ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

A Survey on Privacy Policy Inference of Uploaded Images on Content Sharing Sites

Akshada Thorgule, Deeksha Maheshwari, Surbhi Singh, Vaidehi Raju, Prof. Anuja Zade

B.E. Student, Dept. of Information Technology, RSCOE, Savitribai Phule Pune University, India

B.E. Student, Dept. of Information Technology, RSCOE, Savitribai Phule Pune University, India

B.E. Student, Dept. of Information Technology, RSCOE, Savitribai Phule Pune University, India

B.E. Student, Dept. of Information Technology, RSCOE, Savitribai Phule Pune University, India

Asst. Professor, Dept. of Information Technology, RSCOE, Savitribai Phule Pune University, India

ABSTRACT: Social Network is an emerging e-service for content sharing sites (CSS). It is an emerging service which provides reliable communication. Though this communication is a new attack ground for data hackers; they can easily misuse the data through these media. Some users over CSS affect user's privacy on their personal contents, where some users keep on sending unwanted comments and messages by taking advantage of the user's inherent trust in their relationship network. Towards 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 user's 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 user's social features. Over time, the generated policies will follow the evolution of user's 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 per cent.

KEYWORDS: Social media; Content sharing sites, Privacy, Meta data, CSS, APP.

I.INTRODUCTION

Social Networking (SN) is an improving technology with hundreds of millions of people participating inswapping their content through text, media like image, audio, video, etc. Social media (SM) become one of the most important parts of our daily life as it allows us to communicate with a group of people. It assists an exterior of self-expression for users, and assists them to entertain and exchange content with other users through social media's e-service. Some of the social networkslike Friendster.com, Tagged.com, Xanga.com, Live Journal, MySpace, Facebook, Twitter and LinkedIn have developed on the Internet over the past several years. It provides a content sharing mechanism and connects people across the world. Users of social media can define a personal profile and modify it as they wish. This feature is allowed by the SM. Through this SM, users may engage with each other for various purposes like business, leisure, and knowledge sharing. People use social networks to get in touch with further people, and create and contribute content that includes personal information, images, and videos. The service providers have admission to the content presented by their users and have the right to collect data and share them to unauthorized users. A very familiar service provided in SN is to produce proposition for finding new friends, groups, and events using mutual filtering techniques. The success of the SN based on the number of users it attracts, and cheering users to add more users to their circle and to share data with other users in the SN. So the information will go across the world [1]. End users are nevertheless often not aware of the size or nature of the spectators accessing their data and the sense of understanding created by organism among digital friends often leads to disclosures that may not be suitable in a public forum. Due to such an open accessibility of exposed data in SN, the users face a number of security and privacy risks.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

In spite of the fact that content sharing represents one of the important features of existing Social Network sites, Social Networks yet do not sustain any mechanism for collaborative execution of privacy settings for shared content [2]. Social Networking sites are used by a huge number of users all over the world. It provides different features to the customers like chatting, posting comments, image sharing, video chatting etc. Images are now one of the key enablers of user's connectivity.Sharing takes place both among previouslyestablished groups of known people or social circles (e.g., Google+, Flickr or Picasa), and also increasingly withpeople outside the users social circles, for purposes ofsocial discovery-to help them identify new peers andlearn about peers interests and social surroundings.However, semantically rich images may reveal content sensitiveinformation [2]. Consider a photo of a student's2012 graduation ceremony, for example. It could beshared within a Google+ circle or Flickr group, but mayunnecessarily expose the students. Sharing images within online content sharing sites, therefore, may quickly lead to unwanted disclosure and privacy violations [3], [4]. 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 [3], [2], [4]. The aggregated information can result in unexpected exposure of one's social environment and lead to abuse of one's personal information.

II.RELATED WORK

Consider a photo of a student's 2012 graduation ceremony, for example. It could be shared within a Google+ circle or Flickr group, but may unnecessarily expose the student's family members and other friends. Sharing images within online content sharing sites, therefore, may quickly lead to 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.

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.

Our work is related to some existing recommendation systems which employ machine learning techniques. Chen et al. [7] proposed a system named Sheep Dog 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. [42] proposed an automated recommendation system for a user's images to suggest suitable photo-sharing groups. Usage of social media's increased noticeably in today world facilitate the user to distribute theirpersonal information like images with the other users. This enhanced technology leads to privacy disobedience where theusers are allocated large volumes of images across additional number of people. To provide security for theinformation, mechanical explanation of images are introduced which aims to create the meta data information about theimages by using the novel approach called Semantic interpret Markovian Semantic Indexing(SMSI) for repossess theimages [1]. The proposed system automatically interprets the images using hidden Markov model and features areextorted by using colour histogram and Scale Invariant Feature Transform (SIFT) descriptor method. After interpretingthese images, semantic retrieval of images can be done by using Natural Language using a tool named Word Netfor measuring semantic comparison of annotated images in the database. Experimental results make available enhancedretrieval performance when evaluated with the existing system.

III.PROBLEM STATEMENT

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

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

access to their shared content is apparent. Towards 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.

IV.SYSTEM OBJECTIVE

Our purpose is related to works on privacy setting configuration in social sites, recommendation systems, and privacy analysis of online images.

V.FEW RECENTLY IMPLEMENTED TECHNIQUES FOR PRIVACY OF UPLOADED IMAGES

A.An improved Privacy of User Data and Images on Content Sharing Sites using BIC:

Social Network is an emerging e-service for content sharing sites (CSS). It is emerging service which provides a reliable communication. Though this communication is a new attack ground for data hackers; they can easily misuse the data through these media. Some users over CSS affect users privacy on their personal content, where some users keep on sending unwanted comments and messages by taking advantage of the user's inherent trust in their relationship network. By this, privacy of the user's data may be lost. For this issue, this paper handles the most prevalent issues and threats targeting different CSS recently. This proposes a privacy policy prediction and access restrictions along with blocking scheme for social sites using data mining techniques. To perform this, the system utilizes APP (Access Policy Prediction) and Access control mechanism by applying BIC algorithm (Bayesian Information Criterion).

B.Privacy-Aware Image Classification:

Training neutral networks is a computationally intensive task that is best suited for massively parallel machines like GPUs or server farms, and as such, users realistically would have to give their data to the cloud for building the classifier. When giving this data away to be processed, it is at risk of being taken by non-intended parties. In this work, we propose modifying the data being sent; in this case images, such that if it were intercepted, it would be difficult to re-construct the original image. We propose multiple methods for achieving this privacy strategy and show the trade-offs encountered in these scenarios.

C.*Recommending Flickr groups with social topic model:*

The explosion of multimedia content in social media networks raises a great demand to develop tools in order to facilitate producing, sharing and viewing media content. Particularly, Flickr groups, self-organized communities with declared, common interests, are able to help users to conveniently participate in social media network. In this paper, we address the problem of automatically recommending groups to users. We propose to simultaneously exploit media contents and link structures between users and groups. To this end, we present a probabilistic latent topic model to model them in an integrated framework, expecting to jointly discover the latent interests for users and groups and simultaneously learn the recommendation function. We demonstrate the proposed approach on the dataset crawled from Flickr.com.

D.*The PViz Comprehension Tool for Social Network Privacy Settings:*

User's mental models of privacy and visibility in social networks often involve natural subgroups, or communities, within their local networks of friends. Such groupings are not always explicit, and existing policy comprehension tools, such as Facebook's Audience View, which allows the user to view her profile as it appears to each of her friends, are not naturally aligned with this mental model. In this paper, we introduce PViz, an interface and system which corresponds more directly with the way user's model groups and privacy policies applied to their networks. PViz allows the user to understand the visibility of her profile according to natural sub-groupings of friends, and at different levels of granularity. We conducted an extensive user study comparing PViz to current privacy comprehension tools (Facebook's Audience View and Custom Settings page). Despite requiring users to adapt to new ways of exploring

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

their social spaces, our study revealed that PViz was comparable to Audience View for simple tasks, and provided a significant improvement for more complex, group based tasks.

E.*Contextual Dynamics of Group-Based Sharing Decisions:*

In this paper we investigate how decisions made while using a granular access control mechanism for sharing photographs are influenced by contextual factors and properties relating to the identities of contacts. We develop analytical models using logistic regression to understand relationships between variables that affect sharing decisions. We also investigate how predefined, static groups for privacy control cope with the challenge of sharing large amounts of content associated with numerous different contexts, and whether they need to be adjusted to suit particular contexts.

F.*Analyzing Facebook Privacy Settings: User Expectations vs. Reality:*

The sharing of personal data has emerged as a popular activity over online social networking sites like Facebook. As a result, the issue of online social network privacy has received significant attention in both, research literature and the main stream media. Our overarching goal is to improve defaults and provide better tools form an aging privacy, but we are limited by the fact that the full extent of the privacy problem remains unknown. There is little quantification of the incidence of incorrect privacy settings or the difficulty users face when managing their privacy.

G.*Tag, You Can See It! Using Tags for Access Control in Photo Sharing:*

Users often have rich and complex photo-sharing preferences. But properly configuring access control can be difficult and time consuming. In an 18-participant laboratory study, we explore 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. We find that (a) tags created for organizational purposes can be repurposed to create efficient and reasonably accurate access-control rules; (b) users tagging with access control in mind develop coherent strategies that lead to significantly more accurate rules than those associated with organizational tags alone; and (c) participants can understand and actively engage with the concept of tag-based access control.

H.*Using Web Co-occurrence Statistics for Improving Image Categorization:*

Object recognition and localization are important tasks in computer vision. The focus of this work is the incorporation of contextual information in order to improve object recognition and localization. For instance, it is natural to expect not to see an elephant to appear in the middle of an ocean. We consider a simple approach to encapsulate such common sense knowledge using co-occurrence statistics from web documents. By merely counting the number of times nouns (such as elephants, sharks, oceans, etc.) co-occur in web documents, we obtain a good estimate of expected co-occurrences in visual data. We then cast the problem of combining textual co-occurrence statistics with the predictions of image-based classifiers as an optimization problem. The resulting optimization problem serves as a surrogate for our inference procedure. Albeit the simplicity of the resulting optimization problem, it is effective in improving both recognition and localization accuracy. Concretely, we observe significant improvements in recognition and localization rates for both Image Net Detection 2012 and Sun 2012 datasets.

I.*Personalized Portraits Ranking:*

Portraits, also known as images of people, constitute an important part of consumer photos. Existing methods manage portraits based on either explicit objectives, e.g., a specified person or event, or aesthetics, i.e., the aesthetic quality of portraits. This paper presents a novel system for personalized portraits ranking. First, four kinds of personalized features, i.e., composition, clothing style, affection and social relationship are proposed to quantify user's intent. Then, example-based and sketch-based user interfaces (UI) are developed, which are capable of capturing user's personal intent hardly described by queries or aesthetics. Finally, portrait ranking is implemented by combing these features together with the developed user interfaces. Experimental results show that the system performs well in providing personalized preferences and the proposed features are effective for portrait ranking. From the user study, our system gets promising results.

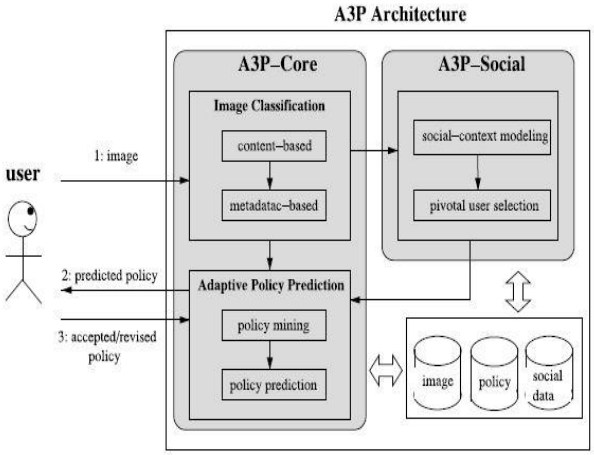
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

VI.SYSTEM ARCHITECHTURE

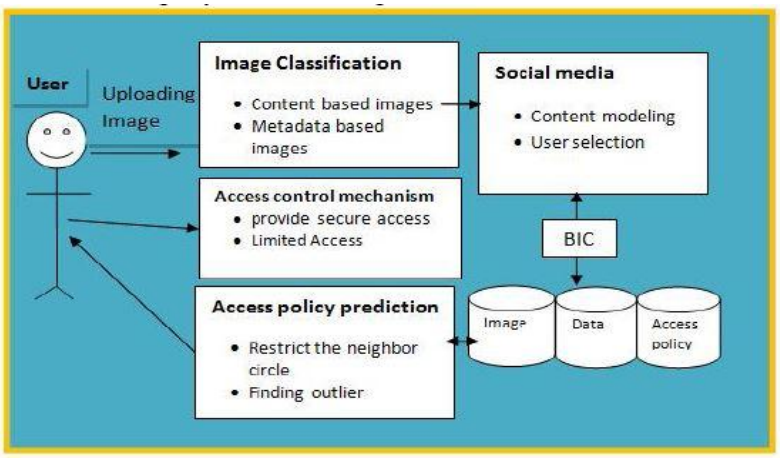


Explanation:
We propose an Adaptive Privacy Policy Prediction (A3P) system which aims to provide users a hassle free privacy settings experience by automatically generating personalized policies.The A3P system handles user uploaded images, and factors in the following criteria that influence one's privacy 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 user's privacy preferences.

- The role of image content and metadata. In general, similar images often incurred similar privacy preferences, especially when people appear in the images.

VII.MODIFIED ARCHITECHTURE



Architecture Explanation:

Some users over CSS influence user's privacy on their private contents, where the users keep on distributing superfluous comments and messages by attractive advantage of the user's intrinsic trust in their connection network. The overall architecture of the proposed work has given in figure 1.0. This paper switches the most widespread issues

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

and threats objective different CSS freshly. In CSS, privacy is frequently a key apprehension by the users. As millions of people are willing to interrelate with others, it is also a new harass ground for image misuses. They are dispersing the images and contents. This paper will demonstrate and argue the most widespread issues and threats targeting different CSS today. Finally, it finds just the right privacy policy scheme for that privacy. This proposes a privacy policy forecast and access boundaries along with overcrowding scheme for social sites using data mining techniques. It helps to detect and defend distrustful activates, which violates user's privacy in CSS by making an allowance for the following parameters: i) Text annotation, which emerge in the uploaded contents; ii) Image and policy descriptions; iii) Detection of superfluous commends and to perform this, the system utilizes APP (Access Policy Prediction) and Access control mechanism by applying BIC algorithm (Bayesian Information Criterion).

VIII.MODULE IMPLEMENTATION

8.1. 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 analysed for 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. 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 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 need 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 were to be introduced, the whole learning model would need to change.

8.2. Image Classification:

To obtain 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. For example, the content-based classification creates two categories: "landscape" and "kid". Images C, D, E and F are included in both categories as they show kids playing outdoors which satisfy the two themes: "landscape" and "kid". These two categories are further divided into subcategories based on tags associated with the images. As a result, we obtain two subcategories under each theme respectively. Notice that image G is not shown in any subcategory as it does not have any tag; image shows up in both subcategories because it has tags indicating both "beach" and "wood".

8.3. Policy Mining:

We propose a hierarchical mining approach for policy mining. Our approach leverages association rule mining techniques to discover 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. 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 in the policies containing both popular subjects and conditions.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

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

IX.CONCLUSION

This paper describes privacy policy techniques for user uploaded data images in various content sharing sites. Based on the user social behaviour and the user uploaded image, the privacy policy can applied. A3P system in used, which provide users easy and properly, configured privacy setting for their uploaded image. By using this we can easily prevent unwanted discloser and privacy violations. Unwanted disclosure may lead to misuse of one's personal information. Users automate the privacy policy settings for their uploaded images with the help of adaptive privacy policy prediction (A3P). Based on the information available for a given user the A3P system provides a comprehensive framework to infer privacy preferences. A3P system is a practical tool. An improvement over current approaches to privacy is offer by A3P.

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.