

# Jian Gao |

Xiyuan Ave No.2006, West Hi-Tech Zone – Chengdu 611731, P.R. China

☎ Mobile: (86) 136-9944-5766 • ✉ gaojian08@hotmail.com

🌐 Homepage: gaocn.net • Skype: gaojianuestc

**University of Electronic Science and Technology of China (UESTC)**

October 19, 2015

CompleX Lab, Web Sciences Center, School of Computer Science and Engineering

I am a Second-year Ph.D. Student in Computer Science at UESTC with an interest in understanding phenomena in social and economic systems through empirical and theoretical analysis. I am working with Prof. Tao Zhou in solving a variety of problems, including information spreading, economic structure, reputation evaluation, etc. The deep involvement in the full process of doing a creative project, from initiating ideas to co-authoring publications, has increased my capacity for independent research. I serve as a Co-PI of a national social science foundation granted project and the investigators of three natural science foundation granted projects. My works have been published in some peer-reviewed journals, such as *Sci. Rep.*, *EPL*, *Physica A* and *Chin. Phys. Lett.*

The majority of my current research applies complex networks, graph mining, data modeling, social and economic systems. Five projects that I mainly contributed to are summarized as follows:

- Local Economic Structure. The growths of companies are direct reflections of the economy development. I intend to extract the local economic structure from data of 24 million companies in China. The basic information of company data contains listing date, delisting date, registered address, industry category, registered capital, capital structure, governance structure, etc. As a preliminary study, I am interested in the dynamic pattern of the economic structure and the interdependency of sectors. I construct the industry space based on the data and find it appears to have a core-periphery structure with two density cores, like a dumbbell. The labor/capital-intensive and technology-intensive industries respectively occupy one core. Further results show that provinces who are centralized located could have the similar industry structure of labor/capital-intensives instead of capital-intensive. This ongoing project is significant to macro-industrial layout adjustment. I am currently working on a paper and looking forward to cooperating with others on addressing this issue.

[Jian Gao, Tao Zhou, *et al.* Symbiosis and monopoly of the local industry structure in China. Oct 2015].

- Regional Economic Complexity. The economic complexity emerges from the interactions between individual activities. Companies to economic is just like cells to human. I am interested in how companies contribute to the regional economy complexity. As a startup, I try to quantify the complexity of a province's economy by characterizing the structure of "province-industry" network, which is built based on thousands of listed companies in China. Preliminary results show that diversified provinces tend to occupy less ubiquitous industry classes. Further results show that regional economic complexity is positively correlated with GDP and negatively correlated with income inequality in provinces of China. This work help us to better understand local economic complexity and its effects on society and I am currently working on a paper in this topic.

[Jian Gao, Tao Zhou, *et al.* Modeling local economy complexity via company data analysis. Oct 2015].

- Spatial Social Network. Previous empirical studies of online social networks have confirmed a spatial scaling law, namely, the probability density function (PDF) of an individual to have a friend at distance  $r$  scales as  $P(r) \sim r^\alpha$ ,  $\alpha \approx -1$ . Recent study suggested the optimization of information collection as a possible explanation for the origin of this spatial scaling law. I investigate how the spatial organization affects information spreading in spatial social network by studying bootstrap percolation (a representative networked spreading process) on Kleinberg network (a typical spatial network model). Numerical results suggest that the self-organization of spatial structure of online social networks is to some extent relevant to the principle of least effort in human behaviour. Partial results of this ongoing project have been published in *Sci. Rep.* and the poster of this paper won the Best Poster Award in the 2nd Conference on Complexity Science (Top 3), 2014.

[Jian Gao, Tao Zhou, and Yanqing Hu. Bootstrap percolation on spatial networks. *Sci. Rep.*, 5:14662, 2015].

- Online Reputation Evaluation. Individual reputation plays the role of fundamental blocks in building up online ecosystems. Meanwhile, new challenges arise that how to evaluate online reputation? Our approach is to

group users based on their rating similarities, and calculate user reputation according to the corresponding group sizes. Further, I introduce an iterative reputation-allocation process into this method, where users with higher reputation have higher weights in dominating the corresponding group sizes. The reputation of users and the corresponding group sizes are iteratively updated until they become stable. This project highlights the positive role of grouping behavior in better reputation evaluation. This project has been closed after I contributed one paper to *EPL* and submitted a manuscript to *Knowl.-Based Syst.*.

[Jian Gao, Yu-Wei Dong, Ming-Shing Shang, Shi-Min Cai, and Tao Zhou. Group-based ranking method for online rating systems with spamming attacks. *EPL*, 110: 28003, 2015]. [Jian Gao and Tao Zhou. Evaluating user reputation in online rating systems via an iterative group-based ranking method. arXiv:1509.00594, 2015].

- Intelligent Human Resource. Employees are the soul of enterprises. To predicate the trends of employees is especially significant for the human resources departments. I collect anonymous employees' work-related interactions and social connections from a social network platform developed and used by a Chinese company consisting of over a hundred employees, and further build a coupled network (AN-SN) to extract employees' behavioural trends. Results show that the in-degree in AN is the most relevant indicator for promotion while the k-shell index in AN and in-degree in SN are both very predictive to resignation. This project provides a novel and actionable understanding of enterprise management. In this ongoing project, I have published two papers in *Physica A* and *Big Data Research*, and contributed a book chapter to *Social Physics*.

[Jian Yuan, Qian-Ming Zhang, Jian Gao, *et al.* Promotion and resignation in employee networks. *Physica A*, 2015]. [Jian Gao, Linyan Zhang, Qian-Ming Zhang and Tao Zhou. Big data human resources. in *Social Physics*, Beijing: Science Press, 2014].

I have accumulated a amount of knowledge in network science, statistic physics and data analysis after over three years researching in both empirical and theoretical projects. Also, I am now practical in processing and analyzing the data in using MATLAB/Python, visualizing the results in using OriginLab, Gephi and D3, writing the papers in using  $\text{\LaTeX}$ . Past works amazed me to do further research in interdisciplinary areas and my background in mathematics and computer science should do well to my advanced study. As a Ph.D. student, I am particularly interested in figuring out the mechanisms behind the observations and providing valuable insight into the laws that govern the social and economic systems.

As future works, I am particularly interested in the following topics: the adjustment of macro-industrial layout in China, the gender bias in academic achievement, the optimization of information spreading on social networks, the adoption of social behaviors in reality, etc. We already have a wealth of data resources, such as millions of company data, millions of resume data, and hundreds of thousands of paper data. I also have plans with my co-authors at UESTC and other collaborators on future projects:

- The Way of Marie Curie. Previous results showed that women are underrepresented in many areas, such as job hunting and personal research funding application evaluations. I intend to extract the pattern of women researchers' career development and figure out the bottleneck. My approach is to explore the co-authorship network with twenty thousands authors after gender identification, focusing on the productivity and influence.
- The Urban Viscosity. As areas perception has a significant impact on the talent recruitment and site selection, I intend to focus on the investigation of the urban viscosity, which indirectly contributes to the development of local economic development. My approach is to extract the pattern of place changing for employees from millions of resume data, considering the places where they were born, grew up, attended school and worked.
- Pairwise Vertex Similarity. Pairwise vertex similarity is a fundamental index for many network functions and dynamics, which has found applications in link predication, community detection, spreading activation, etc. I intend to summary previous literatures and further give an empirical review by studying the characteristics of the existed pairwise vertex similarity indices, such as the stability, similarity, classification, etc.

To summarize, at this point in my studying, my primary interests are in understanding the underlying mechanisms behind the phenomena in social and economic systems. I have pursued my interests to date by accumulating a amount of knowledge in network science, statistic physics and data analysis. In the immediate future I would like to go further with graph mining by initiating both empirical and theoretical projects. In the long term I see myself working on the interdisciplinary field of network science, statistic physics and data mining. To be a distinguished young scholar, I am looking forward to exploring new and exciting areas.