

Preface

We are proud to present the edited book that contains extended versions of a selected set of papers presented at the *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2012*, which was held in Istanbul at Kadir Has University in August 2012. Recently, use of social networks has become a normal activity for a huge number of ordinary people of all ages and as computer scientists we need to find new ways of making their experiences more productive and enjoyable. As a result, social networks and mining have become one of the hottest research fields in computer science and attracts the attention of researchers with different backgrounds. The field is fruitful and offers interesting challenging problems as reflected by our book contents. In the conference there were several top quality papers; however, in this book we are able to include 17 of them written by experts. The studies included in this book consider subjects ranging from practical issues to theoretical considerations such as discovering the political structure in the Italian Parliament, new event detection in tweets, prediction of stock market trends, and finding clique structures in social networks. In the following, we present a short summary of the papers covered in our book. We are confident that the collected efforts presented here will open new horizons for both practitioners and theoreticians working in the field of social networks and mining.

The first chapter: “[A Randomized Approach for Structural and Message Based Private Friend Recommendation in Online Social Networks](#)” by Bharath K. Samanthula and Wei Jiang proposes a two-phase private friend recommendation protocol for a particular user by using network structure and message interaction between users. They compute recommendation scores by preserving the privacy of users and for better security they propose an extended version of the proposed protocol using randomization technique. They provide the complexity and security analysis of their protocol and study its applicability based on different parameters.

The second chapter: “[Context Based Semantic Relations in Tweets](#)” by Ozer Ozdikkis, Pinar Senkul, and Halit Oguztuzun proposes methods to extract semantic relationships among terms in tweets and use them to detect events with higher accuracy, with larger time span, and in a user-friendly form. For discovering semantic associations they use co-occurrence-based statistical methods. They improve their previous work by using similarity scores instead of thresholds and

constant multipliers for semantic expansion and identify context-dependent associations by evaluating terms in specific time windows. Daily event clusters are determined by an outlier analysis.

The third chapter: “[Fast Exact and Approximate Computation of Betweenness Centrality in Social Networks](#)” by Miriam Baglioni, Filippo Geraci, Marco Pellegri, and Ernesto Lastres introduces a novel approach for the betweenness centrality computation. It speeds up considerably Brandes’ algorithm. Their approach exploits the natural sparsity of the data to algebraically and efficiently determine the betweenness of those node forming trees (tree-nodes) in the social network. They also give a fast sampling-based algorithm that computes an approximation of the betweenness centrality.

The fourth chapter: “[An Agent-Based Modeling Framework for Social Network Simulation](#)” by Enrico Franchi proposes an agent-based modeling framework for simulations over social networks. It is created to deal with large simulations and to work effortlessly with other social network analysis toolkits. In order to allow people without a strong programming background to write simulations easily, the framework comes with an internal Domain-Specific Language (DSL) embedded in Python. Their experience with their approach shows that it is successful in providing a user-friendly environment to perform agent-based simulations over social networks.

The fifth chapter: “[Early Stage Conversation Catalysts on Entertainment-Based Web Forums](#)” by James Lanagan, Nikolai Anokhin, and Julien Velcin studies conversation of forum users’ posts for television series episodes, and analyzes comments to obtain a description of the principal point of interest, which is referred to as peak. They focus on this peak and evaluate the behavior of users within it compared with during the entire conversation life cycle. They show by their approach that it is possible to identify those users within the forum who act as conversation catalysts for subsequent analysis, hence they do not lose a significant amount of information that has been produced by such members of the community.

The sixth chapter: “[Predicting Users Behaviours in Distributed Social Networks Using Community Analysis](#)” by Blaise Ngonmang, Emmanuel Viennet, and Maurice Tchuente studies the churn prediction problem, i.e., predicting the tendency of a user to stop using a social platform. They present a novel algorithm, which can deal with pathological cases, to accurately detect overlapping local communities in social graphs. They show that using the graph attributes it is possible to design efficient methods for churn prediction. They present experiments and a successful implementation of their approach using the Hadoop Hbase framework.

The seventh chapter: “[What Should We protect? Defining Differential Privacy for Social Network Analysis](#)” by Christine Task and Chris Clifton studies the use of differential privacy; which is an alternative privacy model, popular in data-mining over tabular data, that uses noise to obscure individuals’ contributions to aggregate results, in social networks. They present a practical introduction to the application of differential privacy to social networks, review two existing standards for adapting differential privacy to network data and analyze the feasibility of several common social-network analysis techniques under these standards.

The eighth chapter: “[Complex Network Analysis of Research Funding: A Case Study of NSF Grants](#)” by Hakan Kardes, Abdullah Sevincer, Mehmet Hadi Gunes, and Murat Yuksel discovers interesting complex network structures from the NSF funding data, and derive the collaboration among researchers in obtaining federal funding. Collaboration patterns at different periods are extracted at researcher, institution, and state levels.

The ninth chapter: “[A Density-Based Approach to Detect Community Evolutionary Events in Online Social Networks](#)” by Muhammad Abulaish and Sajid Yousuf Bhat presents a density-based community mining method for tracking overlapping community evolution in online social networks. It adapts a preliminary community structure toward dynamic changes in social networks using a novel density-based approach for detecting overlapping community structures and automatically detects evolutionary events. Their method does not require the neighborhood threshold parameter to be set by the users; rather it automatically determines it for each node locally. They show that the proposed method is computationally efficient and naturally scales to large social networks.

The tenth chapter: “[@Rank: Personalized Centrality Measure for Email Communication Networks](#)” by Paweł Lubarski and Mikołaj Morzy uses a large dataset of email communication within a constrained community to discover the importance of actors in the underlying network as perceived independently by each actor. For this purpose they use the simple notion that people are more likely to quickly respond to emails sent by people whom they perceive as important. They propose several methods for building the social network from the email communication data and introduce various weighting schemes that correspond to different perceptions of importance and compare their results with a ground truth to verify their method.

The eleventh chapter: “[Twitter Sentiment Analysis: How to Hedge Your Bets in the Stock Markets](#)” by Tushar Rao and Saket Srivastava studies identifying relationships between Twitter-based sentiment analysis of a particular company/index and its short-term market performance. For this purpose they use more than four million tweets. They show that negative and positive dimensions of public mood carry strong cause–effect relationship with price movements of individual stocks/indices. Their analysis of individual company stocks indicate strong correlation values with Twitter sentiment features of that company. They also investigate other features such as how previous week sentiment features affect the next week’s opening.

The twelfth chapter: “[The Impact of Measurement Time on Subgroup Detection in Online Communities](#)” by Sam Zeini, Tilman Göhnert, Tobias Hecking, Lothar Krempel, and H. Ulrich Hoppe studies the community detection problem. They consider methods that allow for the detection of overlapping clusters, they are the Clique Percolation Method and Link Community detection. They use these two methods to analyze data like mailing lists from some open source developer communities and compare the results for varied time windows of measurement. They observe that certain minimal window size is needed to get a clear image with enough “light” (i.e., dense enough interaction data) to detect subgroup membership.

The thirteenth chapter: “[Spatial and Temporal Evaluation of Network-Based Analysis of Human Mobility](#)” by Michele Coscia, Salvatore Rinzivillo, Fosca Giannotti, and Dino Pedreschi uses complex network techniques to mine and analyze a large dataset of human trajectories of a large GPS dataset of vehicles. They build a multiresolution spatial grid and map the trajectories to several complex networks, by connecting the different areas. They analyze different temporal slices of the network, the structural properties of the temporal and geographical slices, and their human mobility predictive power. They provide pointers regarding the significance of their results in understanding of the data transformation process that is needed to connect mobility with social network analysis and mining.

The fourteenth chapter: “[An Ant Based Particle Swarm Optimization Algorithm for Maximum Clique Problem in Social Networks](#)” by Mohammad Soleimani-pouri, Alireza Rezvani, and Mohammad Reza Meybodi uses the Particle Swarm Optimization algorithm to enhance the performance of the ant colony optimization algorithm for finding the maximum clique in social network graph. It is known that finding a maximum clique is essential for analysis of certain groups and communities in social networks. The simulation results on popular social network datasets and some additional graphs show improved performance with respect to the standard ant optimization algorithm.

The fifteenth chapter: “[XEngine: An XML Search Engine for Social Groups](#)” by Kamal Taha introduces a collaborative filtering recommender system called XEngine. It categorizes social groups based on various user characteristics such as age, ethnicity, religion, etc. The output generated by XEngine is ranked according to the user preferences and system inferred user social groups. The experimental comparison with three other systems and statistical tests shows that it provides a marked improvement and in terms of execution time efficiency it is comparable with other systems. A demo version of the engine is made accessible on the Web.

The sixteenth chapter: “[Size, Diversity and Components in the Network Around an Entrepreneur: Shaped by Culture and Shaping Embeddedness of Firm Relations](#)” by Maryam Cheraghi and Thomas Schott examines the causal scheme of culture, personal networking, and business networking, where attributes of the entrepreneur and the firm are included as controls. They conceptualize properties of the personal network around the entrepreneur, specify hypotheses about cultural effects on the personal network, test them, and then turn to hypothesizing and testing subsequent effects on the business network around the entrepreneur’s firm. Their conclusions indicate that entrepreneurs’ personal networks and firms’ business networks have consequences for their performance, specifically their innovation and expectations for growth of their firms.

The seventeenth chapter: “[Content Mining of Microblogs](#)” by M. Özgür Cingiz and Banu Diri considers the problem of microblog classification. In their work, they use Multinomial Naive Bayes and Support Vector Machines and show that the first one significantly outperforms the second. They also present several other experimental observations and possible uses of their results that may help other researchers in their studies.

Before concluding, we would like to thank our authors for their significant contributions and referees for their insightful constructive criticism. Finally, we would like to acknowledge the invaluable support of several Springer people who made this book possible.

Fazli Can
Tansel Özyer
Faruk Polat

State of the Art Applications of Social Network Analysis

Can, F.; Özyer, T.; Polat, F. (Eds.)

2014, XII, 372 p. 137 illus., 94 illus. in color., Hardcover

ISBN: 978-3-319-05911-2