FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY

Tree Similarity Measure-based Recommender Systems

Dianshuang Wu

A thesis submitted for the Degree of

Doctor of Philosophy

University of Technology, Sydney
October, 2014

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.

ACKNOWLEDGEMENTS

I would like to express my earnest thanks to my principal supervisor, Professor Guangquan Zhang, and my co-supervisor, Professor Jie Lu. Their comprehensive guidance has covered all aspects of my PhD study, including research methodology, research topic selection, experiments, academic writing skills and thesis writing, and even the sentence structure and formulas. Their critical comments and suggestions have strengthened my study significantly. Their strict academic attitude and respectful personality have benefited my PhD study and will be a great treasure throughout my life. Without their excellent supervision and continuous encouragement, this research could not have been finished on time. Thanks to you all for your kind help.

I am grateful to all members of the Decision Systems and e-Service Intelligent (DeSI) Lab in the centre for Quantum Computation and Intelligent Systems (QCIS) for their careful participation in my presentation and valuable comments for my research. I especially thank PhD students Mr Mingsong Mao, Mr Wei Wang, Master student Mr Yushi Zhou and other students for their contributions in the development of the recommender system softwares related to this study.

I would like to thank Ms. Barbara Munday and Ms. Sue Felix for helping me to correct English presentation problems in my publications and this thesis.

I am grateful to FEIT Travel fund, Vice-Chancellor's Postgraduate Conference Fund, and the International Postgraduate Research Scholarship (IPRS) and Australian Postgraduate Award (APA) scholarship.

Last but not least, I would like also to thank my family members. Thanks to my mother and father for their conscious encouragement and generous support.

ABSTRACT

The rapid growth of web information provides excellent opportunities for developing e-services in many applications but also caused increasingly severe information overload problems whereby users are not able to locate relevant information to exactly meet their needs efficiently by using the current Internet search functions. A personalised recommender system aims to handle this issue.

A big challenge in current recommender system research is: the items and user profiles in many recommender system applications nowadays, such as the e-business and e-learning recommender systems, are so complex that they can only be described in complicated tree structures. Therefore, the item or user similarity measure, as the core technique of the recommendation approach, becomes a tree similarity measure, which existing recommender systems cannot provide. Another challenge is that in many real life situations, online recommendations to customers in selecting the most suitable products/services are often made under incomplete and uncertain information, which needs fuzzy set theory and techniques to deal with. Thus, how to use fuzzy set techniques to handle data uncertainty issues in tree-structured items or user profiles needs to be investigated.

This research aims to handle these two challenges in both theoretical and practical aspects. It first defines a tree-structured data model, which can be used to model tree-structured items, user profiles and user preferences. A comprehensive similarity measure on tree-structured data considering all the information on tree structures, nodes' concepts, weights and values is then developed, which can be used to compute the semantic similarity between tree-structured items or users, and the matching degree of items to tree-structured user requests. Based on the tree-structured data model, the tree-structured

items and user requirements are modelled as item trees and user request trees respectively. An item tree and user request tree-based hybrid recommendation approach is then developed. To model users' fuzzy tree-structured preferences, a fuzzy preference tree model is proposed. A fuzzy preference tree-based recommendation approach is then developed. Experimental results on an Australian business dataset and the Movielens dataset show that the proposed recommendation approaches have good performance and are well-suited in dealing with tree-structured data in recommender systems. By use of the proposed tree similarity measure and recommendation approaches based on that, two real world applications, a business partner recommender system, Smart BizSeeker, and an e-learning recommender system, ELRS, are designed and implemented, which demonstrate the applicability and effectiveness of the proposed approaches.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES