Analyzing the Structure & Dynamics of Multilayer Networks: An Application Centric Study

Soumajit Pramanik

Abstract

Real-world systems are mostly comprised of multiple inter-related subsystems and hence, can be accurately represented as multilayer networks. Understanding different characteristics of these multilayer networks is of prime importance as it is essential in accurately discerning the true behavior of the underlying real-world systems. In this thesis, we investigate the following structural and dynamical behavior of multilayer networks through the lens of data science and network science.

(a) Community structure: We propose a community detection algorithm for multilayer networks. The crux of this algorithm is based on a novel multilayer modularity index Q_M. The proposed algorithm is parameter-free, scalable and adaptable to complex network structures. More importantly, it can simultaneously detect communities consisting of only single type, as well as multiple types of nodes (and edges). We evaluate the performance of the proposed community detection algorithm both in the controlled environment (with synthetic benchmark communities) and on the empirical datasets (Yelp, Meetup and Digg datasets); in both cases, the proposed algorithm outperforms the competing state-of-the-art algorithms.

(b) Information diffusion: In this work, we investigate the information diffusion dynamics in the context of multilayer networks. We specifically concentrate on the popular online microblogging network Twitter where information propagates via two modalities - retweeting and mentioning. We develop an analytical framework for cascade formation considering both retweet and mention activities into account. The proposed framework $\mathcal{C}_F^M$ analytically computes the cascade size, depicting tweet popularity and discovers the presence of a critical retweet rate, under which mentioning in a tweet significantly helps in cascade formation. Additionally, taking cues from the model, we propose a mention recommendation system *Easy-Mention* which outperforms the state-of-the-art mention recommendation strategies.

(c) Node movement across layers: We conduct an empirical study to understand the dynamics of node movement across the layers of multilayer networks. For this purpose, we concentrate on the knowledge social network where each layer contains the researchers working in a particular field, as nodes and a link among a pair of nodes signify collaboration between the corresponding researchers. Due to shift of research interest, researchers often tend to migrate from one field (layer) of research to another field (layer), which depict the node movement across the layers of the multilayer network. We investigate the key factors regulating a researcher's (node's) decision to migrate to a specific research field (layer) and observe the effect of such migration on her career and the respective field (layer). We observe that in general publication quantity & quality, collaborator profile, fields' popularity contribute to a researcher's decision of field-migration.

(d) Entity recommendation: Most of the classical recommendation systems perform poorly while recommending entities in systems with multiple interacting entities, which we observe in most real systems. For instance, in event based social networks such as Meetup, a collection of entities (say, events, groups, venues and members) interact with each other both online and offline. In such a scenario, we present a deep learning

based venue recommendation system *DeepVenue* which provides context driven venue recommendations for the Meetup event-hosts to host their events. For hosting an event, the proposed *DeepVenue* model computes a score for each candidate venue and returns the list of top k venues ranked by these scores. Our rigorous evaluation shows that *DeepVenue* significantly outperforms the baselines algorithms. Precisely, for 84% of events, the correct hosting venue appears in the top 5 of *DeepVenue* recommended list.

Summarizing, in this thesis we analyze the multilayer networks from the perspective of different structural and functional properties with the help of state-of-the-art network science and data science techniques. Furthermore, we develop several applications such as *Easy-Mention*, *DeepVenue* for the benefit of the users belonging to multiple real multilayer networks.

Keywords: multilayer networks, community, information diffusion, mention recommendation, entity recommendation, venue recommendation, node movement, scientific migration, Twitter, Meetup, knowledge social network.