International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 12 Ver II || Dec 2018 || PP 09-13

Enhanced Privacy Policy Prediction for User-Uploaded Images on Content Sharing Sites

Rupali Ashok Narkhede, Madhuri R. Zawar.
[1]*(Computer Engineering, G's COE, Jalgaon , North Maharashtra University, India)*
[2]*(Computer Engineering, G's COE, Jalgaon , North MaharashtraUniversity, India)*
Corresponding Author:Rupali Ashok Narkhede

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:*Adaptive privacy policy prediction, content sharing sites, metadata,online information services,web based services.*

Date of Submission: 28-11-2018 Date of acceptance:14-12-2018

I. INTRODUCTION

The Internet and online social networks, in particular, are a part of most people's lives. eMarketer.com reports that in 2011, nearly 150 million US Internet users will interface with at least one social networking site per month. eMarketer.com also reports that in 2011, 90 percent of Internet users ages 18-24 and 82 percent of Internet users ages 25-34 will interact with at least one social networking site per month. This trend is increasing for all age groups. As the young population ages, they will continue to leverage social media in their daily lives. In addition, new generations will come to adopt the Internet and online social networks. These technologies have become and will continue to be a vital component of our social fabric, which we depend on to communicate, interact, and socialize.

Not only are there a tremendous amount of users online, there is also a tremendous amount of user profile data and content online. For example, on Facebook, there are over 30 billion pieces of content shared each month. New content is being added every day; an average Facebook user generates over 90 pieces of content each month. This large amount of content coupled with the significant number of users online makes maintaining appropriate levels of privacy very challenging.

In addition, it measures the human effects of our improvements. It introduces three new improvements to privacy management models:

1. Assisted Friend Grouping—an incremental improvement to traditional group-based policy management.
2. Same-As Policy Management—a new paradigm improvement over traditional group-based policy management.

3. Example Friend Selection—an incremental improvement to Same-As Policy Management.

The report leverages traditional group-based policy management as our baseline and progressively improve upon this privacy management model. With each new enhancement, we measure their human effects including cluster/user defined relationship group alignment, user privacy sentiment, efficiencies and user perceptions. The report introduces a user-assisted friend grouping mechanism that enhances traditional group-based policy management approaches. Assisted Friend Grouping leverages proven clustering techniques to aid users in grouping their friends more effectively and efficiently. It introduces a new privacy management model that is an improvement over traditional group-based policy management approaches. The new paradigm leverages a user's memory and opinion of their friends to set policies for other similar friends, which we refer to as Same-As Policy Management. Users associate the policy with an example friend and in doing so have this friend in the forefront of their mind. This allows users to be more selective and careful in assigning permissions. Users are thinking of people, not groups. Using a visual policy editor that takes advantage of friend recognition and minimal task interruptions, Same-As Policy Management demonstrated improved performance and user perceptions over traditional group-based policy management approaches.

II. RELATED WORK

Several recent works have studied how to automate the task of privacy settings (e.g., [7], [15], [20], [22], [27], [28]). Bonneau et al. [7] proposed the concept of privacy suites which recommend to users a suite of privacy settings that "expert" users or other trusted friends have already set, so that normal users can either directly choose a setting or only need to do minor modification. Similarly, Danezis [8] proposed a machine-learning based approach to automatically extract privacy settings from the social context within which the data is produced. Parallel to the work of Danezis, Adu-Oppong et al. [15] develop privacy settings based on a concept of "Social Circles" which consist of clusters of friends formed by partitioning users' friend lists. Ravichandran et al. [30] studied how to predict a user's privacy preferences for location-based data (i.e., share her location or not) based on location and time of day. Fang et al. [28] proposed a privacy wizard to help users grant privileges to their friends. The wizard asks users to first assign privacy labels to selected friends, and then uses this as input to construct a classifier which classifies friends based on their profiles and automatically assign privacy labels to the unlabeled friends. More recently, Klemperer et al. [20] studied whether the keywords and captions with which users tag their photos can be used to help users more intuitively create and maintain access-control policies. Their findings are inline with our approach: tags created for organizational purposes can be repurposed to help create reasonably accurate access-control rules.

The aforementioned approaches focus on deriving policy settings for only traits, so they mainly 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 [41] 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 (e.g., [14], [37], [46]), retrieval ([12], [13] are some examples), and photo ranking [35], [40], also in the context of online photo sharing sites, such as Flickr [10], [29], [36]. Of these works, Zerr's work [43] is probably the closest to ours. Zerr 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.

Ales Sandro Acquisti Ralph Gross H. John Heinz[1] Online social networks such as Friendster, MySpace, or the Facebook have experienced exponential growth in membership in recent years. These networks offer attractive means for interaction and communication, but also raise privacy and security concerns. In this study we survey a representative sample of the members of the Facebook (a social network for colleges and high schools) at a US academic institution, and compare the survey data to information retrieved from the network itself. We look for underlyingdemographic or behavioral differences between the communities of the network's members and nonmembers; we analyze the impact of privacy concerns on members' behavior; we compare members' stated attitudes with actual behavior; and we document the changes in behavior subsequent to privacy-related information exposure. We find that an individual's privacy concerns are only a weak predictor of his membership to the network. Also privacy concerned individuals join the network and reveal great amounts of personal information. Some manage their privacy concerns by trusting their ability to control the information they provide and the external access to it. However, we also find evidence of members' misconceptions about the online community's actual size and composition, and about the visibility of members' profiles. Hong-Ming Chen, Ming-Hsiu Chang [9] Online photo albums have been prevalent in recent years and have resulted in more and more applications developed to provide convenient functionalities for photo sharing. In this project, we propose a system named *SheepDog* to automatically add photos into appropriate groups and recommend suitable tags for users on Flickr.

We adopt concept detection to predict relevant concepts of a photo and probe into the issue abouttraining data collection for concept classification. From the perspective of gathering training data by web searching, we introduce two mechanisms and investigate their performances of concept detection. Based on some existing information from Flickr, a ranking-based method is applied not only to obtain reliable training data, but also to provide reasonable group/tag recommendations for input photos. We evaluate this system with a rich set of photos and the results demonstrate the effectiveness of our work.

Munmun De Choudhury, Hari Sundaram[10] We develop a recommendation framework to connect image content with communities in online social media. The problem is important because users are looking for useful feedback on their uploaded content, but finding the right community for feedback is challenging for the end user. Social media are characterized by both content and community. Hence, in our approach, we characterize images through three types offeatures: visual features, user generated text tags, and social interaction (user communication history in the form of comments). A recommendation framework based on learning a latent space representation of the groups is developed to recommend the most likely groups for a given image. The model was tested on a large corpus of Flickrimagescomprising15,689 images. Our method outperforms the baseline method, with a mean precision 0.62 and mean recall 0.69. Importantly, we show that fusing image content, text tags with social interaction features outperforms the case of only using image content or tags.

Kristina Lerman, Anon Plangprasopchok,Chio Wong[21]The social media site Flickr allows users to upload their photos, annotate them with tags, submit them to groups, and also to form social networks by adding other users as contacts. Flickr offers multiple ways of browsing or searching it. One option is tag search, which returns all images tagged with a specific keyword. If the keyword is ambiguous, e.g., ―beetle‖ could mean an in sector a car, tag search results will include many images that are not relevant to the sense the user had in mind when executing the query. We claim that users express their photography interests through the metadata they add in the form of contacts and image annotations. We show how to exploit this metadata to personalize search results for the user, thereby improving search performance. First, we show that we can significantly improve search precision by filtering tag search results by user’s contacts or a larger social network that includes those contact’s contacts. Secondly, we describe a probabilistic model that takes advantage of tag information to discover latent topics contained in the result.

Amit Singhal[31] For thousands of years people have realized the importance of archiving and finding information. With the advent of computers, it became possible to store large amounts of information; and finding useful information from such collections became a necessity. The field of Information Retrieval (IR) was born in the 1950s out of this necessity. Over the last forty years, the field has matured considerably. Several IR systems are used on an everyday basis by a wide variety of users. This article is a brief overview of the key advances in the field of Information Retrieval, and a description of where the state-of-the-art is at in the field. The practice of archiving written information can be traced back to around 3000 BC, when the Sumerians designated special areas to store clay tablets with cuneiform inscriptions. Even then the Sumerians realized that proper organization and access to the archives was critical for efficient use of information. They developed special classifications to identify every tablet and its content.

Kenan Xu, Hossam Hassanein[33]Online social networking communities such as Facebook and My Space are extremely popular. These sites have changed how many people develop and maintain relationships through posting and sharing personal information. The amount and depth of these personal disclosures have raised concerns regarding online privacy. We expand upon previous research on users’ under-utilization of available privacy options by examining users’ current strategies for maintaining their privacy, and where those strategies fail, on the online social network site Facebook. Our results demonstrate the need for mechanisms that provide awareness of the privacy impact of users’ daily interactions.

Xiwang Yang[34] Recommendation plays an increasingly important role in our daily live. Recommender systems automatically suggest to a user items that might be of interest to her. Recent studies demonstrate that information from social networks can be exploited to improve accuracy of recommendations. In this project, we present a survey of Collaborative Filtering(CF) based social recommender systems. We provide a brief overview over the task of recommender systems and traditional approaches that do not use social network information.

1.Problem Definition:

In this work, we present an overhauled version of A3P, which includes an extended policy prediction algorithm in A3P-core (that is now parameterized based on user groups and also factors in possible outliers), and a new A3P-social module that develops the notion of social context to refine and extend the prediction power of our system. We also conduct additional experiments with a new data set collecting over 1,400 images and corresponding policies, and we extend our analysis of the empirical results to unveil more insights of our system’s performance.

2.Motivation:

We then present how social network information can be adopted by recommender systems as additional input for improved accuracy. We classify CF–based social recommender systems into two categories: matrix factorization based social recommendation approaches and neighborhood based social recommendation approaches. For each category, we survey and compare several representative algorithms. Most of the surveyed algorithms are trained and tested offline. One of the next steps will be to test and improve their performance in real online social networks, with real-time user experience feedback. Finally, privacy in online social networks has attracted more and more user awareness. Privacy-preserving social recommender systems are another interesting direction for future work.

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.Sharing images within online content sharing sites,therefore,may quickly leadto 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 project, 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 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.

System Architecture

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.

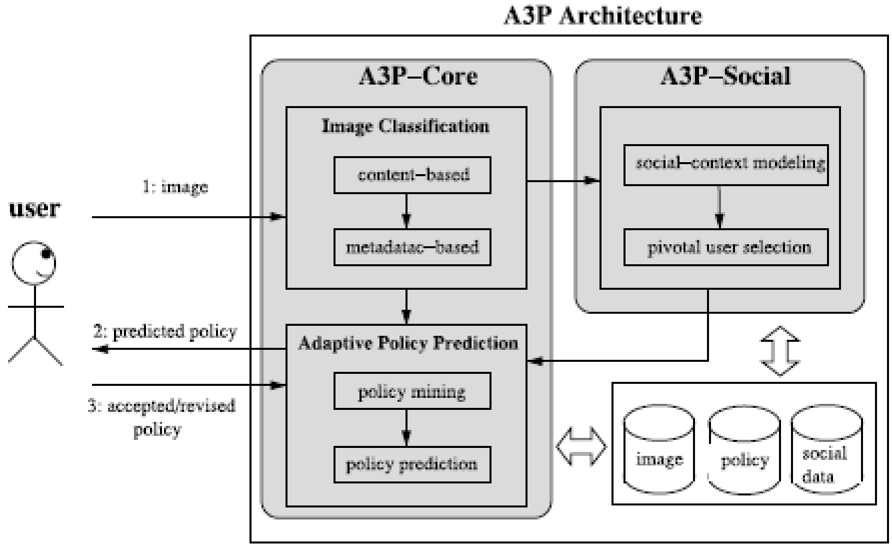

Fig:A3P Architecture

The A3P system consists of two main components: A3P-coreand A3P-social. The overall data flow is the following.When a user uploads an image, the image will be first sentto the A3P-core. The A3P-core classifies the image anddetermines whether there is a need to called the A3P-social.In most cases, the A3P-core predicts policies for the usersdirectly based on their historical behavior. If one of the followingtwo cases is verified true, A3P-core will invoke A3Psocial:(i) The user does not have enough data for the typeof the uploaded image to conduct policy prediction; (ii) TheA3P-core detects the recent major changes among the user'scommunity about their privacy practices along with user'sincrease of social networking activities (addition of newfriends, new posts on one's profile etc). In above cases, itwould be beneficial to report to the user the latest privacypractice of social communities that have similar backgroundas the user. The A3P-social groups users into social communitieswith similar social context and privacy preferences,and continuously monitors the social groups. When theA3P-social is invoked, it automatically identifies the socialgroup for the user and sends back the information about thegroup to the A3P-core for policy prediction. At the end, thepredicted policy will be displayed to the user. If the user isfully satisfied by the predicted policy, he or she can justaccept it. Otherwise, the user can choose to revise the policy.The actual policy will be stored in the policy repository ofthe system for the policy prediction of future uploads.

V. 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.

REFERENCES

Journal Papers:

[1]. A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, and privacy on the facebook," in Proc. 6th Int. Conf. Privacy Enhancing Technol. Workshop, 2006, pp. 36–58.

[2]. R. Agrawal and R. Srikant,"Fast algorithms for mining association rules in large databases," in Proc. 20th Int. Conf. Very Large Data Bases, 1994, pp. 487–499.

[3]. S. Ahern, D. Eckles, N. S. Good, S. King, M. Naaman, and R. Nair, "Over-exposed?: Privacy patterns and considerations in online and mobile photo sharing," in Proc. Conf. Human Factors Comput. Syst., 2007, pp. 357–366.

[4]. M. Ames and M. Naaman, "Why we tag: Motivations for annotation in mobile and online media," in Proc. Conf. Human Factors Comput. Syst., 2007, pp. 971–980.

[5]. A. Besmer and H. Lipford, "Tagged photos: Concerns, perceptions, and protections," in Proc. 27th Int. Conf. Extended Abstracts Human Factors Comput. Syst., 2009, pp. 4585–4590.

[6]. D. G. Altman and J. M. Bland ,"Multiple significance tests: The bonferroni method," Brit. Med. J., vol. 310, no. 6973, 1995.

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

[8]. J. Bonneau, J. Anderson, and G. Danezis, "Prying data out of a social network," in Proc. Int. Conf. Adv. Soc. Netw. Anal. Mining, 2009, pp.249–254.

[9]. H.-M. Chen, M.-H. Chang, P.-C. Chang, M.-C. Tien, W. H. Hsu, and J.-L. Wu, "Sheepdog: Group and tag recommendation for flickr photos by automatic search-based learning," in Proc. 16th ACM Int. Conf. Multimedia, 2008, pp. 737–740.

[10]. M. D. Choudhury, H. Sundaram, Y.-R. Lin, A. John, and D. D. Seligmann, "Connecting content to community in social media via image content, user tags and user communication," in Proc. IEEE Int. Conf. Multimedia Expo, 2009, pp.1238–1241.

[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.

[12]. 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.

[13]. 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]. Available: http://portal.acm.org/citation.cfm?id=1888150.1888157

Theses:

[14]. 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.

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

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

[17]. 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

[18]. A. Kaw and E. Kalu, Numerical Methods with Applications: Abridged., Raleigh, North Carolina, USA: Lulu.com, 2010.

[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. Conf. Human Factors Comput. Syst., 2012, pp. 377–386.

Rupali Ashok Narkhede"Enhanced Privacy Policy Prediction for User-Uploaded Images on Content Sharing Sites"International Journal of Engineering Science Invention (IJESI), vol. 07, no. 12, 2018, pp 09-13