

UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Engineering and Information Technology

**NON-IID RECOMMENDER SYSTEMS:
A MACHINE LEARNING APPROACH**

by

Liang Hu

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

Sydney, Australia

2018

Certificate of Original Authorship

I, Liang Hu declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Advanced Analytics Institute at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This thesis is the result of a research candidature conducted jointly with Shanghai Jiao Tong University as part of a collaborative Doctoral degree.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by an Australian Government Research Training Program Scholarship.

Production Note:
Signature removed
Signature: prior to publication.

Date: **March 28, 2019**

ABSTRACT

NON-IID RECOMMENDER SYSTEMS: A MACHINE LEARNING APPROACH

by

Liang Hu

A recommender system (RS) comprises the core software, tools, and techniques that effectively and efficiently cope with information overload as well as locate information that is genuinely required. As one of the most widely used artificial intelligence (AI) systems, RSs have been integrated into daily life over the past two decades. In recent decade, the machine learning approach has dominated AI research in almost all areas. Therefore, modeling advanced RSs using the machine learning approach forms the basic methodology of this thesis.

Current RSs suffer from many problems, such as data sparsity and cold start, because they fail to consider the non-IIDness in data, which includes the heterogeneities and coupled relations within and between users and items, as well as their interactions. Thus, we propose non-IID recommender systems by modeling the non-IIDness in recommendation data with the machine learning approach. Specifically, we study non-IID RS modeling techniques from three perspectives: users, items, and interactions. This research not only promotes the design of new machine learning models and algorithms in theory, but also extensively influences the evolution of technology and society.

To construct the non-IID RS from a user perspective, we jointly model two aspects: (1) the heterogeneities of users and (2) the coupling between users. Specifically, we study the non-IID user modeling in two representative RSs: (1) a group-based RS (GBRS) and (2) a social network-based RS (SNRS). First, we perform an in-depth analysis of existing GBRSs and demonstrate their deficiencies in modeling

the heterogeneity and coupling between group members for making group decisions. A deep neural network is designed to learn a group preference representation, which jointly considers all members' heterogeneous preferences. Second, we model an SNRS by modeling the influential contexts that embed the influence of relevant users and items, because a user's selection is largely influenced by other users with social relationships.

To construct the non-IID RS from an item perspective, we target two modeling aspects: (1) the heterogeneities of items and (2) the coupling between items. Specifically, we study the non-IID item modeling in two representative RSs: (1) a cross-domain RS (CDRS) and (2) a session-based RS (SBRs). First, existing CDRSs may fail to conduct cross-domain transfer because of domain heterogeneity; thus, we propose an irregular tensor factorization model, which can more effectively capture the coupling between heterogeneous domains with learning the domain factors for each domain. Second, we construct an effective and efficient personalized SBRs to more effectively capture the couplings between items by modeling intra- and inter-session contexts.

To construct the non-IID RS from an interaction perspective, we target two modeling aspects: (1) the heterogeneities of interactions and (2) the coupling between interactions. Specifically, we study the non-IID interaction modeling in two representative RSs: (1) a multi-objective RS (MORS) and (2) an attraction-based RS (ABRS). First, we study an MORS to tackle the challenges of recommendation for users and items in the long tail. Subsequently, a coupled regularization model is proposed to jointly optimize two objectives: the credibility and specialty. Existing content-based RSs can recommend new content according to similarity; however, they are not capable of interpreting the attraction points in user-item interactions. Therefore, to construct an interpretable content-based RS, we propose attraction modeling to learn and track user attractiveness.

In the last section, we summarize the contributions of our work and present the future directions that can improve and extend the non-IID RS.

Dedication

To my parents, my wife, and my son.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Longbing Cao for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of my research papers and this thesis. His advice on both research as well as on my career have been priceless. Prof. Longbing Cao shows strong motivation, ambition, and dedication in his work and research, which impressed me very much. Moreover, I'd like to thank my co-supervisor Prof. Guandong Xu. With his invitation, I first time came to Australia and visited Advanced Analytics Institute (AAI), University of Technology Sydney. His strong sociability shows me a great example of career development.

I would also like to thank my committee members, Professor Ling Chen, Professor Dong Xu, and Professor Jinyan Li for serving as my committee members even at hardship. I also want to thank you for letting my candidate assessment be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

A special thanks to my family. Words cannot express how grateful I am to my mother, and father for all of the sacrifices that you have made on my behalf. Your prayer for me was what sustained me thus far. In the end, I would like express appreciation to my wife who was always my strongest back support. Thanks for her great efforts to and take care of my son. This thesis would not have been possible without their warm love, continued patience, and endless support.

I would like to thank all my lab mates for their continued support. I appreciate the discussion with Shoujin Wang, visiting Ph.D. student Xin Li and Sheng Luo, and visiting Professor Wenpeng Lu. Moreover, I thank all my coauthors in the lab, including Shoujin Wang, Qianqian Chen and Wei Cao. Especially, I'd like to thank Songlei Jian for coauthoring top-rank papers and co-tutoring in top-rank

conferences. I am impressed by her brilliance, diligence, optimism and great comprehension, and we got a lot of inspirations from our discussion. I hope all of them have a happy life and successful career.

Last but not least, there are still many other people that I would like to express my gratitude to them for their contribution during my Ph.D. study.

Liang Hu
Sydney, Australia, 2018.

List of Publications

Journal Papers

- J-1. Hu, L., Chen, Q., Jian, S., Cao, L., and Cao, J. Neural Cross-session Filtering: Next-item Prediction under Intra- and Inter-session Context. *IEEE Intelligent System*, vol. 33, pp. 57-67, 2018.
- J-2. Hu, L., Cao, L., Cao, J., Gu, Z., Xu, G., and Wang, J. Improving the Quality of Recommendations for Users and Items in the Tail of Distribution. *ACM Trans. Inf. Syst.* 35, 3, 1-37, 2017.
- J-3. Hu, L., Cao, L., Cao, J., Gu, Z., Xu, G., and Yang, D. Learning Informative Priors from Heterogeneous Domains to Improve Recommendation in Cold-Start User Domains. *ACM Trans. Inf. Syst.* 35, 2, 1-37, 2016.
- J-4. Hu, L., Cao, J., and Gu, Z. . Service Discovery and Recommendation in Rough Hierarchical Granular Space. *Advanced Science Letters*, 7(1), 135-139, 2012.
- J-5. Yang, D., Guo, J., Wang, Z.-J., Wang, Y., Zhang, J., Hu, L., Yin, J., and Cao, J. FastPM: An approach to pattern matching via distributed stream processing. *Information Sciences* 453, 263-280, 2018.
- J-6. Cao, J., Xu, W., Hu, L., Wang, J. and Li, M.. A social-aware service recommendation approach for mashup creation. *International Journal of Web Services Research (IJWSR)*, 10(1), pp.53-72, 2013.
- J-7. Hu, L., Jian, S., Cao, L., and Chen, Q.. Attraction Modeling: Enhancing Content-based Recommender Systems with Statistical Interpretability. (To be submit to AIJ)

Conference Papers

- C-1. Hu, L., Jian, S., Cao, L., Chen, Q., and Z. Gu, "HERS: Modeling Influential Contexts with Heterogeneous Relations for Sparse and Cold-start Recommendation," in AAAI-19, 2019.
- C-2. Hu, L., Jian, S., Cao, L., and Chen, Q. Interpretable Recommendation via Attraction Modeling: Learning Multilevel Attractiveness over Multimodal Movie Contents. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, 3400-3406, 2018.
- C-3. Hu, L., Cao, L., Wang, S., Xu, G., Cao, J., and Gu, Z. Diversifying Personalized Recommendation with User-session Context. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 1858–1864, 2017.
- C-4. Hu, L., Cao, W., Cao, J., Xu, G., Cao, L., and Gu, Z. Bayesian Heteroskedastic Choice Modeling on Non-identically Distributed Linkages. In Data Mining (ICDM), 2014 IEEE International Conference on, 851-856, 2014.
- C-5. Hu, L., Cao, J., Xu, G., Cao, L., Gu, Z., and Cao, W. Deep Modeling of Group Preferences for Group-Based Recommendation. In Twenty-Eighth AAAI Conference on Artificial Intelligence, 1861-1867, 2014.
- C-6. Hu, L., Cao, J., Xu, G., Cao, L., Gu, Z. and Zhu, C. Personalized recommendation via cross-domain triadic factorization. In Proceedings of the 22nd international conference on World Wide Web (pp. 595-606). ACM, 2013.
- C-7. Hu, L., Cao, J., Xu, G., Wang, J., Gu, Z. and Cao, L. Cross-domain collaborative filtering via bilinear multilevel analysis. In International Joint Conference on Artificial Intelligence. IJCAI-13, 2013.
- C-8. Hu, L., Cao, J., Xu, G. and Gu, Z.. Latent Informative Links Detection. In KES (pp. 1233-1242), 2012.
- C-9. Hu, L., Cao, J. and Gu, Z.. Service Discovery and Recommendation in Rough Hierarchical Granular Computing. In Services Computing Conference (APSCC), 2010 IEEE Asia-Pacific (pp. 191-198). IEEE, 2010, December.

- C-10. Jian, S., Hu, L., Cao, L., Lu, K., and Gao H., "Evolutionarily Learning Multi-aspect Interactions and Influences from Network Structure and Node Content," in AAAI-19, 2019.
- C-11. Jian, S., Hu, L., Cao, L., and Lu, K. Metric-Based Auto-Instructor for Learning Mixed Data Representation. In AAAI-18, 2018.
- C-12. Wang, S., Hu, L., Cao, L., Huang, X., Lian, D. and Liu, W.. Attention-based transactional context embedding for next-item recommendation. AAAI, 2018.
- C-13. Wang, S., Hu, L., and Cao, L. Perceiving the Next Choice with Comprehensive Transaction Embeddings for Online Recommendation. In Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18-22, 2017.
- C-14. Cao, W., Hu, L., and Cao, L. Deep Modeling Complex Couplings within Financial Markets. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2518-2524, 2015.
- C-15. Chen, Q., Hu, L., Xu, J., Liu, W., and Cao, L. Document similarity analysis via involving both explicit and implicit semantic couplings. In Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on, 1-10, 2015.
- C-16. Xu, W., Cao, J., Hu, L., Wang, J. and Li, M., June. A social-aware service recommendation approach for mashup creation. In 2013 IEEE 20th International Conference on Web Services (pp. 107-114). IEEE, 2013.
- C-17. You, Y., Huang, G., Cao, J., Chen, E., He, J., Zhang, Y. and Hu, L.. Geam: A general and event-related aspects model for twitter event detection. In International Conference on Web Information Systems Engineering (pp. 319-332). Springer, Berlin, Heidelberg, 2013, October.
- C-18. Shangguan, Q., Hu, L., Cao, J. and Xu, G.. Book Recommendation Based on Joint Multi-relational Model. In Cloud and Green Computing (CGC), 2012 Second International Conference on (pp. 523-530). IEEE, 2012, November.

Contents

Certificate	ii
Abstract	iii
Dedication	v
Acknowledgments	vi
List of Publications	viii
List of Figures	xxi
Abbreviation	xxvi
Notation	xxviii
I Research Background	1
1 Introduction	2
1.1 Background	3
1.1.1 Recommender Systems	4
1.1.2 Recommendation Data Modeling with Machine Learning	8
1.2 Challenges in Modeling RSs	11
1.2.1 Challenges from Data Characteristics	11
1.2.2 Challenges from User Behavior	12
1.2.3 Challenges from Information Coupling	14
1.3 Research Objectives	15
1.3.1 Non-IID Modeling Aspects on Recommendation Data	16

1.3.2	Non-IID RS Modeling	19
1.4	Thesis Organization	28
2	Literature Survey	31
2.1	Basic Recommendation Techniques	31
2.1.1	CF	31
2.1.2	CBF	34
2.1.3	Hybrid CF and CBF	38
2.2	Modeling Implicit Feedback and Long-tail Distribution	40
2.2.1	Implicit Feedback	40
2.2.2	Long Tail	41
2.3	SNRSs	42
2.3.1	Summary of the Challenges to Target in an SNRS	43
2.3.2	Social Media Data	44
2.3.3	Trust Relation Modeling	44
2.3.4	Deep Learning Models	45
2.4	GBRSs	46
2.4.1	Summary of the Challenges to Target in GBRSs	46
2.4.2	Aggregation Approach	46
2.4.3	Deep Learning Approach	48
2.5	CDRSs	48
2.5.1	Summary of the Challenges to Target in CDRSs	49
2.5.2	Latent Factor Models	50
2.5.3	Deep Learning Models	51
2.6	SBRSSs	52

2.6.1	Summary of the Challenges to Target in SBRs	53
2.6.2	Markov Models	53
2.6.3	Deep Learning Models	54
2.7	MORs	55
2.7.1	Summary of the Challenges to Target in MORs	56
2.7.2	Multi-criteria Ratings	56
2.7.3	Multi-objective Ranking	57
3	Preliminaries	60
3.1	Neighborhood Method	60
3.2	Latent Factor Models	60
3.2.1	Matrix Factorization	60
3.2.2	Tensor Factorization	62
3.3	Deep Learning Models	64
3.3.1	Restricted Boltzmann Machines	64
3.3.2	Recurrent Neural Networks	65
3.4	Evaluation Metrics	68
3.4.1	Rating Prediction	68
3.4.2	Ranking Prediction	69
3.4.3	Diversity	70
3.5	Baseline Methods for Experiments	70
3.5.1	Heuristic Methods	71
3.5.2	CF Methods	71
3.5.3	Factorization Machines	71
3.5.4	Content-based Methods	71

3.5.5	Relation-based Method	72
3.5.6	Session Methods	72
3.5.7	Cross-domain Methods	72
II	Non-IID RS: Modeling Non-IIDness on Users	73
4	A Group-based Recommender System for Modeling Group Preference over Member Decisions	74
4.1	Introduction	74
4.2	Problem Formulation	76
4.3	Model and Inference	77
4.3.1	Disentangling Collective and Individual Embedding	77
4.3.2	A Comprehensive Representation of Group Preferences	80
4.3.3	Recommendation for a Group	83
4.4	Experiments	83
4.4.1	Data Preparation	84
4.4.2	Comparison Methods	84
4.4.3	Results	85
4.5	Summary of Contributions	89
5	A Social Network-based Recommender Systems for Modeling User and Item Influential Contexts	90
5.1	Introduction	90
5.2	Problem Formulation	93
5.3	Model and Learning	94
5.3.1	Architecture	94

5.3.2	Influential-context Aggregation Unit	95
5.3.3	User’s Influential Context Embedding	96
5.3.4	Item’s Influential Context Embedding	99
5.3.5	User-item Interaction Ranking	100
5.3.6	Training Procedure	101
5.4	Experiments	101
5.4.1	Data Preparation	102
5.4.2	Experimental Settings	103
5.4.3	Recommendation Performance	104
5.4.4	Recommendation for Cold-start Users and Items	106
5.4.5	Visualization and Interpretation	108
5.5	Summary of Contributions	108

III Non-IID RS: Modeling Non-IIDness on Items 111

6 A Cross-domain Recommender System for Modeling the Couplings over Heterogeneous Item Domains 112

6.1	Introduction	112
6.1.1	Leveraging Cross-Domain Information	113
6.1.2	Modeling Domain Factors	116
6.1.3	Irregular Triadic Relation	116
6.2	Problem Formulation	118
6.2.1	Weighted Regularized Matrix Factorization	118
6.2.2	Learning Priors from Cross-Domain Feedback	119
6.2.3	Learning Posterior on Target Domain for Fine Tuning	120

6.3	Weighted Irregular Tensor Factorization	121
6.3.1	Transformation	121
6.3.2	Parameter Learning	125
6.3.3	Weight Matrix Configuration	128
6.4	Remarks	131
6.4.1	Tricks on Sparse Weight Matrices for Complexity Reduction .	131
6.4.2	Training with Additional Noisy Examples for Improving Generalization	132
6.4.3	Post-Learning	134
6.5	Experiments	135
6.5.1	Comparison Methods	135
6.5.2	Evaluation Metric	137
6.5.3	Rating Prediction on Epinions.com	137
6.5.4	Click Prediction on Tmall.com	144
6.5.5	Time Period as Domain on Movie Rating Prediction	153
6.6	Summary of Contributions	158
7	A Session-based Recommender System for Modeling the Intra- and Inter-sessions Context	161
7.1	Introduction	161
7.2	Problem Formulation	163
7.3	Neural Cross-session Filtering	164
7.3.1	Historical Session Encoding	165
7.3.2	Current Session Encoding	166
7.3.3	Intra- and Inter-session Context Encoding	167

7.3.4	Objective Function	167
7.3.5	Learning and Prediction	168
7.4	Empirical Study	169
7.4.1	Data Preparation	169
7.4.2	Comparison Methods	170
7.4.3	Diversity Evaluation	173
7.5	Summary of Contributions	176

IV Non-IID RS: Modeling Non-IIDness on Interactions177

8	A Multi-objective Recommender System for Modeling Specialty and Credibility for Long-tail Recommendation	178
8.1	Introduction	178
8.1.1	Challenges of Tail Users and Items	179
8.1.2	Optimizing Specialty and Credibility	182
8.2	Preliminaries	185
8.2.1	Reputation Modeling	185
8.2.2	Explicit and Implicit Rating	189
8.3	Model and Learning	190
8.3.1	Overview of RMRM	190
8.3.2	Learning Regularized C-HMF Model	194
8.3.3	Learning Regularized S-HMF Model	203
8.4	Algorithm and Prediction	206
8.5	Discussion	209
8.5.1	Multi-objective Optimization	209

8.5.2	Social Regularization from PoGE Perspective	210
8.6	Experiments	211
8.6.1	A Comparison of the State-of-the-Art Methods	212
8.6.2	Explicit Rating Data Evaluation	213
8.6.3	Implicit Rating Data Evaluation	223
8.7	Summary of Contributions	228
9	Attraction-based Recommender Systems for Capturing and Interpreting User Attraction in Content	231
9.1	Introduction	231
9.2	Attraction Model	234
9.2.1	Content Representation Module	235
9.2.2	Attraction Filter Module	236
9.2.3	Attractiveness Score Module	237
9.3	Hierarchical Attraction Model on Textual Content	238
9.3.1	Multilevel Attraction	239
9.3.2	Bidirectional Gated Recurrent Units	239
9.3.3	Model Architecture	240
9.3.4	Multilevel Attraction Filters	241
9.3.5	Hierarchical Attractive Content Encoder	243
9.3.6	Objectives and Parameter Learning	246
9.4	Multimodal Attraction Model on Movies	247
9.4.1	Multimodal Attraction	249
9.4.2	Model Architecture	249
9.4.3	Story Attraction Module	250

9.4.4	Cast Attraction Module	253
9.4.5	Multimodal Movie Attraction Scoring	254
9.4.6	Objective and Parameter Learning	254
9.5	Experiments	255
9.5.1	Comparison Methods	256
9.5.2	Attraction on Books for Classification	257
9.5.3	Attraction on Academic Papers for Ranking	261
9.5.4	Multimodal Attraction on Movies	266
9.6	Summary of Contributions	272

V Summary and Prospect 273

10 Conclusion 274

10.1	Non-IID RS Modeling on Users	274
10.1.1	GBRS Modeling: Learning Comprehensive Group Preference Representation	274
10.1.2	SNRS Modeling: Learning Comprehensive Group Preference Representation	275
10.2	Non-IID Recommender Systems Modeling on Items	275
10.2.1	CDRS Modeling: Leveraging Knowledge Across Heterogeneous Domains	275
10.2.2	SBRS Modeling: Modeling Selection with Influential Users and Items	276
10.3	Non-IID RS Modeling on Complex Interaction	277
10.3.1	MORS Modeling: Optimizing Credibility and Novelty for Long-tail Recommendation	277

10.3.2 ABRS Modeling: Capturing and Interpreting Attraction

Points in Content 278

11 Open Challenges and Future Directions 279

11.1 New Evaluation Methods and Metrics 279

11.2 SNRSs 280

11.3 GBRs 281

11.4 CDRs 282

11.5 SBRs 283

11.6 MORs 284

11.7 ABRSs 285

11.8 Unified and Ubiquitous RSs 286

Bibliography 288

List of Figures

1.1	Overview of the non-IID RSs targeted in this thesis. Each modeling item for these systems is marked with one or two icons. Each icon corresponds to one non-IID modeling aspect shown in the right-hand section.	20
1.2	Overview of machine learning methods used to model non-IID RSs in this thesis. Each listed item for these systems is marked with one or two icons. Each icon corresponds to one machine learning approach listed in the right-hand section.	21
3.1	The demonstration of CP factorization on a third-order tensor	63
3.2	The structure of restricted Boltzmann machine	65
3.3	Long short-term memory [161]	66
3.4	Gated recurrent unit	67
4.1	Left: Overview of the two-layer collective DBN used to disentangle high-level collective and individual embeddings. Right: More detailed structure of the collective RBM at the top layer where the collective embedding is connected to the member embedding w.r.t. each member.	78
4.2	A dual-wing RBM is placed on the top of DBN, which jointly models the group choices and collective embedding to learn the comprehensive embedding of group preference.	81

4.3	MAP w.r.t. 2-member groups vs. 2^+ -member groups	88
5.1	A multi-relational RS consists of user-user relation, user-item relation and item-item relation. Each user-item interaction is influenced by the user context and the item context.	91
5.2	The architecture of ICE-MRS for modeling user's and item's influential contexts	94
5.3	Influential-context Aggregation Unit (ICAU): A two-stage aggregation model to construct influential context embedding (ICE) .	97
5.4	Item recommendation for cold-start users of Delicious and Lastfm. . .	107
5.5	User recommendation for cold-start items of Delicious and Lastfm. . .	108
5.6	The visualization of influential contexts of a sampled user selection on an artist in the Lastfm dataset. The artists in item network are labeled by their names and the anonymous users in user network are labeled with IDs. The thickness of edges specifies the significance of influence.	109
6.1	Demonstration of the occurrence of the blinder-transfer issue in CDMF	114
6.2	The feedback from the majority of users in each domain is scant due to the power law distribution, and users have different unfamiliar domains due to differences in interests.	115
6.3	Irregular triadic relation and regular triadic relation.	117
6.4	WITF transforms the slices with heterogeneous domain-specific items into a regular third-order tensor containing an identical virtual item set.	122
6.5	Rating distributions of items on Kids & Family and Hotel & Travel show obvious long tails.	138
6.6	RMSEs comparison over user groups with different number of ratings.	141

6.7	RMSEs comparison over different number of noisy examples for each user.	143
6.8	Mean APs over different values of confidence parameter c	149
6.9	Recall@5~20 of comparison methods on target domain D1 and D2. .	151
6.10	Time period as domain to model the heterogeneities of user preferences on movies	154
6.11	MAE comparison with α_k increasing.	157
7.1	The motivation of this work: (a) Classic recommendation without considering session context; (b) Two sequences adding milk and bread into the cart; (c) Rigid-order session vs. relaxed-order session; (d) First-order Markov chain model vs. recurrent model; (e) Next item prediction based on both intra- and inter-session contexts. . . .	162
7.2	The architecture of NCSF. In the left part of split line, historical sessions are used to build inter-session context. In the right part of split line, the items in the current session are used to build intra-session context.	164
7.3	REC@5 - REC@50 on test cases <i>LAST</i> and <i>LOO</i>	173
7.4	F1 scores on test cases <i>LAST</i> and <i>LOO</i>	175
8.1	Items (left) and users (right) are ranked by the number of their ratings (truncated from 0 to 100) on Rich Epinions Dataset; they are both clearly distributed with short heads and long tails.	180
8.2	A recurrent mutual regularization process couples S-HMF and C-HMF using the user and item-factors learned from one another as the empirical priors to couple the objectives <i>specialty</i> and <i>credibility</i> .	185

8.3	The screen snapshots of reviews from Amazon.com (left) and Ciao.com (right), where the red, dotted boxes show the helpfulness score.	188
8.4	The graphical representation of the RMRF framework, where S-HMF and C-HMF are recurrently regularized by the empirical priors, induced from one another.	191
8.5	The geometric illustration of the recurrent mutual regularization process, where the estimates of S-HMF and C-HMF are recurrently regularized by the empirical priors induced from one another.	193
8.6	Long-tail distributions for the number of ratings of items and users (truncated from 0 to 500).	214
8.7	The distributions for the number of helpful scores w.r.t. items and users (truncated from 0 to 200).	215
8.8	MAEs of rating prediction for the long-tail item distribution.	217
8.9	MAEs of rating prediction for the long-tail user distribution.	218
8.10	MAEs varying the numbers of involved (a) user trusters, and (b) system experts.	221
8.11	MAEs for head items and tail items with shilling attack.	223
8.12	Long-tail distributions over the number installations w.r.t. users and items (truncated).	224
8.13	Recall@20-50 of tail-item recommendations for users.	227
8.14	Recall@20-50 of tail-user recommendations for users.	228
9.1	Users with different background may have quite different attractive points on an article [199]	232
9.2	Architecture of Attraction Model	235
9.3	The architecture of hierarchical attraction model	241

9.4	The architecture of hierarchical attraction model over movies with two modalities: Cast (left) and Story (right)	250
9.5	Attractiveness visualization on the abstract of <i>Invitation to the Game</i> w.r.t. sentences and words in the most attractive sentences for two users. The larger size and deeper color of font denotes the larger attractiveness weight is assigned.	260
9.6	Recall@5-50 on the Published Paper testing set and the New Paper testing set	264
9.7	Demonstration of the different of group attraction on the abstract of the paper <i>Exploring complex networks</i> . We list the most attractive word in each sentence for each group to illustrate their difference. . .	266
9.8	Demonstration of two exemplary users' most attractive papers in the testing set. We list the most attractive word in each sentence of the recommended papers for interpretation.	266
9.9	R@5-50 on the Released Movies and the New Movies	270
9.10	Two comparative testing samples of <i>User 182</i> : the left movie <i>Wild America</i> obtains a high attraction score because of the cast members in red appear in user's watched movies while the cast members of <i>Bogus</i> never appear in user's movie list.	270
9.11	Statistical attractiveness on movie <i>Election (1999)</i> w.r.t. sentences, words in the most attractive sentences and cast members from the perspectives of User 156 and User 2163. The larger size and deeper color of font denote the larger attractiveness weight is assigned. . . .	271

Abbreviation

RS - Recommender System

AI - Artificial Intelligence

CF - Collaborative Filtering

CBF - Context-based Filtering

NLP - Natural Language Processing

CV - Computer Vision

CNN - Convolutional Neural Networks

RNN - Recurrent Neural Networks

TSA - Time Series Analysis

CDRS - Cross-domain Recommender Systems

SNRS - Social Network-based Recommender Systems

CARS - Context-aware Recommender Systems

MORS - Multi-objective Recommender Systems

SBRS - Session-based Recommender Systems

GBRS - Group-based Recommender Systems

ABRS - Attraction-based Recommender Systems

MF - Matrix Factorization

TF - Tensor Factorization

RBM - Restricted Boltzmann Machines

GRU - Gated Recurrent Units

LSTM - Long Short-Term Memory

MAE - Mean Absolute Error

RMSE - Root Mean Square Error

AP - Average Precision

MRR - Mean Reciprocal Rank

nDCG - Normalized Discounted Cumulative Gain

AUC - Area Under the ROC Curve

Nomenclature and Notation

$(.)^\top$ denotes the transpose operation.

\mathbf{I} is the identity matrix.

\mathbb{R} , \mathbb{R}^+ denote the field of real numbers, and the set of positive reals, respectively.

\mathcal{X} denotes a tensor (boldface script letter)

$\mathcal{X}^{I \times J \times K}$ denotes The dimensionality of each mode

\mathbf{X} denotes a matrix (boldface capital letter)

$\mathbf{X}_{i,:}$ denotes the i th row of matrix \mathbf{X}

$\mathbf{X}_{:,j}$ denotes the j th column of matrix \mathbf{X}

$\mathbf{X}_{i,j}$ denotes the entry (i, j) of matrix \mathbf{X}

\mathbf{x} denotes a vector (boldface lower-case letter)

$\mathbf{X}_{(n)}$ denotes mode- n matricization of a tensor

\otimes denotes Kronecker product

\odot denotes Khatri-Rao product

\cdot^* denotes Hadamard product

$\ ./$ denotes element-wise division

\circ denotes outer product

$\|\mathcal{X}\|$ denotes the norm of a matrix or a tensor \mathcal{X}

$\mathcal{X} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ denotes the factorization form of a tensor

$\text{vec}\mathbf{X}$ denotes vectorization of matrix