech in Information Technology from Qis Institute of Technology, Ongole, Affiliated to Jawaharlal Nehru Technological University, Kakinada in 2016-20 respectively.

Ms Y Susmitha pursuing B Tech in Information Technology from Qis Institute of Technology, Ongole, Affiliated to Jawaharlal Nehru Technological University, Kakinada in 2016-20 respectively.

Ms D Manasa pursuing B Tech in Information Technology from Qis Institute of Technology Ongole, Affiliated to Jawaharlal Nehru Technological University, Kakinada in 2016-20 respectively.