An Improved Privacy of User Data and Images on Content Sharing Sites

SANDELA MAMATHA[#1] and K.Ranjith Kanna[*2]

[#]*M.Tech Scholar, Department of Computer science and engineering, Ganapathy Engineering College,Warangal., India*

[*] *Assistant Professor, Department of Computer science and engineering, Ganapathy Engineering College, Warangal, 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. This improved technology leads to privacy violation 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 examined as a measure of user's privacy preferences. The Framework determines 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.**

***Index Terms*— Social media, CSS, Privacy Data, Adaptive Privacy Policy Prediction (A3P) and privacy policy**

I. Introduction

Social media is the two way communication in Web 2.0 and it means to communicate, share, and interact with an individual or with a large audience. Social networking websites are the most famous websites on the Internet and millions of people use them every day to engage and connect with other people. Twitter, Facebook, LinkedIn and Google Plus seems to be the most popular Social networking websites on the Internet [1].

Today, for every single piece of content shared on sites like Facebook—every wall post, photo, status update, and video—the up loader must decide which of his friends, group members, and other Facebook users should be able to access the content. As a result, the issue of privacy on sites like Facebook has received significant attention in both the research community [1] and the mainstream media [2]. Our goal is to improve the set of privacy controls and defaults, but we are limited by the fact that there has been no in depth study of users' privacy settings on sites like Facebook. While significant privacy violations and mismatched user expectations are likely to exist, the extent to which such privacy violations occur has yet to be quantified.

Images are now one of the key enablers of users' connectivity. Sharing takes place both among previously established groups of known people or social circles (e.g., Google+, Flickr or Picasa), and also increasingly with people outside the users social circles, for purposes of social discovery to help them identify new peers and learn about peers interests and social surroundings [3]. With the increasing volume of images users share through social sites, 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 access to their shared content is apparent [2].

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 present by their users and have the right to progression collected data and share them to unauthorized [4]. 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 goes across the world [3]. 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. Such an open accessibility of data exposes in SN, the users obtain a number of security and privacy risks. 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 executive of privacy settings for shared content [5]. 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.

Users regularly sharing the data and images in SN by this happening the privacy of the images may lock with the un-wanted parties. Hackers can chop the images through these social media so the privacy of the user images may loss. Today, for every single quantity of content sharing sites like Facebook—every wall post, photo, status update, and video—the up loader must settle on which of his friends, group members, and other Facebook users should be intelligent to access the content. As a result, the problem of isolation on sites like Facebook has received significant concentration in both the research society [3] and the mainstream media. Our goal is to improve the set of privacy controls and defaults, but we are restricted by the reality that there has been no in-depth study of users' privacy settings on sites like Facebook. While significant privacy disobedience and mismatched user expectations are likely to exist, the extent to which such privacy disobedience arises has yet to be quantified [4].

A. Contribution and plan of 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.

II. LITERATURE SURVEY

Some previous systems shows different studies on automatically assign the privacy settings. One such system which Bonneau et al.[2] proposed shows the concept of privacy suites. The privacy 'suites' recommends the user's privacy setting with the help of expert users. 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. [6] Prediction of the users privacy preferences for location-based data (i.e., share the location or no) was studied by Ravichandran et. Al [2]. 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.

Peter F. Klemperer [5] developed a tag based access control of data shared in the social media sites. An approach that produces access-control policies from photo management tags. Every photo is included with an access network for mapping the photo with the participant's friends. The contributor can choose apposite preference and access the data. Photo tags can be classified as managerial or unrestrained based on the user needs. There are several significant limitations to our study design. First, our outcomes are limited by the participants we conscript and the photos they offered. A second set of limitations apprehension our use of machine generated access-control rules. The algorithm has no admittance to the context and significance of tags and no approaching into the policy the contestant proposed when tagging for access control. As an outcome, some rules become visible strange or random to the contributor, potentially pouring them in the direction of explicit policy-based tags like "private" and "public.

Fabeah Adu-Oppong developed the privacy settings depends on the model of social circles [6]. It facilitates a web based explanation to defend personal information. The technique named Social Circles Finder; automatically construct the friend's list. It is a process that studies the social circle of a person and categorizes the concentration of relationship and as a result social circles offer a meaningful labelling of friends for surroundings privacy policies. The relevance will recognize the social circles of the subject but not show them to the subject. The subject will then be asked questions about their motivation to share a piece of their individual information. Based on the respond the function finds the visual graph of users.

SergejZerr proposes a approach Privacy-Aware Image Classification and Search [7] to robotically detect private images, and to facilitate privacy-oriented image search. It coalesce textual meta data images with assortment of visual features to facilitates security strategy. In this the chosen image features (edges, faces, color histograms) which can help differentiate between natural and man-made objects/prospect (the EDCV feature) that can indicate the existence or absence of meticulous objects (SIFT). It uses different classification models qualified on a large scale dataset with isolation assignments achieved through a social explanation game.

Anna Cinzia Squicciarini developed an Adaptive Privacy Policy Prediction (A3P) [8] system, a free privacy settings system by robotically produces personalized policies. The A3P system levers user uploaded images based on the person's individual characteristics and images content and metadata. The A3P system consists of two components: A3P Core and A3P Social. When a user uploads a data like image, the image will be first sent to the A3P-core. The A3Pcore organizes the image and resolves whether there is a need to appeal to the A3P-social. The disadvantage is mistaken privacy policy production in case of the lack of Meta data information about the images. Also guide creation of Meta data log data information direct to imprecise classification and also contravention privacy.

K. Strater et al [9] had discussed about the Strategies and struggles with privacy in an online social networking community. These sites have changed how many people develop and maintain relationships through posting and sharing personal information. The amount and depth of these personal disclosures have raised concerns regarding online privacy. We expand upon previous research on users' under-utilization of available privacy options by examining users' current strategies for maintaining their privacy, and where those strategies fail, on the online social network site Facebook. Our results demonstrate the need for mechanisms that provide awareness of the privacy impact of users' daily interactions.

In the past years an incredible growth on Online Social Networks [1,9] like Facebook, Orkut and Twitter is seen. These OSNs not only propose gorgeous means for virtual social communications and data sharing, but also elevate a number of security issues. Although OSNs allow a single user to admission to her or his data, they presently do not provide any device to implement privacy protection over data connected with large number of users, departure privacy contravention largely unanswered and leading to the probable confession of information that at least one user proposed to keep private. This paper analyses an assortment of privacy and security issues in OSNs. OSNs come across different types of attacks such a fake identity, Sybil harass, uniqueness clone attacks, The main aim is to augment the privacy and security in OSNs which is one of the Quality of Service (QoS) issues and thus declining the attacks and problems. This paper is a survey which is more detailed to representation the various attacks and privacy models in OSNs with deference to augmentation of security and privacy [10].

III. 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 [11].

A. Issues in Existing System

1. Sharing images within online content sharing sites, therefore, may quickly lead to unwanted disclosure and privacy violations.
2. Further, the persistent nature of online media makes it possible for other users to collect rich aggregated information about the owner of the published content and the subjects in the published content.
3. The aggregated information can result in unexpected exposure of one’s social environment and lead to abuse of one’s personal information.

IV. PROPOSED SYSTEM

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.

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.

A. Advantages of proposed system

- The A3P-core focuses on analyzing each individual user’s own images and metadata, while the A3P-Social offers a community perspective of privacy setting recommendations for a user’s potential privacy improvement.
- We design the interaction flows between the two building blocks to balance the benefits from meeting personal characteristics and obtaining community advice.

V. 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 A3P-social: (i) The user does not have enough data for the type of the uploaded image to conduct policy prediction;

(ii) The A3Pcore 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 new friends, 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.

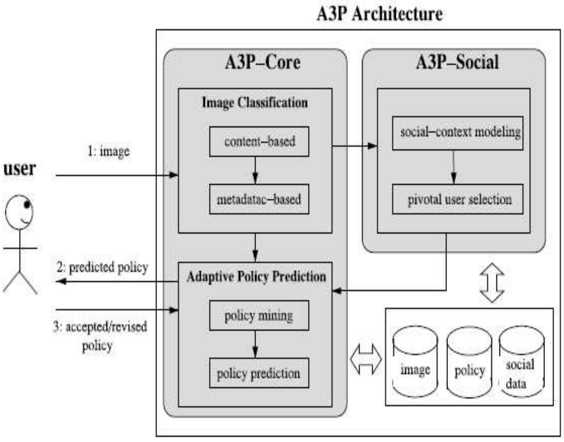


Fig. 1 System Architecture

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

VI. IMPLEMENTATION

The proposed system of this paper is divided into four major modules and described as below.

- System Construction Module
- Content-Based Classification
- Metadata-Based Classification
- Adaptive Policy Prediction

A. MODULES DESCRIPTION

1) System Construction Module

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) 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 new friends, new posts on one's profile etc).

2) Content-Based 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.

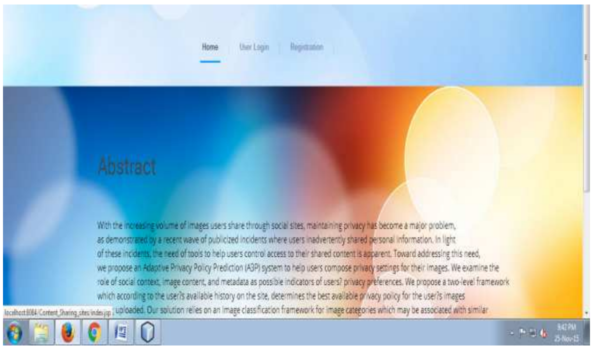
Our approach to content-based classification is based on an efficient and yet accurate image similarity approach. Specifically, our classification algorithm compares image signatures defined based on quantified and sanitized version of Haar wavelet transformation. For each image, the wavelet transform encodes frequency and spatial information related to image color, size, invariant transform, shape, texture, symmetry, etc. Then, a small number of coefficients are selected to form the signature of the image. The content similarity among images is then determined by the distance among their image signatures.

3) Metadata-Based Classification

The metadata-based classification groups images into subcategories under aforementioned baseline categories. The process consists of three main steps. The first step is to extract keywords from the metadata associated with an image. The metadata considered in our work are tags, captions, and comments. The second step is to derive a representative hypernym (denoted as h) from each metadata vector. The third step is to find a subcategory that an image belongs to. This is an incremental procedure. At the beginning, the first image forms a subcategory as itself and the representative hypernyms of the image becomes the subcategory's representative hypernyms.

4) Adaptive Policy Prediction

The policy prediction algorithm provides a predicted policy of a newly uploaded image to the user for his/her reference. More importantly, the predicted policy will reflect the possible changes of a user's privacy concerns. The prediction process consists of three main phases: (i) policy normalization; (ii) policy mining; and (iii) policy prediction. The policy normalization is a simple decomposition process to convert a user policy into a set of atomic rules in which the data (D) component is a single-element set.



Abstract

With the increasing volume of images users share through social sites, 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 access to their shared content is apparent. 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 uploaded. Our solution relies on an image classification framework for image categories which may be associated with similar

Fig. 2 Welcome Page

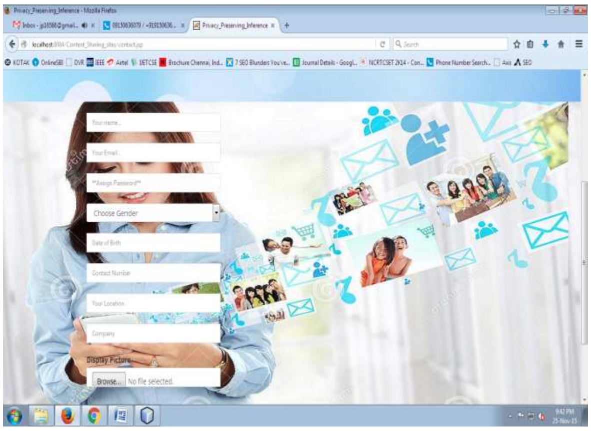
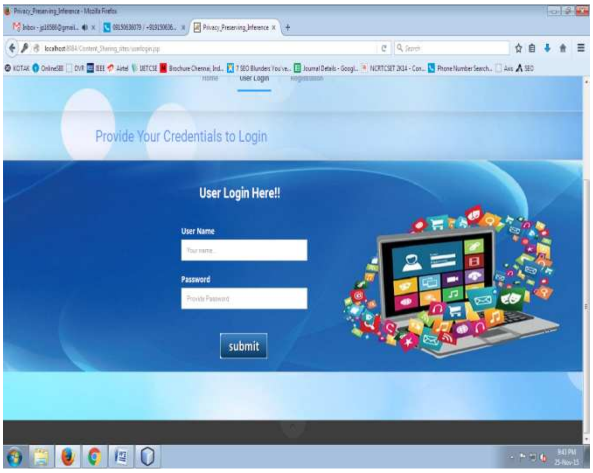

Privacy_Preserving_Inference - Mozilla Firefox
localhost:8084/Content_Sharing_sites/contact.jsp
Your name..
Your Email..
Assign Password
Choose Gender
Date of Birth
Contact Number
Your Location
Company
Display Picture
Browse... No file selected.

Fig.3 Registration form

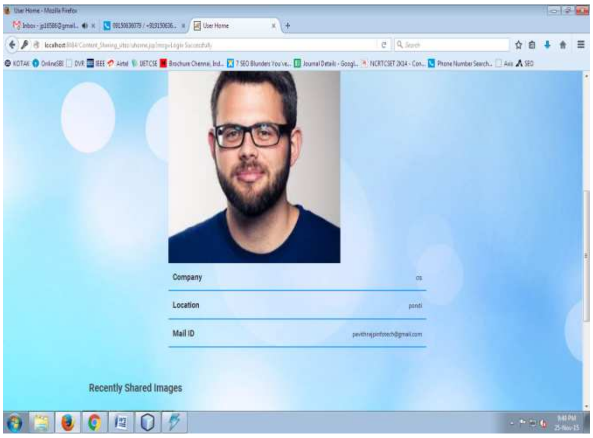

Privacy_Preserving_Inference - Mozilla Firefox
localhost:8084/Content_Sharing_sites/userlogin.jsp
Home
User Login
Registration

Provide Your Credentials to Login

User Login Here!!

User Name
Your name..
Password
Provide Password
submit

Fig.4 User Login


User Home - Mozilla Firefox
localhost:8084/Content_Sharing_sites/uhome.jsp?msg=Login Successfully

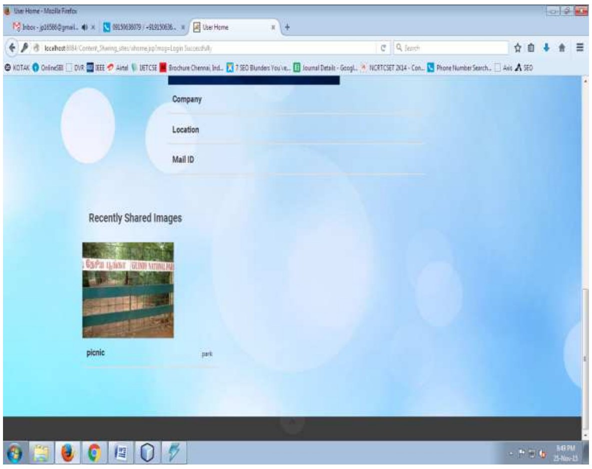
Company cs

Location pondi

Mail ID pavithrajpinfotech@gmail.com

Recently Shared Images

Fig .5 User details


User Home - Mozilla Firefox
localhost:8084/Content_Sharing_sites/uhome.jsp?msg=Login Successfully

Company

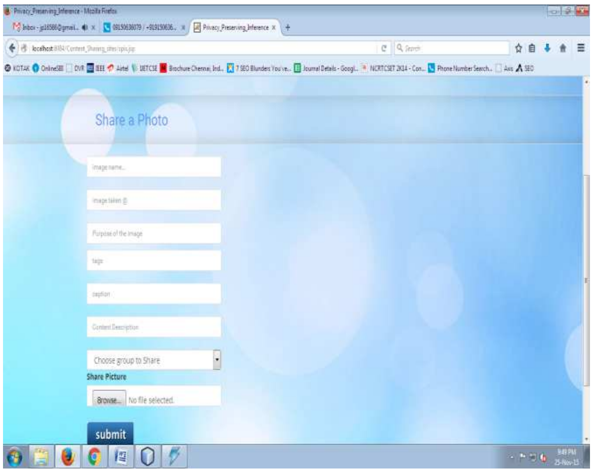
Location

Mail ID

Recently Shared Images

picnic park

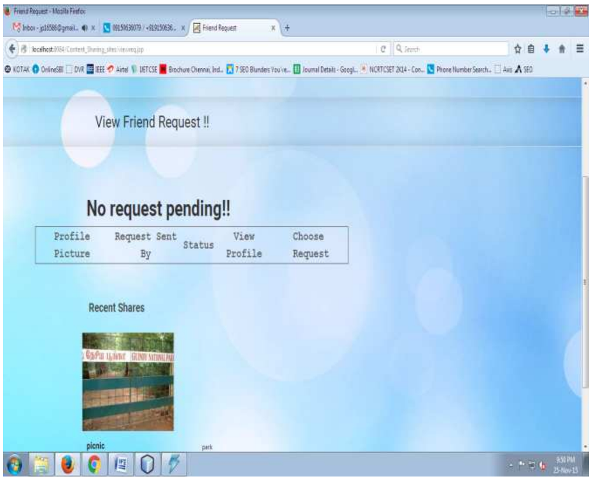
Fig .6 Image upload by User


Privacy_Preserving_Inference - Mozilla Firefox
localhost:8084/Content_Sharing_sites/upic.jsp

Share a Photo

image name..
image taken @
Purpose of the image
tags
caption
Content Description
Choose group to Share
Share Picture
Browse... No file selected.
submit

Fig. 7 Sharing a photo


Friend Request - Mozilla Firefox
localhost:8084/Content_Sharing_sites/viewreq.jsp

View Friend Request !!

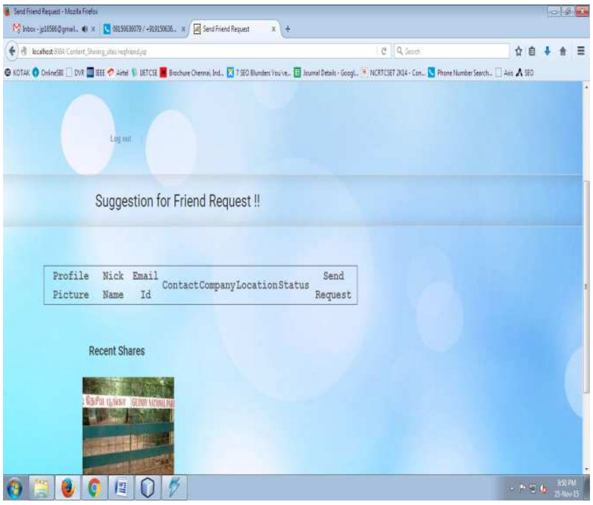
No request pending!!

Profile Picture	Request Sent By	Status	View Profile	Choose Request

Recent Shares

picnic park

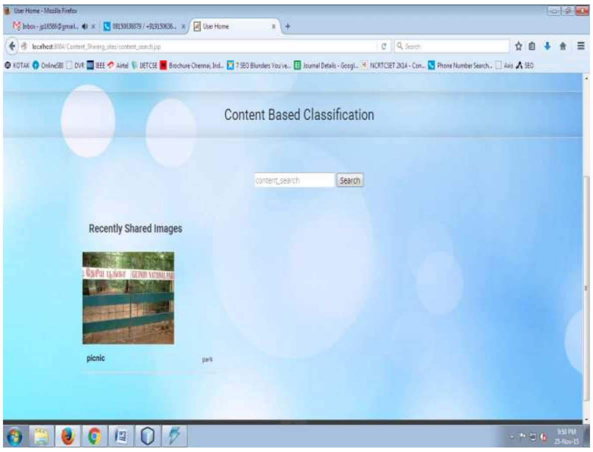
Fig. 8 View Friend Request



Suggestion for Friend Request !!

Profile Picture	Nick Name	Email Id	Contact	Company	Location	Status	Send Request

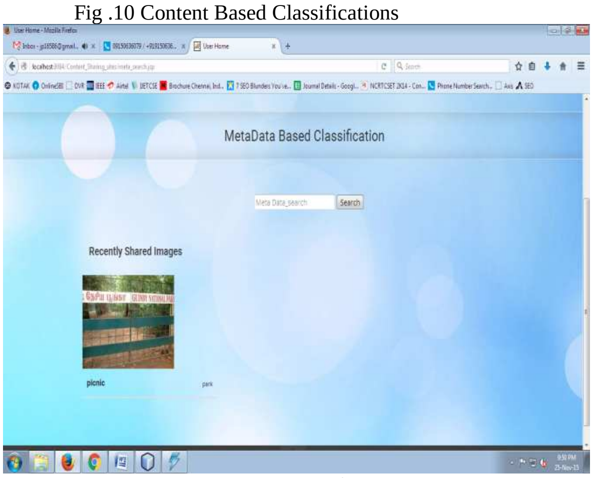
Fig .9 Suggestions for friend request



Content Based Classification

Recently Shared Images

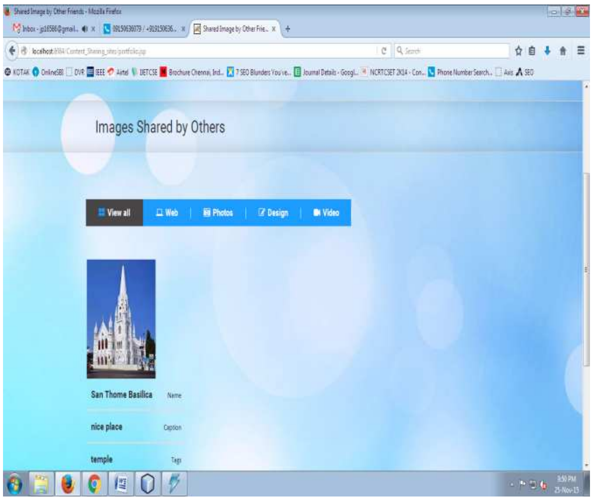
Fig .10 Content Based Classifications



MetaData Based Classification

Recently Shared Images

Fig .11 Metadata Based Classification



Images Shared by Others

San Thome Basilica Name

nice place Caption

temple Tags

Fig.12 Images shared by others

VII. CONCLUSION

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 issues, we 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.

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 Informatica Teorica 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.