

Models and Algorithms for Social Influence Analysis

Jimeng Sun[†] and Jie Tang[‡]

[†]IBM TJ Watson Research Center, USA

[‡]Department of Computer Science and Technology, Tsinghua University, China
jimeng@us.ibm.com, jietang@tsinghua.edu.cn

ABSTRACT

Social influence is the behavioral change of a person because of the perceived relationship with other people, organizations and society in general. Social influence has been a widely accepted phenomenon in social networks for decades. Many applications have been built based around the implicit notation of social influence between people, such as marketing, advertisement and recommendations. With the exponential growth of online social network services such as Facebook and Twitter, social influence can for the first time be measured over a large population. In this tutorial, we survey the research on social influence analysis with a focus on the computational aspects. First, we introduce how to verify the existence of social influence in various social networks. Second, we present computational models for quantifying social influence. Third, we describe how social influence can help real applications. In particular, we will focus on opinion leader finding and influence maximization for viral marketing. Finally, we apply the selected algorithms of social influence analysis on different social network data, such as twitter, arnetminer data, weibo, and slashdot forum.

Categories and Subject Descriptors

J.4 [Computer Applications]: Social Behavioral Sciences; H.1 [Information Systems]: Models and Principles

General Terms

Algorithms, Experimentation, Measurement

Keywords

Social influence, Social network, Influence Maximization

1. OUTLINE

As social networks have become popular in many domains, more and more people make decisions based on their interactions from social networks. For example, people often pick what restaurants to go to based on recommendations and reviews from Yelp. As the growth of social networks in all domains, such behaviors of social influence become more and more prevalent. More and more people make decisions and changes influenced by their social networks.

In this tutorial, we survey the theories, algorithms and applications on social influence analysis.

First, we give the definition of social influence and introduce related concepts such as homophily, conformity, and selection.

Second, we describe methodologies for verifying the existence of influence in various social networks. The methods include shuffle test and randomization test. We will give real world examples to demonstrate how the social influence behaves in different social networks.

Third, we present models and algorithms for quantitative analysis of social influence, which goes beyond macro-level analysis. Employing several large social networks (including Twitter and Weibo) as the examples in our study, we introduce how social influence affect individual behaviors and change the network structure.

Finally, we survey applications of social influence. In particular, we introduce opinion leader finding and influence maximization for viral marketing, which have many important applications in the real world. In particular, we focus on node-specific statistics for opinion leader finding. Then we describe structural effects of both edges and nodes for influence maximization. In addition, we will present methods for predicting customer behavior and online advertising through viral marketing.

To conclude, the tutorial has the following outline.

- Preliminaries
 - Definition of social influence
 - Homophily
 - Influence and Selection
 - Other related concepts
- Existential Test for Social Influence
- Computational models for Social Influence
 - Learning influence probability
 - Influence and action dynamics
 - Influence and interaction
 - Influence maximization models
- Influence Applications
 - Opinion leader finding
 - Influence maximization for viral marketing

2. BIOGRAPHY

Jimeng Sun is a research staff member at Healthcare Transformation group, IBM T. J. Watson lab. He received a Bachelor and MPhil in Computer Science from Hong Kong University of Science and Technology in 2002 and 2003. After that, he obtained a MS and PhD degree in Computer Science from Carnegie Mellon University in 2006 and 2007. His research interests include data

mining for streams and networks, databases and healthcare analytics. He has received the best research paper award in ICDM 2008, the best research paper award in SDM 2007. He has published over 40 refereed articles and two book chapters. He filed four patents and has given five tutorials. He has served as a program committee member of KDD, ICDM, SDM, CIKM, and SIGMOD, senior committee member for SDM 2010-2011, 2013, for KDD 2013.

Jie Tang is an associate professor at the Department of Computer Science and Technology, Tsinghua University. He serves as the director of Scientific Office of the Department of Computer Science. His main research interests include social network analysis and data mining. He has been a visiting scholar at Cornell University, Chinese University of Hong Kong, Hong Kong University of Science and Technology, and Leuven University. He has published over 100 research papers in major international journals and conferences including: KDD, IJCAI, WWW, SIGMOD, ACL, Machine Learning Journal, TKDD, and TKDE. He serves as Local Chair of SIGKDD'12, Publications Co-Chairs of SIGKDD'11, Program Chair of ADMA'11, Program Chair of SocInfo'12, and also serves as the PC member of more than 50 international conferences.

Acknowledgements

Jie Tang is supported by the Natural Science Foundation of China (No. 61222212, No. 61073073, No. 61170061), Chinese National Key Foundation Research (No. 60933013, No.61035004), a special fund for Fast Sharing of Science Paper in Net Era by CSTD, and National Basic Research Program of China (No. 2011CB302302).

3. REFERENCES

- [1] N. Agarwal, H. Liu, L. Tang, and P. S. Yu. Identifying the influential bloggers in a community. In *Proceedings of the 1st ACM International Conference on Web Search and Data Mining (WSDM'08)*, pages 207–217, 2008.
- [2] A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'08)*, pages 7–15, 2008.
- [3] S. Aral, L. Muchnika, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009.
- [4] E. Bakshy, B. Karrer, and L. A. Adamic. Social influence and the diffusion of user-created content. In *EC '09: Proceedings of the tenth ACM conference on Electronic commerce*, pages 325–334, New York, NY, USA, 2009. ACM.
- [5] P. Bonacich. Power and centrality: a family of measures. *American Journal of Sociology*, 92:1170–1182, 1987.
- [6] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489:295–298, 2012.
- [7] W. Chen, Y. Wang, and S. Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'09)*, pages 199–207, 2009.
- [8] R. B. Cialdini and N. J. Goldstein. Social influence: compliance and conformity. *Annu Rev Psychol*, 55:591–621, 2004.
- [9] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'08)*, pages 160–168, 2008.
- [10] P. W. Eastwick and W. L. Gardner. Is it a game? evidence for social influence in the virtual world. *Social Influence*, 4(1):18–32, 2009.
- [11] S. M. Elias and A. R. Pratkanis. Teaching social influence: Demonstrations and exercises from the discipline of social psychology. *Social Influence*, 1(2):147–162, 2006.
- [12] T. L. Fond and J. Neville. Randomization tests for distinguishing social influence and homophily effects. In *Proceeding of the 19th international conference on World Wide Web (WWW'10)*, 2010.
- [13] M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring Networks of Diffusion and Influence. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'10)*, pages 1019–1028, 2010.
- [14] A. Goyal, F. Bonchi, and L. V. Lakshmanan. Learning influence probabilities in social networks. In *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining (WSDM'10)*, pages 207–217, 2010.
- [15] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'03)*, pages 137–146, 2003.
- [16] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang. Mining Topic-Level Influence in Heterogeneous Networks. In *Proceedings of the Nineteenth Conference on Information and Knowledge Management (CIKM'2010)*, pages 199–208, 2010.
- [17] L. Liu, J. Tang, J. Han, and S. Yang. Learning Influence from Heterogeneous Social Networks. In *Data Mining and Knowledge Discovery (DMKD)*, 25(3):511–544, 2012.
- [18] M. E. J. Newman. A measure of betweenness centrality based on random walks. *Social Networks*, 2005.
- [19] J. Sun and J. Tang. A Survey of Models and Algorithms for Social Influence Analysis. *Social Network Data Analytics*, Aggarwal, C. C. (Ed.), Kluwer Academic Publishers, pages 177–214, 2011.
- [20] C. Tan, J. Tang, J. Sun, Q. Lin, and F. Wang. Social action tracking via noise tolerant time-varying factor graphs. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'10)*, pages 807–816, 2010.
- [21] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li. User-level sentiment analysis incorporating social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'11)*, pages 1397–1405, 2011.
- [22] J. Tang, J. Sun, C. Wang, and Z. Yang. Social influence analysis in large-scale networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'09)*, pages 807–816, 2009.
- [23] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, pages 440–442, Jun 1998.