# Cross-domain Recommender System Through Tag-based Models

**Peng Hao**

A thesis submitted for the Degree of

*Doctor of Philosophy*

Faulty of Engineering and Information Technology

University of Technology Sydney

March 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 candidate:

Production Note:
Signature removed prior to publication.

Date:

05/04/2018

# Acknowledgements

This thesis is the result of four years hard work, during which I received a lot of help from many people. So in there I would like to show my gratitude to all of them.

First, I would like to thank my principle supervisor, A/Prof. Guangquan Zhang, for offering me an opportunity to conduct my research in the Decision Systems and e-Service Intelligence (DeSI) Lab, University of Technology Sydney (UTS), Australia. I can still remember the excitement I had when I got the official offer from UTS. At the beginning of the research, he encouraged me to choose the topic that I am interested in. He also taught me how to approach a research problem in general, and was always enthusiastic to help solve the difficulties in my life. I would also like to express my thanks to my co-supervisor, Distinguished Professor Jie Lu. I have learned a lot from her over these years, not only the methodology of doing research but also the skills of writing a scientific paper. Her comments and suggestions have strengthened this thesis significantly. Her continuous hard work and generous personality have influenced me deeply, and will be a great treasure in my future research and work. I feel very lucky to have both of them as my supervisors. Without their excellent supervision and continuous encouragement, this research could not be finished on time.

Special thanks also goes to Prof. Luis Martinez for welcoming me and providing all the necessary help during my visit to the University of Jaen. I had spent a wonderful time with the collaboration of him. The beautiful scene and quiet life have attracted me to visit there again in the future.

I would also express my appreciation to all the members of the DeSI Lab in the Centre of Artificial Intelligence (CAI), for their active participation and valuable comments in every presentation I made during my study.

I also wish to thank the financial support I received from the China Scholarship Council (CSC) and UTS, which support me to finish my study.

Finally, I would like to thank my family members, including my wife, father, mother and mother-in-law. Without their great love, conscious encouragement and infinite compassion, my dream of pursuing a Ph.D. degree will not be achieved. Very special thanks to them all!

# Abstract

Nowadays, data pertaining to clients are generated at such a rapid rate it is completely beyond the processing ability of a human, which leads to a problem called *information explosion*. How to quickly and automatically provide personalized choices for someone from a large collection of resources has become a key factor in determining the success of many commercial activities. In this context, *recommender systems* have been developed as a type of software that aims to predict and suggest items which are relevant to a specific user by analyzing the user's previous interaction data with certain items. Recommender systems have a broad application in our daily life, such as product recommendation in Amazon, video and movie recommendation in Youtube, music recommendation in Spotify.

A fundamental brick in building most recommender systems is the *collaborative filtering*-based model, which has been widely adopted due to its outstanding performance and flexible deployment. However, this model together and its variations suffer from the so-called *data sparsity* problem, which results when user sonly rate a limited number of items. With the development of the *transfer learning* technique in recent years, cross-domain recommendation has emerged as an effective way to address data sparsity in recommender systems. The principle

of cross-domain recommendation is to exploit knowledge from auxiliary source domains to assist recommendation making in a sparse target domain.

In the development of cross-domain recommender systems, the most important step is to build a bridge between the domains in order to transfer knowledge. This task becomes more challenging in disjoint domains where users and items in both domains are completely non-overlapping. In this respect, tags are studied and utilized to establish explicit correspondence between domains. However, how to effectively exploit tags to increase domain overlap and ultimate recommendation quality remains as an open challenge which needs to be addressed.

This thesis aims to develop novel tag-based cross-domain recommendation models in disjoint domains. First, it review the existing state-of-the-art techniques related to this research. It then provides three solutions by exploiting domain-specific tags, tag-inferred structural knowledge and tag semantics, respectively. To evaluate the proposed models, this thesis conducts a series of experiments on public datasets and compare them with state-of-the-art baseline approaches. The experimental results show the superior performance achieved by our models in different recommendation tasks under sparse settings. The findings of this research not only contribute to the state-of-the-art on cross-domain recommender systems, but also provide practical guidance for handling unstructured tag data in recommendation tasks.

# Table of contents

# List of figures

# List of tables