

# Using Graph Database for File Recommendation in *PAD* Social Network

Fattane Zarrinkalam, Mohsen Kahani, Samad Paydar

Dept. of Computer Engineering  
Ferdowsi University of Mashhad  
Mashhad, Iran

[fattane.zarrinkalam@stu-mail.um.ac.ir](mailto:fattane.zarrinkalam@stu-mail.um.ac.ir), [kahani@um.ac.ir](mailto:kahani@um.ac.ir), [samad.paydar@stu-mail.um.ac.ir](mailto:samad.paydar@stu-mail.um.ac.ir)

**Abstract**—In this paper, a file recommender system is introduced which is used in *PAD*, an academic social network at Ferdowsi University of Mashhad. Considering the large number of files shared in *PAD*, the proposed system is aimed at helping users in finding related files. It uses content based and collaborative filtering techniques, where the former is based on automatic tagging of file names, and the later is based on the users' activities. Further, in order to improve efficiency, Neo4J graph database engine is employed at the data layer of the recommender system. The experimental evaluations, mainly based on the users' feedbacks, demonstrate that the proposed system has very good performance and it provides good quality recommendations.

**Keywords**—recommender system; social network; automatic tagging; graph database; Neo4j

## I. INTRODUCTION

During the last decade, social networks have emerged as the killer application of Web 2.0 in which the web acts as a medium for social activities, in addition to its previous appearances as a software platform or information dissemination infrastructure.

Social networks and resource recommender systems can benefit from each other: on the one hand, social networks have some public data and metadata like tags, user relationships, comments and ratings which enrich the underlying layer of a recommender system and consequently increase its performance; on the other hand, different types of resources, like files, wall posts, groups, and forums are created in a social network, and as the users' involvement increases, the greater becomes number of these resources which leads to the issue of how to utilize existing resources for providing added-value for the users. A basic solution which is integrated in most social networks is some resource recommendation facility [1] [2].

In this paper, a recommender system is introduced which is used in *PAD*<sup>1</sup>, a social network deployed at Ferdowsi University of Mashhad [3]. The goal of this recommender system is to help the user in finding files similar to the file which he is visiting. As a result, a graph based hybrid algorithm which uses a combination of tagging information and user activities is proposed.

Among different techniques which are used in the literature for the purpose of representation, storage and retrieval at the data layer of a recommender system, the proposed recommender system models its required data by utilizing graph database capabilities. The reason is that graph databases provide an effective and flexible mechanism for managing connected data, the kind of data that exactly fits social networks data [4]. Experimental evaluations demonstrate that the proposed recommender system is promising in this regard and it helps users in exploration of the files existing in *PAD*.

The rest of the paper is organized as follows. In section II, some related works are briefly reviewed. Then, the proposed recommender system is described in Section III. The experiments and some analysis are presented in Section VI. Finally, Section V concludes the paper by describing some directions for future works.

## II. RELATED WORK

The proposed recommender system is a graph based hybrid recommender system which is integrated an academic social network. As a result, this section reviews the related work in two categories, i.e. social network based recommender systems and graph based recommender systems.

### A. Social network based recommender systems

In some works, different types of social network information are utilized within traditional recommendation algorithms to improve performance and accuracy of a recommender system.

The three main techniques used in the traditional recommendation algorithms are collaborative filtering, content-based filtering and hybrid techniques [5]. In collaborative filtering, the recommended resources are those that are used by users similar to the user for whom the recommendations are being made. Here, some metric for determining neighboring users are required. Collaborative filtering has the benefit that it has a global view over the activities of the users [6], however, it does not perform effectively until enough information is generated about the users' interactions. This problem is known in the literature as the cold start issue [7].

In case of this type of recommendation algorithms, it is shown that incorporating social networking information like friendship and group membership data for calculating nearest

<sup>1</sup> <http://pad.um.ac.ir>