Project Title:

Social Network Analytics: Adoption, Persuasion, Recommendation, and Knowledge Refreshing

Project Summary:

Social networks such as those facilitated by social media, online games, or mobile devices have attracted increasing attention from both academia and industry that explore how to leverage such networks for greater business and societal benefits. Toward that end, we develop novel models, theories, and methods that mine massive social network data for business purposes. In this project, we focus on a unique phenomenon in social networks – the diffusion of adoption behavior (e.g., adoption of a product, service, or opinion) from one social entity to another. Specifically, we investigate three critical and related problems concerning this phenomenon: adoption, persuasion, and link recommendation. That is, the diffusion of adoption behavior is initiated by persuaders and reached to adopters through the linkage structure of a social network. Accordingly, we study the following problems: how to predict adoption probabilities in a social network? how to predict top persuaders in a social network? and how to recommend links for a social network? We follow design science principles to design, build, and evaluate models, theories, and methods that solve these problems. Let us take the problem of predicting adoption probabilities as an illustration. Adoption probability refers to the probability that a social entity will adopt a product, service, or opinion in the foreseeable future. Building on relevant social network theories, we identify key factors that affect adoption decisions: social influence, structural equivalence, entity similarity, and hidden factors. The principal challenge thus is how to predict adoption probabilities in the presence of hidden factors that are generally unobserved. To address this challenge, we develop a Bayesian learning method on the basis of the expectation-maximization framework. Using data from two large-scale social networks, we demonstrate that the developed method significantly outperforms prevalent existing methods. The empirical results also offer two interesting observations: existing methods that exclusively use social influence to predict adoption probabilities seem ineffective; hidden factors appear to play a significant role in adoption probability predictions.

We also study a fundamental problem in social network analytics in particular and big data analytics in general. Data (e.g., social network data) are dynamic in real world as new data are continuously generated. Newly generated data could bring in new knowledge and invalidate part or even all of the earlier discovered knowledge. As a result, a fundamental problem in social network analytics as well as big data analytics is how to maintain the currency of knowledge discovered from rapidly evolving data sources, namely the problem of knowledge refreshing. We model the knowledge refreshing problem as a Markov decision process and show that the optimal knowledge refreshing policy is monotonically increasing in the system state. We further show that the problem of searching for the optimal knowledge refreshing policy can be reduced to the problem of finding the optimal thresholds and propose a method for computing the optimal knowledge refreshing policy. The effectiveness and the robustness of the proposed method are examined through extensive empirical studies addressing a real-world knowledge-refreshing problem.

In this project, we design models, theories, and methods to solve fundamental problems in social network analytics. Our design of IT artifacts is rigorous because (1) our design is grounded in theories (e.g., social network theories); (2) we develop theories that govern our design of IT artifacts (e.g., the theorem on the monotonicity of the optimal knowledge refreshing policy); and (3) we rigorously evaluate

our designed IT artifacts using real-world data in real-world settings. According to the knowledge contribution framework of design science research[1], our project makes contributions in both problem formulation and solution development. In each study of our project, we formulate a new research problem. For example, we first propose to predict adoption probabilities in the presence of hidden factors. As another example, we introduce the problem of knowledge refreshing. In each study, we also develop novel, effective, and rigorous IT artifacts to solve the formulated problem.

Verification:

The project is entirely led and driven by university based faculty and students for educational purposes.

Supporting Documents:

Publications in Refereed Journals

1. Fang, X and Hu, P. Top Persuader Prediction for Social Networks. ***MIS Quarterly*** (Forthcoming).
2. Li, Z., Fang, X., Bai, X. and O. R. Liu Sheng. Utility-based Link Recommendation for Online Social Networks. ***Management Science*** (Forthcoming).
3. Fang, X., Hu, P., Li., Z., and W. Tsai. 2013. Predicting Adoption Probabilities in Social Networks. ***Information Systems Research,*** 24(1), pp. 128-145.
4. Fang, X., Liu Sheng, O. R., and P. Goes. 2013. When Is the Right Time to Refresh Knowledge Discovered from Data? ***Operations Research,*** 61(1), pp. 32-44.

Selected Media Coverage of Our Project

1. MIT Technology Review (October 2013): Viral Marketing Successfully Modeled By Network Theorists

   https://www.technologyreview.com/s/519786/viral-marketing-successfully-modeled-by-network-theorists/
2. MIT Sloan Management Review (August 2013): Dose Data Have a Shelf Life?

   http://sloanreview.mit.edu/article/does-data-have-a-shelf-life/
3. KSL Radio, Utah (11/15/2012): Interviewed about our research on predicting adoption probabilities in social networks.

[1] Gregor, S. and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337–355.