International Journal of Research
Available at https://edupediapublications.org/journals

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 01
January 2015

Privacy using Adaptive Privacy Policy Prediction (A3P) for User Uploaded Images on Content Sharing Sites

P Pavan Kumar[1]; M Suresh[2]; K Rama Krishniah[3] & D Sarath Babu[4]

[1]M.Tech Student, Dept of CSE, R K College of Engineering, Krishna-521456,A. P., India
[2]Assistant Professor, Dept of CSE, R K College of Engineering, Krishna-521456,A. P., India
[3]Professor & Principal, Dept of CSE,R K College of Engineering, Krishna-521456,A. P., India
[4]Assistant Professor, Dept of CSE, N V R College of Engineering, Krishna-522201,A. P., India

Abstract:
With the mounting quantity of images users split during social Web sites, maintaining seclusion has turn into a most important crisis, as confirmed by a current wave of revealed incidents anywhere users reluctantly shared private information. In beam of these incidents, require of tackle to aid users organize admittance to their common contented is evident. We recommend an Adaptive Privacy Policy Prediction (A3P) system to aid users compose privacy settings for their images. Social context, image content, and metadata are examined as they are potential indicators of users' isolation preferences. We suggest a two-level structure based on *the user's available history on the site and determine the best available privacy policy for the user's images being uploaded. Our way out relies on an image arrangement frame for image categories which may be linked with parallel policies, and on a policy forecast algorithm to repeatedly make a policy for each recently uploaded illustration, also according to users' social features. Over time, the generated policies will follow the evolution of users' privacy attitude. We have used over 5,000 policies, which exhibit the effectiveness of our system, with prediction accuracies over 90 percent.*

Index Terms: Online information services; web-based services.

I. INTRODUCTION

IMAGES are at this time one of the contribution enablers of users' connectivity. Distribution takes place together between earlier recognized groups of identified citizens or social circles (e. g., Google+, Flickr or Picasa), and also more and more with people exterior the users social circles, for purposes of communal discovery-to help them categorize new peers and be qualified about peers good and social environment. However, semantically loaded images may expose content perceptive information. Consider a snap of a student's 2012 commencement observance, for example. It might be collective surrounded by a Google+ circle or Flickr group, but may without cause expose the students BA pos relations and other friends [1]. Sharing images inside online content sharing sites, consequently, may hastily lead to redundant confession and privacy violations. Online media makes it possible to extract the aggregated information about the owner and the subjects in the available content [2].Aggregated sequence results to an unexpected exposure of one's social environment and lead to misuse of personal information [3]. In this paper, we advise an Adaptive Privacy Policy Prediction (A3P) system which automatically generates the personalized policies and aims to provide users a hassle free privacy settings. When images are uploaded by the user A3P system handles factors in the following criteria that influence one's privacy settings of images. Social context of users, such as their profile information and relationships with others helps in getting the users' privacy preferences. For example, users interested in photography may like to share their photos with other non-professional photographers. Users having several family members in their social contacts may share the pictures related to family events among them. However, using common policies across all users or across users with similar property may very simple and might not satisfy individual preferences. Users may have drastically different opinions even on the same type of images. Suppose, person who doesn't need privacy may be willing to share all his personal images in contrary where a more conservative person may just want to share personal images with his family members [4]. In light of these considerations, it is important to find the balancing point between the impact of social environment and individual characteristics of user in order to predict the policies that match each user needs. As time progresses, individuals may change their overall attitude toward privacy. In order to develop system which recommends personalized policies, changes on privacy opinions should be carefully considered with his family members. Considering these points it is important to find the balancing point between the impact of social environment and characteristics of individual user in order to predict the policies that match each individual's needs.

II. LITERATURE SURVERY

Work in this paper is related to privacy setting configuration in social sites, recommendation systems, and privacy analysis of images uploaded in online.

Privacy Setting Configuration: Recent works have studied how to automate the task of privacy settings

Concept of Privacy suites was proposed by "Bonneau which recommend to users a suite of privacy settings like trusted friends have already set, then users can do minor modification or directly choose a setting. Similarly, Machine-learning based approach was proposed by "Danezis to automatically extract privacy settings from the social context where data is produced with in it. Danezis, Adu-Oppong et al. have done a work parallel to the same to develop privacy settings based on clusters of friends formed by partitioning users' friend lists" which is the concept of "Social Circles. Work related to user's privacy preferences for location-based data (i.e., share her location or not) based on location and time of day has been done by "Ravi Chandran". A privacy wizard was proposed by Fang to help users grant privileges to their friends. The wizard presents users to first assign privacy labels to selected category, and based on the input classifier is constructed which classifies category based on their profiles and privacy labels are assigned to the unlabeled category. More recently, Klemperer et al. studied whether the keywords and captions used by the users to tag their photos can be used to help users more intuitively create and maintain access-control policies [5]. Their findings are in line with our approach: tags created for organizational purposes can be reused to help create reasonably accurate access-control rules. The mentioned approaches focus on deriving policy settings for only those who 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 have presented an expressive language for images uploaded in social sites [6]. This work is complementary to ours as we do not deal with policy expressiveness, but depend on common forms policy specification for our predictive algorithm.

III. SYSTEM ANALYSIS

Existing system: Currently content sharing websites allow users to enter their privacy preferences. Unfortunately, recent studies have revealed that users struggle to set up and maintain such privacy settings. This process can be tedious and error-prone for the given amount of shared information. Therefore, many have acknowledged the need of policy recommendation systems which can help the users to easily and properly configure privacy settings. However, existing proposals for determining privacy settings automatically appear to be inadequate as they do not address the unique privacy needs of images, because of the amount of information implicitly carried within images, and their relationship with the online environment where they are exposed.

Proposed System: In projected System an Adaptive Privacy Policy Prediction (A3P) system that lends hand users computerize the isolation policy settings for their uploaded images. The A3P system provides a wide-ranging support to assume isolation preferences based on the in sequence obtainable for a given user. We also efficiently tackled the concern of cold-start, leveraging community framework information. Our investigational revision proves that our A3P is a realistic tool that offers important improvements more than existing approaches to privacy. Preserve both competence and high forecast accurateness of a system.

IV. A3P FRAMEWORK

Preliminary Notions: Users can set the privacy preferences to their content using privacy policies. Our policies are inspired by popular social sites (i.e., Facebook, Picasa, Flickr) [2], although the actual implementation varies from each content-management site structure and implementation. We define privacy policies according to the below definition:

Definition: A privacy policy P of user u consists of the following components:
Subject (s): A set of users socially connected to user.
Data (d): A set of data items shared by user.
Action (a): A set of actions granted by user to s on d
Condition (c): A boolean expression which must be satisfied in order to perform the granted actions.

System Architecture: The A3P system consists of two main components: A3P-core and A3P-social. The overall data flow is as follows:

1. When a user uploads an image, the image will be first sent to the A3P-core.
2. A3P-core classifies the image and determines whether if there is a need to invoke the A3P-social.
3. A3P-core predicts policies for the users directly based on their historical behavior in most of the cases.
4. A3P-core will invoke A3Psocial:
   (i) If the user does not have enough data for policy prediction of the uploaded image

(ii) The A3P-core detects the recent major changes in privacy practices among the user's community along with user's increase of social networking activities (like new friends, new posts on one's profile etc).

In above cases, the A3P-social groups users into social communities with similar social context and privacy preferences by continuously monitoring the social groups.

When the A3P-social is invoked, the social group is automatically identified for the user and the information about the group is sent back to the A3P-core for policy prediction.

The predicted policy will be displayed to the user at the end. If the user is fully satisfied by the predicted policy, he or she can just accept it. Else, the user can choose to revise the policy.

The actual policy will be stored in the policy repository of the system and will be used for policy prediction of future uploads.

A3P-CORE:

Image classification and Adaptive policy prediction are two major components of A3P-core.

For every user:

1. Classify the images based on content and metadata. Then, privacy policies of each category of images are analyzed for the policy prediction.
2. A two-stage approach is adopted for more suitable policy recommendation than applying one-stage data mining approach to mine both image features and policies together.
3. By implementing a two stage approach, the system initially classify the image, then finds and assigns the right set of policy recommendation. This approach is not possible in one-stage mining as the policies are not available yet.
4. In order to make the system independent of specific syntax of the policy, it is crucial to implement two stage approach. This also helps to segregate the learning model from being revisited when there are changes in the policies.

Image Classification

A hierarchical image classification initially classifies images first based on their contents and then based on metadata it further refines each category into subcategories. Images that have no metadata will be grouped by content. Therefore, hierarchical classification gives a higher priority to image content minimizes the influence of missing tags. Of course, there's a exception that images having typical content features or metadata will be included in multiple categories.

Content-Based Classification

Content-based classification is based on an efficient and yet accurate image similarity approach. Classification algorithm which we propose compares image signatures defined based on quantified and sanitized version of Haar wavelet transformation. or The wavelet transform encodes frequency and spatial information related to image color, size, invariant transform, shape, texture, symmetry, etc for each image

A small number of coefficients are selected to form the signature of the image. Similarity among images is then determined by the distance.

The policy prediction algorithm presents a predicted policy of a newly uploaded image to the user for his/her reference. All possible changes of a user's privacy concerns are addressed in the predicted policy. The prediction process consists of three main phases: (1) policy normalization; (2) policy mining; and (3) policy prediction. The policy normalization is a simple decomposition process which converts a user policy into a set of atomic rules in which the data (D) component is a single-element set among their image signatures. After verifying the accuracy of the classifier, lets discuss its use in the context of the A3P core.

1. When a user uploads an image, it is considered as an input query image. The signature of the newly uploaded image is compared with the signatures of images in the current image database. And the class of the uploaded image is determined
2. We find its first m closest matches. The class of the uploaded image is then calculated based on the class to which majority of them images belong.
3. If no predominant class is found, a new class is created for the image. If the predicted policy for this new image turns out correct, the image will be inserted into the corresponding image category in our image database later, in order to refine future policy prediction. In our current prototype, m is set to 25 which is obtained using a small training data set.
4. The distance between the image and the subcategory is computed as a weighted sum of the edit distance between corresponding pair of representative hypernyms as shown in Equation

$$Dist_m = w_n \cdot D\left(h_n, h_n^c\right) + w_a \cdot D\left(h_a, h_a^c\right) + w_v \cdot D\left(h_v, h_v^c\right). \qquad (1)$$

Note that $w_n + w_a + w_v = 1$, and $w_n \;\; > \; w_a \;\; > \; w_v$. In Equa-

Adaptive Policy Prediction

Policy Mining

We propose a hierarchical mining approach for policy mining. Association rule mining techniques are used to discover popular patterns in policies. As the images in the same category are more likely under the similar level of privacy protection., Policy mining is carried out within the same category of the new image. The idea of the hierarchical mining is to follow an order where 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[7].

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 of the policies containing both popular subjects and conditions. [8]

Suppose that there are six images in the same category of the newly uploaded image "park.jpg" and the corresponding policies are P2, P5, P9, P13, P18 and P22. Table 1 shows what subjects are mentioned in each policy. Mining data in Table 1 may return a best association rule like Rsub 1 : {family} ➔ {friends} meaning that when the user specifies a policy for his family members, he tends to grant the same access right to his friends.

TABLE 1
Example of Subject Component

PolicyID	family	friend	coworker	others
P_2	0	0	1	0
P_5	1	1	0	0
P_9	1	1	0	0
P_{13}	1	1	0	0
P_{18}	0	1	1	0
P_{22}	1	0	0	0

Policy Prediction

The policy mining may generate several candidate, policies while our goal is to return the most promising one for the user. We present a user's privacy tendency approach to choose the best candidate policy that.

Let us define a notion of strictness level to define user's privacy policy The strictness level is a quantitative metric that describes how "strict" a policy is. In particular, a stricter 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 (a), where l is determined by the combination of subject and action in a policy, and is determined by the system using the condition component.

l is obtained via Table 1. In Table 1, 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 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, we define a as the percentage of people in the specified subject category who satisfy the condition in the policy. For example, a user has five family members documented in the system and two of them are kids.

TABLE 4 Major Level Look-Up Table

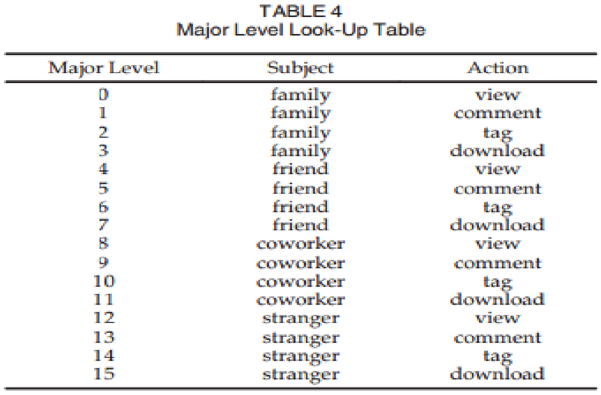


TABLE 4
Major Level Look-Up Table

Major Level	Subject	Action
0	family	view
1	family	comment
2	family	tag
3	family	download
4	friend	view
5	friend	comment
6	friend	tag
7	friend	download
8	coworker	view
9	coworker	comment
10	coworker	tag
11	coworker	download
12	stranger	view
13	stranger	comment
14	stranger	tag
15	stranger	download

A3P-SOCIAL

The A3P-social generates representative policies by employing a multi-criteria inference mechanism by leveraging key information related to the user's social context and his likely attitude toward privacy. As discussed earlier, A3Psocial will be triggered by the A3P-core in two scenarios.

One is when the user is a newbie of a site, and does not have enough images stored for the A3P-core to match the meaningful and customized policies.

The other is when the system notices significant changes of privacy settings in the user's social circle, which may be of interest for the user to possibly adjust his/her privacy settings correspondingly. In what follows, we first present the types of social context considered by A3P-Social, and then recommend policy process.

Let R1; ... ; Rn denote the n types of relationships observed among all users. Let Nu Ri denote the number of user U's contacts belonging to relationship type Ri. The connection distribution (denoted as Conn) is represented as below:

$$Conn : \left\{ \frac{N^u_{R_1}}{\sum_{i=1}^{n} N^u_{R_i}}, \ldots, \frac{N^u_{R_n}}{\sum_{i=1}^{n} N^u_{R_n}} \right\}.$$

Identifying Social Group

We know introduce the policy recommendation process based on the social groups obtained from the previous step.

Suppose that a user U uploaded a new image and the A3P-core invoked the A3P-social for policy recommendation.

The A3P-social will find the social group which is most similar to user U and then choose the representative user in the social group along with his images to be sent to the A3P-Core policy prediction module to generate the recommended policy for user U. Given that the number of users in social network may be huge and that users may join a large number of social groups, it would be very time consuming to compare the new user's social context attributes against the frequent pattern of each social group. In order to speed up the group identification process and ensure reasonable response time, we leverage the inverted file structure to organize the social group information. The inverted file maps keywords (values of social context attribute) occurring in the frequency patterns to the social groups that contain the keywords. Specifically, we first sort the keywords (except the social connection) in the frequency patterns in an alphabetical order. Each keyword is associated with a link list which stores social group ID and pointers to the detailed information of the social group.

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

Fig 1 shows the architecture of the A3P Framework which we designed.

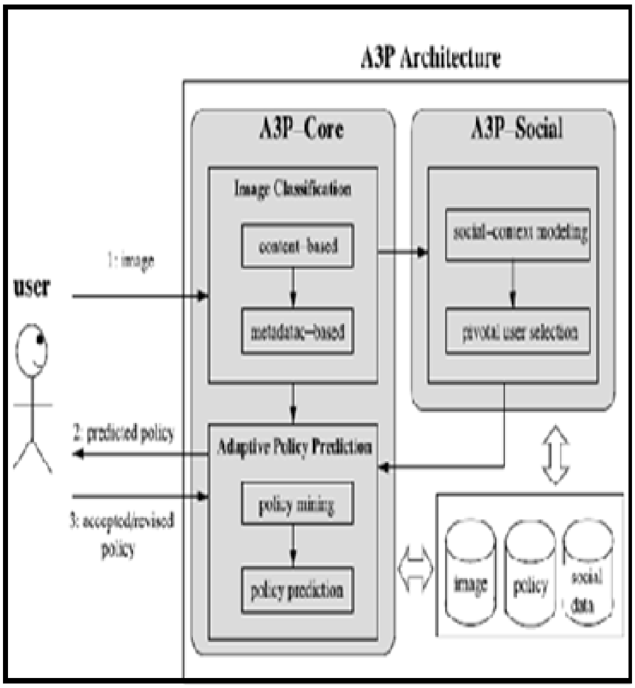
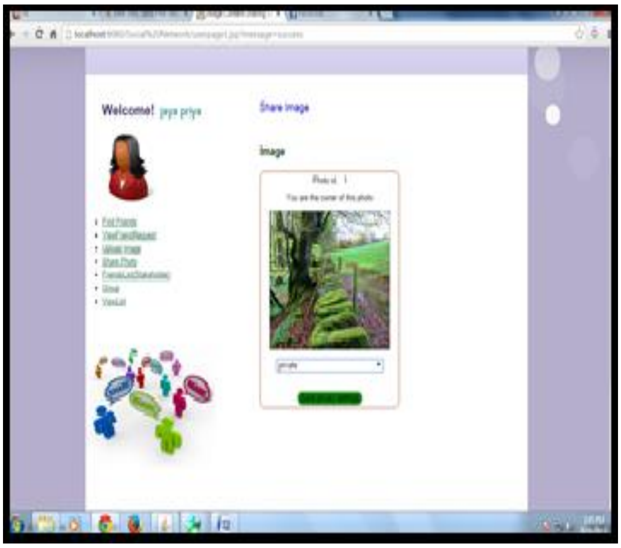


Fig 1. Architecture for A3P

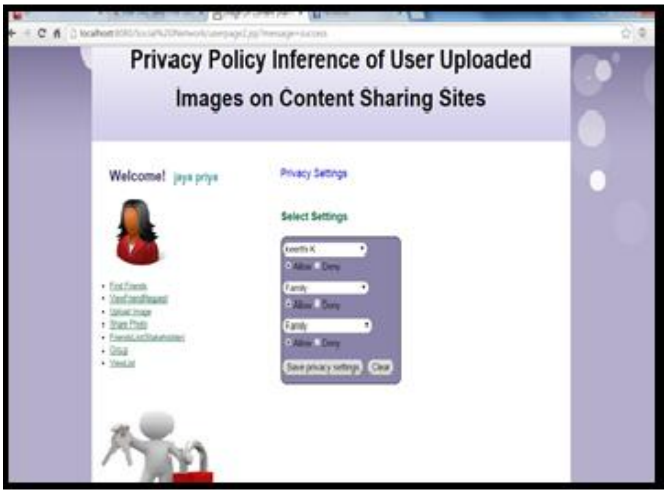
V. EXPERIMENTAL RESULTS

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods A3P-CORE,A3P-SOCIAL .

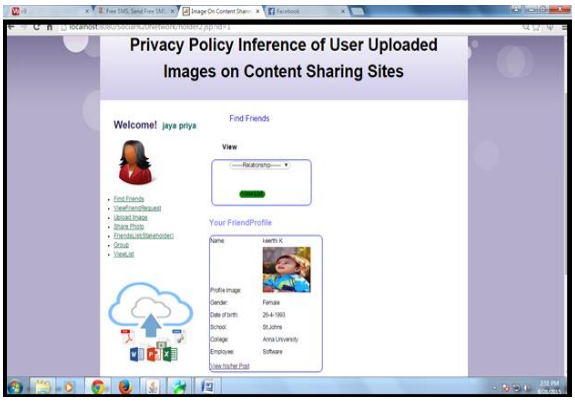
Step 1: Login, upload image and make the privacy setting as private



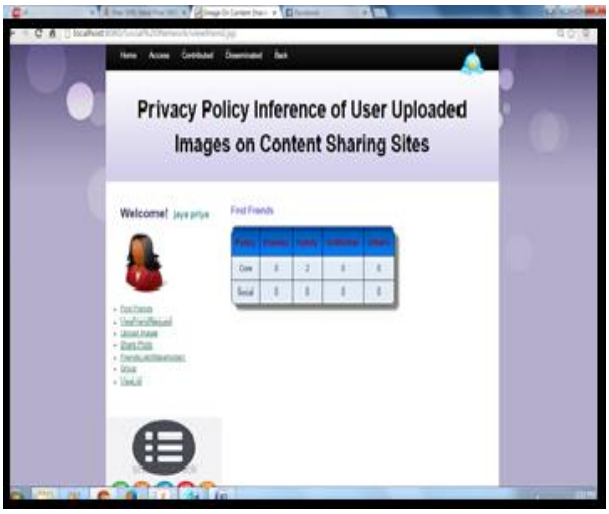
Step 2: Select the privacy settings



Step 3: View friends list and profile



Step 4: Display List of Policies


Privacy Policy Inference of User Uploaded Images on Content Sharing Sites
Welcome! jaya priya
Find Friends

VI. CONCLUSION

We have projected an Adaptive Privacy Policy Prediction (A3P) scheme that helps users computerize the privacy policy settings for their uploaded images. The A3P structure provides a wide-ranging structure to suppose privacy preferences based on the in order available for a given user. We also successfully tackled the subject of cold-start, leveraging social circumstance information. Our investigational study proves that our A3P is a matter-offact tool that offers considerable improvements over present approaches to privacy.

VII. REFERENCES:

[1] Facebook Developers. http://developers.facebook.com/.

[2] Facebook Privacy Policy. http://www.facebook.com/policy.php/.

[3] Facebook Statistics. http://www.facebook.com/press/info.php?statistics.

[4] Google+ Privacy Policy. http://http://www.google.com/intl/en/+/policy/.

[5] A. Besmer and H. Richter Lipford. Moving beyond untagging: Photoprivacy in a tagged world. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1563–1572. ACM, 2010.

[6] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda. All your contacts are belong to us: automated identity theft attacks on social networks. In *Proceedings of the 18th international conference on World wide web*, pages 551–560. ACM, 2009.

[7] B. Carminati and E. Ferrari. Collaborative access control in online social networks. In *Proceedings of the 7th Internationaattacl Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pages 231–240. IEEE, 2011.

[8] B. Carminati, E. Ferrari, and A. Perego. Rule-based access control for social networks. In *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, pages 1734–1744. Springer, 2006.