Future Generation Computer Systems 78 (2018) 413–418

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Editorial

Recommender Systems for Large-Scale Social Networks: A review of challenges and solutions

Magdalini Eirinaki [a], Jerry Gao [a], Iraklis Varlamis [b,*], Konstantinos Tserpes [b]

[a] *Computer Engineering Department, San Jose State University, San Jose, CA, USA*
[b] *Department of Informatics and Telematics, Harokopio University of Athens, Athens, Greece*

ARTICLE INFO

Keywords:
Recommender systems
Social networks
Scalability
Big data

ABSTRACT

Social networks have become very important for networking, communications, and content sharing. Social networking applications generate a huge amount of data on a daily basis and social networks constitute a growing field of research, because of the heterogeneity of data and structures formed in them, and their size and dynamics. When this wealth of data is leveraged by recommender systems, the resulting coupling can help address interesting problems related to social engagement, member recruitment, and friend recommendations.

In this work we review the various facets of large-scale social recommender systems, summarizing the challenges and interesting problems and discussing some of the solutions.

1. Introduction

The fields of computational intelligence and knowledge management have made significant advances over the past decades. The potential ability to create intelligence from the analysis of raw data has been successfully applied to diverse areas such as business, industry, sciences, social media etc. The coupling of online social networks with recommender systems created new opportunities for businesses that consider the social influence important for their product marketing, as well as the social networks that want to improve the user experience by personalizing the content that is provided to each user and enabling new connections. At the same time, these changes have created new challenges for researchers in the area of recommender systems and social network analysis. The large volume of social network interactions that expand the size of the social graph with increased velocity, the variety of information provided in the form of written reviews, ratings or permanent and volatile relations, and the veracity of data expressed in the form of trust or distrust between users who become product reviewers or opinion influencers, are only some of the factors that make social networks and the associated recommender systems an ideal case of big-data research.

In this work we focus on large-scale recommender systems that take advantage of the characteristics of the underlying social network and focus on the variety and volatility of social bonds, tackle the problems of size and speed of change of social graphs, test the scalability of traditional recommender systems and/or present solutions that can take recommender systems to the next level.

2. Recommender systems and social networks

The most popular techniques used in recommender systems are Content-based (CB) filtering and Collaborative Filtering (CF). Each user in such systems is typically represented by a user profile, that includes the "items" this user has rated (or purchased). In content-based filtering, new items are recommended based on their similarity to items already present in the user's profile. To achieve this, more details on each item are needed. On the other hand, collaborative filtering approaches are "agnostic" to the items. Instead, they employ the ratings assigned by the user, to find other similar users (or items) based on their rating patterns. The generic nature of collaborative filtering systems was the reason of their broad success, as they are used to recommend a wide range of products such as movies, music, news, books, research articles, search queries, social tags, etc.

While a bipartite user-item graph containing the ratings used to be the sole input of such systems, such a structure is no longer sufficient to represent all available information, such as content, context, social information, and metadata. As a result, many attempts have been made to incorporate more information about the user, his/her context (spatial, temporal, social etc) and its evolution over time.

* Corresponding editor.
E-mail addresses: magdalini.eirinaki@sjsu.edu (M. Eirinaki), jerry.gao@sjsu.edu (J. Gao), varlamis@hua.gr (I. Varlamis), tserpes@hua.gr (K. Tserpes).
URLs: http://www.sjsu.edu/people/magdalini.eirinaki/ (M. Eirinaki), http://www.sjsu.edu/people/jerry.gao/ (J. Gao), https://www.dit.hua.gr/∼varlamis/ (I. Varlamis), https://www.dit.hua.gr/∼tserpes/ (K. Tserpes).

http://dx.doi.org/10.1016/j.future.2017.09.015
0167-739X/

2.1. Beyond user-item ratings: Context-aware recommender systems

Context-aware recommender systems (CARS) generate more relevant recommendations by adapting them to the specific context of the user. The contextual factors that must be considered by a recommender system relate to the time, location, and purpose of the targeted user. According to [1], the user context can be static or change over time. CARS assume a pre-filtering step, where context information is used to select the set of relevant items, a classic recommendation step that ranks relevant items according to the predicted ratings and a contextual post-filtering step, that re-ranks and filters the output of the traditional recommender [2].

Context-aware systems combine information from multiple sources within the social network in order to refine the context space and solve major recommender system issues such as "scalability" and the "cold-start" problem [3]. For example, in [4] contextual information is encoded in or reflected by the user-specific and item-specific latent factors. The user-item rating matrix is split into partitions, by grouping users and items with similar contexts and matrix factorization is applied to the generated sub-matrices.

A context-aware multimedia recommender system, which considers in tandem user preferences (in previous items' metadata), opinions (textual comments in user reviews), behavior (past items observations and actions) and feedbacks (expressed in the form of ratings) within the same framework is presented in [5].

2.1.1. Time-aware recommender systems

Time-aware recommender systems (TARS) can be considered as specialized CARS focusing on exploiting contextual information in the form of time. They assume that user preferences are drifting over time and user taste is evolving as new items become available and new trends appear. This in turns affects item popularity, which is constantly changing, bringing old items to the long tail of user preferences and moving new items to the head [6].

Handling the temporal dynamics of user preferences in recommender systems raises new challenges since the change on each individual user interests is different from the concept drift problem [7]. In a social network with multiple users and items, many different features are changing simultaneously, and influence each other, whereas in the general concept drift problem, only a single concept is tracked. Using sliding windows and preference decay functions increases the sparsity of an already sparse problem (since past information is discarded or lost) and is usually avoided. In [8] Koren extends the static matrix factorization model and the associated baseline predictors with functions that capture the gradual drift of user and item bias and introduce the timeSVD++ algorithm, which outperforms its predecessors.

Several techniques have been used in the literature, to adopt CF algorithms to temporal changes, by boosting recent ratings and penalizing older ratings, such as discrete time windows [9] or continuous decay function [10]. Recently, an algorithm based on association rule and community identification approach has been proposed to handle the drift problem in recommender systems [11]. Bayesian Probabilistic Tensor Factorization [12] has also been proven an appropriate temporal CF model.

2.1.2. Location-aware recommender systems

Location-aware recommender systems (LARS) have become popular mainly through place recommendation systems in the travel and tourism industry, as discussed by Chen and Tsai [13]. The advent of location-based social networks (LBSNs), such as Facebook Places and Foursquare, increased the available data and challenges for researchers [14]. LARS exploit location ratings when partitioning user-location bi-partite rating graphs with spatial criteria, so that locations that are spatially close to the users are employed and those in a distance are ignored, in a manner that maximizes system scalability while not sacrificing recommendation quality [15]. LARS consists of two components, an offline modeling component, which learns the interest of each individual user and the local preference of each individual location, by capturing item co-occurrence patterns and exploiting item contents and an online recommender component, which automatically combines the learned user's local preferences and produces the top-k recommendations [16,17].

Time is a crucial factor in Location-aware recommender systems, so recent approaches in location or activity recommendations employ unified spatio-temporal frameworks [18–20].

2.1.3. Community-aware or Social recommender systems

Community-aware or "Social recommender systems", have gained the attention of researchers since they leverage social relationships in order to improve the recommendation process. Authors in [21] give a narrow definition of social recommendation as "*any recommendation with online social relations as an additional input, i.e., augmenting an existing recommendation engine with additional social signals*". A broader definition, by [22] refers to recommender systems targeting social media domains.

The main premise in this line of research is that users' preferences are influenced more by the preferences of their friends, than these of unknown users. Such attempts enhance the typical recommendation process with social data, assuming that item ratings are available and that some form of influence/trust propagation exists within the user network. For example, a common approach is to enhance the memory-based collaborative filtering process by forming the user's neighborhood using similarities derived from the users' ratings and/or their social relationships, focusing on trust.

Community-aware systems employ user preferences, user connectedness, or any other social information, in order to detect user clusters and consequently partition the recommendation problem into smaller problems. Based on the concept of *homophily* in social networks, that a user's preferences are likely to be similar to, or influenced by these of her friends, such systems manage to fill the gap in cold-start problem and find similarities between users [23]. This can be done through co-factorization, where the assumption is that the users share the same preference vector in both the rating and the social spaces (e.g. [24]), ensemble methods, where the resulting recommendation is derived by the linear combination of two systems (e.g. [25]), or regularization, where priority is given to the social-based ratings (e.g. [26]). For example, in [27] authors propose a preference-aware community detection method to group users based on their social relations, whereas in [28], the users' social information (user-to-user friendship network) is used to partition the large user-to-item bipartite graph into smaller partitions and perform collaborative filtering in a narrower social context. When users belong to more than one community (i.e. we have overlapping communities), multi-label propagation based methods are used on the user-user graphs [29].

2.2. Beyond simple item recommendations

Recommender systems (RSs) have become popular since they can personalize the user experience by providing automated recommendations. They first appeared in e-commerce sites [30], used to recommend individual items, products of potential interest for customers to purchase. However, their usage has span multiple other domains in the past few years, from digital collections (e.g. news and research articles [31], to database queries [32,33] or even web services [34,35]).