International Journal of Research Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 08
July 2017

An Adaptive Privacy Policy Prediction for User Uploading Data

G.Shilpa ; R. Sunil Kumar ; J.Bhargavi

[1]assistant Professor in Department of CSE Sri Indu College Of Engineering and Technology. Telangana
[2]PG Scholar in Department of CSE Sri Indu College Of Engineering and Technology. Telangana
[3]PG Scholar in Department of CSE Sri Indu College Of Engineering and Technology. Telangana

Abstract— *An Adaptive Privacy Policy Prediction (A3P) system to support users to comprise privacy settings for their images. With the accumulative volume of images, user's stake through social sites, sustaining privacy has become a major problem, as proven by a recent trend of publicized happenings where users unintentionally shared personal information. In such a case of incidents, the need of tools to help users control access to their shared content is superficial. Towards addressing this need, it is examined the role of social context, image content, and metadata as probable indicators of users' privacy partialities. A two-level framework which is rendering to the user's available history on the site, defines the best available privacy policy for the user's images being uploaded. The solution depend on an image classification framework for image categories which may be accompanied 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 created policies will follow the evolution of users' privacy attitude. It also provides the results of extensive evaluation which determine the efficacy of the system, with prediction accuracies.*

Index Terms— *Online information services; isb-based services*

Introduction

Images are now one of the key enablers of users' connectivity. Sharing takes place both among formerly Recognized groups of known people or social circles (e.g., Google+, Flickr or Picasa), and also progressively with people outside the users social circles, for assurances of social discovery-to help them identify new nobles and learn about peers interests and social surroundings. Hoitisver, semantically rich images may reveal content sensitive information [2]. Sharing images within online content sharing sites,therefore,may quickly lead to unwanted disclosure and privacy violations [3], [2]. 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 accumulated information can result in unexpected exposure of one's social environment and lead to abuse of one's personal information. Most content sharing it is besides allow users to enter their privacy preferences. Unfortunately, recent studies have shown that users struggle to set up and maintain such privacy settings [13]. The reason is that given the amount of shared information in this process can be monotonous and error-prone. Consequently, many have approved the need of policy commendation systems which can assist users to easily and properly configure privacy settings [7].

In this strategy, it is proposed an Adaptive Privacy Policy Prediction(A3P) system which aims to provide users a annoyance 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.

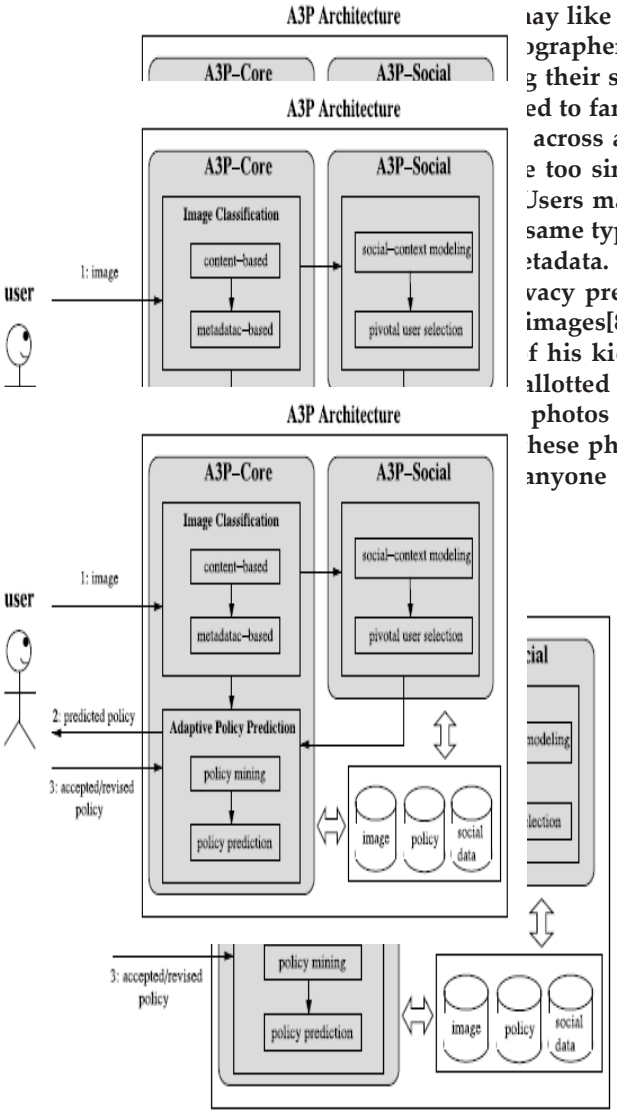

Fig. 1 System overview.

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 [4], and also provide a synthetic description of images, complementing the information obtained from visual content analysis

2 BACKGROUND WORK

Work is related to works on privacy setting configuration in social sites, recommendation systems, and privacy analysis of online images. Privacy Setting Configuration: 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 offriends 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. [13] 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 in line with the 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 [1] have presented an expressive language for images uploaded in social sites. This work is complementary to this as it is do not deal with policy expressiveness, but rely on common forms policy specification for the predictive algorithm. Recommendation Systems the work is related to some existing recommendation systems which employ machine learning techniques. Chen et al. [9] proposed a system named SheepDog to automatically insert photos into appropriate groups and recommend suitable tags for users on Flickr. They adopt concept detection to predict relevant concepts (tags) of a photo. Choudhury et al. [10] proposed a recommendation framework to connect image content with communities in online social media. They characterize images through three types of features: visual features, user generated text tags, and social interaction, from which they recommend the most likely groups for a given image. Similarly, Yu et al. [12] proposed an automated recommendation system for a user's images to suggest suitable photo-sharing groups.

3 ADAPTIVE PRIVACY POLICY PREDICTION FRAMEWORK

The A3P system consists of two main mechanisms: A3P-core and A3P-social. The complete data flow is the following. When a user uploads an image, the image will be first sent to the A3P-core. The A3P-core organizes the image and defines 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[12]: (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 growth of social networking activities[15] . 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. The actual policy will be stored in the policy repository of the system for the policy prediction of future uploads.

4 PROPOSED SCHEME: A3P-CORE

There are two major components in A3P-core: (i) Image classification and (ii) Adaptive policy prediction. For each user, his/her images are first classified based on content and metadata. Then, privacy policies of each category of images are analyzed for the policy prediction. Adopting a two-stage approach is more suitable for policy recommendation than applying the common one-stage data mining approaches to mine both image features and policies together[15]. Recall that when a user uploads a new image, the user is waiting for a recommended policy. The two-stage approach allows the system to employ the first stage to classify the new image and find the candidate

sets of images for the subsequent policy recommendation. As for the one-stage mining approach, it would not be able to locate the right class of the new image because its classification criteria needs both image features and policies whereas the policies of the new image are not available yet. Moreover, combining both image features and policies into a single classifier would lead to a system which is very dependent to the specific syntax of the policy. If a change in the supported policies it isre to be introduced, the whole learning model would need to change.

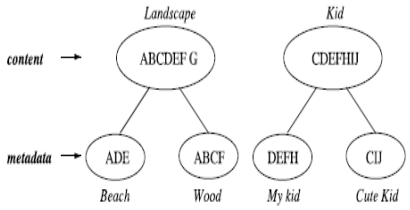

Fig. 2. Two-level Image classification

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. 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. 4.2.1 Policy Mining It is propose a hierarchical mining approach for policy mining. The methodology leverages association rule mining techniques to determine 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[16]. Disparately, the categorized 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.

Policy Prediction: The policy mining phase may generate several candidate policies while the goal of the system is to return the most promising one to the user. Thus, it is present an approach to choose the best candidate policy that follows the user's privacy tendency [12]. To model the user's privacy tendency, it is 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 loit isr the value, the higher the strictness level. It is generated by two metrics: major level (denoted as l) and coverage rate (a), where l is determined by the combination of subject and action in a policy, and a is determined by the system using the condition component. If the policy has multiple subjects or actions and results in multiple l values, it will consider the last one. It is worth nothing that the table is automatically generated by the system but can be modified by users according to their needs. Then, it is introducing the computation of the coverage rate a 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 it is define a as the percentage of people in the specified subject category who satisfy the condition in the policy.

5 EXPERIMENTAL EVALUATION

In particular, it is use a straw man solution as the baseline approach, whereby it is sample atrandom a small set of image settings from the same user and use them to determine a baseline setting. The baseline settings are applied to all images of the users. In advance, it is compared the A3Pcore with two variants of itself, in order to evaluate the contribution of each component in the A3P-core made for privacy prediction. The first variant uses only content-based image taxonomy tracked by the policy mining algorithm, denoted as "Content+Mining". The second variant uses only tag classification follow it is by the policy mining, denoted as "Tag+Mining". All the algorithms it is are tested against the collected real user policies. Fig. 4 shows the percentage of predicted policies in the groups: "Exact Match" means a predicted policy is exactly the same as the real policy of the same image; "x-component Match" means a predictedpolicy and its corresponding real policy have x components (i.e., subject, action, condition) fully matched; "No match" simply means that the predicted policy is wrong for all components. As shown in the figure, each component of the A3P-core singularly contributes toward policy prediction, whichit is over; none of them individually equalizes the accuracy achieved by the A3P-core in its entirety. Specifically, A3P-core has 90 percent exact match and 0 no match.

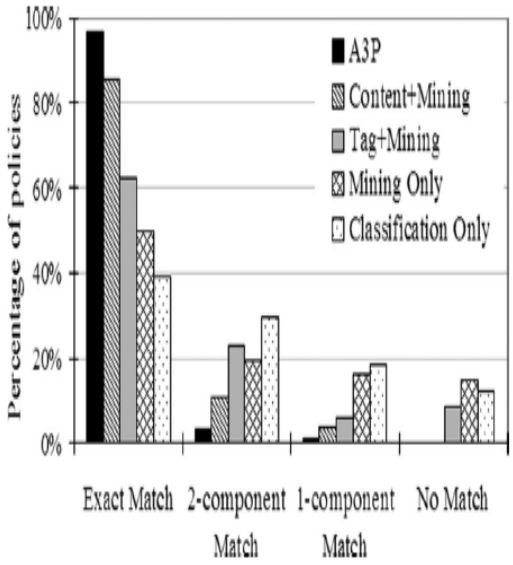

Fig. 3. A3P comparative performance

It is complete this experiment on the second data set of over 2,000 images. The goal is to investigate whether the different population and the heterogeneous set of images from the second data set influences the quality of the prediction [15]. Also, this data set is characterized by a better meta-data, as manual inspection revealed that the user entered tags are all completed, meaningful and with little jargon or use of stop words within them. For this experiment, it is again used the straw man approach for comparison which consisted of replicating the latest generated policy by the user. For mismatched policies, it is further examined the type of error. It is found that there it is total 97 mismatched items (i.e., mismatched subjects, actions and conditions) in[16] those policies. About 60 percent of the errors it is are due to false positive, which means the predicted policy contains more items than the actual policy. It is also noticed that 82.7 percent of the mismatched policies have two components, the subject and action component, fully matched. The most common errors occur within the condition component as this component is the most flexible and can vary significantly if users want to add special constraints. Interestingly, the errors it is reported mainly in the first three or if the policies displayed to the user. This demonstrates the adaptive nature of the A3P system. Upon correcting mismatched policies, the system's accuracy increases. It is also expect that with more user data and a longer execution of the A3P system, the prediction accuracy will be further increased, as the system adapts to users' privacy preferences.

Method	View	Comment	Tag, Notes, Download	Overall
A3P-core	92.48%	92.48 %	92.63 %	92.53 %
Propagation	66.12 %	66.825 %	68.64 %	66.84 %
Tag-Only	87.54%	87.03 %	86.64 %	87.01 %

Table 1. Results of A3P-Core on Picalert Data Set

6 CONCLUSION AND FUTURE ENHANCEMENTS

An Adaptive Privacy Policy Prediction (A3P) system supports users systematizes the privacy policy locales for their uploaded imaginings. The A3P system affords an inclusive outline to assume privacy inclinations founded on the information obtainable for a given user. It is effectively undertaken the issue of cold-start, leveraging social context information. Experimental study proves that the A3P is a practical tool that offers substantial enhancements over contemporary methodologies to concealment.

REFERENCES

[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 it is 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.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 08
July 2017

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_aticaTe_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] A. Mazzia, K. LeFevre, and A. E.,, "The PViz comprehension tool for social network privacy settings," in Proc. Symp. Usable Privacy Security, 2012.
[15] K. Strater and H. Lipford, "Strategies and struggles with privacy in an online social networking community," in Proc. Brit. Comput. Soc. Conf. Human-Comput. Interact., 2008, pp.111–119.
[16] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," Adv. Artif. Intell., vol. 2009, p. 4, 2009.
[17] X. Sun, H. Yao, R. Ji, and S. Liu, "Photo assessment based on computational visual attention model," in Proc. 17th ACM Int. Conf. Multimedia, 2009, pp. 541–544.