Computer Communications 49 (2014) 18–32

Contents lists available at ScienceDirect

Computer Communications

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

# Modeling online social network users' profile attribute disclosure behavior from a game theoretic perspective

Jundong Chen *, Ankunda R. Kiremire, Matthias R. Brust, Vir V. Phoha

*Center for Secure Cyberspace, Louisiana Tech University, Ruston, LA 71272, USA*

ARTICLE INFO

*Article history:*
Received 4 October 2013
Received in revised form 30 April 2014
Accepted 2 May 2014
Available online 14 May 2014

*Keywords:*
Game theory
Social network
Privacy settings
Nash equilibrium
Replicator dynamics

ABSTRACT

Privacy settings are a crucial part of any online social network as users are confronted with determining which and how many profile attributes to disclose. Revealing more attributes increases users' chances of finding friends and yet leaves users more vulnerable to dangers such as identity theft. In this paper, we consider the problem of finding the optimal strategy for the disclosure of user attributes in social networks from a game-theoretic perspective.

We model the privacy settings' dynamics of social networks with three game-theoretic approaches. In a *two-user* game, each user selects an ideal number of attributes to disclose to each other according to a utility function. We extend this model with a basic *evolutionary* game to observe how much of their profiles users are comfortable with revealing, and how this changes over time. We then consider a weighted evolutionary game to investigate the influence of attribute importance and the network topology in selecting privacy settings.

The two-user game results show how one user's privacy settings are influenced by the settings of another user. The basic evolutionary game results show that the higher the motivation to reveal attributes, the longer users take to stabilize their privacy settings. Results from the weighted evolutionary game show that users are more likely to reveal their most important attributes than their least important attributes regardless of the risk. Results also show that the network topology has a considerable effect on the privacy in a risk-included environment but limited effect in a risk-free environment.

Motivation and risk are identified as important factors in determining how efficiently stability of privacy settings is achieved and what settings users will adopt given different parameters. Additionally, the privacy settings are affected by the network topology and the importance users attach to specific attributes. Our models indicate that users of social networks eventually adopt profile settings that provide the highest possible privacy if there is any risk, despite how high the motivation to reveal attributes is. The provided models and the gained results are particularly important to social network designers and providers because they enable us to understand the influence of different factors on users' privacy choices.



## 1. Introduction

Concerns regarding the privacy in social networks have received worldwide attention and led to frequent public debates [1,2]. Social networks contain large amounts of information that can be used to uniquely identify their users as well as provide information on their habits, interests, and history [3]. On the positive side, this information enables the users to identify potential new "friends" and find old friends [4]. However, revealing information also makes it accessible to potential criminals leaving the users vulnerable to dangers such as identity theft, sexual predators, stalkers, and inference by defrauders [5]. The risk to user privacy has caused so much concern that over 60% of social network users employ privacy increasing measures such as deleting friends and concealing profile attributes from other social network users [6]. The benefits and risks create a dilemma that every user of a social network faces: reveal more attributes to attract more friends, or reveal less attributes to become less vulnerable.

In this paper, we propose three game-theoretic models to study the dynamics of privacy settings between users in a social network. These models include a two-user model and two evolutionary game models and are built on a novel analytical definition of *risk* and *motivation* in a social network.

* Corresponding author.
*E-mail addresses:* jdc074@latech.edu (J. Chen), ark010@latech.edu (A.R. Kiremire), mbrust@latech.edu (M.R. Brust), phoha@latech.edu (V.V. Phoha).

http://dx.doi.org/10.1016/j.comcom.2014.05.001
0140-3664/

The *two-user game* models the interactions between two users in both risk-free and risk-included scenarios. We use this model to understand how the privacy choices of one user affect the privacy choices of another user. For example, given a network in which Alice and Bob are "friends" with an identical number of profile attributes, the two user game investigates whether a strategy by Alice to withhold 30% of her attributes would make Bob withhold or reveal more of his attributes.

The *evolutionary game* is an extension of the two-user game to model the interactions of multiple users over time with the utility function of the evolutionary game derived from the utility function of the two-user game. In the basic evolutionary game, all the users are allowed to change their strategy over time in order to maximize the benefit of friendship establishment and minimize the risk to their privacy. The users' choice of strategy at any point in time is dependent on the strategies currently employed by all the users in the network. Informally, given that Alice is part of a large social network, the evolutionary game investigates whether she would change her decision to withhold 30% of her attributes if she knew that 60% of the network's users were revealing all their attributes. The evolutionary game also investigates whether and how her new privacy strategy would affect the rest of the network users. This iterative process is repeated until the entire population reaches an equilibrium state. The equilibrium states as well as the dynamics of the network provide insights into understanding the privacy preferences of social network users.

The weighted model considers different types of networks and different types (weights) of attributes in order to investigate two concepts. First, it investigates what influence, if any, the type of network has on the privacy strategy of the users of the network given the benefit of friendship enhancement. The network types considered include random networks, scale-free networks, and small-world networks to model different social network properties [7]. For example, if Alice decided to reveal 30% of her attributes in a social network exhibiting small-world characteristics, would she make the same decision if the network exhibited random graph characteristics instead? Second, this model investigates whether the importance of the revealed and hidden attributes plays a role in the decision. By weighting the attributes, this model considers the possibility that some attributes have a higher impact than others in either self-disclosure or privacy. This model investigates whether Alice revealing attributes such as her religion and sexual preferences would affect the network more than her revealing that she likes playing soccer and watching movies.

As results, we present the Nash equilibria for the proposed two-user game model [8] as well as the population dynamics for the evolutionary models. In our models, the Nash equilibrium refers to the optimal strategies taken by the users of the network. The strategies are optimal because the users cannot achieve a higher benefit by unilaterally changing their strategy. We also present the population dynamics for the evolutionary game showing the popularity of different strategies as different users change their privacy over time.

For the two-user game, we find that the pure strategy is for at least one of the players to disclose no attributes at all if there is an element of risk. Surprisingly, removing the risk element does not mean that all players will disclose all their attributes.

For the basic evolutionary game, we discover that the dominant strategy is to disclose no attributes if there is an element of risk. By dominant strategy, we refer to the strategy employed by most of the users in the social network. On the other hand, if the risk factor is ignored, the dominant strategy is to disclose all attributes. Revealing all but one attributes is also a common initial strategy in a risk-free network. Additionally, we find that networks where the risk factor is considered achieve equilibria faster than networks where risk was ignored. Our results indicate that users in a network will eventually select the highest possible privacy settings regardless of the motivation and benefits of less private settings as long as there is an element of risk in the social network.

Using the weighted evolutionary game model, we observe a tendency by users to reveal their most important attributes more than their least important attributes. By important attributes, we refer to those attributes which have a larger impact on the social capital of a user. Additionally, users in random and scale-free networks are more likely to reveal their attributes than users in small-world networks. Interestingly, we find that the type of network topology has a limited effect on privacy settings of a social network in the risk-free case and yet have a considerable effect on the privacy in the risk-included scenario.

While some of the results that we present, such as the final stable states of the models, might seem intuitive, the manner of the transition of the network towards this equilibrium is not obvious, and yet equally important because it allows us to see how the privacy settings in the network evolve. We use the models to investigate the influence of certain factors on the stability of a social network and the privacy strategies employed by its users. Additionally, our models can be used to understand and predict the dynamics of a social network based on attribute disclosure. This is particularly important to social network providers, designers, and users in determining how to maximize the self-disclosure in a network while keeping the privacy risk under a certain threshold.

The remainder of this paper is as follows. We discuss related work in the next section and specify the definition and strategies used in our model in Section 3.1. Our game-theoretic models are described in Section 3.2 and analyzed in Section 4. We then present the results and highlight the significance of our approach in Section 5 and conclude this paper with a discussion of our findings in Section 6.

## 2. Related work

A considerable amount of research has been done in understanding online social networks and the factors that contribute towards their success. Online social networks are built on the concept of *self-disclosure* [9], which is positively affected by factors like *relationship-building* and *platform enjoyment*. In contrast, perceived privacy risk is a factor with a negative effect on self-disclosure [9]. The benefit of relationship-building is linked to the number of friends a user stands to gain by disclosing personal information. The link between number of potential friends and revealed information is based on the *homophily* principle more commonly expressed as "birds of a feather flock together" [10]. In the context of a social network, this principle translates to users with similar attributes being more likely to establish a friendship [10,11]. On top of the similarity in attributes, the number of revealed attributes also positively affects relationship-building. Lampe et al. [12] find that the number of friends that a user has is exponentially related to the size of the set of attributes that user reveals. This is because sharing more profile attributes allows more users to establish common ground that promotes interaction and encourages "friendship" [13]. However, *profile disclosing* increases the privacy risk to social network users [9]. Profile disclosing is defined as the amount of a user's profile that is visible to a third party [9,14].

Therefore, each user in a social network weighs both the risks and benefits to determine how many profile attributes to reveal. Additionally, the privacy settings of one user affect the choice of privacy settings of another user. However, little work has been done to show how all these factors are linked together. Consequently, there is a need to model the interaction of users in a generic social network to understand how privacy risk and relationship-building both influence the level of self-disclosure