Applications of Social Network Analysis for User Modeling

Tasos Spiliotopoulos
Madeira Interactive Technologies Institute
University of Madeira
9000-390 Funchal, Portugal
tspiliot@gmail.com

Ian Oakley
Madeira Interactive Technologies Institute
University of Madeira
9000-390 Funchal, Portugal
ian.r.oakley@gmail.com

ABSTRACT
This position paper makes the case for the use of social network analysis for user modeling in social media. It argues that a network perspective, in combination with more traditional approaches for user modeling, can allow the creation of more effective user models. Towards this end, we outline two ongoing studies that utilize social network analysis techniques for improving understanding of user behavior in social media, with the goal of instigating discussion in the workshop.

Keywords
Social network analysis, user modeling, Facebook, Twitter, social network sites, social media

INTRODUCTION
Adaptive systems are advanced software artifacts capable of automatically customizing themselves to the needs of their users. In order to achieve this they construct and maintain user models, typically based on a mix of explicit profile information, declared preferences, and observed behavior. In contrast to more traditional media, based on a one-way flow of information from content creators to users, social media facilitate and encourage the creation and sharing of content by the users themselves. These actions and relationships provide a rich source of observable user behavior (for example, related to what, and who with, content is shared) that modeling approaches can leverage.

In this position paper, we argue that network perspectives on users' behavior in social media can provide important inputs and explanatory insights that can aid user-modeling processes. We present two ongoing studies that utilize social network analysis techniques for improving understanding of user behavior in social media, with the goal of instigating discussion in the workshop. Although these studies have not been designed with the explicit goal of enabling or supporting the creation of user models, we are interested in exploring the extent to which they can be applied to this domain.

FROM SOCIAL NETWORK ANALYSIS TO USER MODELING

Social network analysis emphasizes structural relations as its key orienting principle. Its objectives are to measure and represent these structural relations accurately, and to explain both why they occur and what their consequences are. In other words, social network analysis focuses on the regularities of the patterns of relations, instead of the attributes of the entities [1]. Proponents of this approach argue that an individual's position in a network can provide a better understanding of "what's going on" or "what's important" than that person's individual attributes. This paper argues that such analyses will also provide valuable information for modeling users.

Social network analysis involves the calculation of a number of network metrics that can assist the understanding of the network, depending on a particular network's definition (i.e. what the nodes and edges represent exactly). In a network such as the egocentric Facebook friendship network pictured in Figure 1, we can identify nodes with specific network characteristics. For example, high betweenness centrality suggests that a node is very important in spanning a number of groups or spheres of activity, and thus may be jointly involved in several activities with ego. Similarly, community detection algorithms can effectively identify clusters or groups of friends and, thus, enable narrowcasting of information.

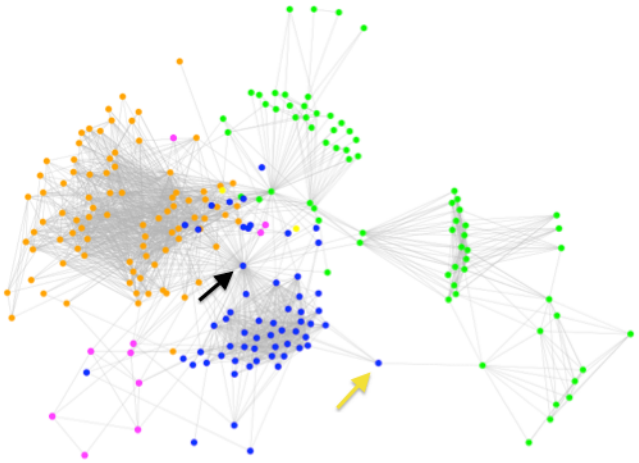

Figure 1. Visualization of a typical egocentric Facebook network (with ego removed). The yellow arrow points to a node with high betweenness centrality, while the black arrow points to a node with high closeness centrality. The different-colored nodes denote different communities.

International Workshop on User Modeling from Social Media / IUI 2012, February 14, 2012, Lisbon, Portugal.

Social network sites currently deploy simple social network analysis to perform basic user adaptation, such as suggesting friends based on the number of common friends, a feature derived from triadic closure. However, current techniques are limited in scope and this paper argues there is room for more sophistication. Some notable efforts by researchers in this area include predicting tie strength with information collected from Facebook [2], investigating user influence on Twitter based on in-degree, mentions and retweets [3], and demonstrating that two structural characteristics, transitivity and mutuality, are significant predictors of the desire to form new ties [4].

In addition to calculating metrics, social network analysis enables the visualization of networks, which makes possible a more exploratory analysis. Most popular social network sites have an API available, which can be employed for mining social network data, which in turn can be fed into a social network analysis package for analysis.

ONGOING WORK

We currently have a number of ongoing studies aimed at investigating users' behavior in online social networks through social network analysis. Below we outline two of them, with the goal of stimulating discussion in the workshop with regards to the extent that the findings of these studies can provide useful insights for user modeling approaches in social media.

Facebook and Twitter

We have deployed an online survey aimed at capturing uses and gratifications for Facebook and Twitter, similar to the one described by Joinson [5]. Uses and gratifications refer to the "how" and "why" of media use, and as such they can classify the users according to their motives and activities and they may be considered as aiding the creation of user models on their own behalf. As an additional step, however, participants in our survey also install a Facebook application that we developed, which, with the users' consent, captures their 1.5-degree Facebook friendship network (i.e. their friends and the mutual friends among them), some basic profile information, and several summary statistics regarding their activity on the network. The activity captured includes such information as the number of posts made, likes, comments, number of groups they are member of, number of photographs they have posted or have been tagged in. From the user's egocentric network we can calculate a number of meaningful network metrics, such as network density, average path length, average degree, and average coefficient, along with the distributions for these metrics for the nodes. Similarly, from Twitter we capture some network information, such as the 1.5-degree follower network and the number of accounts followed, as well as some activity information, such as the number of posts made, the rate of posting, and the number of mentions and retweets received.

Our main consideration in this study is mapping particular uses with gratifications. The structure of a user's Facebook network, which can be obtained without the user's explicit involvement, may indicate particular uses and gratifications for the user, which in turn can help in building a user model and can lead to a customized social media experience. In addition, by examining the uses and gratifications of both Facebook and Twitter for the same user, we can explore the ways that the two networks complement each other and get insights regarding the extent to which user models are transferable across social media.

Twitter hashtags

A second ongoing study involves examining the networks created by tweets mentioning specific hashtags. More specifically, we collected all the tweets across the duration of a month that contained as a hashtag the name of one of three specific European cities. We gathered from the Twitter API all the publicly accessible information for the account that made each tweet in the dataset and we constructed three directed networks for each city, based on the "follows", "replies to" and "mentions" relationships among the accounts.

We performed a preliminary analysis of the data, where we calculated a number of metrics for all the nodes in each of the 9 networks in total, such as in- and out-degrees, betweenness, closeness and eigenvector centralities, clustering coefficients, as well as identified clusters in the network. This type of analysis allows the identification of influential nodes, hubs, and nodes that act as bridges in the network. Identifying correlations between specific usage patterns, profile information, and these network metrics can lead the way for the construction of user models in this dataset.

Further analysis of these data suggests that the Twitter community defined by the members of each city network shows strong signs of a virtual community. Therefore, it might be reasonable to expect members of the network to assume roles similar to those detectable in virtual communities and perhaps attempt to model certain users according to these roles.

ACKNOWLEDGMENTS

The work reported in this paper is supported by FCT research grant SFRH/BD/65908/2009.

REFERENCES

1. Knoke, D. and Yang, S. *Social Network Analysis*. Sage Publications, Inc, California, 2008.
2. Gilbert, E. and Karahalios K. Predicting tie strength with social media. In *Proc. CHI 2009*, ACM Press (2009), 211-220.
3. Cha, M., Haddadi, H., Benevenuto, F., and Gummadi, K.P. Measuring user influence in twitter: The million follower fallacy. In *Proc. ICWSM 2010* (2010).
4. Golder, S.A. and Yardi, S. Structural predictors of tie formation in twitter: Transitivity and mutuality. In *Proc. SocialCom 2010*, IEEE (2010).
5. Joinson, A.N. "Looking at", "looking up" or "keeping up with" people?: Motives and use of facebook. In *Proc. CHI 2008*, ACM Press (2008), 1027-1036.