

# Discovering cohesive subgroups from social networks for targeted advertising

Wan-Shiou Yang<sup>\*</sup>, Jia-Ben Dia

*Department of Information Management, National Changhua University of Education, No. 1, Jin-De Road, Changhua 500, Taiwan, ROC*

## Abstract

In this paper, we propose a framework that utilizes the concept of a social network for the targeted advertising of products. This approach discovers the cohesive subgroups from a customer's social network as derived from the customer's interaction data, and uses them to infer the probability of a customer preferring a product category from transaction records. This information is then used to construct a targeted advertising system. We evaluate the proposed approach by using both synthetic data and real-world data. The experimental results show that our approach does well at recommending relevant products.

© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Social network; Targeted advertising; Recommender system; Knowledge discovery

## 1. Introduction

The effectiveness of targeting a small portion of customers for advertising has long been recognized by businesses (Armstrong & Kotler, 1999) for two main reasons. First, the amount of product/service information available to customers is ever-increasing, and hence it is desirable to help customers wade through the information to find the product/service they want. Second, understanding the needs of current and potential customers is an essential part of customer-relationship management. The ability to accurately and efficiently identify the needs of customers and subsequently advertise products/services that they will find desirable will increase customer-retention, growth, and profitability of a business (Armstrong & Kotler, 1999).

The traditional approach to targeted advertising is to (manually) analyze a historical database of previous transactions and the features associated with the (potential) customers, possibly with the help of some statistical tools, and identify a list of those customers who are most likely to respond to the advertisement of the product. The advent

of new technologies has led to automatic tools being advocated for identifying potential customers (Hayes, 1994), with many recommender systems having emerged over the past few years whose basic idea is to advertise products according to users' preferences as obtained by ratings either explicitly stated by the users or implicitly inferred from previous transaction records, Web logs, or cookies.

The first type of recommendation technique was the *content-based approach* (Cohen, 1992), in which recommendable products are characterized by a set of content features, and customers' interests are represented by a similar feature set. Content-based approaches select target customers whose interests have a high degree of similarity to the product's content profile. To establish an accurate content profile for a product, the detailed description of a product must be parsable (e.g., as text), and a set of content features are extracted by some information-extraction or summarization techniques (Mooney & Roy, 2002). The content-based approach is inappropriate for products whose content is not electronically available, or is based on subjective factors such as quality, style, or point of view. Furthermore, since the content features of an individual are derived purely from the products in which s/he has

<sup>\*</sup> Corresponding author. Tel.: +886 4 7232105x7611.

E-mail address: [wsyang@cc.ncue.edu.tw](mailto:wsyang@cc.ncue.edu.tw) (W.-S. Yang).

shown an interest, the content-based approach is inappropriate to advertise products to a new customer and will result in a low probability of providing surprising advertisements.

Another type of recommendation technique is the *collaborative approach* (or sometimes called the *social-based approach*) (Cohen, 1992). The collaborative approach looks for relevance among customers by observing their ratings assigned to products in a small training set. The nearest-neighbor customers are those that exhibit the strongest relevance to the target customer. These customers act as “recommendation partners” for the target customer, and collaborative approaches advertise products to the target customer that appear in the profiles of these recommendation partners but not in that of the target customer. Whilst this approach has demonstrated usefulness in many applications, it still has limitations such as an inability to advertise newly introduced products that have yet to be rated by customers, an inability to advertise products to a new customer who has yet to provide rating data, and poor predictions if the data are sparse.

In order to remedy these problems, in this research we propose a framework that utilizes the concept of a social network to facilitate the automatic construction of a targeted advertising system. Our approach applies data mining techniques to gather social interaction data so as to cluster customers into a set of cohesive subgroups. Based on the set of cohesive subgroups, we infer the probabilities of a customer preferring a product category from transaction records. This information is crucial to the successful online promotion of products to customers. Utilizing such information allows targeted advertising to be conducted on a broader scope, by advertising new products to new customers.

This paper is organized as follows. In Section 2, we introduce the concept of a social network and define our problem. Section 3 describes the algorithm designed for discovering cohesive subgroups from a social network. Section 4 derives the algorithm used to select the customers to which to promote a given product. The proposed approach is evaluated first using synthetic data and then using empirical data obtained from our university email logs and library-circulation data. We report the evaluation results in Section 5. Section 6 describes related work, and Section 7 concludes with a summary and a discussion of future research directions.

## 2. The problem

A social network is a set of individuals connected through socially meaningful relationships, such as friendship, coworking, or information exchange (Wasserman, Faust, Iacobucci, & Granovetter, 1994; Wellman, 1996). Social networks are formed when people interact with each other (Garton, Haythornthwaite, & Wellman, 1997) and thus can be seen in many aspects of everyday life. Social network theory traditionally views social relationships in

terms of nodes and links (Wasserman et al., 1994), where the nodes are the individual actors within the networks and the links are the relationships between the actors. In its most simple form, a social network is a map of all of the relevant links between the nodes being studied. These concepts are often displayed in a social network diagram, with nodes indicated as points and links indicated as lines (Wasserman et al., 1994).

The strength of ties between nodes in a real-world social network is an important theoretical issue. As noted by Milgram (1967), the strength of a tie between two actors is much greater if they have another mutual acquaintance. In other words, the probability of two friends of an individual knowing each other is much greater than the probability of two people chosen randomly from the population knowing each other (Guare, 1990; Newman, Watts, & Strogatz, 2002; Watts & Strogatz, 1998). Actors with strong ties usually have some sort of common ground on which they establish their relationships (Preece, 2002; Wellman & Gulia, 1997), and thus often constitute a subgroup. Because of the common ground, actors with strong ties – and hence representing a subgroup – often share common interests, needs, or services that provide a reason for the subgroup (Preece, 2002; Schwartz & Wood, 1993; Wellman & Gulia, 1997).

In reality, a customer’s decision to buy a product is often strongly influenced by his/her friends, acquaintances, or business partners, etc. A classic example is the Hotmail free email service, whose growth from zero to 12 million users within 18 months is attributed to the inclusion of a promotional message containing the service’s URL in every email sent using it (Jurvetson, 2000). Because of the mutual influences, customers with strong ties are apt to exhibit similar purchasing behavior. This observation initiates the interesting idea that identifying subgroups – comprising customers with strong ties and hence common interests – from a customer’s social network may facilitate the targeting of a small portion of customers for advertising.

The advent of information techniques, especially communication systems (e.g., email, bulletin boards, and messaging) and contacting systems (e.g., MSN (<http://www.msn.com>), ICQ (<http://www.icq.com>), Orkut (<http://www.orkut.com>), and Friendster (<http://www.friendster.com>)), has led to a rapid growth in computer-mediated social networks. The human activities on these networks generate a huge amount of data suitable for analyzing the social networks and facilitating the automatic construction of a targeted advertising system.

Therefore, in this research we investigate the problem of targeted advertisement in an environment with the following features:

(1) *A database that contains customer’s connectedness is available*

This database can be obtained by links either explicitly stated by customers (e.g., friendships stated on contacting sites) or implicitly inferred from previous interaction data (e.g., email logs). We transform the connectedness database into a customer’s social network and represent the network

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات