Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Identifying hidden social circles for advanced privacy configuration

Anna Squicciarini [a,*], Sushama Karumanchi [a], Dan Lin [b], Nicole DeSisto [a]

[a] Information Sciences & Technology, Pennsylvania State University, USA
[b] Department of Computer Science, Missouri University of Science & Technology, USA

ARTICLE INFO

Article history:
Received 8 May 2013
Received in revised form
27 July 2013
Accepted 29 July 2013

Keywords:
Privacy
Social groups
Access control
Social networks
Cliques
Policies

ABSTRACT

With the dramatic increase of users on social network websites, the needs to assist users to manage their large number of contacts as well as providing privacy protection become more and more evident. Unfortunately, limited tools are available to address such needs and reduce users' workload on managing their social relationships. To tackle this issue, we propose an approach to facilitate online social network users to group their contacts into social circles with common interests. Further, we leverage the social group practice to automate the privacy setting process for users who add new contacts or upload new data items. We evaluate our approach using real-world data collected through a user study. The study also includes an analysis of the properties that are most critical for privacy related decisions.

1. Introduction

Social networking sites are proliferating fast with an increasing number of users and increasingly complicated social relationships among users. Micro-managing this large amount of personal data has shown to be a very burdensome task for many users, as acknowledged by a growing number of research studies and news articles (Acquisti and Gross, 2006; Blog, 2008; Chronicle, 2008; Hogben, 2007; Kandias et al., 2013a, 2013b; Irvine, April 2008; Zheleva and Getoor, 2009; Ziegele and Quiring, 2011). It is even more challenging to configure proper privacy settings for data being shared in social networking sites. Security unconscious users typically follow an open and permissive default policy. As a result, the potential for unwanted information leakage is great.

In this paper, we tackle the above problems by introducing an approach to assist users in managing their social relationships as social groups. Social groups denote circle of users with similar characteristics, be these related to their interests, locations, or other demographics. We leverage this notion of social groups to provide privacy setting recommendation for users.

Our approach builds on the following rationale. As confirmed by the most recent social network platforms, social circles in modern social networks can act as the foundation of user and privacy management. For instance, Facebook provides an optional mechanism that allows users to create custom lists to organize friends and set privacy restrictions accordingly. Facebook also recently announced smart lists which automatically group friends who live near by or attend the same school. Similarly, the newly released Google+ creates four default circles for users: Friends, Family, Acquaintances, and Following. A user can remove/rename any of the default circles or add

[*] Corresponding author.

E-mail addresses: asquicciarini@ist.psu.edu (A. Squicciarini), sik5273@ist.psu.edu (S. Karumanchi), lindan@mst.edu (D. Lin), ndesi2@brockport.edu (N. DeSisto).

0167-4048/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cose.2013.07.007

new circles. For privacy management, users in Google+ can selectively share information with a specific set of their circles, all their circles, their extended circles or with the public (everyone).

While the idea of social circles is very interesting and promising (Javed and Shehab, 2012), existing social network platforms have not fully explored the full benefits of this concept and their related systems are at primitive stage with no or limited support on circle formation and privacy management. As an advancement in this direction, we design a multi-criteria model that takes into account multiple aspects of users' profiles, and automatically groups each user's contacts into social circles with common characteristics. Users in the same social circle (group) have similar behavior, such as similar education background, hobbies, and similar privacy preferences. We consider a possibly large number of unique characteristics that can serve as criteria for group identification, taking into account the importance of certain demographics (such as privacy preferences or interests) for privacy decision making. Given the obtained grouping information, we further propose an approach to recommend privacy policies for newly uploaded data items or newly added contacts. In particular, when a user uploads an object (a data item or a contact), our system selects the social group which is most likely to deal with the object in the similar way as the user, and then a carefully extracted average of the privacy settings adopted by the selected group is considered as the base for predicting policies for the new object. Fig. 1 gives an overview of our approach.

Notice that our approach has the ability to identify hidden groups which may play an important role for privacy management, but may not be explicitly considered by users. For example, within a user's friend list, there may exist a subgroup which includes mainly close friends with whom the user shares a large amount of data; in a user's family group, there may be direct family members with whom the user shares family pictures, events (e.g., anniversary, notes). Privacy settings are likely to be different in each such hidden sub-group, and hence identifying these sub-groups will help enhance and simplify users' privacy practices. Our approach's effectiveness has been tested via an empirical evaluation on data collected from real-world users.

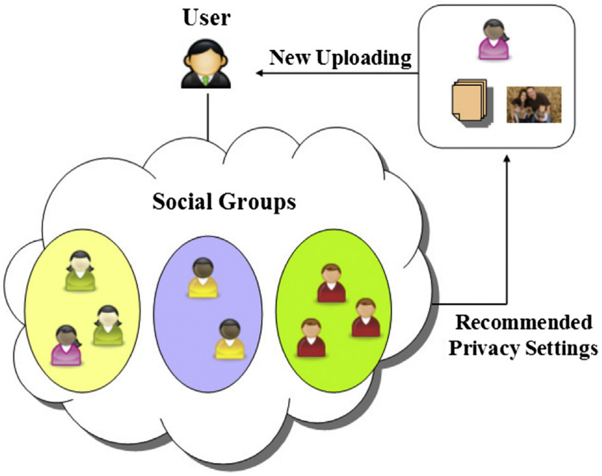

Fig. 1 – Policy recommendation using social groups.

A preliminary version of this work is (Squicciarini et al., 2012). In this paper, we make the following new contributions. First, we propose an incremental social group maintenance algorithm which provides a solution to the problem of user feature or property updates. Second, we increase the number of features used to group the users, and check the prediction accuracy. Further, we perform several additional experiments, that include feature analysis where in we find out which features of the users are the most critical or useful in predicting the policies for the content of a user.

The rest of the paper is organized as follows. Section 2 discusses related works with respect to social circles and associated privacy management issues. Section 3 introduces notations and defines the problem. Section 4 presents the detailed algorithms for social grouping, followed by Section 5 which leverages the grouping information for policy prediction. Then, Section 6 reports experimental results. Finally, Section 7 concludes the paper.

2. Related work

Work on social networking privacy enhancing technologies is nowadays proliferating. In particular, several recent works have studied how to automate the task of privacy settings (Alessandra Mazzia and LeFevre, 2011; Baden et al., 2009; Bonneau et al., 2009a, 2009b; Ravichandran et al., 2009; Maximilien et al., 2009).

Bonneau et al. (2009a) proposed the concept of privacy suites for social network sites, based on the idea that most users currently stick with default privacy settings. In particular, they recommend to users a suite of privacy settings that expert users or other trusted friends have already set, so that normal users can either directly choose a setting or only need to do minor modification to available settings. Along similar lines, Danezis (Bonneau et al., 2009b) proposed a machine-learning based approach to automatically extract privacy settings from the social context within which the data is produced. Parallel to the work of Danezis, Adu-Oppong et al. (Fabeah Adu-Oppong et al., 2008) develop privacy settings based on a concept of social circles which consist of clusters of friends formed by partitioning users friend lists. Ravichandran et al. (2009) studied how to predict a user privacy preferences for location-based data (i.e., share her location or not) based on location and time of day.

In general, semi-supervised learning has been used in social networks to infer users' private information from the public labeled and unlabeled data using graph based semi-supervised learning, e.g. (Javed and Shehab, 2012). For instance, Fang et al. (Fang and LeFevre, 2010) proposed a privacy wizard to help users grant privileges to their friends. The wizard asks users to first assign privacy labels to selected friends, and then uses this as input to construct a classifier which classifies friends based on their profiles and automatically assign privacy labels to the unlabeled friends. Subsequently, the same research group (Alessandra Mazzia and LeFevre, 2011) introduced a policy visualization tool which displays privacy settings for user specific subgroups of friends