Proposal title : Influence and Diffusion-Aware Group Recommendation in Social Media

Internship location: Laboratoire de Recherche en Informatique, Paris-Sud (LRI)

Keywords: diffusion networks, group recommendation, communities, influence maximization

Supervisors: Prof. Bogdan Cautis (cautis@lri.fr), Assoc. Prof. Silviu Maniu (maniu@lri.fr)

Abstract: Recommendation is an ubiquitous task on the Web, where *users* are recommended *items* to *adopt*. In the common scenarios that are targeted by recommendation systems, people have always been influenced by their social circles to various degrees, by *word-of-mouth*. But this phenomenon has taken a whole new dimension with the advent of social networking on the Web. With a simple click, re-tweet, or "like", one can now instantly exert influence and diffuse information to her peers. Therefore, a deeper understanding of the way items are diffused, endorsed, and used can greatly improve the effectiveness of recommendations. However, due to the highly complex and uncertain nature of data in real-world social media, this raises new conceptual and technical challenges. Better adapted *data models* and *algorithms* are needed to address them, for a new generation of recommendation mechanisms, which can be influence-aware and social network-aware. In this context, the internship project aims at advancing the state-of-the-art in the area of *group recommendation* techniques, by taking into account key ingredients such as influence, diffusion patterns, and the community-based, modular structure of social networks. Group recommendation is important in many scenarios, such as those targeting movies, travel destinations, restaurants, online games, music playlists, etc. Our goal is to devise new models and algorithms, and to empirically validate them in diverse applications.

Detailed description

Context of Research Recommendation is an ubiquitous task in applications on the Web. Examples are recommendations of friends on Facebook, of movies to watch on Netflix, of books to buy on Amazon, of blog posts in Tumblr, of restaurants to go to in TripAdvisor, of online games to purchase in Google Play, of music playlists in Spotify, etc. In particular, in recent years, online social media advertising has seen a tremendous growth; for one example, Facebook announced more than $26 billions in advertising revenue in 2016. Generally, the recommendations are based on *collaborative filtering*, i.e., detecting co-occurring patterns between similar users, often by exploiting also an underlying graph (social) structure. Motivated by the need for effective viral marketing strategies, *influence estimation* and *influence maximization* (IM) have become important research problems, at the intersection of data mining and social sciences. The former aims at understanding how users may influence each other, who are the heavy influencers or the early-adopters, etc. The latter aims at selecting a set of nodes from a given diffusion graph, such that the spread is maximized.

One important area in modern recommendation applications, and the focus of this project, is the one of *group recommendation*, where a recommendation targets a *group*, or a *community*, instead of individual users. In particular, we focus on scenarios where the recommended items are to be "consumed" in groups. In the context of social networks, another way to assess the effects of performed recommendations is by looking at the evolution in time of adopted items (friends, movies, books, games): *How do the new recommendations propagate in the social network of the adopter(s)?* If, instead of looking at one adoption only, we anticipate the long-term impact in

such a setting (the *influence* of the recommendation), we become interested in recommendations that are more long-term effective and have a more complex target.

An important ingredient for the long-term impact of recommendation are *communities*; it is well-known that the spread of information inside social communities is faster than between them. In this context, models and algorithms for the group recommendation problem can have as starting point recent work from several research domains: (i) spread of information in social networks (ii) community-based influence models(iii) group recommendation approaches, (iv) influence estimation and diffusion network inference

Objectives The purpose of this project is to study models and algorithms for the *influence-aware group recommendation* problem. The mission of the student is to study ways in which group recommendation can be performed in settings where information propagates through like-minded communities. Starting with the study of previous approaches in group recommendation and influence maximization, the student will be tasked with

1. identifying ways to adapt / extend these ideas to influence-based group recommendation systems, and
   a) proposing an *influence model that is community-based*, where links are not between users, but mostly between communities, and where influence probabilities are parameters of communities, rather than pairs of users,
   b) devising algorithms that are at the same time closer to real-world assumptions and effective for the studied problem.

In terms of *applications*, we intend to focus on diverse applications, including Twitter adoption predictions, academic or research topics recommendation, online game recommendation, blog post recommendations, restaurant or trip recommendation, etc.

Work Environment During the internship, the student will be part of the LaHDAK team in LRI laboratory of University Paris-Sud. The team's research orbits around knowledge, data management, and Web mining. Previous relevant research by the two supervisors includes [LCCM17, LMM+, HCC+17]. This internship can be the starting point of a PhD thesis.

Required Skills Solid bases in Web data management and mining, graph theory and algorithms, as well as programming in well-known languages (C++, Java, Python).

References

[HCC+17] Shoubo Hu, Zhitang Chen, Bogdan Cautis, Laiwan Chan, Yanhui Geng, and Xiuqiang He. Model-free inference of diffusion networks using reproducing kernel hilbert space embeddings. *To appear*, 2017.

[LCCM17] Paul Lagrée, Olivier Cappé, Bogdan Cautis, and Silviu Maniu. Large scale online influence maximization. In *IEEE Int. Conference on Data Mining (ICDM)*, 2017.

[LMM+] Siyu Lei, Silviu Maniu, Luyi Mo, Reynold Cheng, and Pierre Senellart. Online influence maximization. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015*.