

MULTI-COLLABRATIVE FILTERING TRUST NETWORK
FOR ONLINE RECOMMENDATION SYSTEMS

by

CHEN WEI, Jenny

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science in E-Commerce Technology

Faculty of Science and Technology
University of Macau

2011

Approved by _____
Supervisor

Date _____

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Macau, I agree that the Library and the Faculty of Science and Technology shall make its copies freely available for inspection. However, reproduction for this thesis for any purposes or by any means shall not be allowed without my written permission. Authorization is sought by contacting the author at

Telephone: (853) 6224 8078
E-mail: jennywei1987@gmail.com

Signature _____

Date _____



Abstract

Nowadays, Recommendation Systems (RS) play an important role in the e-Commerce business and they have been proposed to exploit the potential of social networks by filtering information and offering useful recommendations to customers. As the personalization service is built to present the users with highly relevant set of items, the customer loyalty of the web companies can be improved.

Collaborative Filtering (CF) is believed to be a suitable underlying technique for recommendation systems based on social networks, since it harvests information both from similar products and from peer users to infer a suggested item out of many for a user. Meanwhile, social networks provide the needed collaborative social environment.

CF and its variants have been studied extensively in the literature on online recommender, marketing and advertising. However, most of the works were based on Web 1.0 where all the information necessary for the computation is by default assumed to be always available, as if it were readily stored in a database. In the distributed environment of Web 2.0 such as social networks, the required information by CF may either be incomplete or scattered over different sources. This poses new computational changes for Web 2.0 recommender.

The system we proposed here is the Multi-Collaborative Filtering Trust Network Recommendation System, which combined multiple sources, by using MovieLens, Delicious and Facebook datasets, measured trust, temporal relation and similarity factors. After series of experiments, we found that the performance of recommendation system with considering above four aspects is much better than considering any other single/combined aspects.

Keywords: *Recommendations, Multi-Collaborative Filtering Trust Network*

Acknowledgments

Firstly, I would like to express my gratitude to my thesis supervisor, Dr. Simon Fong, for his patience and guidance throughout the whole process of my research. I am also grateful to him for lending me the reference books, related topics of papers and journals. Furthermore, he is a great friend. He provided with many useful ideas to me, and gave me confidence in proceeding with my work and encouraged me when I felt frustrated. Without his advice and efforts, I would not be able to finish this thesis.

Also thank University of Macau, which provides a good environment and facilities for developing. In addition I am grateful to other professors and lecturers who have taught me during my study. They have all been kind and eager to help.

Finally, I would like to thank my parents, who have always been supported me.

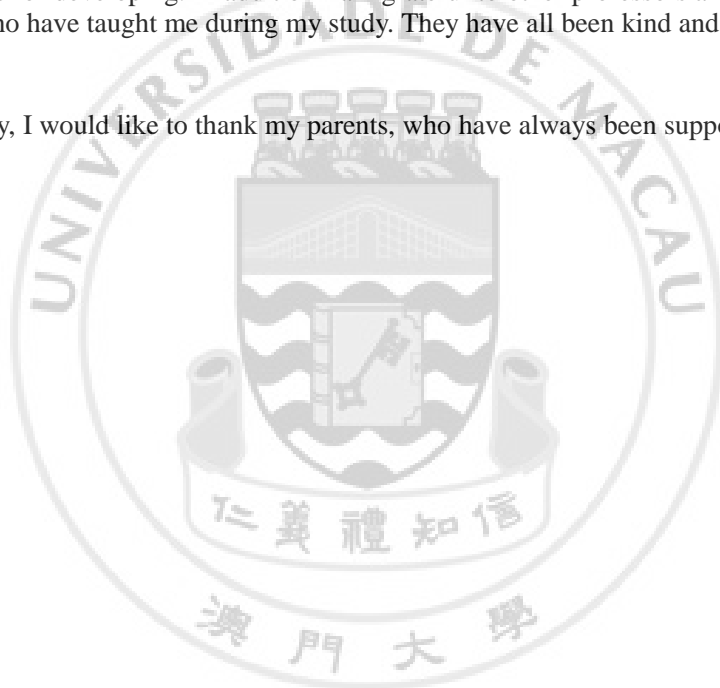
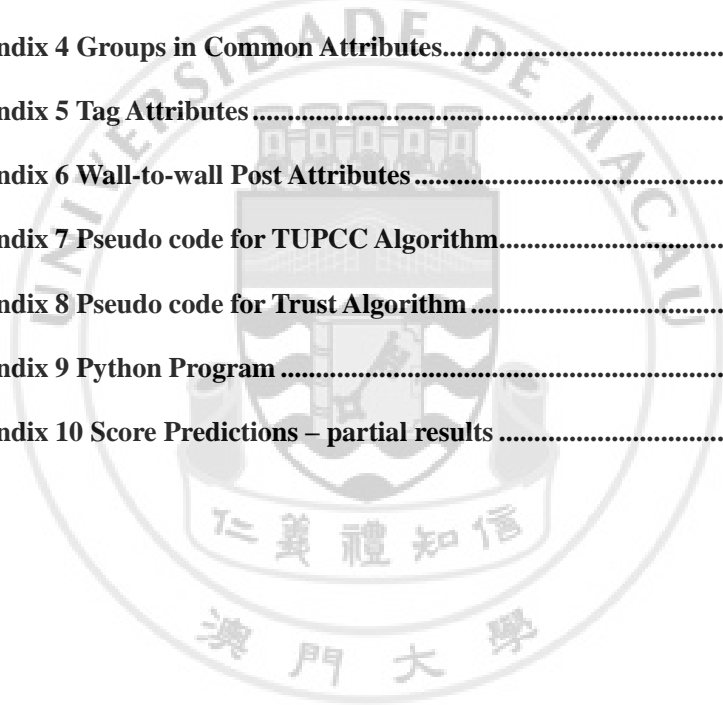


Table of Contents

Chapter 1 Introduction	1
1.1 Motivation and Objectives	1
1.2 Background	1
1.2.1 Recommendation Systems	1
1.2.2 Social Network.....	3
Chapter 2 Literature Review	8
2.1 Current Work Comparison.....	8
2.2 Critical Challenges	11
Chapter 3 Application Scenario.....	13
3.1 Workflow.....	13
3.2 Application Scenario	14
Chapter 4 MCFTN Model and MCFTN Algorithm	25
4.1 Similarity	25
4.1.1 UPCC and IPCC.....	26
4.1.2 Similarity with Temporal Relation.....	27
4.2 Trust.....	28
4.2.1 Trust by relation	28
4.2.2 Trust by reputation	29
4.2.3 Estimating the weights	30
4.3 Proposed Model and Final Algorithm	37
Chapter 5 Empirical Analysis	39
5.1 Experimental Setup	39
5.1.1 Data Preparation.....	39
5.1.2 Metrics	41
5.2 Assumptions	41
5.3 Experiment Evaluation	42
5.3.1 Similarity with/without temporal relation	42
5.3.2 Similarity with temporal relation with/without trust.....	46
5.3.3 Influence of parameters α, β	49
5.3.4 Influence of parameters w_s, w_T	52
5.4 Limitations	56
Chapter 6 Conclusion and Future Works.....	58

6.1 Summary	58
6.2 Future Works	58
References.....	59
Publication.....	61
Appendix 1 Movie Lens Dataset Attributes.....	62
Appendix 2 Personal Information Attributes.....	63
Appendix 3 Number of Friends Attributes.....	64
Appendix 4 Groups in Common Attributes.....	65
Appendix 5 Tag Attributes	66
Appendix 6 Wall-to-wall Post Attributes	67
Appendix 7 Pseudo code for TUPCC Algorithm.....	68
Appendix 8 Pseudo code for Trust Algorithm	69
Appendix 9 Python Program	70
Appendix 10 Score Predictions – partial results	83



List of Figures

Figure 1- 1 Recommendation Systems example.....	2
Figure 1- 2 Visualization of friends on Facebook by strengths of relations	4
Figure 1- 3 Selected social network sites unique visitor trend.....	6
Figure 3- 1 Workflow.....	13
Figure 3- 2 Like button plugin code	14
Figure 3- 3 Connection to Facebook.....	15
Figure 3- 4 Facebook connects to Yahoo!.....	16
Figure 3- 5 Goiaba apps example	17
Figure 3- 6 Goiaba Apps basic information	18
Figure 3- 7 heroku cloud platform for apps	19
Figure 3- 8 Host your site with Heroku	20
Figure 3- 9 Goiaba request for Permission	21
Figure 3- 10 Apps Scenario	22
Figure 3- 11 Post to your wall through your apps.....	23
Figure 3- 12 News Feed on my Facebook account.....	24
Figure 4- 1 The distribution of the predictive power of the seven tie strength dimensions as part of the how-strong model. Source: [14]31	31
Figure 4- 2 Hierarchical trust metrics model for Facebook	32
Figure 4- 3 Lift charts of the decision tree models by different groups of factors	35
Figure 4- 4 A MCFTN model that shows interactivity relations between users and products in a site internally and also shows external relations of users and products across other websites.....	38
Figure 5- 1 MAE Performance Comparisons-with/without temporal relation	43
Figure 5- 2 MAE Improvement by considering temporal relation	44
Figure 5- 3 RMSD Performance Comparisons-with/without temporal relation	45
Figure 5- 4 RMSD Improvement by considering temporal relation	46
Figure 5- 5 MAE Performance- without trust.....	48
Figure 5- 6 RMSD Performance- without trust	49
Figure 5- 7 Impact of Parameters α, β on MAE.....	51
Figure 5- 8 Impact of Parameters α, β on RMSD.....	52
Figure 5- 9 Impact of Parameters w_S, w_T on MAE	55
Figure 5- 10 Impact of Parameters w_S, w_T on RMSD	56

List of Tables

Table 1- 1 Facebook Growth of Top 10 Social Network markets Worldwide.....	5
Table 1- 2 Visitations to Twitter.com by Global Regions	6
Table 2- 1 Areas of Work Comparison 1-by work	10
Table 2- 2 Areas of Work Comparison 2 –by algorithms.....	11
Table 4- 1 Users and Attributes table for calculating similarity.....	26
Table 4- 2 Different weights assigned to each attributes	26
Table 4- 3 An example mapping of JUMP counts and tagging tiers.....	28
Table 4- 4 Trust by Reputation Attributes.....	30
Table 4- 5 Attributes used in dataset	34
Table 5- 1 Statistics of Dataset MovieLens	40
Table 5- 2 Attributes of dataset MovieLens.....	40
Table 5- 3 Personal information attributes.....	40
Table 5- 4 Number of Friends attributes	40
Table 5- 5 Groups in Common attributes.....	40
Table 5- 6 Wall-to-wall attributes	41
Table 5- 7 Tag attributes	41
Table 5- 8 MAE Performance Comparisons.....	43
Table 5- 9 RMSD Performance Comparisons	44
Table 5- 10 MAE Performance Comparisons.....	47
Table 5- 11 RMSD Performance Comparisons.....	47
Table 5- 12 Impact of Parameter on MAE.....	50
Table 5- 13 Impact of Parameter on MAE.....	50
Table 5- 14 Impact of Parameter on RMSD	51
Table 5- 15 Impact of Parameter on RMSD	52
Table 5- 16 Impact of Parameter on MAE.....	53
Table 5- 17 Impact of Parameter on MAE.....	53
Table 5- 18 Impact of Parameter on MAE.....	53
Table 5- 19 Impact of Parameter on RMSD	54
Table 5- 20 Impact of Parameter on RMSD	54
Table 5- 21 Impact of Parameter on RMSD	54

List of Abbreviations

CF	Collaborative Filtering
M-CFTN	Multiple Collaborative Filtering Trust Network
KNN	K-nearest neighbors
UPCC	User-based Pearson Correlation Coefficient
IPCC	Item-based Pearson Correlation Coefficient
TUPCC	Temporal related User-based Pearson Correlation Coefficient
RWR	Random Walk Recommender
RWRR	Random Walk Restart Recommender
PrefAggr	Preference Relation based approach
MRR	Mean-Reciprocal Rank
ARP	Average Relative Position
Time SVD	Time-Aware Matrix Factorization Model
TARM	Trust-Aware Recommendation Model

