Dynamic Implicit Social Recommendation

Qin Zhang

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

Doctor of Philosophy

January 2018

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student

Production Note:
Signature removed prior to publication.

Acknowledgements

I would like to express my earnest thanks to my supervisor, Professor Chengqi Zhang, and co-supervisors, Doctor Peng Zhang, Doctor Guodong Long and Doctor Jing Jiang. They have provided tremendous support and guidance for my research and my life in the past three and half years.

I would like to express my sincere appreciation to Professor Chengqi Zhang, for his wise guidance and supervisions during my Ph.D study. Professor Chengqi Zhang provided me an opportunity to study in the Center for Artificial Intelligence (which used to be the Center for Quantum Computation and Intelligent Systems), where I met and learnt a lot from many smart and sharp people. I benefited significantly from his unselfish help and valuable suggestion on my research career.

I would like to thank Doctor Peng Zhang who encouraged me staying in academia after I got my Master degree from the University of Chinese Academy of Sciences. Doctor Peng Zhang got me start to love research and taught me how to write an academic paper step by step with his endless patience. His great passion on research also impacted me so much and made me longing to learn new technologies. I could remember the time we discuss the research details, mainly because his broad knowledge background often motivated me to target some interesting fields, which were always exciting experiences.

I also would like to thank Doctor Guodong Long, who gave me a different point of view on how to observe the world, how to know more about myself and how to plan my career. The conversations with him always benefited me so much and gave me a new perspective

on the problems. His vision, creativeness and enthusiasm in solving challenging problems both in research and in projects have greatly encouraged me and inspired my works. I also want to give thanks to him as a friend of mine, for giving me invaluable instructions and suggestions during my life.

I would like to thank Doctor Jing Jiang who gave me a lot of help even before I came to UTS. From her, I learnt a lot on time and energy management. I benefited so much from various discussions and communications with her. She gave me advices, encouragement on the difficult problems in research and in life.

I also wish to express my sincere appreciation to Professor Wei Wang for all the support and help during my visiting in University of California, Los Angeles (UCLA). The guidance and support from her and her team benefited me so much. Her persistent enthusiasm on research and life impressed me so much. I learnt a lot from her, not only on research, but also on how to explore the real-would problem for great ideas, on how to keep firm and tenacious on what we love, and on how to manage well a big team.

I would like to give special thanks to Professor Ivor Tsang. I was so fortunate to spend time and cooperated with him. Discussing a problem and writing a paper with him have been always a pleasure and eye-opening experience. He gave me sufficient freedom and encouragement to think and explore my research interests.

I would also like to thank all the people that had a positive influence on my day-to-day enjoyment of the job: Jia Wu, Shuirui Pan, Haishuai Wang, Maoying Qiao, Tongliang Liu, Zhibin Hong, Shaoli Huang, Sujuan Hou, Zhiguo Long, Zhe Xu, Bozhong Liu, Bo Du, Daokun Zhang, Chaoyue Wang, Shan Xue, Xiaolin Zhang, Ruizhi Zhou, Yue Ma, Li Gao, Guoqing Xu, Stephanie Li, Grace Lin, Christina Tong, FuChien Tseng, Paulina Rivera, Teigan Valenzuela. They are the ones who have given me support during both joyful and stressful times, to whom I will always be thankful. Without their endless patience,

generous support, and constant guidance, this thesis could not have been accomplished.

Finally, and above all, I want to thank my family for their continuous support. I especially thank my parents for their unconditional encouragement and support, both emotionally and financially. No words could possibly express my deepest gratitude for their endless love, self-sacrifice and unwavering help. To them I dedicate this dissertation.

Publications

In Journals

1. **Qin Zhang**, Peng Zhang, Guodong Long, Wei Ding, Chengqi Zhang and Xindong Wu. Online learning from trapezoidal data streams. *IEEE Transactions on Knowledge and Data Engineering (**TKDE**)*, vol. 28(10), pp. 2709-2723, 2016. (Australia ERA Ranked **A**, Core ranked **A**)

2. **Qin Zhang**, Jia Wu, Peng Zhang, Guodong Long and Chengqi Zhang. Discriminative subsequence mining for time series clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)*, **Under third-round review**, 2018. (Australia ERA Ranked **A***, Core ranked **A***)

3. Qinzhe Zhang, Jia Wu, **Qin Zhang**, Peng Zhang, Guodong Long and Chengqi Zhang. Dual influence embedded social recommendation. *World Wide Web Journal (**WWWJ**)*, pp. 1-26, 2017. (Australia ERA Ranked **A**, Core ranked **A**)

In Conferences

4. **Qin Zhang**, Jia Wu, Hong Yang, Yingjie Tian and Chengqi Zhang. Unsupervised feature learning from time series. *26rd International Joint Conference on Artificial Intelligence (**IJCAI**)*, pp. 2322-2328, 2016. (Australia ERA Ranked **A**, Core ranked **A***)

5. **Qin Zhang**, Jia Wu, Peng Zhang, Guodong Long, Ivor W. Tsang and Chengqi Zhang. Inferring latent network from cascade data for dynamic social recommendation. *International Conference on Data Mining (**ICDM**)*, pp. 669-678, 2016. (Australia ERA Ranked **A**, Core ranked **A***)

6. **Qin Zhang**, Peng Zhang, Guodong Long, Wei Ding, Chengqi Zhang and Xindong Wu. Towards mining trapezoidal data streams. *International Conference on Data Mining (**ICDM**)*, pp. 1111-1116, 2015. (Australia ERA Ranked **A**, Core ranked **A***)

7. Qinzhe Zhang, **Qin Zhang**, Guodong Long, Peng Zhang and Chengqi Zhang. Exploring heterogeneous product networks for discovering collective marketing hyping behavior. *Pacific-Asia Conference on Knowledge Discovery and Data Mining (**PAKDD**)*, pp. 40-51, 2016. (Australia ERA Ranked **A**, Core ranked **A**)

8. Li Gao, Yao Lu, **Qin Zhang**, Weixue Lu and Yue Hu. Query expansion for exploratory search with subtopic discovery in community question answering. *International Joint Conference on Neural Networks **IJCNN***, pp. 4715-4720, 2016. (Australia ERA Ranked **A**, Core ranked **A**)

Abstract

Due to the potential value of social relations [125] [34], social recommendation has attracted a lot of attention recently in the research communities.

Modelling time drifting data is a central problem in this area. Temporal changes, such as the emergence of new products or services and the changes in users' purchase behaviour pattern, bring unique challenges. The need for modelling time changes at the level of each individual significantly reduces the amount of available data for detecting such changes. Thus we should resort to more accurate techniques that suffice the dynamic nature of social recommendations.

Besides, the essence of social recommendation methods is utilizing users explicit social connections to improve recommendation results. However, explicit social connection information is not always available in real-world recommender systems. Only few Web sites have implemented the social or trust mechanisms, like Epinions and Douban. Lacking social recommendation data greatly limits the impact and utilization of social recommendation methods. Fortunately, in case that we do not have explicit social information, we can always compute a set of implicit social information to improve the recommendation performance.

To this end, we propose a new Implicit Social Recommendation(ISR) model from cascade data (a variant of time series) in Chapter 3, which makes recommendations based on the inferred latent social network. It can sufficiently mine the information contained in time by mining the cascade data and identify the dynamic changes in the users in time by using the latest updated social network to make recommendations.

Further, since the temporal information is crucial in social recommendation, we explore the temporal behaviour pattern from users' time series data in Chapter 4. The main challenge is finding the discriminative and explainable features (shapelets) that can best represent the raw time series data. We build an economical shapelet learning model that can automatically-learn shapelets without labels. Specifically, a novel Unsupervised Discriminative Subsequence Mining (UDSM) model that automatically-learn shapelets from unlabelled time series data.

Another challenge in social recommendation is the continually appeared new users and new items. In real-world application, this problem is unavoidable. How to involve this kind of trapezoidal data streams and how to handle this problem smoothly must be considered. Most of the existing research either builds a general model in a sample scenario such as ignoring the dynamic nature of the system and the problem of new users and items, or only focuses on some specific fields, such as dynamic restaurant recommendation and timely news recommendation. We propose a new Online Learning with Streaming Features (OLSF) algorithm and its two variants OLSF-I and OLSF-II for mining trapezoidal data streams in Chapter 5. OLSF and its variants combine online learning and streaming feature (i.e. continuously appeared new items) together to handle the double-streaming data.

Finally, we introduce a novel General Online Dynamic Social Recommendation(GODSR) model in Chapter 6, which combines network inference from cascade data, double-streaming users and items, and collaborative filtering together in an iterative process. By inferring the latent dynamic core social networks from cascade data, identifying the drift of a users preferences and involving new users and items in the online learning process, GODSR method reacts rapidly and makes accurate recommendations to users when new data arrive.

Contents

List of Figures

List of Tables