International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 25 Issue 1 – JANUARY 2018.

Privacy Policy Using Content and Metadata Based Search with Image Security

A.Sudha Prasanna[#1] and Dr. K.V. Sambasiva Rao[*2]

[#]*Research Scholar, Dept. of Computer Science & Engg., NRI Institute of Technology, Agiripalli, India*

[*] *Dean, Dept. of Computer Science & Engg., NRI Institute of Technology, Agiripalli, India*

***Abstract*— Social Network is an emerging E-service for content sharing sites (CSS). It is emerging service which provides a reliable communication, through this communication a new attack ground for data hackers; they can easily misuses the data through these media. Some users over CSS affects users privacy on their personal contents, where some users keep on sending unwanted comments and messages by taking advantage of the users' inherent trust in their relationship network. 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.**

***Index Terms*— Adaptive Privacy Policy Prediction (A3P), A3P- Core, A3P- Social, Polar Fourier Transform (PFT)**

I. INTRODUCTION

Pictures are presently one of the key empowering agents of clients' network. Sharing happens both among already set up gatherings of known individuals or groups of friends (e. g., Google+, Flickr or Picasa), furthermore progressively with individuals outside the clients groups of friends, for reasons for social revelation to help them distinguish new companions and find out about associates interests and social environment. In any case, semantically rich pictures may uncover content sensitive data.

Sharing pictures inside online substance sharing sites, therefore, may rapidly lead to undesirable exposure what's more, protection infringement. Further, the industrious nature of online media makes it workable for different clients to gather rich amassed data about the proprietor of the distributed substance and the subjects in the distributed substance.

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. Users who have several family members among their social contacts may share with them pictures related to family events. However, using common policies across all users or across users with similar traits may be too simplistic and not satisfy individual preferences. Users may have drastically different opinions even on the same type of images. For example, a privacy adverse person may be willing to share all his personal images while a more conservative person may just want to share personal images with his family members. In light of these considerations, it is important to find the balancing point between the impact of social environment and users' individual characteristics in order to predict the policies that match each individual's needs.

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

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.

II. RELATED WORK

Our work is related to works on privacy setting configuration in social sites, recommendation systems, and privacy analysis of online images. Our continuous research looks at security in online social organizing locales, meaning to enhance the security and security administration of individual data. As a first step, we are analyzing and proposing changes to current security systems.

III. LITERATURE SURVEY

Privacy Suites [1] is proposed by Jonathan Anderson which allows users to easily choose ―suites" of privacy settings. Using privacy programming a privacy suite can be created by an expert. Privacy Suites could also be created directly through existing configuration UIs or exporting them to the abstract format. To the members of the social sites the privacy suite is distributed through existing distribution channels. Transparency is the main goal, which is essential for convincing influential users that it is safe to use. The disadvantage of a rich programming language is less understandability for end users. To verify a Privacy Suite sufficiently high-level language and good coding practice, motivated users are able.

Privacy-Aware Image Classification and Search [2] is a technique to automatically detect private images, and to enable privacy-oriented image search introduced by Sergej Zerr. To provide security policies technique combines textual meta data images with variety of visual features. It uses various classification models trained on a large scale dataset with privacy assignments obtained through a social annotation game. 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 ca3pan indicate the presence or absence of particular objects (SIFT).

A tag based access control of data [3] is developed by Peter F. Klemperer. It is 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. A suitable preference can be selected by participants and access the information. Based on the user needs photo tags can be categorized as organizational or communicative.

There are several important limitations .First, our results are limited by the participants recruited and the photos provided by them. Machine generated access-control rules are the second limitation. Algorithm used here has no access to the context and meaning of tags and no insight into the policy the participant intended when tagging for access control. Hence, some rules appeared strange to the participants who makes them to tag explicitly like ―private‖ and ―public

A decentralised authentication protocol [4], is a access control system proposed by Ching-man Au Yeung based on a descriptive tags and linked data of social networks in the Semantic websites. Here users can specify access control rules based on open linked data provided by other parties and it allows users to create expressive policies for their photos stored in one or more photo sharing.

Adaptive Privacy Policy Prediction (A3P) [5] system is introduced by Anna Cinzia Squicciarini. Personalized policies can be automatically generated by this system. It makes use of the uploaded images by users and a hierarchical image classification is done. Images content and metadata is handled by the A3P system .It consists of two components: A3P Core and A3P Social. The image will be first sent to the A3P-core, when the user uploads the image. The A3P-core classifies the image and determines whether there is a need to invoke the A3P-social. When meta data information is unavailable it is difficult to generate accurate privacy policy. This is the disadvantage of this system. Privacy violation as well as inaccurate classification will be the after effect of manual creation of meta data log information.

IV. PROBLEM STATEMENT

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 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, retrieval, and photo ranking, also in the context of online photo sharing sites. Of these works, probably the closest to ours. 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.

V. 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. However, existing proposals for automating privacy settings appear to be inadequate to address the unique privacy needs of images, due to the amount of information implicitly carried within images, and their relationship with the online environment wherein they are exposed.

VI. PROPOSED SCHEME

In proposed System 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. Our experimental study proves that our A3P is a practical tool that offers significant improvements over current approaches to privacy.

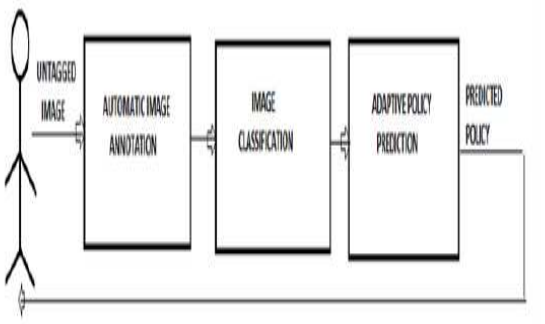
A. Adaptive Policy Prediction:

The Adaptive Policy Prediction consists of two following sub-parts:

1. Policy Mining
2. Policy Prediction

B. Policy Mining:

A hierarchical mining approach for policy mining is used. Policy mining is carried out within the same category of the new image. The basic idea of this is to follow a natural order in which a user defines a policy. 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 in the policies containing both popular subjects and conditions. **Policy Prediction:** It is an approach to choose the best candidate policy that follows the user's privacy tendency. To model the user's privacy tendency, define a notion of strictness level. The strictness level is a quantitative metric that describes how "strict" a policy is. a strictness level L is an integer with minimum value in zero, wherein the lower the value, the higher the strictness level.



Proposed System

C. Advantages:

Maintain both efficiency and high prediction accuracy of a system.

VII. SYSTEM MODEL

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:

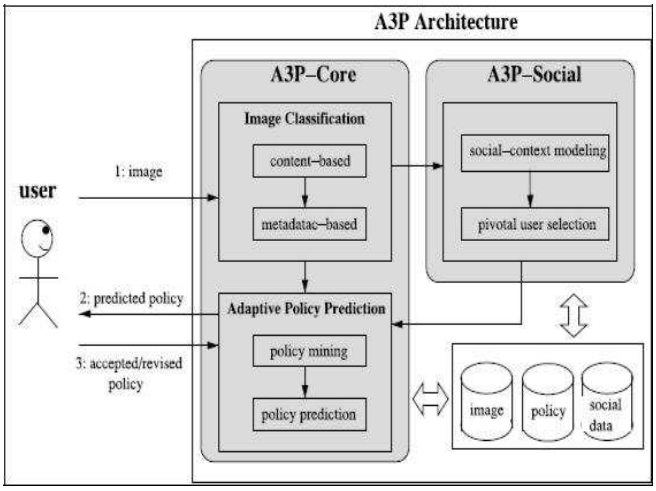
(i) The user does not have enough data for the type of the uploaded image to conduct policy prediction;

(ii) (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 newfriends, new posts on one's profile etc).

In above cases, it would be beneficial to report to the user the latest privacy practice of social communities that have similar background as the user. 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.



System Model

VIII. IMPLEMENTATION

1. A3P-CORE
2. A3P-SOCIAL

A. 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

Image classification: 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.

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.

1) Policy normalization: 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.

2) Policy mining: 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 in the policies containing both popular subjects and conditions.

3) Policy prediction: The policy mining phase may generate several candidate policies while the goal of our system is to return the most promising one to the user. Thus, we present an approach to choose the best candidate policy that follows the user's privacy tendency. To model the user's privacy tendency, we define a notion of strictness level. The strictness level is a quantitative metric that describes how "strict" a policy is.

B. A3P-SOCIAL:

The A3P-social employs a multi-criteria inference mechanism that generates representative policies by leveraging key information related to the user's social context and his general attitude toward privacy. As mentioned earlier, A3Psocial will be invoked 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 infer meaningful and customized policies.

Social Context Modeling: The social context modeling algorithm consists of two major steps. The first step is to identify and formalize potentially important factors that may be informative of one's privacy settings. The second step is to group users based on the identified factors.

IX. CONCLUSION AND FUTURE WORK

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. Our experimental study proves that our A3P is a practical tool that offers significant improvements over current approaches to privacy.

Social network is an upgrading media for information sharing through internet. It provides a content sharing like text, image, audio, video, etc… With this emerging E-service for content sharing in social sites privacy is an important issue. It is an emerging service which provides a reliable communication, through this a new attack ground from an un-authored person can easily misuses the data through these media. For this issue our proposed systems use the BIC algorithm to classify the attackers and the users with the help of the Access Policy Prediction and Access control mechanism. These provide a privacy policy prediction and access restrictions along with blocking scheme for social sites and improve the privacy level for the user in social media.

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 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.

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

[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] A. Kaw and E. Kalu, Numerical Methods with Applications: Abridged., Raleigh, North Carolina, USA: Lulu.com, 2010.
[20] 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.
[21] K. Lerman, A. Plangprasopchok, and C. Wong, "Personalizing image search results on flickr," CoRR, vol. abs/0704.1676, 2007.
[22] H. Lipford, A. Besmer, and J. Watson, "Understanding privacy settings in facebook with an audience view," in Proc. Conf. Usability, Psychol., Security, 2008.
[23] D. Liu, X.-S. Hua, M. Wang, and H.-J. Zhang, "Retagging social images based on visual and semantic consistency," in Proc. 19th ACM Int. Conf. World Wide Web, 2010, pp.1149–1150.
[24] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing facebook privacy settings: User expectations vs. reality," in Proc. ACMSIGCOMMConf. Internet Meas. Conf., 2011, pp. 61–70.
[25] D. G. Lowe, (2004, Nov.). Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. [Online]. 60(2), pp. 91–110. Available: http://dx.doi.org/10.1023/B: VISI.0000029664.99615.94
[26] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 8, pp. 959–973, Aug. 2003.
[27] E. M. Maximilien, T. Grandison, T. Sun, D. Richardson, S. Guo, and K. Liu, "Privacy-as-a-service: Models, algorithms, and results on the Facebook platform," in Proc. Web 2.0 Security Privacy Workshop, 2009.
[28] A. Mazzia, K. LeFevre, and A. E.,, "The PViz comprehension tool for social network privacy settings," in Proc. Symp. Usable Privacy Security, 2012.
[29] M. Rabbath, P. Sandhaus, and S. Boll, "Analysing facebook features to support event detection for photo-based facebook applications," in Proc. 2nd ACM Int. Conf. Multimedia Retrieval, 2012, pp. 11:1–11:8.
[30] R. Ravichandran, M. Benisch, P. Kelley, and N. Sadeh, "Capturing social networking privacy preferences," in Proc. Symp. Usable Privacy Security, 2009.
[31] A. Singhal, "Modern information retrieval: A brief overview," IEEE Data Eng. Bullet., Speicial Issue on Text Databases, vol. 24, no. 4, pp. 35-43, Dec. 2001. [32] A. C. Squicciarini, S. Sundareswaran, D. Lin, and J. Wede, "A3p:
[32] Adaptive policy prediction for shared images over popular content sharing sites," in Proc. 22nd ACM Conf. Hypertext Hypermedia, 2011, pp.261–270.