Computer Communications 41 (2014) 11–21

Contents lists available at ScienceDirect

Computer Communications

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

Using community structure to control information sharing in online social networks

Amin Ranjbar *, Muthucumaru Maheswaran

Advanced Networking Research Lab, School of Computer Science, McGill University, Montreal, QC H3A 2A7, Canada

ARTICLE INFO

Article history:
Received 2 December 2011
Received in revised form 7 January 2014
Accepted 8 January 2014
Available online 18 January 2014

Keywords:
Online social networks
Confidentiality
Information sharing
Access control
Information leakage

ABSTRACT

The dominant role of social networking in the web is turning human relations into conduits of information flow. This means that the way information spreads on the web is determined to a large extent by human decisions. Consequently, information security lies on the quality of the collective decisions made by the users. Recently, many access control schemes have been proposed to control unauthorized propagation of information in online social networks; however, there is still a need for mechanisms to evaluate the risk of information leakage within social networks. In this paper, we present a novel community-centric confidentiality control mechanism for information flow management on the web. We use a Monte Carlo based algorithm to determine the potential spread of a shared data object and to inform the user of the risk of information leakage associated with different sharing decisions she can make in a social network. By using the information provided by our algorithm, the user can curtail sharing decisions to reduce the risk of information leakage. Alternatively, our algorithm can provide input for a fully- or semi-automatic sharing decision maker that will determine the outcomes of sharing requests. Our scheme also provides a facility to reduce information flowing to a specific user (i.e., black listing a specific user). We used datasets from Facebook and Flickr to evaluate the performance of the proposed algorithms under different sharing conditions. The simulation results indicate that our algorithm can effectively control information sharing to reduce the risk of information leakage.

1. Introduction

As online social networks (OSNs) increase in size and more people use them as their primary Internet portal, the volume of information shared in OSNs keeps on growing. Information is created by different sources in OSNs including people posting information in their profile pages, relational information generated by people initiating connections among themselves, and data feeds generated by sensing people's activities such as gaming and purchasing. In any sharing activity, OSNs store and process different pieces of information: picture files, relationships among people, and sharing preferences regarding data objects. This means that the way OSNs are architected and the security primitives built into them play key roles in defining information security in social web. Consequently, many research thrusts have examined wide-ranging security issues in the context of OSNs [1–6].

While information sharing is vital for socializing online, many security and privacy issues have been raised such as confidentiality and integrity violations of shared data objects. The main issue is to ensure users that their privacy and access control requirements are preserved when information sharing occurs within OSNs. Recently, users in OSNs have started to become more aware of the risk of unauthorized propagation of their information through social networking sites. To partially answer users concern, several topology-based access control mechanisms for OSNs were proposed in order to identify authorized users by specifying some constraints on the social graph [2,3,7–10]. In these schemes, to regulate information sharing, access control rules are defined by identifying the relationships that users must have in order to access the shared data.

Because existing techniques only deal with information release, a user might not be able to precisely identify who is authorized to have access to her data. Even in small social networks, it is difficult for a user to keep track of the topology of her constantly changing social network and to identify users who are actually authorized even with simple access rules such as "friends-of-friends". In addition, the user requirements for privacy are constantly changing [11]. This may potentially lead users to losing control of their shared data and to generating risks of unauthorized information dissemination related to their access control decisions. A user does not have complete knowledge about a set of users authorized by her access control settings and their potential malicious behaviors

* Corresponding author. Tel.: +1 (514) 398 1465; fax: +1 (514) 398 3883.
E-mail addresses: amin.ranjbar@mail.mcgill.ca (A. Ranjbar), maheswar@cs.mcgill.ca (M. Maheswaran).

0140-3664/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comcom.2014.01.002

in releasing her data to unauthorized users [12]. Therefore, it is necessary to have new access control mechanisms in OSNs in order to evaluate the potential risks and to make users fully aware of the possible consequences of their decisions in specifying access rules.

There are two main challenges found in defining new access control techniques and controlling confidentiality of information on OSNs. The primary challenge is about enforcing usage conditions. The typical corporate information networks [13] such as a course management network in a university use predefined roles (e.g. professor, teaching assistant, and student) and policies to regulate information flow. For example, a student does not have access to assignments or exams of other students, while students appointed as teaching assistants have access to all exams and assignments of specified courses. The employment conditions of the teaching assistants require them to keep certain information confidential. Information flow control in such a network breaks down if the users fail to abide by the usage conditions. The sharing problem in OSNs, however, is not governed by precise usage policies. The second challenge is that information sharing in OSNs is not automatically coupled with the level or the direction of interactions. In analog social networks [14], physical contacts remain as the dominant mechanism for sharing information between users. Therefore, people can implicitly control information sharing by avoiding contact with undesirable friends. The explicit controls are necessary in OSNs to avoid information sharing with undesirable friends. Therefore, there is a need for novel access control techniques that work with minimal user intervention.

In this paper, we address the two challenges identified above by presenting a novel community-centric confidentiality control scheme for OSNs. We develop a new strategy where the eventual information distribution is shaped by the initial release of objects into the network. Because the initial release is completely controlled by the owner, she could shape the information distribution by making appropriate release decisions to minimize possible information leakage. To the best of our knowledge, this is the first work proposing an algorithm to closely approximate the exact risk of information leakage associated with user access control decisions. Our scheme uses a Monte Carlo method to compute the set of potential users who could receive the data objects belonging to a data owner. Our algorithm can provide input to a fully- or semi- automatic sharing decision maker that will determine the consequences of accepting or rejecting sharing requests. In addition, we provide algorithms for preventing information from reaching certain users by shaping the initial release set. Using datasets from Facebook and Flickr, we simulate sharing situations in social settings and estimate information leakage values considering that our algorithm controls information sharing.

Section 2 discusses the secure information sharing problem and associated security challenges in social networking. In Section 3, a new community-centric confidentiality control mechanism for online systems is presented. Section 4 presents an analysis of the experiments performed on information sharing patterns in Facebook and Flickr. Related work is described in Section 5.

2. Information sharing model

One of the important characteristics of OSNs is the private information space it provides for the users joining a network. After joining, users, at their discretion provide access to their friends using simple mechanisms provided by the OSN operators. Most OSN operators provide facilities to restrict access to subset of friends, friends, friend-of-friends, or public. These controls only deal with information release and expect the user to detect any misuse and modify the release conditions (for example, block an offending user from accessing data) [15,16].

Information sharing in OSNs takes place without any formal usage restrictions or guidelines, which is an acceptable and preferable approach for OSNs users as shown in Fig. 1. This survey was conducted to find the value of information sharing on OSNs among McGill University students from various fields of study. Only 24 percent of participants like explicit sharing conditions when they receive data from their friends whereas majority of the users prefer to attach specific constraints when they provide the information. This makes policy-based access control less suitable for OSNs sharing situations. Because information sharing is not carried out under strict usage conditions in the social networks, information leakage can occur widely. If an unauthorized user has access to the shared data, that object is said to be *leaked* or that *information leakage* occurred. Therefore, it is necessary to be able to compute the risk of information being leaked. Here, we consider two different use-cases of computing the risk of unauthorized information leakage:

First, if a user needs to share some information with a subset of her friends (S_1), she should simply set the access control to S_1. However, there is a risk of information leakage associated to her decision for which she does not possess any knowledge. If there are intense and frequent interactions between S_1 and another subset of the user's friends named S_2, then the chances that the information shared with S_1 will leak to the members of S_2 is quite high. Therefore, there should be a mechanism to compute the risk of information leakage related to the user's sharing decisions and to provide the subset of the user's friends who will eventually have access to the shared information. Based on this information, the user can shape her access control decisions properly.

Second, if a user attempts to black list a specific user (her adversary), the only thing she can do in existing OSNs is to add the

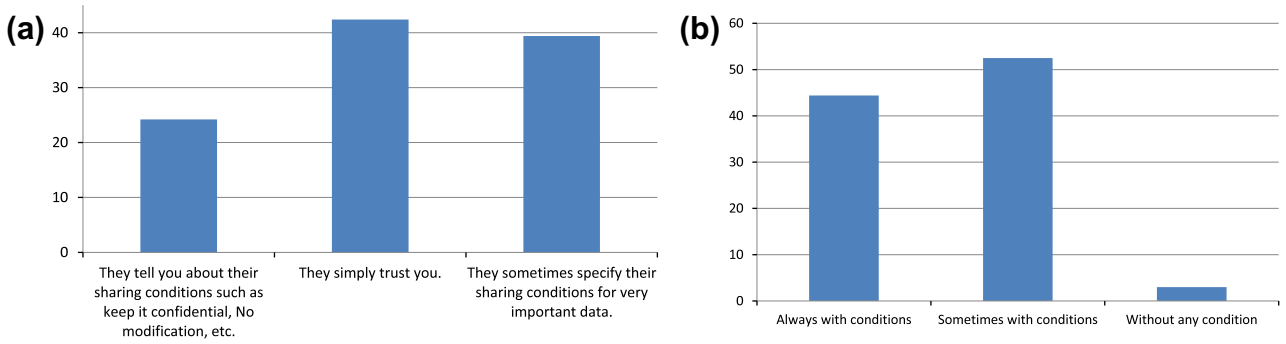

Fig. 1. Survey of information sharing on OSNs: (a) When friends share data with you, do you prefer when; (b) would you like to share information under specific conditions or under no conditions?