Information & Management 52 (2015) 789–800

Contents lists available at ScienceDirect

Information & Management

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

Personalized recommendations based on time-weighted overlapping community detection

Haoyuan Feng [a], Jin Tian [a,*], Harry Jiannan Wang [b], Minqiang Li [a]

[a] College of Management and Economics, Tianjin University, Tianjin 300072, PR China
[b] University of Delaware, Newark, DE 19716, USA

ARTICLE INFO

Article history:
Received 17 September 2014
Received in revised form 22 January 2015
Accepted 14 February 2015
Available online 25 February 2015

Keywords:
Personalized recommendations
Recommender system
Dynamic user interests
Overlapping community
Time-weighted association rules

ABSTRACT

Capturing and understanding user interests are an important part of social media analytics. Users of social media sites often belong to multiple interest communities, and their interests are constantly changing over time. Therefore, modeling and predicting dynamic user interests poses great challenges to providing personalized recommendations in social media analytics research. We propose a novel solution to this research problem by developing a temporal overlapping community detection method based on time-weighted association rule mining. We conducted experiments using MovieLens and Netflix datasets, and our experimental results show that our proposed approach outperforms several existing methods in recommendation precision and diversity.

1. Introduction

Capturing, understanding, and predicting user interests are critical tasks for social media sites to provide better product designs [1], accurate targeted advertising, and personalized services [12,55] as well as to enhance users' continuance and sense of belonging to these social media sites [32]. Twitter launched a new push notification feature in late 2013 to deliver personalized recommendations for tweets and accounts to follow after implementing the long-running @MagicRecs experiment to monitor users' following and re-tweeting behavior. Recommender systems have been developed to predict user preferences and have been applied to various applications and websites, such as Amazon, Netflix, Pandora, and LinkedIn.

Recommendation algorithms are at the core of recommender systems, which can be divided into three primary categories, i.e., content-based recommendation approaches, collaborative filtering (CF) recommendation approaches, and hybrid approaches [6]. Content-based approaches aim to recommend items to a user based on the item description and the user's browsing history and interest profile. A lack of diverse recommendations is the drawback of content-based approaches [13]. Collaborative filtering approaches provide recommendations to a user by collecting preference information from many other users, but these approaches often have limitations such as rating sparsity, scalability, and efficiency [36]. Studies have shown that the quality of recommender systems can have a great impact on overall customer satisfaction [17].

Some recommendation methods are developed based on information from all users [26,30]. Given the huge amount of data on users and their item preferences, these approaches often face performance issues due to time-consuming data modeling and processing [37]. Users of social media sites often form special interest groups/communities, e.g., Pinterest users use different boards to form groups such as fashion, home decor, and gardening, and Twitter users use hash tags to group tweets based on special topics. Researchers have found that modeling user preferences based on users' community or group associations can enhance the efficiency of recommendation algorithms [8,15,21]. Some social media analytics approaches use existing virtual communities on social websites, such as online knowledge communities or discussion forums [2,3,49]. However, these communities cannot be easily categorized in many scenarios. Thus, there is a large variety of community detection methods. Overlapping community is a type of community structure in which a user may belong to more than one community. Clustering users into overlapping communities can not only help with the data sparsity problem but also enhance recommendation diversity by obtaining recommendations from different communities [14,42].

* Corresponding author. Tel.: +86 22 27404796.
E-mail address: jtian@tju.edu.cn (J. Tian).

http://dx.doi.org/10.1016/j.im.2015.02.004
0378-7206/

Some early user interest models in recommender systems usually assumed that users' interests were static and did not change over time. Actually, users' interests are dynamic rather than static. For example, users may prefer different clothing styles in different seasons and/or at different ages, and users' interests in digital products such as music and movies are more likely to change over time [5,7,20,27].

Fig. 1 provides an example from the MovieLens dataset to show the percentage of movies in different categories rated by a user over 8 months, from September 1997 to April 1998. Fig. 1 clearly illustrates that this user's interest in different types of movies changes significantly over time. For example, the user seems to like watching horror movies before February, but this interest disappears in the last three months. Static recommender systems may capture users' previous interests nicely, but they fail to update the interests over time, resulting in inaccurate recommendations and a poor user experience. Association rules (AR) have been widely adopted to represent user interests in many recommendation models due to their ability to scale to large datasets and achieve high precision [13,38,48]. Time decay functions in AR and other methods, such as Singular Value Decomposition (SVD) and Hidden Markov, have been used to model user interest drift over time [16,22,34,43,45]. However, most of the existing approaches model temporal user interest drift at the individual user level, without considering the effects of related changing overlapping communities.

In this paper, we propose a novel recommendation method called Temporal Overlapping community detection using Time-weighted Association Rules (TOTAR). The key thrust of TOTAR is that temporal factors of user interests are fully incorporated in both overlapping community detection and association rules generation. More specifically, the main contributions of this paper are threefold: (1) the development of a temporal method to detect overlapping communities based on a user–user graph with time-weighted links; (2) the design of a new time-weighted association rule mining algorithm based on temporal overlapping communities to model user interest drift over time; and (3) the proposal of a new recommendation model based on users' dynamic temporal interests and multi-memberships in their overlapping communities.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant literature. The details of the proposed TOTAR approach are elaborated in Section 3. Section 4 presents experiments using the MovieLens and Netflix datasets and discusses the results. Finally, Section 5 summarizes the key points of the paper and discusses future research.

2. Literature review

2.1. Overlapping community detection

Overlapping community detection methods provide alternatives for representing the variability and diversity of user interests. Early efforts at community detection assumed that communities are non-overlapping or disjointed. To increase detection speed and quality, Newman proposed the Fast Newman algorithm [39], which adopted the Modularity function for global optimization. This algorithm has been widely used in community detection. Palla et al. presented the first overlapping community detection algorithm, the Clique Percolation Method (CPM), which performed well in recognizing overlapping frameworks in complex networks [40]. Lancichinetti and Fortunato showed that the Modularity function did not fit the overlapping community context very well, and it had a problem with resolution limit and extreme degradation under overlapping conditions [24]. More recently, overlapping community detection methods based on local optimization have been proposed. Lancichinetti et al. proposed a local fitness maximization (LFM) algorithm that took advantage of both overlapping and the hierarchy of communities [25]. Meo et al. emphasized enhancing existing community detection algorithms by adding a pre-processing step in which links were weighted according to their centrality [41]. Furthermore, Lee et al. found that a phenomenon called pervasively overlapping communities existed in many networks, which meant that nearly all nodes belong to multiple communities [28]. Pervasively overlapping communities may contain more external than internal links and thus violate the properties of network communities. An expanded LFM algorithm has been proposed to handle the pervasively overlapping community problem [28,29].

2.2. Association rule mining

Association rules have been successfully used to represent user interests in static recommendation models [48]. Generally, user interests can be provided by AR in the form '$A \rightarrow B$' (where A and B are user interests or items), which means that users interested in 'A' are likely interested in 'B'. For dynamic interest models, Li et al.

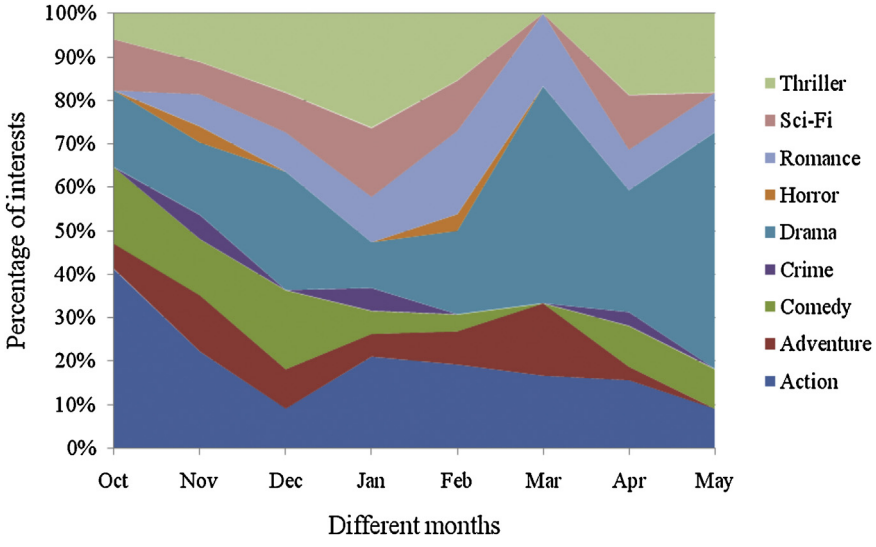

Fig. 1. Percentage of a user's interest during different months.