An A3P approach towards Image Privacy on Social Sites

Mr. Rishiraj C. Bhosale
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering
Lohegaon,Pune,India
rbrajbhosale@gmail.com

Mr. Harsh A.Patel
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering
Lohegaon,Pune,India
akvnh7996@gmail.com

Prof. P. T. Borse
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering
Lohegaon,Pune,India
Pborse386@gmail.com

Ms. Kasturi K. Ghonsikar
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering
Lohegaon,Pune,India
k.ghonsikar@gmail.com

Mr. Taher K. Lakdawala
Department of Computer Engineering
Dr. D. Y. Patil School of Engineering
Lohegaon,Pune,India
taherlakdawala9@gmail.com

Abstract**—** Usage of social media's has been considerably increasing in today's world which enables the user to share their personal information like images with other users. This improved technology leads to privacy desecration where the users can share large number of images across the network. To provide security for the information, we put forward this paper consisting Adaptive Privacy Policy Prediction (A3P) framework to help users create security measures for their images. The role of images and its metadata are studied as a measure of user's privacy preferences. The Framework defines the best privacy policy for the uploaded images. It includes an Image classification framework for association of images with similar policies and a policy prediction technique to automatically generate a privacy policy for user-uploaded images.

Keywords*- A3P, Data Mining, Image Classification, Image-Sharing, Privacy Policy.*

*****

I. INTRODUCTION

Images are shared widely now a days on social sharing sites . Sharing takes place between friends and associates daily. Image Sharing may lead to exposure of personal information and privacy desecration. This gathered information can be misused by malevolent users.

To prevent such kind of unwanted disclosure of personal images, flexibility in the privacy settings are required. In recent years, such settings for privacy are made available but setting up and conserving these measures is a complex and prone-to-error process. Therefore, recommendation system is required which provide user with easy and a flexible assistance for providing privacy settings in much easier way.

In this paper, we are implementing an Adaptive Privacy Policy Prediction (A3P) system which will provide users a bother free privacy settings experience by automatically generating personalized policies.

II. LITERATURE SURVEY

Some traditional systems shows different studies on automatically assign the privacy settings. One such system which Bonneau et al.[2] shows the concept of privacy suites. The privacy 'suites' alots the user's privacy setting with the help of experts using the system. The expert users are trusted friends who already set the settings for the users.

Similarly, Danesiz [4] proposed an automatic privacy extraction system with a machine learning approach from the data produced from the images. Based on the concept of "social circles" i.e. forming clusters of friends was proposed by Adu-Oppong et al. [3]Prediction of the users privacy preferences for location-based data (i.e., share the location or no) was studied by Ravichandran et. Al[6]. This was done on the basis of time of the day and location.The study of whether the keywords and captions used for tagging the users photos can be used more efficiently to create and maintain access control policies was done by Klemperer et al.

III. SYSTEM ARCHITECTURE

A. A3P Framework

Privacy policies are the changes or settings made by user other than normal preferences for the security of the content disclosure to other connected users.

Privacy policies is defined as follows:
Definition: A Privacy policy 'P' can be described for user 'U' by

Subject(S) : A Set of users socially connected to user U.
Data (D) : A set of data items shared by U.
Action (A) : A set of actions granted by U to S on D.
Condition (C) : A boolean expression which must be satisfied in order to perform the appointed actions.

In the above definition, Subject(S) can be socially connected people on websites like , relations such as family, friend, co-workers, etc. and organizations.

Data(D) is the collection of image uploaded by user till date. Action(A) consists of four factors: View, Comment, Tags and Download.Condition(C) specifies whether the actions are effective or not.

Example 1. 'A' wants to allow her friends and family
to view and comment on images in the album named
"birthday_album" and the image named "cake.jpg"
before year 2015.The policy for her privacy preference will be
P: [{friend, family}, {birthday_album, cake.jpg},
{view ,comment}, (date< 2015)].

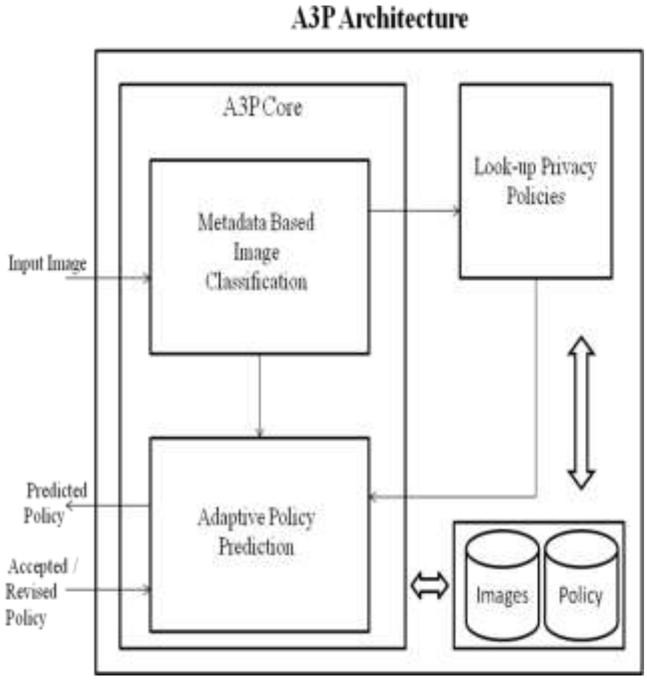
B. A3P Architecture

A3P (Adaptive Privacy Policy Prediction) is a framework used for defining new privacy preferences policies for users and to make the experience flexible and secure at the time. The A3P Architecture consists of followings blocks:

1. A3P Core.
   - Metadata based Image classification.
   - Adaptive policy prediction.
2. Look-Up Privacy Policies
3. Database

A3P Core is used for classification of images with the help of metadata of the image and also provides the new predicted policy based on the behaviour of user.

The Look-up Privacy Policy block provides the user with the information whether same image exists in the database and if it does then provide the same policy predicted previously . Otherwise, the image is stored as new for further help in policy prediction.



C. A3P Core

The A3P Core consist of two major blocks:

i. Metadata based Image Classification
ii. Adaptive Policy Prediction

In metadata based classification the user uploaded images are compared and classified with the use of metadata, with this approach of metadata-based-classification the policy recommendation becomes easy and more accurate. Based on the Classification through metadata, the policies are applied to the right class of images. Metadata classification plus policy prediction will provide better and efficient policies for users.

D. Metadata Based Image Classsification

As mentioned, the metadata based Image classification are divided into sub-categories with the help of following three steps.

Step 1 of this process allows to extract keywords from the metadata of the image. Tags, Comments and Captions are the types of metadata through which the keywords are obtained. After the keywords are obtained, our task is to identify different properties like nouns, verbs and adjectives and store them into a metadata vector such as

$T_n=\{t_1,t_2,t_3,\ldots..,t_k\}$, $T_v=\{t_1,t_2,t_3,\ldots..,t_j\}$,
$T_a=\{ t_1,t_2,t_3,\ldots..,t_l\}$ where k,j and l are the total number of nouns, verbs and adjectives respectively.

Step 2 of this process is to have a similar hypernym from each vector. The hypernym is denoted by ‘h’ and first retrieved for every ‘t_i’. This hypernym can be represented as “h={(v_1,f_1), (v_2,f_2),….}”.Here ‘v’ are hypernyms and ‘f’ is for frequency. For example, consider a metadata vector T={, ”Promotion”, ”Job”,”Party”}.By this set we can learn that Job and Promotion are with same hypernym ‘work’ but Party has a hypernym ‘Activity’. Hence, we can show the hypernm list as h={(work,2),(Activity,1)}.From this list we select the hypernym with the maximum frequency.

Step 3 of this process is to show and learn the subcategory in which the image fits in. The incremental procedure in which the first image forms a subcategory and the hypernyms of the image are also allotted to their respective subcategory.The closeness between these hypernyms and each category is computed to define a subcategory for that image.

E. Adaptive Policy Prediction

This part deals with the privacy concerns of the user by deriving the privacy policies for the images. The Adaptive Policy Prediction consists of two following sub-parts:

i. Policy Mining
ii. Policy Prediction

Policy mining deals with mining of policies for images with same categories and Policy prediction applies prediction algorithm to predict the policies.

1) *Policy Mining:* The privacy policies are the privacy preferences expressed by the users. Policy mining features out these policies by applying association rules and methodology. It follows the sequence in which a user has to define a policy and decides what rights are applicable to the images. This hierarchical mining approach initiates by looking the important subjects and their popular actions in the policies and finally goes for the conditions. It can be approached with the help of following steps.
Step 1 of this process focuses on Association rule mining on the subject components of the image and its policies. With the association rule mining, the best rules are written according to one of the interestingness measure i.e., support and confidence giving the most popular subjects in policies.
Step 2 of this process applies the rules on the action components. Similar to the first step we will select the best rules which will give the best combinations of action in policies.
Step 3 of this process carries out the mining on the condition component in each policy set. The rules giving the best outcomes are selected which gives us a set of attributes which often appear in policies.

2) *Policy Prediction:* The policy mining phase may provide us with many policies but our system needs to choose the best one to the user. Thus, this approach is used to get the best policy for the user on the bases of strictness level. The Strictness level decides how “strict” a policy is by giving an integer value. This value should be least to acquire high strictness. The strictness can be discovered by major level and coverage rate. The major level is calculated with the help operations on subject and action in a policy and coverage rate is determined using the condition. Different

range values are assigned based on the strictness to the combinations and for data with multiple combinations we will select the lowest rate. It provides a fine-grained strictness level which adjusts the major level obtained earlier. For example a user has five friends and two of them are females. Hence if a policy as “friends”=male is applied, then the coverage rate can be calculated as (3/5)=0.6. Hence, the restricted on the image is less if the coverage rate value is high.

IV. IMPLEMENTATION

A. *Flow of Image Uploading System*

1. START
2. Select an image to upload.
3. Enter Caption and Tags for the selected image.
4. Proceed to upload the image.
5. Call method to get image id which is having most similar caption and tags.
   (Algorithm of Privacy Policy Prediction)
6. Get privacy policies already set for the result image id.
7. Show predicted policies to user.
8. If user is satisfied with predicted policies then continue to upload image.
9. If user is not satisfied with predicted policies then allow user to set privacy
   policy for the image and continue to upload.
10. END.

B. *Algorithm of Privacy Policy Prediction*

INPUT: Caption & Tags.

OUTPUT : Relevant Image Id.

1. Get Caption and Tags from front-end.
2. Execute SQL query to search for image having exact same caption and tags.

```
Resultset matchedId=
executeQuery(ExactMatch(WholeCaption && AllTags));
If(matchedId is not null){
        Return matchedId;
}Else{
        Resultset matchedId=
        executeQuery(ExactMatch(WholeCaption || AllTags));
        If(matchedId is not null){
                Return matchedId;
        }Else{
                Resultset matchedId=
                executeQuery(ExactMatch(PartialCaption && MinimumTags));
                If(matchedId is not null){
                        Return matchedId;
                }Else{
                        Resultset matchedId=
                        executeQuery(ExactMatch(PartialCaption || MinimumTags));
                        If(matchedId is not null){
                                Return matchedId;
                        }Else{
                                Return 0;
                        }
                }
        }
}
```

V. FUTURE SCOPE

The future scope for the given A3P approach rests by employing a mechanism that generates policies by key information based on Social context of the users and their general attitude towards the privacy. Activities performed between different users create a Social group which can be modeled to provide the privacy policies.

Also an external domain like Image processing can be used to classify the image more accurately without being dependent on the metadata collected from the user.

VI. CONCLUSION

We have studied and approached towards an adaptive privacy policy prediction in this paper that assists users for maintaining the privacy of their uploaded images by automatically recommending privacy policies. This system provides a framework which deduces privacy preference based on the history of the users proclivity. this help user to set hassle free and flexible policy selction.

VII. ACKNOWLEDGEMENT

We would like to thank Dr. D. Y. Patil School Of Engineering for providing us with all the required amenities. We are also grateful to Prof. S. S. Das, Head of Computer Engineering Department, DYPSOE, Lohegaon, Pune for their indispensable support, suggestions and motivation during the entire course of the project

VIII. REFERENCES

[1] Anna Cinzia Squicciarini, Member, IEEE, Dan Lin, Smitha Sundareswaran and Joshua Wede, ”Privacy *Policy Inference of User-Uploaded Images on Content Sharing sites*”.IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING,VOL. 27,NO. 1, JANUARY 2015

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

[3] 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

[4] 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.

[5] 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.

[6] R. Ravichandran, M. Benisch, P. Kelley, and N. Sadeh, “Capturing social networking privacy preferences,” in Proc. Symp. Usable Privacy Security, 2009.