

PERSONALIZED RECOMMENDATIONS BASED ON USERS' INFORMATION-CENTERED SOCIAL NETWORKS

by

Danielle Lee

BA, Sangmyung University, 1996

MS, Sangmyung University, 1998

MS, Syracuse University, 2001

Submitted to the Graduate Faculty of
School of Information Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH
SCHOOL OF INFORMATION SCIENCES

This dissertation was presented

by

Danielle Lee

It was defended on

November 9, 2012

and approved by

Brian Butler, Associate Professor, School of Information Studies, University of Maryland,

Daqing He, Associate Professor, School of Information Sciences,

Stephen Hirtle, Professor, School of Information Sciences,

Titus Schleyer, Associate Professor, School of Dental Medicine, and

Dissertation Advisor: Peter Brusilovsky, Professor, School of Information Sciences

Copyright © by Danielle Lee

2013

**PERSONALIZED RECOMMENDATIONS BASED ON USERS' INFORMATION-
CENTERED SOCIAL NETWORKS**

Danielle Lee, PhD

University of Pittsburgh, 2013

The overwhelming amount of information available today makes it difficult for users to find useful information and as the solution to this information glut problem, recommendation technologies emerged. Among the several streams of related research, one important evolution in technology is to generate recommendations based on users' own social networks. The idea to take advantage of users' social networks as a foundation for their personalized recommendations evolved from an Internet trend that is too important to neglect – the explosive growth of online social networks. In spite of the widely available and diversified assortment of online social networks, most recent social network-based recommendations have concentrated on limited kinds of online sociality (i.e., trust-based networks and online friendships). Thus, this study tried to prove the expandability of social network-based recommendations to more diverse and less focused social networks. The online social networks considered in this dissertation include: 1) a watching network, 2) a group membership, and 3) an academic collaboration network. Specifically, this dissertation aims to check the value of users' various online social connections as information sources and to explore how to include them as a foundation for personalized recommendations.

In our results, users in online social networks shared similar interests with their social partners. An in-depth analysis about the shared interests indicated that online social networks have significant value as a useful information source. Through the recommendations generated

by the preferences of social connection, the feasibility of users' social connections as a useful information source was also investigated comprehensively. The social network-based recommendations produced as good as, or sometimes better, suggestions than traditional collaborative filtering recommendations. Social network-based recommendations were also a good solution for the cold-start user problem. Therefore, in order for cold-start users to receive reasonably good recommendations, it is more effective to be socially associated with other users, rather than collecting a few more items. To conclude, this study demonstrates the viability of multiple social networks as a means for gathering useful information and addresses how different social networks of a novelty value can improve upon conventional personalization technology.

TABLE OF CONTENTS

1	INTRODUCTION	19
	PURPOSE OF THE STUDY AND RESEARCH QUESTIONS	23
	CONTRIBUTIONS OF THE STUDY	26
	STRUCTURE AND OUTLINE.....	27
2	BACKGROUND AND LITERATURE REVIEW	31
	2.1 ONLINE SOCIAL NETWORKS AND THE INTEREST SIMILARITY	32
	2.2 RECOMMENDATION TECHNOLOGY: STATE OF ART	38
	2.3 SOCIAL NETWORK-BASED RECOMMENDATIONS	44
	2.3.1 Definition of Social Network-based Recommendations	46
	2.3.2 Limited Kinds of Social Networks in Personalized Recommendations.....	53
	2.3.2.1 Trust Network-based Recommendations	55
	2.3.2.2 Friendship-based Recommendations.....	59
	2.3.2.3 Recommendation based on Other Kinds of Social Networks.....	61
	2.3.3 Acceptability of Recommendations Suggested by Social Connections	63
	2.4 TAGS AS USER INTEREST INDICATORS.....	66
3	SOCIAL NETWORKS IN CONSIDERATION AND THE DATA SOURCE	75
	3.1 WATCHING NETWORK: SOCIAL CONNECTIONS CENTERED ON OBJECTS... 78	
	3.2 GROUP MEMBERSHIP: SOCIAL CONNECTIONS CENTERED ON ONE TOPIC . 81	

3.3	COLLABORATION NETWORK: SOCIAL CONNECTIONS OF PERSONAL ACQUAINTANCE AND SHARED INTERESTS	84
3.4	THE DATA SOURCES.....	86
4	RECOMMENDATION TECHNOLOGIES.....	88
4.1	INFORMATION SIMILARITY MEASURES.....	88
4.1.1	Similarity Measure for Unary Ratings.....	89
4.1.2	Similarity Measure for Textual Metadata.....	92
4.1.3	Similarity Measure for Numeric Ratings	93
4.2	RECOMMENDATION TECHNOLOGIES.....	96
4.2.1	Recommendations based on K-Nearest Neighbors.....	97
4.2.2	Matrix Factorization Recommendations.....	97
4.3	EVALUATION OF RECOMMENDATION QUALITY.....	101
4.3.1	Evaluation Method for Unary Ratings.....	102
4.3.2	Evaluation Method for Numeric Ratings	104
5	WATCHING NETWORK-BASED RECOMMENDATIONS.....	107
5.1	DATA SOURCE FOR WATCHING NETWORK: CITEULIKE.....	108
	Information Usage Patterns of Users according to Their Social Status.....	111
5.2	FEASIBILITY OF WATCHING NETWORK AS A USEFUL INFORMATION SOURCE	114
5.2.1	Information Similarity Measures.....	116
5.2.1.1	Number of Co-bookmarks and Jaccard Coefficient.....	116
5.2.1.2	Log-likelihood Similarity	117
5.2.1.3	Popularity Weight of Co-bookmarked Items.....	118
5.2.1.4	Metadata-based Similarity	119
5.2.1.5	Social Tag-based Similarity	121

5.2.2	Information Similarity of User Pairs according to their Social Distances	123
5.2.3	Information Similarity vs. Social Features.....	126
5.2.4	Social Connections vs. Top N Anonymous Peers	129
5.2.5	The Best Similarity Measures for Watching Network	133
5.3	EVALUATION OF WATCHING NETWORK-BASED RECOMMENDATIONS.....	140
5.3.1	Recommendation Algorithms.....	144
5.3.1.1	Watching Network-based Recommendation and Hybrid Recommendations.....	144
5.3.1.2	Content Similarity Weights	147
5.3.1.3	Content Boosted Recommendations	149
5.3.2	Experimental Evaluation of Recommendations.....	151
5.3.3	Social Network-based Recommendations using Various Settings.....	152
5.3.4	Comparison of Various Recommendation Approaches	157
5.3.5	Watching Network-based Recommendation for Cold-start Users	161
5.4	CONCLUSION	165
6	GROUP MEMBERSHIP-BASED RECOMMENDATIONS.....	167
6.1	DATA SOURCE FOR GROUP MEMBERSHIP: CITEULIKE	168
	Information Usage Patterns of Users according to Their Social Status.....	172
6.2	FEASIBILITY OF GROUP MEMBERS AS A USEFUL INFORMATION SOURCE	176
6.2.1	Members' Contributions on Group Library.....	178
6.2.2	Information Similarity of Group Members	185
6.2.3	Group Co-members vs. Top N CF Peers	186
6.2.4	Group Library vs. Co-members' Personal Repository	188
6.2.5	Two Social Networks in Citeulike: Watching Network vs. Group Membership	191
6.2.6	The Best Similarity Measures for Watching Network	195

6.3	EVALUATION OF GROUP MEMBERSHIP-BASED RECOMMENDATIONS.....	199
6.3.1	Recommendation Algorithms.....	203
6.3.1.1	Group-based Recommendations.....	203
6.3.1.2	Recommendations based on Two Social Networks and Hybrid Recommendations.....	205
6.3.1.3	Adding Metadata Information of Items in Recommendations	206
6.3.1.4	Community vote-based Recommendations	208
6.3.2	Experimental Evaluation of Recommendations.....	209
6.3.3	Social Network-based Recommendations using Various Settings.....	210
6.3.4	Comparison of Various Recommendation Approaches	216
6.3.5	Comparison of Group Membership-based Recommendations with other Social Recommendation Approaches	219
6.3.6	Recommendation Quality according to the Ratio of Group Members' Contributions	221
6.3.7	Group Membership-based Recommendation for Cold-start Users.....	226
6.4	CONCLUSION	230
7	COLLABORATION NETWORK-BASED RECOMMENDATIONS	233
7.1	DATA SOURCE FOR COLLABORATION NETWORK: CONFERENCE NAVIGATOR 3	235
7.2	FEASIBILITY OF RESEARCH COLLABORATORS AS A USEFUL INFORMATION SOURCE	244
7.2.1	Data Analysis Methods	245
7.2.2	Relevance of Users' Research Profiles with Contents of their Favorite Talks ...	248
7.2.3	Information Similarity of User Pairs according to the Kinds of their Social Links	250
7.2.4	The Best Similarity Measures for Collaboration Network	254

7.3	EVALUATION OF COLLABORATION NETWORK-BASED RECOMMENDATIONS	260
7.3.1	Recommendation Algorithms.....	264
7.3.1.1	Basic Recommendation Approaches.....	264
7.3.1.2	Collaboration Network-based Recommendations	266
7.3.1.3	Content Similarity Weights	268
7.3.1.4	Profile-based Recommendations.....	269
7.3.1.5	Community Vote Recommendations	270
7.3.2	Experimental Evaluation of Recommendations.....	270
7.3.3	Social Network-based Recommendations using Various Settings	271
7.3.4	Comparison of Social Recommendations with Typical Approaches	278
7.3.5	Collaboration Network-based Recommendations for Cold-start Users.....	283
7.4	CONCLUSION	286
8	CONCLUSION AND DISCUSSION.....	289
8.1	FINDINGS OF RESEARCH QUESTIONS	291
8.1.1	Findings of the First Part: Feasibility of Social Networks as Useful Information Source.....	291
8.1.2	Findings of the Second Part: Evaluation of Social Network-based Recommendations.....	294
8.1.3	Conclusion	297
8.2	LIMITATIONS OF THE STUDY AND FURTHER RESEARCH DIRECTION.....	299
	APPENDIX A.....	303
	APPENDIX B	325
	BIBLIOGRAPHY	326

LIST OF TABLES

Table 1. Summary of Social Network-based Recommendations	50
Table 2. Classification of the Possible Outcomes of the Recommendations	102
Table 3. Descriptive Statistics about the Cituelike Watching Dataset.....	109
Table 4. Comparison of Item-based Similarity Depending on the Distances of Watching Relations	124
Table 5. Comparison of Metadata-based Similarity Depending on the Distances of Watching Relations	126
Table 6. The Optimal Setting of Each Recommendation Approach.....	153
Table 7. Differences of Recommendation Precision Depending on Metadata Properties of Items.....	154
Table 8. Differences of Recommendation Recall Depending on Metadata Properties of Items	154
Table 9. Differences of Recommendation Precision between Different Methods to Compute Prediction Probability	156
Table 10. Differences of Recommendation Recall between Different Methods to Compute Prediction Probability	157
Table 11. Differences of Recommendation F1 between Different Methods to Compute Prediction Probability	157
Table 12. Descriptive Statistics about the Dataset of Citeulike Group Memberships	170
Table 13. Summary of Differences according to Users' Clusters.....	175
Table 14. Item-based Similarity of Group Co-members	185
Table 15. Metadata-based Similarity of Group Co-members.....	186

Table 16. Metadata-based Similarity Comparison between Co-members vs. Top N Peers.....	188
Table 17. Comparison of Information Similarity between Group vs. Co-members	190
Table 18. Metadata-based Similarity Comparison between Co-members vs. Top N Peers.....	191
Table 19. Comparison of Item-based Similarity between Watching Relations and Group-based Relations	192
Table 20. Comparison of Metadata-based Similarity between Watching Relations and Group-based Relations	193
Table 21. Similarity Comparison between Watching Connections & Group Co-members for the Same Users	194
Table 22. The Optimal Setting of Each Recommendation Approach.....	211
Table 23. Differences of Recommendation Precision Depending on Metadata Properties of Items.....	212
Table 24. Differences of Recommendation Recall Depending on Metadata Properties of Items	213
Table 25. Differences of Recommendation Precision depending on Probability Computation Algorithms	215
Table 26. Differences of Recommendation Recall depending on Probability Computation Algorithms.	215
Table 27. Descriptive Statistics of Conference Navigator Dataset	241
Table 28. Differences of Recommendation Precision according to the Kind of Preferences	274
Table 29. Differences of Recommendation Recall according to the Kind of Preferences	275
Table 30. Differences of Recommendation Precision Depending on Content Properties of Items.....	275
Table 31. Differences of Recommendation Recall Depending on Content Properties of Items	276
Table 32. Differences of Recommendation Precision Depending on Probability Computation Algorithms	277
Table 33. Differences of Recommendation Recall Depending on Probability Computation Algorithms	278

LIST OF FIGURES

Figure 1. The Structure of Dissertation	24
Figure 2. Various Social Networks Considered in this Dissertation.....	78
Figure 3. Data Sources for Each Social Network.....	87
Figure 4. Summary of Recommendation Algorithms.....	97
Figure 5. Distribution of Users according to the Number of their Bookmarks	110
Figure 6. Distribution of Articles according to the Number of The Bookmarks	110
Figure 7. Distribution of users according to how many users they are watching	110
Figure 8. Distribution of users according to how many users they are being watched by	110
Figure 9. Bookmark distribution of users who involve in watching relations contrasted with the one of users not involved	112
Figure 10. Bookmark distribution of users who are watching others, contrasted with the users who are being watched by others.....	114
Figure 11. Summary of Similarity Measures	123
Figure 12. Metadata-based Similarity Comparison between Watching Connections vs. Top N Peers	132
Figure 13. Results of the Test to Forecast Watching Relations	136
Figure 14. Precisions of Combined Measures for the Prediction of Watching Relations	138
Figure 15. Recalls of Combined Measures for the Prediction of Watching Relations.....	139
Figure 16. Overall Design of Watching Network-based Recommendation Algorithms	141

Figure 17. Revised Design of Watching Network-based Recommendation Algorithms	156
Figure 18. Differences of Recommendation Quality depending on the Kinds of Recommendation Approaches	159
Figure 19. Differences of Precision Depending on Users' Number of Bookmarks	163
Figure 20. Differences of Precision Depending on Users' Number of Bookmarks	164
Figure 21. Distribution of Groups according to the Bookmark Number of their Group Libraries.....	171
Figure 22. Distribution of Group Members according to the Bookmark Number of their Personal Repositories	171
Figure 23. Distribution of Groups according to the Number of Members.....	171
Figure 24. Distribution of Users according to the Number of Group Memberships	171
Figure 25. Bookmark distribution of users who involve in group membership contrasted with the one of users not involved	173
Figure 26. Bookmark distributions of three user clusters	175
Figure 27. Summary of Similarity Measures	178
Figure 28. Citeulike interface enabling users to post favorite items to both their personal repositories and group libraries simultaneously	180
Figure 29. Ratio of Member's Contributions in terms of the Size of Group Library	181
Figure 30. Ratio of Members' Contributions in terms of the size of their Personal Repository	181
Figure 31. Contrast of Members' Contributions in terms of both Group Library and the Members' Personal Repository	182
Figure 32. Members' Contributions vs. Information Similarity of Co-Members.....	183
Figure 33. Members' Contributions vs. the Similarity between a Member and his Group without his Own Contributed Items	184
Figure 34. Results of the Test to Forecast Group Co-members.....	198
Figure 35. Precisions of Combined Measures for the Prediction of Group Co-members.....	198
Figure 36. Recalls of Combined Measures for the Prediction of Group Co-members	199

Figure 37. Overall Design of Group Membership-based Recommendation Algorithms	201
Figure 38. Revised Design of Group Membership-base Recommendation Algorithms	214
Figure 39. Differences of Recommendation Precision depending on the Approaches	217
Figure 40. Differences of Recommendation Recall depending on the Approaches.....	218
Figure 41. Comparison of Various Citeulike Social Recommendations	221
Figure 42. Differences of Precision according to Users' Contribution Patterns	224
Figure 43. Differences of Recall according to Users' Contribution Patterns.....	224
Figure 44. Differences of Precision according to Users' No. of Bookmarks.....	228
Figure 45. Differences of Recall according to Users' No. of Bookmarks	229
Figure 46. Main Page of Conference Navigator.....	236
Figure 47. Distribution of Users according to the Number of their Bookmarks	242
Figure 48. Distribution of Articles according to the Number of their Bookmarks	242
Figure 49. Distribution of Users' Collaborators.....	243
Figure 50. Distribution of Users' Publications.....	243
Figure 51. Summary of Similarity Measures	248
Figure 52. Similarity between Users' Research Profiles and their Keyword Vectors	249
Figure 53. Comparison of Item-based Similarities according to the Kinds of Social Links	251
Figure 54. Comparison of Metadata-based Similarities according to the Kinds of Social Links.....	253
Figure 55. Results of the Test to Predict Collaborator Relations using Individual Similarity Measures .	256
Figure 56. Results of the Test to Predict Collaborator Relations using Combined Measures.....	257
Figure 57. Results of the Test to Predict CN3 Relations using Individual Similarity Measures.....	259
Figure 58. Results of the Test to Forecast CN3 Relations using Combined Similarity Measures	260
Figure 59. Overall Design of CN3 Talk Recommendation Algorithms	261
Figure 60. KNN Recommendations using Different Kinds of User Preferences	273

Figure 61. KNN Recommendations using Better User Preferences	273
Figure 62. Modified Design of CN3 Tal Recommendation Algorithms.....	277
Figure 63. Differences of Recommendation Precision Depending on the Approaches	279
Figure 64. Differences of Recommendation Recall Depending on the Approaches	279
Figure 65. Comparison of Recommendation Quality according to Users' Social Status.....	283
Figure 66. Differences of Precision Depending on Users' Number of Bookmarks	285
Figure 67. Differences of Recall Depending on Users' Number of Bookmarks.....	286
Figure 68. Comparison of KNN Recommendations based on Anonymous Peers (CF).....	304
Figure 69. Comparison of KNN Recommendations based on Direct Watching Relations (Watch)	306
Figure 70. Comparison of KNN Recommendations based on Watching and Watched Relations (ReciWatch).....	307
Figure 71. Comparison of KNN Recommendations based on Direct & Indirect Watching Relations (1hopWatch)	308
Figure 72. Comparison of Matrix Factorization Recommendations based on Users' Anonymous Peers (CFSVD).....	309
Figure 73. Comparison of Matrix Factorization Recommendations based on Direct Watching Relations (WatchSVD)	310
Figure 74. Comparison of Matrix Factorization Recommendations based on Relations in both Directions (reciWatchSVD)	311
Figure 75. Comparison of Matrix Factorization Recommendations based on Relations in Direct and Indirect Social Connections (1hopWatchSVD).....	312
Figure 76. Comparison of KNN Recommendations based on Users' Anonymous Users (CF)	313
Figure 77. Comparison of KNN Recommendations based on Users' Group Co-members (GMem)	315
Figure 78. Comparison of KNN Recommendations based on Co-members and Group Memberships (Gmem).....	316

Figure 79. Comparison of KNN Recommendations based on Watching Relations and Group Memberships (GMemWatch)	318
Figure 80. Comparison of Matrix Factorization Recommendations based on Users' Anonymous Peers (CFSVD).....	318
Figure 81. Comparison of Matrix Factorization Recommendations based on Users' Co-members (GMemSVD)	319
Figure 82. Comparison of Matrix Factorization Recommendations based on Users' Co-members and their library (GroupSVD).....	320
Figure 83. Comparison of Matrix Factorization Recommendations based on Watching Networks and Group Memberships (GMemWatchSVD).....	321
Figure 84. Comparison of Matrix Factorization Recommendations based on Users' Anonymous Peers	322
Figure 85. Comparison of Matrix Factorization Recommendations based on Users' Collaborators	322
Figure 86. Comparison of Matrix Factorization Recommendation based on Users' CN3 Social Connections	323
Figure 87. Comparison of Matrix Factorization Recommendation based on Both Collaborators and CN3 Social Connections	324

이 박사학위 논문을 사랑하는 어머니께 바칩니다.

1 INTRODUCTION

Information on the Web continues to grow very rapidly and is changing dynamically, ever since users have been contributing by continuously creating Web contents such as Wikipedia, blogs, social bookmarking and tagging, micro-blogging, etc. A statistic about Internet blogging suggested that blog posts being created by users daily equal the entire Web pages in mid-1995 in size [54]. The overwhelming amount of information available makes it difficult for users to find useful information. In order to cope with this information glut problem, recommendation technologies emerged. A variety of recommendation technologies have proven their effectiveness in a number of real-life applications such as Amazon.com, Last.fm, Netflix, etc. Because of the great potential for growth, not only industry but also academia paid a lot of attentions on the development of the technology. Among several streams of the related research, one important evolution of this technology is to generate recommendations based on users' own social networks, instead of anonymous peers – a new generation of social network-based recommendations (hereafter, SN recommendations). The idea to take advantage of users' social networks as a foundation of personalized recommendations evolved from an Internet trend which is too important to be neglected – explosively growing online social networks.

Compared with the era when computer users remained in isolation, users on the Web 2.0 have found it easier to find interesting people, as well as interesting information, through online

social network sites. Following the leading online social networking sites (a.k.a. social network sites; SNS in the abbreviation), such as LinkedIn, MySpace, Facebook, etc., various information systems have adopted collective intelligence where users are enacting both information supplier and consumer roles. Social bookmarking systems, image and video sharing sites, review sites of various goods (e.g. books, movies, songs, restaurants, etc.) are a few examples. The systems help users find other users whose interests are similar to theirs or familiar to them and, furthermore, encourage them to connect with these users. The self-defined online social links establish a rich source of information, which is, in turn, used to propagate various kinds of information [11, 18]. The users in these online SNSs acquire numerous useful resources, for instance, movies to watch, books to read, papers to refer to, bookmarks to explore, music concerts to enjoy, from their online connections. According to a recent statistic about online media, Web users in the United States spent 22.5 percent of online time on SNS sites (9.8% for online games, 7.6% for emails and 4.5% for portals) and as of the third quarter of the year 2011, four-fifths of active Web users visit SNSs and blogs. Over 140 million users visited the most famous SNS, Facebook, and over 50 million and 23 million users visited Blogger and Twitter, respectively, just in May, 2011. Naturally, the opinions and information of online connections influence social media users to a large extent. Compared with the average adult Web users, active social media users are more likely to express their opinions about products and services (about 60% more), and a pair of online connections are more likely to follow the same brand (53%) or the same celebrity (32%) [47, 137]. Moreover, according to prominent social science theories (i.e. homophily and social influence), people make social connections selectively with other people who are similar to them and their attitudes, beliefs and behavioral propensities are affected by and adopted to their social

ties. Therefore, with the proliferating online social networking applications and the abundant Web-based sociability, it is legitimate to expect to merge online social links with recommendation technology.

In spite of the widely available and diversified online sociability, however, SN recommendation technology lies still in infant stage. In particular, according to the literature review introduced in the Chapter 2, most of the recent SN recommendations have concentrated on limited kinds of online sociality: trust-based networks and friendships. The relevant data sources are also restricted to very few kinds such as Epinions, Facebook and Last.fm. Due to the trouble to collect the data, sometimes researchers artificially built implicit social networks through machine learning approaches. In order to fill these critical needs, as a component of personalized recommendations, this dissertation aims to extensively explore newly emerged but less focused social network kinds. This study will investigate how to take advantage of users' online social networks as information sources to acquire useful information and how to generate recommendations based on those social networks.

Conclusively, the significance of this dissertation is to check the availability of users' online social connections as an information source and to explore how to include them as a foundation of personalized recommendations. Now, there is one important question to answer. Which social network will be considered in this dissertation? The focus of this study is on users' self-defined online social networks, rather than implicit social networks inferred by a machine learning method. In case of the implicit social networks, depending on which machine learning method is used, the structure of the whole social network and social connections of individual users will be different. However, users' explicit social connections are declared by them, and

there is no chance that the members of social connections will be altered, unless the users change them. Most importantly, users know who their peers are. When a person makes social links by his own choice, it is known that his links affect widely the information he receives, the interactions he makes, and other various decision making tasks [18, 36, 106, 176]. Therefore, the information favored by his social connections could be legitimately useful and effective to be the source of his recommendations. Moreover, users will know that their recommendations are based on preferences of their social connections. By using users' explicit social connections, it is expected that SN recommendations can utilize not only the information about "what users like," but also the information about "whom users are interested in." Additionally, generating and accepting the recommendations is a decision making process. For this reason, it is important to know a source of recommendations (i.e., where the recommended information comes from). Once users know the source of their recommendations, it is known that user satisfaction about the suggested information increases [20, 164, 169].

As a more elaborated question, which explicit social network will be considered in this dissertation? In spite of the explosive popularity of Facebook, every Web 2.0 application doesn't have the same kind of friendships as Facebook offers. Depending on the domain each system belongs to, the way to develop the sociability on the system, and the purpose of the sociality (i.e. how the sociality to be used on the system), there are several kinds. In order to prove expandability of SN recommendations to more diverse social networks, this study will consider three kinds of online social networks having high level of object-centered sociality: 1) watching network, 2) group membership and 3) academic collaboration network. Except for the academic collaboration network, the other two social networks exist online. Compared with offline

networks, they are relatively new and less bounded. Therefore, in spite of the continuously growing popularity, it is doubtful about the significance of the online social networks and whether the presence of the social relations defines the overlapped interests. Without any positive proof about this matter, the idea about the personalized recommendations using users' less focused online social networks is unworkable. The investigation about the shared interests of online social networks and how to measure the interests will be the first part of this dissertation (refer to the Figure 1). Once I find a reasonable degree of similarity in the users' self-defined social networks, the second part will develop recommendations which utilize the social networks and compare the recommendation quality with a conventional recommendation technology. The data source used for each kind of social network is different in terms of the data structure and the way for users to express their interests. Therefore, different recommendation algorithms for each social network kind will be examined.

1.1 PURPOSE OF THE STUDY AND RESEARCH QUESTIONS

The purpose of this dissertation is made up of two parts as shown on Figure 1.

It is important to note that the promise of SN recommendations is based on at least one assumption: users share similar interests with their online social networks. In order to check this assumption, the first objective of this dissertation is to explore online landscape which is intricately comprised of information items and social connections collected by users. Many researchers whose research interests are social media have tried to understand the related online

landscape and to discover its meaningful links with real life. Then, they built practical applications on the basis of these discoveries. This dissertation is ultimately to find out valuable items for users in the haystack of information by utilizing users' online social connections. Therefore, one must understand how users' social network structure is correlated with their information collections, as the preceding investigation of SN recommendations. The first part of this dissertation shown on Figure 1 will address this first objective.

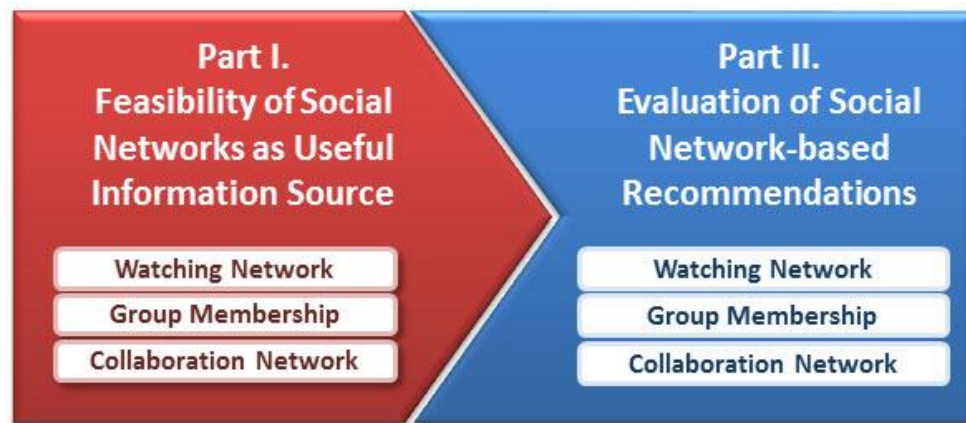


Figure 1. The Structure of Dissertation

The questions will be explored in this part are the following.

- **Q1.** Compared with pairs of users who are not socially associated, is it true that socially connected users share more common interests?
- **Q2.** What is a proper measurement for information similarity of two socially connected people?

- **Q3.** Do users have equivalent degree of common interests with their social connections as they do with their top N anonymous peers, which are chosen by conventional collaborative filtering recommendations (CF recommendations, hereafter)?

The second objective of this dissertation is to develop and evaluate recommendation algorithms using users' online social networks. Through the survey of the existing SN-based recommendation studies (Section 2.3), it is found that SN recommendation technology is still in infant stage. Mainly due to the limitation of available data sources, most of the SN recommendations are limited to two kinds of social networks - trust-based social networks and friendships. For friendship-based recommendations, most of the studies have focused on the acceptability of recommendation directly suggested by friends, rather than the improved quality. On the other hand, while trust network-based recommendations have tried to improve recommendation quality through the social network, the researchers recurrently used the same data set – Epinion.com. Although online social networks and sharing interesting information via the networks are ubiquitous on the current Web applications, we hardly find the attempts to utilize more diverse sociability and their interactions in recommendations. Therefore, this study will demonstrate the viability of less focused online social networks as a means for gathering useful information and the potentials of SN recommendations on various collective intelligence applications. The specific questions to be considered in this part are the following.

- **Q4.** Do SN recommendations produce better suggestions than traditional CF recommendations?
- **Q5.** For a given three kinds of social networks, are all of them useful sources of information?
- **Q6.** Can SN recommendations solve cold-start user problem?

1.2 CONTRIBUTIONS OF THE STUDY

The contribution of this study is twofold. First and foremost, compared with existing SN recommendation studies, the three kinds of social networks considered in this dissertation are relatively underused. They were rarely taken advantage of as a component of established SN recommendations. Therefore, it is expected that this study will demonstrate the potentials of new types of SN recommendations. There has been little research to date that investigates the feasibility and significance of various social networks for the sake of personalizing users' information spaces. Existing studies have focused on limited kinds of social networks. This may be caused by restricted number of available datasets. Otherwise, the researchers put too much emphasis on a few popular social networks, such as trust networks on Epinions.com and friendship on Facebook. However, while the three types of social networks introduced in the study are widely used in several applications, and users actively participate in the social networks, they have been rarely used as a component of personalized recommendations. This study addresses how different social networks of a novelty value can improve conventional personalization technology.

Second, it will add significantly new insight into how feasible the online social networks as sources of useful information. This study intends to reveal in-depth understanding of users' online social landscape where the social structures are intertwined with knowledge cultures. In other words, this study provides a comprehensible picture of diverse online sociability and knowledge sharing patterns. Nowadays, several kinds of new sociability emerge to distribute and acquire information. It is speculated that users on social applications may enjoy the social

interactions around shared information, and some studies prove that this speculation is true [92, 96, 191, 201]. However, the nature of various social networks relating to information similarity among members of the networks has not been identified comprehensively in literature. Most of the recent studies investigated how preeminent social science theories explored through small-size social data are different with large-sized online social data [136]. In addition, like the studies about social recommendations, the studies about online sociality are too much biased to a few popular collective intelligent sites, such as Facebook, Twitter, Wikipedia, etc. [53, 78, 84, 89, 171, 183] With evolving knowledge society around online social network, this study aims to give new insights and understanding about social interaction patterns centered on information objects. Therefore, it is expected that this offers practical insights to social media researchers and social application developers about how to utilize social networks as a critical mechanism to develop users' knowledge and improve their management of information.

1.3 STRUCTURE AND OUTLINE

This section provides an outline for the dissertation starting with the literature review.

Chapter 2 comprises a literature review on the recent studies about the interest similarity existing in online social networks, issues of conventional CF recommendations and the trends of recent SN recommendations. One element that explained in this literature review was the theoretical background describing various online social interactions in psychological view points and discussing how recommender system can use online social networks as a supportive

mechanism for users to manage their information spaces. In addition, by studying the state-of-the-art technologies of CF recommendations, I will explain why SN recommendations are necessary as an alternative of CF recommendations. Lastly, this section illustrates how recent works of literature about SN recommendation have used online social networks as foundations of the information personalization. This investigation of recent literature identifies that it is necessary to expand the SN recommendations into the technologies using more diverse and less focused social networks.

Chapter 3 introduces three kinds of social networks considered in this dissertation and explains the theoretical background about why these social networks are chosen. It provides the description of the concept of “object-centered sociality” and “personal familiarity” as theoretical backgrounds, since online social networks are made up of social structure and knowledge cultures mainly developed through these theoretical backgrounds.

Chapter 4 presents brief descriptions of information similarity measures and recommendation algorithms used in this dissertation. Even though more detailed explanations about similarity measures and SN recommendation algorithms specialized for a social network will be explained in each corresponding chapter of the social network, it is a necessary process to understand the similarity measures and recommendations. In particular, the chapter explains why each measure and recommendation algorithm was chosen among various options.

From chapter 5 to chapter 7 provides the findings regarding each social network.

Chapter 5 describes the results of bookmark recommendations using users’ watching network. In this chapter starting with the description of data source, since this unilateral relationship is newly emerged, the first part of this chapter suggests the proofs that watching

relations share similar interests and the feasibility of watching relations as useful information source. Using the outcome of the first part, the second part will explore the recommendations utilizing watching relations' information. The quality of this watching network-based recommendation was compared with traditional CF recommendations.

Chapter 6 describes bookmark recommendations using users' group memberships. In case of group memberships, there are two references to be considered – group per se and group co-members. When members of a group are able to collaboratively compose their group collection and additionally, each member also manages their personal information collections, it is curious which reference is more useful for users to acquire useful information. The first part of this chapter compares the information similarity of users with their groups or their co-members. The second part of this chapter is to generate recommendations using these two references and compare the quality with traditional CF recommendations.

Chapter 7 presents conference talk recommendations using users' collaboration network. According to the literature review, it is the first attempt to suggest talk recommendations. Collaborative network is a kind of professional relationships and highly correlated with their shared research interests. Therefore, it is expected that the talks that I am interested in could interest my colleagues. Therefore, in the first part of this chapter checks the presence of the shared research interests between colleagues. The second part is to generate recommendations using the collaboration network. In particular, in this study, recommendations are produced not only using the collaboration networks but also another kind of social network existing on the recommendation system. The quality of these two recommendations will be compared alongside CF recommendations.

The final chapter summarizes the major findings of this study. The findings for each question are presented as the answers of the research questions. Through answering research questions, the implications of the final results of the study are discussed. Following the discussions of the implications and research findings, limitations of the study and directions of future study are presented.

2 BACKGROUND AND LITERATURE REVIEW

In this chapter, I address several issues related to the topic of this dissertation. First, since main topic of this dissertation is to tailor users' information spaces based on the preferences of their online social connections, this literature review will start with the introduction of the studies about why users' online social networks can be a useful information source from a social science point of view. Secondly, various recommendation technologies will be explained. Since this study is to utilize users' online social networks as a part of their personalized recommendations, a basic understanding about the recommendation technology cannot also be overlooked. In this section, particularly, the problems of the current recommendation technologies and the needs of an alternative recommendation technology will be mentioned. In the third section, the in-depth understanding of existing recommendations based on social networks will be addressed. With the explosive popularity of online social networking sites, there are several attempts to take into account the online sociality in the information personalization (i.e. personalized recommendations). I present the current trends and explain the value of social network-based recommendations and its potentials as a solution of the current recommendation problems. In addition, I investigate whether the recent research of social network-based recommendations has any missing part to fill in. Lastly, since some of the data sources used in this dissertation are constituent of social tags, I introduced several relevant studies. Unlike numeric rating-based user

preferences, social tags express users' cognitive understandings or users' various view-points about items. Therefore, existing studies which considered social tags as users' preferences will be introduced in the last section.

2.1 ONLINE SOCIAL NETWORKS AND THE INTEREST SIMILARITY

This dissertation aims to assess how users' explicit online social networks are feasible as information sources and further, as foundations of personalized recommendations. The idea to utilize users' social networks in personalized recommendations is based on two preeminent theories about sociality – *homophily* and *social influence*. First, an old proverb saying “birds of a feather fly together” articulates with the definition of Homophily.

“*Homophily* is the principle that a contact between similar people occurs at a higher rate than among dissimilar people ([127, p.416]).”

This theory insists that people make social connections selectively with the other people who are similar to them, for instance, in terms of age, gender, religion, ethnicity, educational and occupational class, social positions, etc. Along with the homophily based on these personal traits, there is homophily based on people's perceived values and their internal states (i.e. *value homophily*). Therefore, it is also called as *social selection* [37]. A researcher suggested that the desire to be connected to the similar people is because of *social comparison* [50] and another researcher suggested that it is related to *constructuralism* [31]. People tend to use the others who are similar to them as a reference group and compare themselves with the reference to get

information or make a decision [50]. On the other hand, there is another suggestion that, due to the ease of communication, shared knowledge and other factors which make the interactions comfortable, people are more likely to construct social connections with the similar people [31, 117]. One psychological study showed that the perceived similarity is related to the attraction, as well [72].

As one of the exemplary studies, Barrtarjav, et al. (2005) shows homophily in online groups. In the study based on Facebook data, the authors focused on the group members' personal traits such as age, gender, religion, living area, political opinions, etc. They built clusters solely based on these traits and found distinguishable characteristics of each group. Depending on the discovered characteristics, they recommended a group to join and their recommendation accuracy was 73% on average. Therefore, similar people get together as a group. However, while they built the clusters, they excluded outlier users whose traits are not similar to the other members, and the outlier users (32% of the users in their data) were not covered by the recommendations. Put differently, for one-third of the users in consideration, the system was unable to find any recommendable group to join because they were classified as noises. Conclusively, people choose a group consisting of other group members who are similar with them. However, not only the similarity, there are other factors affecting the choice of groups as well, perhaps, such as social network structures in a group, usefulness of the information that each group has, etc. [15]

While homophily explains a mechanism how and why people selectively choose their social partners, *social influence* explains how people are affected by their social partners. *Social influence* represents the situation where people's attitudes, beliefs, and behavioral propensities are

affected by and adopted to their social ties [36, 37]. Deutsch and Gerard (1955) distinguished social influences into two distinct processes according to the expected results of the influences – *normative influence* and *informational influence* [43]. *Normative influence* is an influence to conform to the positive expectation of their friends or group members. The authors suggested that this influence process is based on the fact that people are interdependent with the others and seeks the social approval. Therefore, rather than normative influences between two people (e.g. friends), the normative influence among a bunch of people (e.g. group members) are stronger. People under this influence desire to be a member of a group or a small social network and sometimes, this influence transformed into power to urge them to coercive compulsion and compliance [43, 153]. A researcher interpreted this influence is rather emotional and irrational [176]. As another kind of influence, *informative influence* refers to “an influence to accept information from another as (trustworthy) evidence about objective reality [176, p.35].” This influence is motivated by ambiguity and uncertainty of reality. When people are not sure about an accurate view of reality and whether they are acting in a right way or not, they seek conformity from other people who are similar or have expertise and credibility. It is known that the desire of conformity is decreased by the increasing certainty of the reality or situation. This informational influence is also a part of *social comparison* [43]. Therefore, another researcher discriminated this informational influence from the normative influence in the point that the former is formed for a long time and internalized without any outer pressure [177]. While Deutsch and Gerard suggested that two influences occur together, in the current Web 2.0 era of which the information is created by countless anonymous Web users, the informational influence can occur by itself to collect accurate and useful information. As the results, it is critical to

choose the social partners whose information is reliable and accurate. By experiencing various SNSs, it seems that Web users are knowledgeable about the importance of reliable social partners. It is because, through this social influence, social capital is more valuable than ever, and SNSs is one of the most important marketing tools [106].

Brzozowski and the colleagues (2008) examined social influences of three kinds of social connections – friends, allies and foes. In their paper, friends are based on personal familiarity (i.e. strong ties). Allies do not necessarily have personal familiarity but share similar ideologies (i.e. weak ties) (The data source of this study is online forum for political debates; hence users' ideologies are important criterion to consider). Foes are the ones having conflicting ideologies or political views (i.e. negative tie). In the study using an online forum which aims to vote on various controversial political topics and to share the related opinions, they examined what resolves users vote on and how they will vote. They found that users' friends were more influential in the choice of resolves to vote on than allies and foes. That is to say, when their friends voted on a certain resolve, users are more likely to vote on the same resolve. In addition, interestingly, users tend to avoid resolves that they foes already voted on. However, users voted resolves in more similar pattern with their allies than with their friends or foes. Of course, users did the least agree with their foes. They concluded that strong ties play a more important role as a social filter than weak ties or negative ties [28]. However, it seems that all three ties played certain roles as social filters. For instance, strong ties and negative ties are important in the choice of items to consider and allies are important to acquire favorable items.

Singla and Richardson (2008) also tested how frequent contacts affect social connections' interests. The authors computed the correlations between instant messenger logs and the

similarity of search queries. The results demonstrated that search interests of a user pair who exchanged instant messages to each other frequently were more similar than interests of random pairs. Moreover, the longer they talked, the more similar they were [163].

Another study [3] explored semantic similarity among social contacts. They examined word terms on the Stanford and MIT personal home pages and the users' social connections. The terms in homepages, in-links, out-links and mailing lists were analyzed to see how the information similarity predicts the friendship connections. All these four kinds of information appeared to be similar for socially connected users. There was also proportional co-relationship between information dissimilarity and the relationship distance [3].

Yang and Chen's research (2008) was also based on social influence. In an educational P2P system, knowledge and social networks were combined to suggest items to download. Firstly, the system picked up the most knowledgeable and the most preferred users and then suggested the target user to interact with the picked users using instant messenger. As time goes by, the more often a user had interacted with the target user, the more likely his information was recommended to the target user. The authors considered social networks not only through the strength of a tie at a certain time point but using evolving change of the tie. Even though the thoughtful design of recommendation structure, the limitation of this study is too much requirement of human interventions. In order to choose the most knowledgeable user for a query term, the system needed evaluations of human experts. The users' preferences about other users were also based on the explicit ratings. Differently from this educational domain, in general system, it is impossible for human experts to evaluate knowledge of every individual and let users to give ratings for all other users [191].

Although social science researchers theoretically distinguish between the homophily and social influence, the theories suggest one common point – similar interests between social connections. Homophily indicates existing similarity prior to the social relations, and social influence shows evolving similarity through interactions with social relations. In addition, two theories are highly correlated to each other. The cumulative studies about the homophily have showed that the social relations with the similar people affect a person's way of thinking, his attitude, the information for him to choose, and many other his personal perspectives [127]. Furthermore, when two socially associated people are alike to each other, their reciprocal influence is stronger and longer than another pair who is less similar to each other [126]. Therefore, many studies about social dynamics of online social networks have focused on whether a pair of social connections bears similarity or not, regardless if the similarity is an existing one or evolving one.

Ziegler and Golbeck (2007) compared interest similarity between people in a social network in traditional CF context. They used user information and the user's trust-based social networks in the book and movie recommendation domain. For the book data set, to lessen the data sparsity problem, they grouped the items by topics using an existing taxonomy, rather than counting each book. Since this study focused on the similarity of users' topics of interests derived from the taxonomy, not from the books per se, it was hard to see a clear picture of information similarity on the item level. As the authors pointed, the similarity is highly dependent on the taxonomy's design. Yet, in the second study with movie data set, they compared user similarity on the level of individual items and found that the average ratings of each movie became more similar as two users trusted each other more [200].

Referral Web [83] was an information retrieval application utilizing social influence. In this system, the authors extracted social network information automatically through publically available Web information such as co-authorships of papers, organizational charts, or message exchanges in newsgroups. Social network was used as a way to disambiguate search queries and to reorder the search results. The system chose a certain term which the direct or indirect social networks of a searcher are familiar with if there are synonyms and gave a higher priority to the results more related to the searchers' social connections [83].

There are also some social interaction-based games such as Dogear and ESP game. Using knowledge about 'who is interested in what' among our social connections, users guess who might own the suggested information among their social connections. When a user's guess is incorrect, the system asks him whether the friend they guessed could be interested in, and if the answer is positive, it recommend it to the corresponding friend [45, 180].

2.2 RECOMMENDATION TECHNOLOGY: STATE OF ART

The overwhelming amount of information available on the Web makes users difficult to locate the right information on the right time. Personalized recommendation technologies emerged as a solution of the information glut problem and have proven the effectiveness in a number of real-life applications, such as Amazon.com, Netflix, Last.fm, Youtube, and Google, just to name a few. The statement of Jeff Bezos, CEO of *Amazon.com* explained what recommendation systems are for.

“If I have 3 million customers on the Web, I should have 3 million stores on the Web [159].”

Depending on how to derive users’ preferences and how to select candidate items, the technologies are classified into three kinds – collaborative filtering based, content-based and hybrid recommendation [4, 29, 147].

As the first recommendation technology, collaborative filtering (CF) recommendations employ a process of ‘*word of mouth*’ systematically. The power of this technology is based on a relatively simple idea: start with a target user’s ratings, find a neighborhood of users who have similar interests to the target users and recommend items favored by this cohort to the target users. Because the technology mainly analyzes patterns of ratings across users or items, it has been possible to discover presumably favorable items without analyzing any content properties of items. For taste-oriented domains (e.g. music, pictures, videos, movie, jokes, cultural events, etc.) where content information is unavailable, or the tastes and interests are hardly captured by content properties, CF technology is an only viable technology. A common algorithm to find a target user’s cohort is to compare the similarities using K-Nearest Neighbor (KNN) approach. Euclidean distance and Pearson correlation are widely used similarity measures to find the neighborhood [160, pp. 7 ~ 28]. Clustering models, Bayesian Network model, Gibbs sampling, probabilistic relational model, a linear regression, maximum entropy model, and matrix factorization of user-rating matrix are other algorithms used in CF technology [4, 24, 42]. Not only rating similarities among users but similarities of metadata or properties among items are taken into consideration [70].

In spite of the strengths and big success in both academia and industry, questions about CF recommendation quality have been arisen, because personal recommendations are utterly tailored by the tastes of unknown users who are selected by fully automated similarity calculations [159]. When ad-hoc users who have malicious intentions attempt to distort systems or make profits, the CF recommendation systems are totally vulnerable. In an actual incident about Amazon's sex link gaffe¹, a group of ad-hoc users intentionally copied the whole profiles of Amazon.com users who bought a Christian spiritual guide book, and the recommendation system picked these ad-hoc users as perfect peers. Naturally, the recommender system suggested the items favored by the ad-hoc users, as it turned out, which were male sex manuals [194]. A group of ad-hoc users are also able to reinforce their own ratings and shift the recommendation predictions to the directions they intended [140]. Beyond these examples, there are several studies showing the vulnerability of CF technologies to the attack of malicious users [91, 119, 128, 140]. Even for well-intended users, if their tastes are eccentric (so called '*black sheep users*'), it is not easy to find their peers due to the little overlap in tastes with other users [159]. CF recommendations also suffer from other problems like data-sparsity problem, cold-start users, and computational overload. When there are too many items to be rated in comparison with the total number of users or when the items last for a short period of time (e.g. job openings and news articles), the rating data is distributed too sparsely or lasts too short for the system to find sufficiently co-rated items among users [4, 130, 159]. For the cold-start users who just join a system and have insufficient number of rated items, it is also impossible or very difficult to generate CF recommendations [4, 118, 159]. Additionally, because CF recommenders require to

¹ <http://news.cnet.com/2100-1023-976435.html> (accessed on April 18, 2011)

compare one target user's taste with the tastes of all other users, it is computationally very expensive [118]. I suggest that all of these problems and shortcomings of CF technology occur in part because the CF recommendations are relying on fully automated similarity computations in the selection of peers. Although users are the recipients of CF recommendations, the system doesn't allow them to get involved in the recommendation process. Once CF recommender systems choose their peers, users don't know who they are and are unable to include or exclude peers as they want, even though the source of the recommendation is important criterion for judging the quality of recommendations [20]. Therefore, it is required to reconsider the current CF recommendation technologies in order to find a way to make users to participate in their own recommendation process.

As the next recommendation technology, content-based recommendations learn user preferences from the content of their favorite items and similar information are suggested to users based on the content properties [134]. Typical content based recommendations analyze full text content and otherwise case-based recommendations, which is a special kind of content recommendation, analyze semi-structured or fully-structure meta data [123]. Various information retrieval technologies, for example, decision tree, rule induction, nearest neighbor methods, VSM, linear classifier, probabilistic language model, naïve Bayes, and artificial neural networks are the algorithms for content-based recommendation [4, 130]. The strength of content-based recommendation is that cold-start problem can be reduced because the recommendations are available with small number of ratings or even no ratings [147]. However, the weakness of this recommendation technology is to suggest too obvious and boring recommendations [147]. If there is an outbreak of gun violence, for example, it will be constantly on the news and in the

papers and magazines. If users expressed interests on a relevant news article, their recommendations will be plastered with a bunch of the same news about this gun violence due to the same contents.

Finally, hybrid recommendations are to combine more than one recommendation technology. As introduced, each recommendation technology has the pros and cons. In order to make up for the shortcomings and maximize the recommendation capability, various hybrid systems have been introduced. Especially, not only the CF and content-based recommendations, but other less popular recommendation technologies such as knowledge based or demographic based technologies have been hybridized [29]. Ziegler, et al. (2004) introduced possibility of a new hybrid approach through CF technology and metadata –a ‘collaboration via content’ approach. In book recommendation domain, they applied topic-related taxonomy information to users’ favored book list. They matched each book with topics in a taxonomy using descriptive terms such as topic descriptors and then built user profile vector using the matched topics. Therefore, the system was able to generate recommendations for two users who do not have any common book. The user and book information were collected from *AllConsuming* and book-related taxonomy was book category of *Amazon.com*. Especially, since the taxonomy was hierarchical, the decayed weights of super-topics were considered. For the evaluation, the authors compared their taxonomy based CF recommendations with the ‘naïve’ random recommendations and traditional CF recommendations using item-to-item comparison. In the experiment by k-folding evaluation, the precision and recall of the taxonomy based CF outperformed other approaches. They also evaluated the recommendation with 51 subjects and the ratings and satisfactions for taxonomy based CF were better than the others [202].

Melville, et al. (2002) explored a hybrid movie recommendation to address data sparsity and cold-start problem. Users' sparse ratings were converted to pseudo-ratings based on content-based predictions. In specific, users' movie ratings acquired from *EasyMovie* were extensively inferred by movie content information (title, director, cast, genre, plot summary and keywords, user comments, external reviews and awards) collected from *IMDb*. Therefore, when computing user-to-user similarities, even though two users in consideration didn't co-rate the same items, once they rated similar items (i.e. the contents of the rated items are similar), the recommendations system concluded that they have similar tastes and this is the pseudo-ratings. Then in order to ensure the accuracy of the pseudo-ratings, Harmonic Mean weighting (HM weighting) was used. If the pseudo-ratings are based on at least 50 co-rated items, the weight will be the highest, and otherwise the weight is devalued accordingly. For comparison, their content-boosted CF was compared with pure content-based, pure CF and naive hybrid method (which is the recommendations simply averaged out ratings from pure CF and pure content-based predictions). In the evaluation using MAE (Mean Average Error) and ROC (Receiver Operating Characteristic), their recommendations of content-boosted CF were better significantly [130].

Li and Zaiane (2004) also suggested a way to solve a data sparsity problem and a cold-start problem in the recommendations of Web pages using hybridization. User preferences of this study do not rely on users' ratings but are based on users' usage patterns. By analyzing content information of every visited Web page, users' navigational logs were split into missions under the assumption that each session consists of several interests or missions. Therefore, to make content coherent clusters, pages having similar content within a session were grouped together. These clusters of all users were aggregated as navigational model. When a user starts a new

session, the system identified the navigational pattern and tried to match the pattern with the already defined navigational models and recommend the best Web sites to visit. The results from experimental evaluation based on a university department Web site showed that their algorithm performed better than association rule-based system in terms of accuracy and coverage [103]. Implicit preference inferred by users' usage patterns, however, may be incomplete, limited or inaccurate due to the poor design of web sites or incorrect heuristics. Especially, newly added Web pages have little chance to be suggested as recommendation because it is not included in the usage history.

2.3 SOCIAL NETWORK-BASED RECOMMENDATIONS

Due to the effectiveness and technical innovativeness, as explained, CF recommendation approach has succeeded in many real-life applications. However, one of the weaknesses of the current CF recommendation technology is the lack of user involvement. Even though users are the recipients of recommendations and the recommendations come from the information they contribute, the recommender systems don't allow them to get involved in the recommendation process. In addition, due to the utterly automated recommendation process, the selection of peers is a black-box to users. Users don't know who their peers are, even though their recommendations completely tailored by the tastes of the anonymous peers. When users want to include their self-selected peers or to exclude unwanted peers, there is no way to do so.

Even though the current collaborative filtering recommendation is totally relying on the wisdom of crowds from anonymous peers, in early generation of the CF recommendations, users' social networks were core parts of the recommendation process. One of the earliest CF recommendation systems, *Tapestry*, was based on explicit social connections and allowed users to retrieve personalized contents, using annotations added by their friends and colleagues [159]. Another pioneering CF project combined explicit social connections with the active "push" approach: users could directly send interesting research papers to other colleagues [115]. However, these pioneer systems relied on exchange of information within a "small world" and found it difficult to retain users and to keep them actively contributing to the recommendation process. As the CF-based algorithms became mature, automatic recommendations by computation are now dominant. Thanks to the success of various SNS sites, however, we are now able to have abundant information about less bounded and wider ranging online social networks. In addition, with the help of Web 2.0 applications, it became much easier to share our current interests and activities more actively than the era before the Web 2.0. Therefore, in order to let users get involved in the recommendations, rather than asking users' additional input about their preferences, it is effective to utilize their existing online social networks. In this section, I introduce the recent studies about social network-based recommendations and address some lacking points of this research trend.

2.3.1 Definition of Social Network-based Recommendations

When this social network-based recommendation is one of the critical and innovative matters of information personalization research, first, we need to define what the social network-based recommendation is. According to the literature review, there is not one solid definition of this technology. Depending on the purposes of their studies, researchers defined this recommendation technology in broad ways. The following is the varying definitions of social network-based recommendations.

- Recommendations based on users' explicit (i.e. self-defined) social networks
- Recommendations based on users' implicit social networks which are inferred by their information or profile similarity with the other users
- Recommendations based on communication patterns among intelligent agents, not human beings.
- Studies to recommend social partners using traditional recommendation methods.
- Studies to explore how people feel trustworthiness on suggested recommendations
- Studies under the name of 'social recommendation', even though the proposed recommendation was the traditional collaborative filtering approach which is purely based on collective intelligence of social systems or social media without any consideration of users' social networks

The purpose of this dissertation is to explore various recommendations studies utilizing users' explicit social networks. That is, the social networks considered in this dissertation are the ones explicitly defined by user. Therefore, this study is following the first definition. The

explicitly defined social connections are important because they have social influence to each other and embed a value as social capital. Once users started their social relations with own choice, they know who their social partners are and pay attentions to their partners' online activities. Hence, they are easily affected by their social partners. The recommendations using users' explicit social connections take into account not only the information about 'what users like', but also 'whom users are interested in.' The current CF recommendations are assumed that the roles of all users in the recommendations are equal, in turn, 'role uniformity [173].' However, generating recommendations and accepting the suggestions are decision making processes. Every user has different interests, specialized knowledge, and especially different social context and social roles in terms of the kind and strength. For this reason, it is important for users (i.e. the recipients of recommendations) to define the foundations of their recommendations and to know a source of recommendations (i.e., where the recommended information comes from). Table 1 summarizes existing recommendation studies based on users' explicit social networks.

On the other hand, there are several recommendation studies based on users' implicit social connections. Implicit social connections mean that user-to-user links are inferred by automatic computations of a machine. Thus, the choice of machine learning methods and the computation details are the major determinants of the users' implicit social networks. Put differently, depending on what kind of machine learning method is used, users' social partners will be changed. In addition, users are unable to know who their social partners are, and hence it is hard to expect the resultant social interactions such as social influence and social capital. Debnath, et al. [40], Liu and Maes [109], Lumbreras and Gavalda [110], O'Donovan & Smyth [140, 139], Palau [143], Stecher, et al. [165], Victor, et al. [179], Walter, et al. [181], Pitsilis and

Knapskog [150], Xue, et al. [190] are the recommendations studies based on implicit social networks. The practical reason why the researchers chose the implicit social networks instead of explicit ones is the lack of available data sources. Actually, the data sources available for social recommendations (which consist of users' ratings or bookmarks and their social connections) are severely limited. In the following section, I will explained this matter in more details, but since the researcher have a difficulty to find eligible data sources for their recommendations, they just created their own using machine learning methods.

The studies belonging to the third definition are focusing on how to exchange information from one machine to another machine, not the information sharing among human beings. Depending on how one machine is trustworthy to another machine, the trust values are inferred and propagated. Olsson [142], [90] and Shi, et al. [162] are the examples.

As another kind of social network-based recommendations, there are attempts to recommend people (i.e. social connections), instead of favorable information objects. The recommended people could be a male/female to date [5, 88, 151, 152], a person to befriend with [7, 30, 63], a colleague to work with [100, 101, 125, 173, 189] or a community to join [6, 184]. However, this dissertation is to suggest users their presumably favorable items. Hence, the topic of this research trend is out of my focus for now. Nonetheless, the people recommendations are one of the most active areas of research regarding the recommendations and one of the most important research agenda along with the social network-based recommendations of items.

The studies belong to the fifth definition is to explore users' perceived trustworthiness about recommender systems. Since this kind of research focuses on assessing how recommender systems have generated reliable suggestions, it is important aspect we need to take into account,

but the focus is not quite matched with the points that this dissertation intends. Masthof [120] and Wang and Varadharajan [185] are the examples of this kind.

Lastly, there are other studies just borrowed the terminology from ‘social recommendations’. The technology and topic of the studies which belong to the last sixth definition are nothing to do with social recommendations or social network-based recommendations in our consideration. They investigated how to use the traditional collaborative filtering recommendation technology in social applications or social media. Groh, et al. [146], Sanchez, et al. [157], Victor, et al. [178] and Zhou, et al. [198] are the examples.

As mentioned, the main focus of this dissertation is the personalized recommendations based on users’ explicit social networks. Rooted on this definition, the relevant studies are reviewed. The following section will give us basic understanding of the current trends and address the lacking points.

Table 1. Summary of Social Network-based Recommendations

Paper	Input	Recommendation Method	Kinds of Social Networks	Target Items and the Dataset
Trust Network-based Recommendations				
Avesani, et al [14]	Ski resort routes and users' FOAF (Friend-Of-A-Friend)	Selective choice of information entered by reliable users	FOAF and the trust values	Ski Resorts (Moleskiing.it)
Al-Sharawneh & Williams [8]	Item ratings and users' trust ratings	Hybridization of trust values and item ratings	Trust-based networks	General Items (Epinions)
Chia & Pitsilis [35]	Item ratings and users' trust ratings	Filtered peers based on trust ratings and user properties such as experience level and rating similarity	Trust-based networks	General Items (Epinions)
DuBois, et al. [44]	Item ratings and users' trust ratings	Modification of CF recommendation with Clustering coefficients of trust networks (correlation clustering algorithm)	Trust-based networks	Movies (FilmTrust)
Golbeck & Hendler [56]	Item ratings and users' trust ratings	TidalTrust (trust network inferring algorithm according to social distances)	Trust-based networks	Movies (FilmTrust)
Guo, et al. [62]	Item ratings and users' trust ratings	The nearest neighbors were chosen from users' anonymous peers and their trusted connections in a unified way.	Trust-based networks	Movie (FilmTrust & Filxster) and general items (Epinions)
Jamali & Ester [73]	Item ratings and users' trust ratings	1) content similarity computed by ratings and 2) random walk	Trust-based networks	General Items (Epinions)
Jamali & Ester [74]	Item ratings and users' trust ratings	Latent feature and trust propagation induced by a matrix factorization technique	Trust-based networks	General Items (Epinions) & Movies (Flixster)
Ma, et al. [112]	Item ratings and users' trust ratings	Fusing user-to-user social matrix with user-to-item rating matrix using the probabilistic factor analysis	Trust-based Networks	General Items (Epinions)
Matsuo & Yamamoto [122]	Item ratings & users' trusted networks	'Pull'-based recommendation of trust-based networks (when a trusted party posts a new product with positive opinion, the recommender system picks that product for the trusting user as a	Trust-based networks	Cosmetics (@cosme)

		recommendation).		
Mgureanu, et al. [114]	Item ratings and users' trust ratings	Based on user' ratings patterns and their trust statements, the system clustered users and recalculated the trust ratings and infer the trust values to other users.	Trust-based Networks	General Items (Yahoo! Webscope & Epinions)
Massa & Avesani [119]	Item ratings and users' trust ratings	Combination of rating similarity and direct or propagated trust ratings in the selection of peer cohorts	Trust-based networks	General Items (Epinions)
O'Doherty, et al. [141]	Item ratings and users' trust ratings	Hybridization of trust values and item ratings	Trust-based networks & Implicitly inferred trust-based links	Movie (FilmTrust & Filxster), Dating Partners (LibimSeti), General Items (Epinions)
Walter, et al. [182]	Item ratings and users' trust ratings	Prediction of ratings according to the ratings of users' direct and indirect trust-based social networks	Trust-based networks	General Items (Epinions)
Friends-based Recommendations				
Bonhard, et al. [21]	Users' demographic profile and users' ratings	Direct recommendations from friends who know target users well with explanations of recommenders' identity	Friendships	Movies (MovieMatch)
Bourke, et al. [22]	Users' favorite items and choice of friend neighborhood	Aggregation of items favored by users' friend neighborhoods	Friendships	Movies/TV (Facebook)
Pera & Ng [148]	Users' favorite books including the tag clouds and the average ratings, the ratings given by all users and users' social connections	Hybridization of tags assigned to books and users' friendships information (In the choice of candidate items, content information used and the subsequent rating prediction, the ratings given by users' friends were taken into account)	Friendships	Book (Amazon and LibraryThing)
Sinha & Swearingen [164]	Users' buying/browsing history and the names of their friends	Direct recommendations from friends who know target users well.	Friendship	Books (Amazon, Sleeper & RatingZones) and Movies (Amazon, Reel.com, & MovieCritics)
Konstas, et al. [85]	Users' music play-counts, users' tags and	Application of the Random Walk with Restarts to a graph depicting users, tracks, tags and user-to-	Friendship	Music (Last.fm)

	their friends	user relations.		
Groh & Ehmig [59]	Item ratings & users' social networks	CF recommendations with substitution of users' friends for their peer cohorts.	Friendships	Local Clubs (A German Site)
Liu and Lee [108]	Item ratings & user's social connections	Hybridization of CF (choice of social connections, instead of peer users or fusing with peer users by adding different weights.)	Friendships	Online Products (Cyworld)
Other Social Networks-based Recommendations				
Lee & Brusilovsky [93]	Users' bookmarks and their watching network	Hybridizations of watching network-based recommendations with CF approach	Watching Network (i.e. Unilateral Network)	Academic articles (Citeulike)
Lee & Brusilovsky [99]	Users' bookmarks and their group membership	Hybridizations of group membership-based recommendations with CF approach	Group membership	Academic articles (Citeulike)
Guy, et al. [64]	Users' tags, bookmarks, comments and their social networks	Calculation of the number of social connections who give some feedbacks to bookmarked pages, and the relations between users and bookmarked pages and combination of these values.	Professional colleagues	Bookmarks of Internet or Intranet pages (Lotus Connections)
Yuan, et al., [82]	Users' music play-counts & their social network	Two hybridizations of SN and CF (Similarities weighted by social network info vs. similarities fusing with random walk)	Group memberships & friendship	Music (Last.fm)

2.3.2 Limited Kinds of Social Networks in Personalized Recommendations

With the success of online networking sites, the kinds of social connections possibly used for social network-based recommendations widely ranged from undirected networks such as friendships [59, 85], colleagues [115] and co-membership of a group [99] to directed network such as trust networks [32, 73, 162, 181], following/watching relations [51, 93], and email senders and recipients [155]. Table 1 summaries the existing recommendation studies based on users' self-defined social networks. In this review, I have been trying to collect exhaustive list of the related publications. Up to now, as the summary table indicates, most of SN-based recommendations have concentrated on two kinds of social networks; trust-based social networks and friendships. In case of trust network-based recommendations, most of them concentrated on how to solve the current CF related problems or improve the CF recommendation quality by fusing the social context in the traditional CF recommendations. On the other hand, friendship-based recommendations range from simple direct recommendations to hybrid recommendations fusing users' social context with their content properties. Interestingly, the focus of the friendship-based recommendation studies was usually users' acceptability of the recommendations suggested by their friends and usefulness of the explanations, rather than the improved quality via systematic evaluations.

The summary Table 1 now provides clear evidence that the researchers put too much emphasis on a few popular social networks. Among 25 studies reviewed on the table, more than 60% of them are trust network-based recommendations and about 30% of them are friendship-

based recommendations. Among the remaining 4 studies, if my own studies are excluded, one study was based on a professional network and another study was also based on friendship but additionally considered group memberships. Moreover, the data sources used for these social network-based recommendations were severely limited. It is not easy to find trust network-based recommendations not using the Epinions.com data set. In case of friendship-based recommendations, there is no dominantly used data source, but some recent studies [85, 192] started to focus on Last.FM data, since the system provides an API (application programming interface) for researchers to be able to collect the data conveniently. If these studies have used the same evaluation framework (for instance, the same set of target users, the same test settings and the same evaluation criteria) as the Netflix Prize² contest did, then using the same data source for multiple studies is meaningful and worthwhile, but they did not.

Conclusively, even though they are several social network-based recommendations, the kinds of social networks and the number of data sources used in the information personalized are still restricted. Consequently, in order to fill the gap, this dissertation will test the feasibility of diverse social networks as useful information sources and to assess whether the social network-based recommendations can be expanded to more varied types of social connections, using rarely explored datasets. Because they have been explored by many other researchers, I will not consider trust networks and friend networks in this dissertation.

² <http://www.netflixprize.com/>

2.3.2.1 Trust Network-based Recommendations

Golbeck (2009) says that the traditional definition of ‘trust’ is related to security and reliability, but the broader definition of trust for nowadays is related to ‘a matter of opinion and perspective’. She refers to the broader trust as ‘*social trust*’. She suggests that information can be aggregated, sorted and filtered through social trust [54]. In her own study, Golbeck also showed that users prefer recommendations from trusted people to CF recommendations [56].

Massa and Avesani’s study (2004) showed that a user’s trust network can solve the ad-hoc user problem, improve recommendation prediction and attenuate the computational complexity. Using Epinions data set, a trust-based technology generated more precise recommendations than CF technology. In addition, for users with 4 ratings (i.e. cold start user), trust-based technology could make the correct recommendations for 66% of these users, while CF could make recommendation for only 14% of the users with a higher margin of error [119].

The same authors (2007) compared local trust matrix and global trust matrix in information personalization. Local trust matrix refers to the computation of missing trust values based on users’ own personal trust values. In user-to-user trust network matrix, there could be missing trust values which are not explicitly expressed by users. However, it is possible to propagate the trust weights within the trust-based social network. Therefore, local trust matrix is personal and subjective view of users. On the other hand, global trust matrix predicted the missing trust values using reputation values, which approximate how much the whole social network trusts the nodes. It is not personal but standardized. In recommendation, the authors combined the trust values (either local trust or global trust or both of them) with similarity of ratings. They compared the trust network based recommendation with typical CF

recommendations using Epinions.com data set. In evaluations centered on rating prediction and user coverage, trust network-based recommendation always produced better suggestions than CF recommendations [119].

Other studies indicated that a trust network decreases the recommendation error and increases the accuracy as well [140, 182]. For users with a unique taste, their own trusted network could increase the satisfaction of recommendations, since they are able to know where the information came from [174].

Al-Sharawneh and Williams (2010) classified users into two groups: leaders and followers based on credibility. The credibility means a combination of trustworthiness and expertise. Trustworthiness is calculated using explicit trust ratings and their transitive powers (both credibility from direct followers' trusts and from indirect followers' trusts). Expertise is calculated using the distance of each user's rating distributions on the center of average ratings for a corresponding item. That is to say, when the differences of a user's ratings with the average ratings are smaller, the users' expertise is higher. In order to recommend items, they aggregated leaders' ratings for a candidate item and the average ratings were combined with a target user's average ratings. According to their experiment using Epinions dataset, they found that leader-based recommendation is good to recommend items to cold start users and increase recommendation coverage. However, it is in question how the quality can be improved since the recommendation is based on all other users' ratings regardless of the similarities with a target user and hence is very weak personalization. In addition, they added more expertise to the users who rated more items. For an ad-hoc user who tries to distort the system with many false ratings, there is more chance that the ad-hoc user becomes an expert [8].

DuBois and the others (2009) utilized sub-clusters of trust-based social networks in recommendations. Compared with their previous studies which are based on trust-based social networks [55, 56, 200], in this study, the authors took a set of points in a directed social network matrix and groups them on the center of random centroids with a fixed radius of these centers. Then they computed the CF-based recommendation algorithm by substituting for peer cohorts with users' trusted parties. Especially, when the trusted parties are in the same sub-cluster with their target users, the similarities were reinforced more. Cluster-enhanced recommendations performed better than other recommendations not using the sub-clustering structure. However, the qualities of their trust-based recommendations were worse than traditional CF recommendations [44].

Ma, et al. (2008) intertwined two matrices of users' information (user-to-user social network matrix and user-item rating matrix) through the shared user latent feature space, in order to incorporate users' social network context in the typical recommendations which are based on rating similarities. In this paper, the social network is trust-based. They used probabilistic factor analysis and the resultant matrix was filled with the probability of user-rating prediction. Interestingly, in computing trust values, they took into account the frequency of trust interaction. It means that, when a user trusts a lot of users, the confidence of the trust value is decreased. When a user is trusted by a lot of users, the confidence of the trust value is increased. In the experiment using Epinion's dataset, they compared the quality of their so called '*SoRec*' with other matrix factorization-based CF recommendations. Regarding the accuracy and efficiency, the SN-based recommendation, *SoRec* is better than CF recommendations. In addition, for cold-

start users having very few ratings or even having no rating, *SoRec* produced better recommendations [112].

Ma and the colleagues (2011) integrated three aspects of recommendations altogether, users' trust networks, items and users using the probabilistic factor analysis. They built one latent feature matrix about user-user social networks and another latent feature matrix about user-item, separately. Then, they integrated the two latent matrices through the shared user space. Especially, their integrated latent model is based on the assumption of the social influence – users' preferences are influenced by their social networks. In the empirical evaluation using Epinions' dataset, their social network-based recommendations were more accurate and efficient than other typical matrix factorization techniques (e.g. non-negative matrix factorization, probabilistic matrix factorization, and SVD) and user or item-based neighborhood-based CF. As the authors emphasized, the recommendations based on social networks were able to produce fairly good prediction even when users have very few ratings or no ratings [113].

Matsuo and Yamamoto (2007) introduced a simple 'pull' approach-based recommendation using trust-based networks in recommending cosmetics. Once they think another user's product reviews are useful or matched with their opinions, users add the reviewer as their favorites. Then the system actively referred to favorite products and opinions of the trusted reviewers. That is to say, when users' trusted reviewers posted a product with positive opinions, the system pulled the information of that product and showed it to the users, as a recommendation [122]. This recommendation was based on minimum level of computation. Probably due to this little computation, the quality was bad (the precision was just 1.49%). In

addition, there was no baseline to compare with. Hence, it is hard to guess the recommendation quality when it is applied in other domains or other applications.

2.3.2.2 Friendship-based Recommendations

In order to take into account not only users' tastes about music but also their tags and social networks (i.e. friendships), Konstas, et al. (2009) applied the Random Walk by Restart (RWR) algorithm on the graph consisting of users, music tracks, and social tags as nodes and play-counts, tag annotations and friendship relations as edges. RWR provides a good relevance score between two nodes in a weighted graph. For a given graph, it measures how two nodes are closely related to each other. They used the node relatedness weights between a target user and a music track which is not discovered in the target user's play-counts to predict recommendations. The more closer the candidate track is to the target user, the more probable it is to be recommended to him. In the experiment using Last.fm, their RWR approach produced better result than typical CF recommendations in terms of precision and recall [85]. However, the problem of this study is based on very much skewed data source. When the authors collected their data source, they collected top 50 fans of the top 50 musicians who are the most popular at that time. That means they are really active and leading users of the Last.fm system. The information collection and sharing patterns of much passionate users who were chosen as top 50 fans of the most popular musicians would not be same with other users who are randomly chosen. Therefore, the feasibility of this approach on other datasets is in question.

Pera and Ng (2011) introduced a SN-based recommendation algorithm which is usable for book recommendation domain. Their recommendation is a hybrid approach fusing metadata

properties of books with users' social context. First in order to choose candidate books to recommend, the tag-based content similarities between the candidate books and target users' favorite books were counted, rather than the similarities of contents derived directly from the books (such as from the titles, abstracts or the authors' names). In the subsequent process, they aggregated the ratings of the candidate items given by the target users' friends rated. They also computed how the friends' tastes are similar to the target users are. In the experiment using LibraryThing, they contrasted the SN-based recommendation with the CF recommendations provided by Amazon and content-based recommendations provided by LibraryThing, as baseline. As the result, the quality of the hybrid recommendations combining metadata information and friend relations was better than other two baseline approaches in terms of precisions and ranks [148].

Groh and Ehmig (2007) compared a friendship-based recommendation with collaborative recommendations. Users were asked about the tastes of Munich-area bars. The authors compared the taste similarities with between a user and his friends and finally tested the quality of the friendship-based recommendations. As the results, users in social networks tend to share more similar tastes than randomly paired users and furthermore, the larger the similarity is, the more people form cliques. In recommendations, the friendship-based recommendations performed better, especially when the ratings are very sparse, and produced more novel suggestions than CF-based recommendations. As pointed in the discussion, however, the domain in consideration is in highly social context. When we go to a bar, we are usually in company with our friends. Therefore, in different domains the quality of the recommendations could be different. In addition, the proposed friendship-based recommendation performed better in only special

circumstances such as very sparse ratings or novel recommendations. In other cases, it was worse than CF [59].

Liu and Lee (2010) introduced a recommendation purely using users' friendships based on user study with a Korean online networking site, Cyworld. They suggested items based on typical CF approach (based on the nearest neighbors' preferences), SN-based approach (based on friend's preferences), and hybrid approach (based on the combination of both the nearest neighbors and users' friends). Even though they tried to augment the effects of social network more than the peer users with different weights, the naïve combination of peer and social connections performed the best. In addition, social network-based recommendation performed the worst [108].

2.3.2.3 Recommendation based on Other Kinds of Social Networks

Guy and the colleagues tested recommendations based on professional colleague network. They compared familiarity-based and similarity-based recommendations using a social software application suite for organizations. Given that typical CF recommendation was similarity-based approach, familiarity-based approach suggested items that people in my social networks were interested in. They collected users' self-defined social networks using SONAR, a system to aggregate the social connections across different sources (organizational chart relationship, co-authorship of one paper or patent, etc.) within a company. The recommendation algorithm counted the number of social connections who possessed one candidate item and the relation type of the social connections about the item (authorships, commenters, bookmarkers, or taggers). Depending on the relation types between users and items, they added different weights.

According to the resultant values, the recommendations for interesting items (i.e. internet or intranet pages) were made. In the survey study with 290 subject, the people preferred familiarity-based recommendation to the similarity-based recommendation [64].

While the area of recommender systems originally ignored user groups, more recent research in this field paid close attentions to self-defined user groups. The largest research stream under the name of ‘group-based recommendations’ focuses on aggregating preferences of a group of people and suggesting recommendations to a whole group, not an individual member [3, 11, 14]. For example, one of the most cited group recommendation systems, Flytrap recommended music for people in the same physical space. The system gathers music preferences of each user who is identified by RFID. Once users enter a room, the system identifies who they are, aggregates all their preferences together and computes the probabilistic values of all music for recommendations [11]. MusicFX also aims to recommend and play music to a group of people who are exercising at a gym [20]. Along with recommendation systems for music, there are several kinds of group-based recommendations to suggest movies [22], vacation destinations [3], news articles [25], photos [9], restaurants [23], etc. to a group of users.

The study done by Yuan and colleagues (2009) is one of few studies explored how to use an individual user’s group or community membership information in personalizing his own information. They introduced two ways to fuse users’ social network information (friendships and group memberships) with traditional CF using LastFM dataset. The first way is weighted-similarity. They computed user-user similarities based on their listening logs and based on their friendships, memberships of the same group or both of them, separately. And then, they fused these kinds of similarities together with different weights. The second way is to utilize a random-

walk approach. They built a user model not only about users' social networks (i.e. friendships and group memberships) but also about preferences of artists. Then they calculated random walk to find similar peers. As the results, they found that social networks-based recommendations outperformed typical CF recommendations, regardless of the way. Specifically, the second random walk-based recommendation performed best than the first weight-similarities. This study shows that, although including social networks as a part of recommendations can increase recommendation quality, how to utilize is also important [82].

2.3.3 Acceptability of Recommendations Suggested by Social Connections

Some researchers examined how users perceive the recommendations directly suggested by friends with other recommendations coming from machines.

Sinha and Swearingen compared how users feel about recommendations generated by online recommenders and the ones suggested by friends. Rather than their own calculations of automated recommendations, they relied on recommendations provided by third party applications. They tested the recommendation qualities for books and movies. For books, they borrowed the list of recommendations from Amazon, Sleeper and RatingZones. For movies, they used the list of recommendation from Amazon, Reel.com, and MovieCritics. In order to acquire friends' recommendations, the participants provided the names of their three friends who can provide reasonable recommendations to them. They concluded that friends' recommendations were more useful and satisfying than those of the recommender system. The recommendations of friends also are more trustworthy than those of the system. However, in terms of novelty and

serendipity of the recommended items, the typical CF recommendations were better than the friends' [164].

Bonhard, et al. also tested the quality of recommendations provided by their friends. The major focus of the study was to see the correlation of how they feel recommended items and who recommend the items. Specifically, they asked the qualities of three recommendations, which are based on distinct characteristics – familiarity, profile (demographic data, hobbies, interests and music tastes) similarity and rating similarity – to 60 people (12 groups of 5 friends). The participants picked presumably favorable movies for their friend participants and the authors produced other two (profile similarity and rating similarity-based) recommendations. They were asked to rate the recommended movies with an explanation about who recommend each corresponding movie (real name, occupation, hobbies, preferred movie genres, and rating overlaps). The participants much more liked the recommendations from their friends than the ones of the strangers (73% vs. 27%). In the evaluation of two recommendations, of which each was provided by individual friend, they preferred the recommendation of a friend whose rating pattern is more similar to them [21].

Bourke, et al. (2011) executed the similar study with 82 Facebook users. They gave users controls to choose members of their own social neighborhoods or took into account users' communication frequency with other users to select social neighborhoods. On the center of the tastes the social neighborhoods, the recommendations about Movie/TV were generated. For comparison, the authors selected two different neighborhoods using the similarities of favorite items and generated recommendations according to those neighborhoods, as well. The participants were asked to rate the recommended items with and without the explanation about

which neighbors recommended the items. As the results, of course, when they could see the sources of the recommendations, the ratings were higher. However, there was no significant difference on the ratings regarding the sources of the recommendations. From this result, the authors suggested that users were not good at selecting useful friends as recommendation sources [22].

In another study by Bonhard and Sasse (2006), they examined the reasons why people like to know the sources of the recommendations. They suggested that recommendation is a part of decision-making process. Advice-seekers decide the value of suggested items according to the identity of the recommender. Therefore, the relationship between the receiver of recommended information and the source of that information is critical. Their study found that, along with rating overlap, profile similarities such as demographics, preferences and interests play an important role in users' trustworthiness of recommendations [20].

2.4 TAGS AS USER INTEREST INDICATORS

Among three data sources used in this dissertation, two data sources are social bookmarking datasets consisting of users' bookmark histories and the social tags (Detailed explanations about these data sources will be given in the following Chapter 3). Unlike other collective intelligence systems enables users to express their item preferences using numeric ratings, social bookmarking sites ask users to annotate social tags on their favorite items. In order to generate recommendations based on users' bookmarking history and the social tags, it is critical to decide how to compute users' preferences out of users' social tag sets. Therefore, in this sub-section, several studies about how to model users' preferences or interests in social tags will be explored.

Tags are free-text words selected by users to describe their favorite items. Tags were found to be helpful for organizing, retrieving and sharing information. The collection of social tags is often called as 'folksonomy' (which is a compound word of 'folk' and 'taxonomy'[69]) equating it to a collection of conceptual and structured knowledge created by a bunch of people. According to the analysis of tags sharing in Connotea, around 23% of tags out of 3359 unique tags were shared between users [111]. Although it was preliminary result based on initial usage of the system, it shows that tags are being utilized not only as a way to organize information for personal goods, but also as a good way to help discover useful information shared by others.

Users' social tags can serve as rich evidence about their interests, and this capacity can also be applied to construct user profiles for personalized information access. Compared with

item ratings in CF recommendations, item tags not only indicate user interest, but also show their cognitive understandings about how users approach one item with different viewpoints [71, 79, 196]. Consequently, if two users share not only many identical information resources but many identical tags regarding the resources, they are closely related.

Au-Yeung, et al. (2008) started their study from an assumption that users are interested in multiple topics, and hence their personal collection of tags (i.e. personomy) may consist of several clusters. Using 1000 users' data from Delicious, they produced clusters of Web resources tagged by each user's *personomies*, and found that their assumption is true. In order to test the quality of clustered Web pages, they utilized the most oft-annotated tags in a cluster (i.e. signature tags of the cluster) as a query and assessed how much the query can retrieve the relevant pages. As the result, neither too general nor too specific tags could not retrieve a good set of results in terms of precision and recall [13].

A study using Delicious data set [71] explored the correlation between users' tags and content tags. The authors also investigated how well personal tag sets and social contacts' tag sets represent their preferences. As expected, personal tags were more effective to express user preferences. However, user profiles based solely on personal tag sets were not good enough to generate recommendation (the best precision was 21%). As another study exploring whether tags depicts users' preference actually, Firan, et al. (2007) examined music information from Last.fm and the tags. The system retrieved music information according to users' query. Their tag-based profile took into account the similarity of both tags and music tracks. Subjects evaluated that the search results generated by tag-based profile were better than the pure CF recommendations which just compared tracks [48].

Godoy and Amandi (2008) suggested a way to define users' profiles based on users' visited Web pages and the tags. The profile is built upon an incremental, unsupervised conceptual learning algorithm; WebDCC (Web Document Conceptual Clustering). Web pages that users had visited are represented as word vectors, which later are clustered conceptually. The clustered vectors are summarized by each succinct description using automatic extraction of the descriptors. Moreover, unlike other studies based on social tags, the authors organized a hierarchy of vectors and tags [52]. Guan and the other (2010) also generate user profile based on tags and the contents of documents (e.g. scientific articles or Web pages). Specifically their profiles are generated by a graph-based representation learning algorithm. In order to infer users' interests, they took advantage of three bi-partite relations of social tags such as user-document, tag-document, and user-tag. In this situation, however, it is hard to locate semantically related tags to a document. For instance, the documents tagged by "automobile" are treated differently with the ones tagged by "car." As the solution, they utilized an affinity graph of documents, which are built by word vectors. To recommend interesting documents, they set three bi-partite relations and affinity graph in one space. Then they examine how three components (a.k.a, user, document, and tag) are close to each other and the closest documents to a target user is recommended [61] .

Szomszer, et al. (2008) showed that tags added by a user represent his preference and interests across different systems. They tried to test how similar two sets of tags annotated by one user in two separate social tagging systems are. They collected data from *delicious* and *flickr* and using the usernames or real names of the users, they correlate different user accounts between the systems. They found 502 users who have accounts in both systems. After then,

severe tag filtering mechanisms were applied in order to increase the tag alignment between two folksonomies. In the experimental evaluation, *delicious* users' tag sets was compared with all *flickr* users' tags including themselves and calculated who have the most similar tags. The results showed that the system chose the same user's tags as the closest one for more than 80% of users. Especially, tag filtering mechanisms helped improve the tag alignments. It proved that tags express users' interests and it can be possible to build cross-folksonomy and facilitate multi-domain adaptive search or recommendation using tags [172].

Social annotations are descriptive words added by users to describe information resources for the purpose of organizing, retrieving and sharing information in their own way. It does not require any specific skill and produces instant benefit for users without too much overhead [68]. Millen et al. (2005) suggest that tag annotations are a major reason why current social bookmarking systems have enjoyed greater success than social bookmarking system in the earlier time [131]. Not only in the social bookmarking systems, but in the countless information systems in various domains such as social networking, bibliography management, movie review, blogs and so forth, users tag information actively. It is known that tags of a user represent his conceptual understanding or categorization of a resource from a personal point view. It is 'personomy', which is a combination of words 'personal' and 'taxonomy' [68]. The aggregated collection of multiple users' (folk) personomy is equal to the collection of the background concepts (taxonomy) for a corresponding item [172] and it is 'folksonomy' (which is a compound word of 'folk' and 'taxonomy' [69]. In turn, social annotations refer to a power lightweight metadata creation tool [2]. In addition, Golder and Huberman insisted that defining tag is related to sense-making. Based on their own experience, daily practices, needs and

concerns, people perceive a single information item from different senses. The sense-making is also involved with social factors, considering that personal experience is associated with others, culture or community [57]. Several studies have showed that social tags are expression of users' interests or preferences and socially connected or related users share same information resource and especially tags.

FolkRank algorithm aims to retrieve information based on folksonomy by counting tripartite interaction of users, tags and resources simultaneously. It assumes that a resource tagged by important users with important tags is also important. The authors believed that because FolkRank computes topic-specific rankings of information contrasted with adapted PageRank providing one global rankings, tags can serve as a kind of meta-data whether the formal text-based information property (e.g. TF/IDF) is available or not [68, 69]. Therefore, if tags are not only generic concepts of information but the expression of users' interests, it would be possible to verify if users having common tags prefer similar information suggested by FolkRank. In later study, they compared quality of recommendations using FolkRank with two CF based approaches, where one is based on user-resource similarity based CF and another is based on user-tag similarity based CF. Empirical evaluations used three different data sets – *delicious*, *Last.fm*, *Bibsonomy*. Recommendations using FolkRank outperformed other two approaches in terms of precision and recall [79]. The results can be interpreted that it is more effective to calculate tri-partite hypergraph as a whole than to calculate disjoint union of bi-partite graphs (e.g. user-resource and user-tag). However, the data used in this paper was filtered before the experiment. Only users, resources and tags having a certain number of frequencies (users have enough number of resources and tags, resources tagged by many users and tags

added to many resources by many users) were used in this study. This dense distribution of data is quite different with the real nature.

Other researchers (Li, et al., 2008) did their research with an assumption that tags are consistent with information content and further interests. Hence, they thought that frequently co-occurring tags can cluster sets of similar information and similar users. First, it was tested how much of the ‘most important’ keywords of a document are covered by user-generated tags. For top 10 keywords, 74% of documents are fully covered by user tags and 98.2% of them are partially covered. Given that the number of tags was only 7.4% of the total number of keywords, users’ tags represented the topics of the documents very well. In order to cluster users and resources, they counted the frequency of co-occurring tags and a set of tags co-annotated more than a certain number of times was picked as topics. The users and resources having the tags designated as topics are clustered together. In the evaluation, the intra-cluster similarities were significantly higher than inter-cluster similarities. Most of users had the topics as their 5 or 10 tags [104]. That is, users provided useful and accurate tags and common tags among users show that the users agree each other for content and further for interests.

As explained in the Section 2.2.1, one of the biggest problems of CF technology is data sparsity. In the tripartite relations of social annotation – user, resources and tags – the data is sparser. It is known that the usage pattern of tags follows the power law distribution [65, 68, 69]. Small number of tags used very often and large amount of tags used rarely. In consequence, it is more difficult to find the users having common resources and tags. An experiment about information recommendation using tags showed that tags distribution across multiple users are too sparse to find enough level of exact tag matches [196]. However, two users whose resources

and tags are not overlapped do not imply that they did not have common interests [71]. Social tagging is implicit and asymmetrical preference. When users did not tag an information item, it does not mean they are not interested in the item. In order to solve the data sparsity and improve information retrieval using tags, Zanadi and Capra (2008) explored social annotations through two bi-partite graphs (user-tag and tag-resource). If it is available to solve data sparsity problem through tripartite hypergraph as a whole, as explained, the result may be more meaningful or robust but the tripartite hypergraph normally increase data sparsity. Therefore, they split the tripartite relations into bi-partite graphs. They deployed users' tag similarity as a part of search tool. Two kinds of tag similarity were counted – user-tag similarity and tag-resource similarity. First, they computed the tag similarity of two users regardless of the resources. The more common tags they have, the more similar they are. Second, they calculated the tag similarity in terms of the resources regardless who the taggers are. The more resources annotated by two tags together, the more similar the tags are. When users enter a query, the system treats the query as a tag and the query tag is expanded with the similar tags. The search results depend on the relevance of the tags assigned to the resources. After then, the query results are sorted by the user-tag similarity. If the taggers are similar to the searcher, the resource is ranked higher. The experimental results presented that the user-tag similarity improved the retrieval accuracy and tag-resource similarity helped uncover unpopular resources which are less tagged. Therefore, even though two users do not annotate common information resources, their bi-partite graph similarities supported to find relevant information and showed that they are related [193].

In addition, social tags have semantic problems such as polysemy, synonym and basic level of variations. Therefore, disambiguation and elimination of redundancy is necessary [57,

65, 172]. Tags are redundant and erroneous as well. User can misspell words and use idiosyncratic tags which are meaningful only to the taggers and clueless to someone else. Moreover, on the cognitive aspect of hierarchy and categorization, people consider tags at different levels of specificity. For instance, tag 'java' or 'php' is too specific for a user and tag 'programming' may be too general to another user [57]. Various filtering and dimension reduction mechanisms, for instance, lexical dictionary (WordNet), wikipedia, spelling checker, compound words checkers can be used to check the morphological variations and errors [167, 172]. Semantically close tags can be found using Latent Semantic Analysis [41] or naïve Bayesian network [116, pp. 234 ~ 265] by the probability of the occurrence of tags in a same resource. Symeonidis, et al. (2008) suggested a unified framework to model the triplets of social annotation (user, resource, tag) and in order to capture the latent associations among them, applied a tensor reduction algorithm, HOSVD (Higher Order Singular Value Decomposition). That is, even though a user did not annotate a resource, depending on the interaction of the triplets among multiple users and the latent semantic association, how likely he will annotate the resource with the tag is calculated. For instance, there are three users and two of them u_1 and u_2 tagged resource r_1 with tag t_1 . User u_2 also tagged resource r_2 with tag t_2 and u_3 tagged resource r_3 with tag t_3 . Then the system can infer that it is more likely that user u_1 will add tag t_2 to resource r_2 than the resource r_3 and tag t_3 . Not only the directly annotated resources and tags, but the semantically propagated resources and tags were associated. Using two popular Web annotation system data sets, they tested the recommendation quality. As the result, compared with FolkRank and a collaborative tag suggestion algorithm (Penalty-Reward algorithm), the precision and recall of their tensor reduction method was higher [170]. Even though social

annotations are not involved, there are several studies to solve data sparsity problem by background topical structure, for instance, pre-defined taxonomy [103, 130, 200, 202].

According to the Golder and Huberman (2006), after a certain amount of tags are added, the pattern of tagging become stable, since taggers refer to other users' tags and imitate them or share their knowledge. This trend is called as 'social proof [57].' The following statement says well the social Web systems are developing centered on users' own networks.

“Flickr and delicious work so well for me not because they aggregate the world's tags but because they allow me to aggregate my social network' tags, links, and photos [66].”

3 SOCIAL NETWORKS IN CONSIDERATION AND THE DATA SOURCE

In this dissertation, I consider three kinds of social networks on the center of two aspects – personal familiarity and object-centered sociality.

The first aspect – Personal Familiarity – is a very typical foundation on which many kinds of social network are rooted, in particular networks existing offline. The degree of personal familiarity on online social networks depicts how much an online social relationship is based on offline social acquaintance and personal interactions. According to homophily and social influence theories, it is well-known that strong ties which share frequent interactions and whose connections are greatly overlapped to each other tend to be similar in various ways. Beyond this social similarity, however, a highly cited theory about social strength [58] tells that, depending on the strengths of the social networks, there are distinct patterns in the kind of information circulating within the social connections and the way to share information. Strong social ties circulate conventional and similar information, and the flow of information is private. On the other hand, weak ties tend to share new and heterogeneous information and to disseminate information more widely and publicly. Through an exemplary case about job searching, the author suggested that acquaintances whose social strength is relatively weak are more helpful in acquiring useful job opening information or disseminating the job information [58]. Due to the distinct patterns, personal familiarity represented by social strength is an indispensable aspect in

understanding how to acquire and share information. Along with the success and prosperity of online social networks, there are diverse newly-emerging sociabilities having multiple degrees of personal familiarity. However, as the collection of the related literature in the Chapter 2 shows, recent social recommendations have focused on very limited kinds. By exploring more diverse online social networks as a part of recommendations, thus, I expect to prove the feasibility of various sociabilities as important information sources and to demonstrate possibility of new social recommendations.

The second aspect – Object-oriented Sociality – is, relatively speaking, a new and innovative foundation on which social networks are rooted, and the focus mainly lies on social networks existing online. In the contemporary society which is highly driven by information and knowledge, the theory about the object-centered sociality insists that knowledge cultures are inter-stitched with the current social structures. The original theory is rather radical. It suggests that, in modern society, traditional social linkages forming large communities would be blurred out, and human beings as relationship partners would be objectualized according to their knowledge and expertise. However, the main idea is that information-driven culture spills and weaves the fabrics of the modern society, and information objects are social interaction triggers and anchors of communications [25, 33]. Most of the current SNSs and Web 2.0 systems provide active aids to this object-centered sociality on them. They provide interfaces for users to get together around the contents they create, comment on, link to and add annotations on. Without any personal acquaintance or interactions, by the ‘Constructuralism’ mentioned in the previous Section, the systems are allowed for users to get connected with other users once they perceive the usefulness of the other users’ information collections. However, depending on domains of

systems, the degree of object-centered sociality is different. If an application targets on managing information items, such as bibliographic information or Web bookmarks like Delicious and Citeulike, the users usually focus on other users' collected information, rather than their personal traits or daily lives. Therefore, users are seeking other users whose information collections are beneficial to them. On the other hand, if an application aims to build online social networks like Facebook and LinkedIn, the users are mostly interested in their friends' social reach, personal traits, and activities. Therefore, the users are finding social partners who have had offline interactions and have similar personal traits (e.g. same location, same school or company, and similar ages and sex), rather than the quality of their information collections [15, 16].

On the center of these two aspects, I selected three sorts of online social networks – watching network, group membership and collaboration network, as shown on Figure 2. According to the related literature, these three social networks in my consideration have different degrees of personal familiarity and object-centered sociality. In the following, I will introduce each of them and explain the defining characteristics of this taxonomy, according to these two aspects.

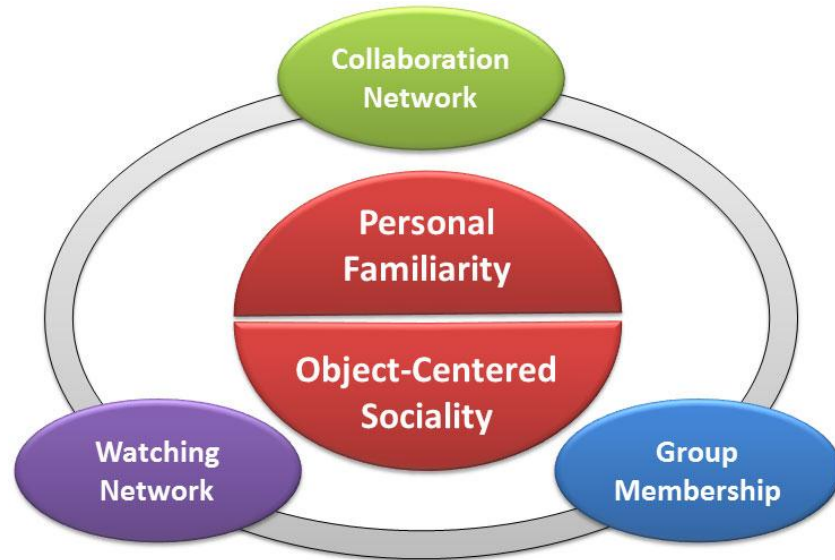


Figure 2. Various Social Networks Considered in this Dissertation

3.1 WATCHING NETWORK: SOCIAL CONNECTIONS CENTERED ON OBJECTS

Watching network is a typical example of newly emerging online social network. Users on the Web 2.0 have found it easier to know who knows what through social networks in virtual space. It is a burden, however, to contact the person who knows the desired knowledge via their personal ties [124]. Unlike social networking systems mainly focusing on linking *mutually agreed-up* friendships, many social tagging systems, which help users to manage and share interesting information online, or blogging systems, which aim to post online journals to express themselves, offer the users this new kind of sociability without any burden to ask a consent to be connected. Once users find other users who have interesting or similar tastes, they are allowed to *watch* the users' information collections continuously. The watching relations don't require any

offline interactions or emotional bonds to make connections or mutual agreements for being connected. It is a *unilateral* relation which forms a special kind of social network, which I named it as ‘watching network’. The most typical examples of watching network are “following” on *Twitter*, “watching” on *CiteULike*, “network” on *Delicious* or “contacts” on *Flickr*. Interestingly, *LiveJournal* is providing this watching relationship, one-directed connections without consent to be connected, but called it as “Friends”. This relation is convenient and purely based on the utility of information possessed by users. In addition, it can compensate for privacy concerns because the watched users can decide what to expose or hide to the strangers by themselves.

Unilateral relations gained attentions along with the success of social tagging and micro-blogging applications [26, 76, 154]. Wellman suggested that, in Web 2.0 era, various new relationships would emerge and the networks are “less bounded [186].” The unilateral relationship is one kind of the new relationships. Unlike befriending in SNS (e.g. Facebook, MySpace or Friendster) where users increase the number of friends simply for fun or curiosity [39], the watching relationships aim to acquire information from the connected people’s collection. When a user unilaterally watches many users, it may cause rapid growth and potential dilution of the item list in the watcher’s collections. Therefore, unilateral relationships require users to be careful to select people to watch, based on the utility of information or taste.

Some researchers may argue that watching relations are not social connections, since they may be not based on social interactions or emotional bonds. However, my preliminary works [92, 94, 97] and the section 5.2 of this dissertation found an important social property in the relationships; homophily. I described the details in the Section 5.2. Specifically, the relations met the similarity attraction hypothesis [132] and held the transitivity power [133]. High degree of

similarity was embedded in unilateral relations and the similarity decreased with the increase of distance. Even though two nodes (i.e. two users) tend not to interact personally or not to share any emotion, they are directed networks on the center of interesting information objects and naturally it is highly object-centered sociality. Breslin and Decker (2007) said that the social networks connecting via items of interests, which are called object-centered sociality, may be more long-lived relations than the relationships not sharing any item of interest. In order to help users develop better relationships, the authors insisted that SNS have to take into consideration people's actions about content such as tagging, blogging, adding comments, etc. to find out the users' people of interests [25]. Even though many social tagging systems support these watching relations, surprisingly, there are few studies about how to utilize this social network for personalizing information. Java and the others (2007) insisted that one of the reasons why users enjoy micro-blogging, such as on the Twitter, is to share information. They also found three kinds of users' intention to socialize on the Twitter (i.e. *following*); not only 1) friendship-wise relationships but also 2) information sharing and 3) information seeking [80].

I interpreted that these watching connections existing on social bookmarking systems mimic the process of bookmarking interesting items. In this context, users bookmark other interesting users. Therefore, the information of users being watched can be used as a part of the watching users' preferences. On the other hand the relations don't require personal interactions or acquaintance necessarily. The watching parties tend not to pay attention to its personal traits. In many situations, they even don't know who their watchers are. Therefore, I declare this kind of relations as highly object-centered sociality but very low personal familiarity.

3.2 GROUP MEMBERSHIP: SOCIAL CONNECTIONS CENTERED ON ONE TOPIC

Group activities are usually centered on one solid topic. Users join a group on the Web in the terms of community of interest or practice, for instance, a fan club of a musician, a community of Hadoop programmers, an online forum for students taking the same class, an online space for members of the same project, etc. The relationships in group networks target to distribute topic-relevant information or contribute topic-related activities. The theory of communal sharing relationships explained the social dynamics of this group membership. Group members think they share common substances. Before online social networks emerged and proliferated, the communal sharing relationships represented very close relationships such as kinship ties [9]. However, in the current Web 2.0 era where relationships are getting flattened and less bounded, the sense of communal sharing can be applied to the online group activities. Even though their relations are not based on strong bonds such as kinship by blood, they still treat information objects as their shared assets. Hence, group members are willing to share what they need and contribute what they can. Members usually don't expect to receive something back as returns of their contributions. In addition, they usually don't pay attention to the portion of contributions made by each individual member. Simply being a member of a group is sufficient to them since they are able to use the resources the group is sharing [9]. According to my own preliminary result, group members didn't pay attention to the information collected in other members' spaces. However, they really pay attention to the group libraries jointly organized by the group members [96]. I suggest that the willingness to share and contribute in groups as the sense of

communal sharing exists online. Therefore, memberships of online groups will be informative in personalizing their information space.

Although I defined that group memberships are self-organized by the members, most of the studies about online group dynamics are about how to derive implicit communities using various machine learning technologies and characteristics of the derived implicit communities which are systematically discovered by pattern mining approaches [10, 17, 49, 144, 197]. As the example, Zhou and the colleagues studied the information similarities in groups using semantic-rich contents. In the study based on Enron email corpus, as the first step, they ran the Bayesian network and chose the latent topics of emails. They extracted a set of implicit communities derived by the latent topics. In addition, by taking into account the associated contacts (i.e. the author and the recipients) of each email, they built another set of implicit communities derived by the contact structure. When they compared the resultant implicit communities with other community formation results from another study as a ground truth, they found that their approach succeeded to generate appropriate groups with high similarity in shared messages [197]. The weakness of this study is that all communities were implicitly inferred by machine learning technology. Even the communities used as the ground truth were implicit communities generated by a machine learning technique.

There are very few attempts to explore social dynamics in self-defined online groups and naturally fewer attempts to utilize users' explicit group memberships in personalizing information. Most of the systems using 'group recommendation' technologies are to aggregate tastes of a group of people into one set and to suggest recommendations for that group of people,

not for an individual group member [12, 34, 39, 121, 138, 145]. Therefore, recommendations for individual members based on their group memberships are innovative and necessary.

On the respect to the two axes of the taxonomy, I suggest that group memberships embed strong object-centered sociality and medium degree of personal familiarity. As explained, users whose interests or purposes are similar to each other get together in an online group. For a group of students who are taking the same class, they discuss assignments or class projects, ask questions to the instructor and post useful resources regarding the class on the group space. For a group of Hadoop programmers, they share exemplary program codes, announce important news regarding the technology, post job openings requiring the Hadoop skills and seek aids for their programming problems. Therefore, interactions on groups are highly concentrated on information about the corresponding topic. Regarding the personal familiarity, in some cases, members in the same group could be familiar to each other before they join the group (e.g. members of the same project or same research lab). Otherwise, after sharing a number of online interactions (e.g. asking and answering a series of questions; participating in the same open source development project), they may be able to develop the relations into more familiar social relations. However, this social evolvement among group memberships is not always the case. That is the reason why I decide the group members as a network having strong object-centered sociality and medium degree of personally familiarity.

3.3 COLLABORATION NETWORK: SOCIAL CONNECTIONS OF PERSONAL ACQUAINTANCE AND SHARED INTERESTS

Recent studies about computer-supported cooperative works have focused on how organizations solve problems in a collaborative way [186]. Due to the effectiveness, several social matching systems for professionals and expert location systems have been in use [83, 102, 105, 124, 191]. The basic idea of the online social networking systems for professionals is to implement offline referral chains onto online spaces. When we confront a problem in a workplace, we usually ask our colleagues who have the relevant knowledge or who know other people who may know topic of the problem well. In an analogous manner, when a user has a specific problem to solve, these online systems help him find individuals who are the most knowledgeable about the problem and are socially close enough to contact.

The Expertise Oriented Search (EOS) system³ [102] is one of examples. It was designed to help users to identify experts of a certain topic. In particular, users' expertise was modified depending on whom they have interacted with professionally. As a way to define users' expertise, the system first draws every user's 20 the most relevant Web pages from Google Search and publication list from the Digital Bibliography and Library Project (DBLP), and Citeseer, respectively. Users' expertise was adjusted by topic relevance of their social connections. The authors suggested that the topic relevance of a user is propagated through his social connections under the assumption that a person's expertise diffuses and is developed

³ <http://www.arnetminer.org/>

through social interactions. Both the original topical expertise and propagated relevance values are taken into account when searching for experts.

Yang and Chen [191] described an educational P2P (peer-to-peer) system using expert location technology. Albeit this system requires significant ongoing human interventions for evaluating users' expertise, when a user queries a topic, it recommends items posted by users with the highest expertise related to the topic or who are most preferred by the target user [191].

In my past research, when researchers seek future collaborators to address newly emerging research topics and engage in inter- and multi-disciplinary, I discovered that they highly relied on their own social networks, rather than the outside of their professional social networks, and cared more about the social distances than about the candidate's expertise [100].

In addition, according to a recent study about why researchers use various online SNSs and how their online activities are related with their tenure promotion, they expected to acquire new and useful tenure-related information from their colleagues through SNSs, and it is the one of the reasons why they keep using the SNS applications. One of the fifth participants answered that the information existing on the SNSs is helpful for their tenure promotion in somehow [60]. Accordingly, in situations where people need new and useful information about their research, they will be willing to refer to their colleagues who they have worked together before and hence, whose expertise and social context are familiar to them [58, 83, 186]. In here, I will take into account the researchers' collaboration networks as plausible information sources. Specifically, they have interacted to each other personally and their relations is centered on their research expertise, the relevant by-products (such as, articles, project proposal, reports, etc.) and furthermore the information of common interests. However, it is rare that two researchers'

interests are perfectly matched. Put differently, some items which interest my colleagues could not be of my interests. Conclusively, I decide that these research collaborators' relations embed medium degree of object-centered sociality and high personal familiarity.

3.4 THE DATA SOURCES

In this dissertation, I used two data sources so as to examine the proposed recommendation algorithms using three social networks – ‘Citeulike’ for watching network and group membership and ‘ConferenceNavigator3’ for collaboration network. Citeulike system was selected because it supported the corresponding social network and allowed the general public to access the social network information. In particular, the system provided a snapshot of their systems for the public to collect the users' information (i.e. RSS feeds). In case of ConferenceNavigator3 (CN3, in short) system, I have been participating in the development of the recommendation module. Hence, I can access the whole data on the system. Especially, rather than the datasets covering wide spectrum of matters on one application such as Twitter, Facebook, MySpace, they were the applications of which the domain concentrate on one kind of information. Citeulike is to manage bibliographic information of scholarly articles and ConferenceNavigator3 (CN3) is to access conference talk information of various academic conferences. Figure 3 shows the list of datasets in my consideration. In the section for each social network-based recommendation, I describe the detailed description of the data.



Figure 3. Data Sources for Each Social Network

4 RECOMMENDATION TECHNOLOGIES

This chapter aims to introduce information similarity measures and recommendation technologies that I used in this dissertation, in general. More detailed approaches which are modified for each social network will be explained in the corresponding chapter.

4.1 INFORMATION SIMILARITY MEASURES

Through this dissertation, I used two datasets (Citeulike and CN3) for three kinds of social network. Users, their preferred items and their online social connections are the main ingredients of these datasets. Generally, users' preferences about the items can be expressed in ways: unary ratings and numeric ratings. Unary rating means that the presence of bookmarks expresses users' interests on corresponding items, but there is no degree of preference. In this case, the absence of the bookmarks doesn't necessarily represent that users are uninterested in or dislike that item, since there is a high possibility for users not to discover the item yet. On the other hand, numeric rating means that the values of users' ratings expressed the degree of their interests or preferences. Depending on the rating schema of systems, the numeric rating values can show positive preferences only or both positive and negative preferences. Both Citeulike (which is used for watching relations and group memberships) and CN3 (which is used for collaboration

connections) datasets have unary rating schema. Hence, in this Chapter, similarity measure for unary rating will be explained mainly. As a reference, similarity measure for numeric ratings will be explained briefly.

For the rest of this paper, I use the following notations. For Citeulike and CN3 datasets which consist of users' bookmarks and the unary ratings, B is the user-item bookmark matrix, $B = \{B_{ui}\}_{l \times n}$ where l and n denote the number of users and items, respectively. $u \in \{u_1, \dots, u_l\}$ represents users and $i \in \{i_1, \dots, i_n\}$ denotes items. The bookmark of user u on item i is b_{ui} . In here, \hat{b}_{ui} denotes the predicted bookmark of item i for user u , which is picked by my recommendation algorithm as presumably favorable and bookmark-able item. The values of all bookmarks are unary rating (the presence of a bookmark).

Another major characteristic of both Citeulike and CN3 datasets is that they have not only the bookmarking history but also various metadata. The target items of both datasets are commonly scientific articles. Even though the target item of CN3 system is conference talk, each talk presented in a conference represents a scientific article. The available metadata of the articles would be titles, authors' names, journal/conference names, etc. Therefore, in these datasets, the similarity of shared information was assessed using *item-based* and *metadata-based* approaches.

4.1.1 Similarity Measure for Unary Ratings

The very basic and traditional way to measure the information similarity using unary ratings is to count the number of items co-bookmarked by two given users. This is the *item-based* approach.

Beyond this number of co-bookmarked items, there are various similarity measures, such as Jaccard similarity, dice similarity and log-likelihood similarity. These measures considered the fact that sizes of item collections varied dramatically from user to user. Hence, whereas the number of co-bookmarked items is absolute measure, these similarities are relative and normalized measures. As the normalized measure, first, I used Jaccard coefficient. The Jaccard coefficient is the proportion of shared items in the union set of two users' collections. Let's assume that user a and user b are in comparison, and variable B_a and variable B_b denote the users' information collections respectively. We can compute the coefficient as the following.

$$\text{Jaccard Power } (u_a, u_b) = \frac{|B_a \cap B_b|}{|B_a \cup B_b|} \quad \text{Eq. 1}$$

The second similarity measure is the log-likelihood similarity, which is another typical measure for the unary ratings. Log-likelihood similarity was initially proposed as a statistical analysis method for text processing. The strength of this similarity is that, unlike other similarity methods which take into account only the cases where an item was co-bookmarked, it takes into account not only the cases of co-bookmarked items, but also the cases where an item was bookmarked by only one person and even the cases where an item was not bookmarked either of them. That is to say, this method measures how unlikely the overlap of interests was occurred just by chance [46, 75].

$$\Lambda(u_a, u_b) = \frac{p_{a,b}}{p_b} \times \frac{p_{\sim a, \sim b}}{p_b} = \left(\frac{n_{a,b}}{N_a} \times \frac{N_b}{n_a} \right) \times \left(\frac{n_{\sim a, \sim b}}{N_b} \times \frac{N_a}{n_b} \right) = \frac{n_{a,b} n_{\sim a, \sim b}}{n_a n_b} \quad \text{Eq. 2}$$

where, $\ln(\Lambda(u_a, u_b))$ denotes the log-likelihood similarity between two user a and user b . $n_{a,b}$ is the number of items bookmarked by both user a and b and n_a is the number of items bookmarked by only user a . n_b is the number of items bookmarked by only user b and $n_{\sim a, \sim b}$ is the

number of items bookmarked by neither of them. The similarity value of 0 indicates that two users in comparison are neither similar nor dissimilar and a large positive similarity means that they are very similar.

In measuring information similarity, due to the popularity bias (i.e. users are more likely to provide feedback on popular items than on rare items [166]), co-bookmarking very popular items doesn't reflect their shared interests as much as the co-bookmarks of rare items. For instance, in measuring movie tastes, the co-bookmarking of a very popular movie like 'Avatar' doesn't show the true shared interests between two users than the co-bookmarking of a unpopular independent film. Therefore, we incorporate the item popularity into information similarity using the average inverse item popularity as the following. We defined this similarity measure as '*popularity weight*' of co-bookmarked items.

$$IP_{ab} = \left(\log \sum_{i \in B_{ab}} \frac{N}{n_i} \right) / n_{ab} \quad \text{Eq.3}$$

where IP_{ab} is the average inverse popularity of co-bookmarks between user a and b . Item i is one of the items that the user a and b commonly bookmarked (i.e. $i \in B_{ab}$), and N is the total number of items in the dataset. n_i is the total number of bookmarks of the item i and n_{ab} is the number of co-bookmarks between the user a and b . Bookmarks containing less popular items are thus assigned a higher popularity weight value.

4.1.2 Similarity Measure for Textual Metadata

Due to the irregular opportunistic nature of the bookmarking process, users with similar interests may not necessarily end up with very similar collections. In this case, similarity of interest might be more reliably measured on the level of item's meta-data. Even though users' preferences were expressed in unary ratings, by taking into account the related metadata of bookmarked items, we are able to analyze why users bookmarked the corresponding items in more cognitive ways. For example, in *Citeulike* domain, bookmarking papers from the same authors or same journal (or conference) can be considered as an indicator of similarity. When titles of two papers in comparison have the similar set of keywords, it would be likely that they contain the similar contents and context. CN3 domain is also very similar. When users have their own published papers, we are able to compare how the metadata of users' own publications (for instance, titles, authors, conference names, etc.) are alike. In addition, on Citeulike, users added free-text social tags when they bookmarked items. Compared with item bookmarks simply showing users' interests on the corresponding items, social tags indicates users' interests on the items in cognitive and more descriptive way [71, 196]. With this reason, various kinds of *metadata-based* similarity were compared.

In order to compute the metadata-based similarities, for every kind of metadata, I used the vector space model [156]. Before computing a metadata-based similarity, we built users' preference model from each metadata. All terms in a metadata of a user's bookmarked papers were aggregated into a bag of keywords. Then for the effective comparison, I applied text processing techniques to the bag of keywords. The terms were case-normalized to the lower-case

letters, and all stop words⁴ were removed. We also applied Porter stemmer so as to reduce word variation to its stems or roots [107]. Stemming reduces word variation by truncating words to their stems or roots. For example, ‘computer’, ‘computers’, ‘computing’ and ‘compute’ are stemmed to ‘comput.’ The processed bag of metadata keywords were transformed into keyword vector consisting of keywords and the TF/IUF (Term Frequency/Inverse User Frequency) values. In this paper, we computed each user as a document. Therefore, the original document frequency value in traditional TF/IDF (Term Frequency/Inverse Document Frequency) formula was changed to the user frequency indicating how many users bookmarked papers of which the titles contain the keywords. As Mori, et al. (2006) suggested, when words in two articles similar, the contextual representations of article pair are also similar. Therefore, this metadata vector-based similarity tries to compute how much two given users are interested in similar contents and contexts.

Depending on the data sources, the available kinds of metadata differ. Hence, the detailed description of metadata-based similarities for each data set will be explained in the corresponding chapter.

4.1.3 Similarity Measure for Numeric Ratings

For a data set having numeric ratings, there are several similarity measures, such as Pearson correlation, cosine similarity, Spearman’s rank correlation coefficient, and mean squared difference. Along with the Pearson correlation, the cosine similarity is widely used. It measures

⁴ <http://jmlr.csail.mit.edu/papers/volume5/lewis04a/a11-smart-stop-list/english.stop>

the similarity between two n -dimensional vectors based on the angle between them as the following. R is the user-item rating matrix. $R = \{R_{ui}\}_{l \times n}$ where l and n denote the number of users and items, respectively. The rating of user u on item i is r_{ui} , and \hat{r}_{ui} denotes the predicted rating value.

$$\text{Adjusted Cosine}(\vec{a}, \vec{b}) = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| * |\vec{b}|} = \frac{\sum_{i \in R_{ab}} n_{ai} n_{bi}}{\sqrt{\sum_{i \in R_a} n_{ai}^2 \sum_{i \in R_b} n_{bi}^2}} \quad \text{Eq.4}$$

\vec{a} and \vec{b} are the rating vectors of user a and user b , respectively. n_{ai} is the numeric rating of item i given by user a . R_{ab} is the set of items co-rated by both user a and b . The problem of this measure is that it doesn't consider the differences in the mean and variance of the ratings made by user a and b . Some users could be optimists who have a tendency to rate items higher than average users and some other users could be pessimists who have a tendency to rate items lower than average users. As a solution of this Cosine measure, there is the adjusted Cosine similarity, which is mainly used item-to-item based similarity, rather than user-to-user similarity. The next measure, the Spearman's rank correlation avoids using the rating values directly.

The Spearman's rank correlation coefficient aims to use the relative differences in users' ratings through ranks of the ratings. For rating sets of a given user, this measure computes the rank of each item, first. When there are the same ratings, these items get the average rank of their spot. Then, the Spearman's rank correlation is computed as the following.

$$\text{SRC}(u_a, u_b) = \frac{\sum_{i \in I_{ab}} (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\sum_{i \in I_{ab}} (r_{ai} - \bar{r}_a)^2 \sum_{i \in I_{ab}} (r_{bi} - \bar{r}_b)^2}} \quad \text{Eq.5}$$

r_{ai} is the rating rank of item i in user a 's list of rated items and \bar{r}_a is the average rank of items rated by a . Even though the purpose of this method is to avoid the problem of rating normalization, when there are a large number of tied ratings and the rating schema ranges a few values, it easily lost the advantage. In addition, this method is computationally expensive.

Mean Square Difference is to evaluate the similarity as the inverse of the average squared difference of co-rated items between user a and b . The weakness of this method is that it doesn't allow computing the negative correlation between user preferences and when there is no co-rated item, it is impossible compute this value.

$$\text{MSD}(u_a, u_b) = \frac{|R_{ab}|}{\sum_{i \in R_{ab}} (n_{ai} - n_{bi})^2} \quad \text{Eq.6}$$

Lastly, the Pearson correlation coefficient is to compute the similarity by removing the effects of mean and variance of users' ratings.

$$\text{Pearson}(u_a, u_b) = \frac{\sum_{i \in R_{ab}} (n_{ai} - \bar{n}_a)(n_{bi} - \bar{n}_b)}{\sqrt{\sum_{i \in R_{ab}} (n_{ai} - \bar{n}_a)^2 \sum_{i \in R_{ab}} (n_{bi} - \bar{n}_b)^2}} \quad \text{Eq.7}$$

According to empirical studies about comparing quality of similarity measures, especially for user-to-user based similarity, the Pearson correlation outperformed other measures [67]. Due to the simple, computationally inexpensive and good quality, this Pearson correlation is the most popular similarity measures used in recommendation systems [42, 77]. Therefore, in this dissertation, I chose this Pearson correlation as a similarity measure for numeric ratings.

4.2 RECOMMENDATION TECHNOLOGIES

In order to generate collaborative filtering (hereafter, CF) recommendations, we generally do the following four steps;

- **Modeling Users Preferences:** By understanding what users' like in the past, recommender systems aim to predict what they are going to like in the future. Hence, the first task of the recommender systems is to model users' preferences based on their rating/bookmarking patterns.
- **Selection of peers:** Once they understand users' preferences, systems based on CF technology pick N peers who have the most similar tastes or interests with their target users.
- **Computation of prediction probability:** Among item of N peers, the system chooses candidate items which are not discovered by our target users yet and, based on the preferences of both target users and N peer, predicts the probability of how much the target users would like each candidate.
- **Suggestion of recommendations:** Depending on the prediction probability, recommender systems choose and suggest the most favorable items as recommendations.

My proposed social network-based (SN) recommendation algorithm mainly aims to make a change at the second and the third stage of the recommendations as shown in the Figure 4. At the second stage, as social network-based recommendations, target users' social connections will substitute for the anonymous peers of traditional collaborative filtering. At the third stage of the recommendations, the way to compute the recommendation probabilities will be considered from K-nearest neighbor approach to matrix factorization. In this section, K-Nearest Neighbors

approach and matrix factorization approach will be described. Detailed descriptions about social network-based recommendations will be explained in each corresponding chapter.

		Computation of Prediction Probability	
		K-Nearest Neighbors	Matrix Factorization
Selection of Peers	Anonymous Peers	CF approach	SVD approach
	Social Networks	SN approach	SN SVD approach

Figure 4. Summary of Recommendation Algorithms

4.2.1 Recommendations based on K-Nearest Neighbors

K-nearest neighbor approach chooses K anonymous peers who are the most similar to our target users and, among K peers' bookmarked items which are not discovered by target users yet, selects recommendable candidate items [159, pp. 291 - 324]. In order to compute the information similarity between a target user and his peers, it is critical to choose the best similarity measure. Particularly, because there is no empirical evidence indicating the best similarity measure for bookmark-based preferences, it is required to test and compare various measures. In addition, several values of K will be examined in this paper.

4.2.2 Matrix Factorization Recommendations

The matrix factorization method is to derive a set of latent (hidden) factors from the users' rating/bookmark patterns and characterize both users and items by such latent factors. For movie recommendations, as an example, the latent factors could be genre, director of the movies or

main actors. For music recommendations, the factors could be music types (e.g., classic, hip-hop, or country) or singers of the music. Even though the latent factors are not interpretable by humans, matrix factorization statistically identifies critical characteristics to make recommendations. For a large number of items, matrix factorization can collapse the whole items into a smaller approximation in which highly correlated and co-occurring items were transformed into small size vectors. In addition, this technology can predict missing ratings as approximate values. Therefore, even though users didn't rate/bookmark a lot of items, it is possible to generate recommendations. The overall recommendation quality is known to be better than traditional collaborative filtering [86, 112, 149]. In this paper, as a matrix factorization method, singular vector decomposition (SVD) with 'Alternating-Least-Squares with Weighted -Regularization (ALSWR) factorizer' was used [129].

Again, for Citeulike and CN3 datasets which consist of users' bookmarks, B is the user-item bookmark matrix, $B = \{B_{ui}\}_{l \times n}$ where l and n denote the number of users and items, respectively. $u \in \{u_1, \dots, u_l\}$ represents users and $i \in \{i_1, \dots, i_n\}$ denotes items. The bookmark of user u on item i is b_{ui} . In here, \hat{b}_{ui} denotes the predicted bookmark of item i for user u , which is picked by my recommendation algorithm as presumably favorable and bookmark-able item.

This SVD starts with a low-rank approximation of the $l \times n$ bookmark matrix B . When f is the number of latent factors to be extracted, we are able to derive two f -dimensional lower rank matrices: $R \approx PQ^T$. One matrix P presents users' latent factor matrix, where $P \in \mathbb{B}^{l \times f}$. The u -th row of P , p_u is the user u 's vector and represents how much a user u is associated with f latent factors. Another matrix Q presents items' latent factor matrix, where $Q \in \mathbb{B}^{n \times f}$. The i -th row of

Q , q_i is the item i 's vector and represents how much an item i possesses f latent factors. By computing the inner product between the two vectors, we can approximate the probability of the missing bookmarks, \hat{b}_{ui} (i.e. how likely a certain user u is going to bookmark item i) [38, 81] as the following.

$$\hat{b}_{ui} = p_u q_i^T \quad \text{Eq. 8}$$

The major challenge of SVD is to minimize the error of prediction, $e_{ui} = b_{ui} - \hat{b}_{ui}$. When there are many missing elements in the bookmark matrix, however, the typical SVD algorithm cannot find P and Q . One way to solve this problem is to fill out the missing values, for instance, with the average number of bookmarks of a user or an item. However, when a system fills out the missing values in a wrong way, even if it is possible to find the P and Q , the results are highly prone to overfitting [86]. As the solution of these data sparsity and overfitting problem in SVD, Zhou, et al. [199] proposed the ALS-WR, which to learn the model by fitting the existing bookmark records. As the first stage, from known bookmark records, we need to model P and Q with a weighted- λ -regularization. The optimal model is defines as

$$(P^*, Q^*) = \min_{P, Q} \sum_{(u,i) \in K} e_{ui}^2 + \lambda(\|p_u\|^2 + \|q_i\|^2) \quad \text{Eq. 9}$$

Where K contains all (u, i) pairs for which the b_{ui} is known (the training set). λ is a constant to control the degree of regularization and usually decided by cross-validation. By learning from the existing bookmark history, this modeling aims to minimize the regularized squared error [86].

Now it is the stage to compute alternating least squares (ALS). ALS aims to minimize noises and possible errors on the above regularized model. The basic idea is, when both P and Q

are unknown, to fix one of the matrices and to compute another matrix again by solving a least squares problem. This algorithm alternates the steps by fixing one of the both matrices. In specific, when Q is fixed, so as to re-compute the P matrix, ALS computes a separate ridge regression for each user. It takes the latent vector (q_i) of items bookmarked by the user u as input variables and the value of his bookmarks (b_{ui}) as output variables.

$$A_u = Q[u]^T Q[u] = \sum_{i \in B_u} q_i q_i^T \quad \text{Eq. 10}$$

$$d_u = Q[u]^T b_u = \sum_{i \in B_u} b_{ui} q_i \quad \text{Eq. 11}$$

Where A_u is the covariance matrix of user u 's item latent factor vector (having the number of factor f), which is the input, and d_u is the input-output covariance vector. Finally it finds the optimal p_u by the ridge regression as the following.

$$p_u = (\lambda l_u E + A_u)^{-1} d_u \quad \text{Eq. 12}$$

E is the f -dimensional identity matrix: $E_f \in \mathbb{B}^{f \times f}$. l_u is the number of items bookmarked by user u . Similarly, when P is fixed, we can re-compute the Q matrix using the following (n_i is the number of user who bookmarked item i) [149].

$$q_i = (\lambda n_i E + A_i)^{-1} d_i \quad \text{Eq. 13}$$

In particular, for this paper, I used the implementation of SVD with ALS-WR recommendation module on Mahout.

4.3 EVALUATION OF RECOMMENDATION QUALITY

Evaluation methods of recommendations are determined by aspects of the assessment. Majority of recommendation-related studies have been focusing on how the recommendation predictions are accurate. Along with the accuracy, there are various aspects to be evaluated. How are the recommendations diverse (diversity)? Are the recommendations novel or obvious (novelty)? Are the recommendations serendipitous to the recipient (serendipity)? Are the systems suggesting trust worthy recommendations (trustworthiness)? Do the recommendation systems produce the suggestions efficiently (scalability)? In e-commerce system, does applying recommendation technology increase the revenue (utility)? As the recommendation technologies are evolving and gaining the popularity, the viewpoints of how to evaluate the technologies are also evolving. In this dissertation, I assessed the quality of recommendation technologies through offline experiments. That is, all experiments in this dissertation are performed using users' existing bookmark histories or ratings histories, and then simulate how the recommendation system can predict the pre-defined bookmarking or rating behaviors. Therefore, rather than evaluation criteria based on the users' perceptions, such as trustworthiness, system utility or scalability, the accuracy of the technology will be the main aspect to be assessed. The evaluation method for the recommendation accuracy is also determined by the kind of predicted user preferences – whether the recommendation is to predict a unary rating (i.e. presence of a bookmark) or a numeric rating. In the following section, evaluation methods for unary ratings and numeric ratings will be explained separately.

4.3.1 Evaluation Method for Unary Ratings

For unary ratings like Citeulike and CN3 datasets, information retrieval-based evaluation methods, such as precision, recall and F-1 measure, are usually used. In an offline experiment, we normally use N-fold cross validation. Users' bookmark sets are randomly split into N equal sized subsets so as to execute N-fold cross validation. For each iteration, one of the N subsets was used as a test set and the remaining N-1 subsets were used as a training set. This process was repeated with a different test set for N times. Then, the accuracy is judged by how many items in a test set appear in the recommendations. In the results, there are four possible outcomes of the recommendations and the test set.

Table 2. Classification of the Possible Outcomes of the Recommendations

	Recommended	Not Recommended
In Test Set	True-Positive (tp)	False-Negative(fn)
Not In Test Set	False-Positive (fp)	True-Negative (tn)

In the classification of the Table 2, the precision is the ratio of 'True-Positive' in the recommended list ($\text{precision} = \text{tp}/(\text{tp}+\text{fp})$) and the recall is the ratio of 'True-Positive' out of the all items in the test set ($\text{recall} = \text{tp}/(\text{tp}+\text{fn})$) [161]. In this paper, I will compute the accuracy of the recommendations according to the recommendation ranks. The recommendations are usually displayed in a ranked list. Users expected that items in a higher rank would be more important than other items in lower ranks. Therefore, it is critical to place correct recommendations in a higher rank, and the position of correct recommendations is a critical evaluation criterion of

recommendations. In all kinds of SN-based recommendation in my consideration, I will use multiple numbers of Top N recommendations (e.g. Top10, Top 5, and Top 2). For the evaluation of recommendations in each rank, precision at point N and recall at point N will be used. The precision at point N (precision@N) is the ratio of the number of correctly predicted items in the top-N recommendation list to N (Eq. 14). Recall at point N (recall@N) is the ratio of the number of correctly predicted items in the top-N recommendation list to the total number of relevant items (Eq. 15).

$$\text{precision@N} = \frac{\text{No. of correct prediction}}{\text{N of top N set}} = \frac{\text{test} \cap \text{top N}}{\text{N}} \quad \text{Eq. 14}$$

$$\text{recall@N} = \frac{\text{No. of correct prediction}}{\text{size of test set}} = \frac{\text{test} \cap \text{top N}}{\text{test}} \quad \text{Eq. 15}$$

Between the precision and recall, it is typical that there is a tradeoff. As the N is larger, it is likely that the recall is getting better and the precision is getting worse, and vice versa. Hence, considering the precision and recall separately, I will use the harmonic value, F1 measure, as the following [82].

$$\text{F1@N} = 2 \cdot \frac{\text{precision@N} \cdot \text{recall@N}}{\text{precision@N} + \text{recall@N}} \quad \text{Eq. 16}$$

Lastly, in order to compute the recommendation accuracy by ranks in more sophisticated manner, I also computed the Mean Average Precision (MAP) as the following [87]. It computes average precision of each test user and averages out the average precision of all users. The formula of average precision (AP) of a test user according to rank N is the following. A variable r denotes a recommended item among the set of Top N recommendations, R ($r \in R$) and another variable Hit_r represents hit of the corresponding item. When a corresponding item is one of the test items, then the hit value is 1 and otherwise it is 0.

$$AP@N = \frac{\sum_{r=1}^R Hit_r / r}{\text{No. of correct prediction}} \quad \text{Eq. 17}$$

4.3.2 Evaluation Method for Numeric Ratings

The difference between unary ratings and numeric ratings is, as explained, the latter schema can represent how much a user like or dislike the corresponding item. Therefore, the evaluation method should judge not only whether the recommendation system can predict favorable items like the above F1 measure do, but also how much the system can predict the rating values correctly or how much the system can predict the relative position of the recommended items on the list of users' rated items. Mean absolute error (MAE) and root mean squared error (RMSE) are the most widely-used measures to compute the recommendation system's predictability of the rating values. On the other hand, the normalized cumulative discounted gain (NDCG) is the most typical measure to compute the predictability according to the relative positions.

First, MAE computes the difference between the known rating values and the predicted rating values as the following.

$$MAE = \sqrt{\frac{1}{|test|} \sum_{(u,i) \in test} |\hat{r}_{ui} - r_{ui}|} \quad \text{Eq.18}$$

where r_{ui} and \hat{r}_{ui} are the known rating value and the predicted rating value of item i for a target user u , respectively. The item i is in the user u 's *test* set. The following equation shows how to compute the RMSE.

$$\text{RMSE} = \sqrt{\frac{1}{|\text{test}|} \sum_{(u,i) \in \text{test}} (\hat{r}_{ui} - r_{ui})^2} \quad \text{Eq.19}$$

Compared with MAE, RMSE penalizes larger errors. Let's assume that there are three items in the test set. In one recommendation, the error of one item is 3 and the errors of the other two items are 0. In one recommendation, the error of two items is 2 and the error of another item is 0. Then the RMSE of the former case is larger than the latter case [161].

Lastly, NDCG is widely used evaluation method not only for information retrieval but also for the evaluation of recommendations. When recommendations are presented to users as ranked list, regardless of the predicted ratings, it considers the relative rank positions of recommended items. This method compares the suggested ranked list with the optimal ranking list. The Discounted Cumulative Gain (DCG) of each user is computed as the following.

$$\text{DCG}(u) = \sum_{i \in \hat{p}_u}^n \frac{2^{r_i}}{\log_2(1 + i)} \quad \text{Eq.20}$$

Where u is the recipient of recommendations (\hat{p}_u) and i denotes an item of i th ranked position in the recommendation list. The r_i is the user's rating on the item i . In particular, using the logarithm base, this measure adds more weight to highly ranked documents and incorporates different level of accuracy through different gain values. Hence this measure emphasizes the correctness of highly-ranked items more. This DCG value is normalized into NDCG by dividing by the ideal DCG which is the optimal ranking list. In the same way, the ideal DCG is computed, but since the ideal DCG is the optimal ranking, the list is reordered such that the most highly

rated items appear first. The average normalized DCG over all users is selected to show the accuracy of recommendation [1, 168].

5 WATCHING NETWORK-BASED RECOMMENDATIONS

This chapter aims to examine how feasible recommendations based on watching network are. As a newly emerging social network, it is hard to find any comparable offline social networks. This means that it is still doubtful about the significance of this new online social network and, in accordance with the intentions defined by the service providers, it is still unsure whether the social network can be really a useful information source. Therefore, the promise of watching network-based recommendations should be based by proof positive that the presence of watching relations defines the similarity of the users' interests. The first part of this chapter will suggest the proof. In particular, using various information similarity measures, it will be examined the degree of information similarity between pairs of watching connections and changes of the similarity according to the social distance. Once we find reasonable degree of similarity between watching relations, the recommendations based on the watching network will be assessed in the second part of this chapter. The quality improvement made by watching network-based recommendations will be assessed by comparison with typical collaborative filtering recommendations. As the data source of this study, Citeulike will be used.

5.1 DATA SOURCE FOR WATCHING NETWORK: CITEULIKE

The data source of the watching network is Citeulike⁵. The system is one of the leading systems for managing and sharing bibliographic information, along with *Bibsonomy*⁶, Mendeley⁷, and *Connotea*⁸. CiteULike is one of the most successful social bookmarking systems and enables the users to engage in watching relations. While other SNSs (e.g. Facebook, Twitter, Google Plus, etc.) which aim to promote online social networking and to encourage sharing various information about users' daily lives, the Citeulike system is to manage and share one solid kind of information – bibliographic information of scholar articles and books – and provides watching network as one method for users to acquire useful information.

Citeulike data in consideration is made and distributed by the Citeulike administrators. Specifically, the dataset used in this dissertation is backed up on May 15, 2011. The data set contains article ids, all users, users' bookmarks, dates and times of the bookmarks and social tags at the time when this dataset was made. Since it doesn't contain information about metadata of articles (i.e. titles, authors, publication journals, abstracts, and publication years) and watching relations, I collected the information separately. The following Table 3 is the descriptive statistics, and Figure 5, and Figure 6 show the distributions of users and articles according to the bookmarks. This dataset has 94,388 Citeulike users and 3,210,960 articles, consisting of about 3.9 million bookmarks. 66.3% of the users ($n = 62,568$) have at most 5 bookmarks and 89.9% of the articles ($n = 2,885,833$) were bookmarked only by one user. The bookmark sparsity of this

⁵ <http://www.citeulike.org>

⁶ <http://www.bibsonomy.org/>

⁷ <http://www.mendeley.com/>

⁸ <http://www.connotea.org/>

Citeulike data set is 0.9999, which is similar with the sparsity of Netflix [77, p. 170]. The users who are in watching relations (whether they are watching or being watched) are 11,439, which are 12.1% of the whole population. They formed 44,847 watching relations totally.

Table 3. Descriptive Statistics about the Cituelike Watching Dataset

Social Networks	No. of Users in Watching Relations	11,439
	No. of Relations	44,847
	No. of Unidirectional Relations	44,381
	No. of Reciprocal Relations	233
	Avg. No. of Watched Partners per Watching User	13.91 ($\sigma = 28.7$)
Information Collections	No. of Distinct Articles	3,210,960
	No. of Users	94,388
	No. of Bookmarks	3,869,993
	Avg. No. of Bookmarks per User	41.0 ($\sigma = 407.5$)
	Avg. No. of Tags per User	143.21 ($\sigma = 2408.3$)
	Avg. No. of Bookmarkers per Article	1.21 ($\sigma = 1.4$)
	Avg. No. of Tags per Article	4.21 ($\sigma = 7.2$)

Figure 7 is the distribution about how many connections each user watches. 3223 users watch other users. Each user watch about 14 users on average and 77.3% of them ($N = 2492$) watch at most 5 connections. Figure 8 displays the distribution about how many users are being watched by other users. 9,897 users are watched by other users, and 4.53 users watch them, on average. Most of them ($N = 8,126$, 82.1%) are watched by at most 5 users, and in particular, 15 users are very popular. They watched by more than 100 users. These users are making 44,381 watching relations. Interestingly, very few of the relations ($N = 233$) are reciprocal. I suggest this majority of unilateral direction in the relations is because Citeulike users are unable to see who are watching them. Before beginning to compute information similarities of watching relations and the watching network-based recommendations, one should understand the heterogeneous

information usage patterns depending on users' social status and the social patterns of watching network.

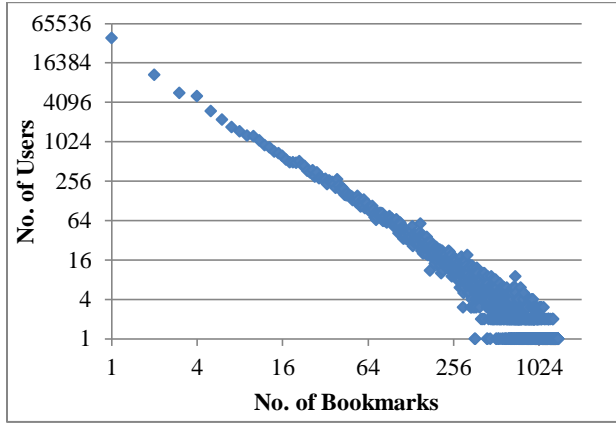


Figure 5. Distribution of Users according to the Number of their Bookmarks

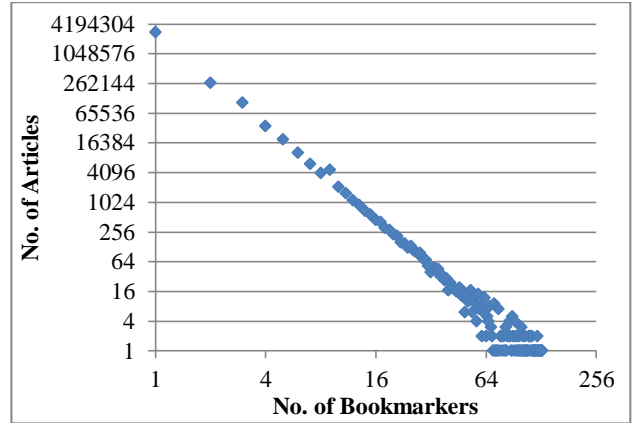


Figure 6. Distribution of Articles according to the Number of The Bookmarks

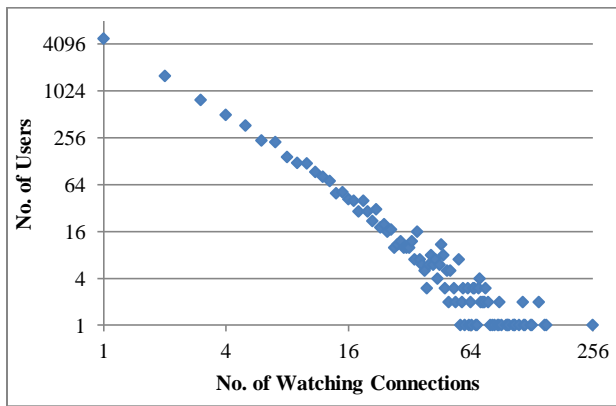


Figure 7. Distribution of users according to how many users they are watching

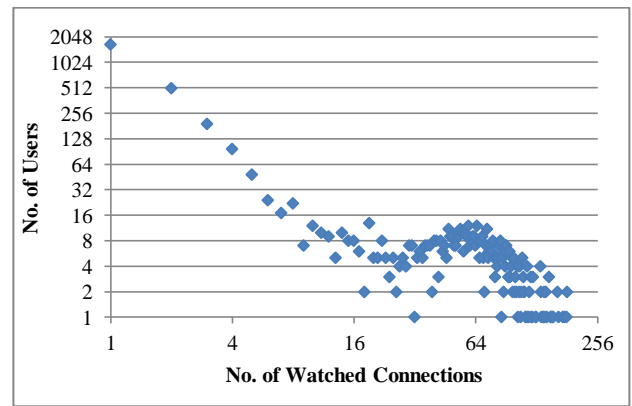


Figure 8. Distribution of users according to how many users they are being watched by

Information Usage Patterns of Users according to Their Social Status

As mentioned, about 12% of users in the dataset (11,439 users) are watching other users or are being watched by other users. That means 88% of the other users are not involved in any watching relations. I examined the differences of these two groups of users in terms of their information usage patterns. The former group of users who are in watching relations (whether they are watching users or watched partners) started to use Citeulike in earlier day of the system lifetime (975.8 days ago when this data set was made, $\sigma = 638.3$ days ago) and have used Citeulike for longer period of time ($M = 455.2$ days, $\sigma = 535.6$ days). On the center of a different date when each bookmark was made, they also used the system more often ($M = 34.3$ times, $\sigma = 70.9$ times). They also collect larger amount of bookmarks, 217.4 bookmarks ($\sigma = 592.9$) on average. On the other hand, the latter group of users not having any social partners started to use Citeulike in relatively later day (801.2 days ago when this data set was made, $\sigma = 567.2$ days) and have used Citeulike for much shorter period of time ($M = 58.6$ days, $\sigma = 191.7$ days). They also used the system less often ($M = 2.8$ times, $\sigma = 8.7$ days). They have about 10 times less bookmarks than the former user group, 16.7 bookmarks ($\sigma = 368.3$). Figure 9 shows how two groups of users have different distributions of bookmarks. Because the difference of the information usage patterns between two user groups is so clear, I didn't run a statistical test. The differences mean that users who are more familiar with and more active in the system tend to engage in watching relations. The Pearson correlation test also shows the same results. The earlier they started to use the Citeulike and the longer a user has used the system, the more social networks he has ($\gamma = .18, p < 0.001$; $\gamma = .05, p < 0.001$). The more bookmarks a user has, the more social partners (either watching users or watched users) he has ($\gamma = .15, p < 0.001$).

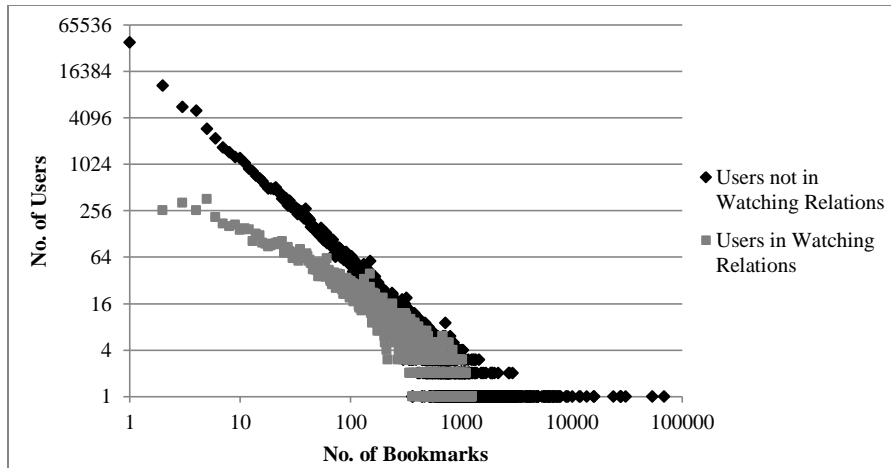


Figure 9. Bookmark distribution of users who involve in watching relations contrasted with the one of users not involved

When there are some differences in usage patterns between users who are in and out of the Citeulike watching network, do watching users and watched users use the system differently? The watching users started to use Citeulike in earlier days of the system lifetime (1,323 days ago when this data set was made, $\sigma = 618.3$ days) than watched users (920.5 days ago, $\sigma = 623.1$ days). They also have used Citeulike for longer period of time ($M = 616.4$ days, $\sigma = 606.1$ days) than the watched users ($M = 474.9$ days, $\sigma = 547.1$ days). However, the numbers of bookmarks between the watching users ($M = 235.7$, $\sigma = 630.6$) and the watched users ($M = 243.4$, $\sigma = 632.1$) are not significantly different. Figure 10 contrasts the bookmark distributions of two user groups. The interesting part is that the more items a user bookmarked, the more users watch him ($\gamma = .37$, $p < 0.001$). The earlier a user started to use the Citeulike and the longer he has used the system, the more users he is watched by ($\gamma = .05$, $p < 0.001$; $\gamma = .25$, $p < 0.001$). In addition, the more item a user bookmarked, the more users he watches and this correlation is significant ($\gamma = .046$, $p < 0.001$). The earlier a user started to use the Citeulike and the longer he has used the system, the

more users he watch ($\gamma = .29, p < 0.001$; $\gamma = .08, p < 0.001$). Through these results, it is clear that users who had experienced the system more tend to watch or be watched. Collecting more bookmarks is also beneficial to involve in watching relations. Since users are unable to see who is watched by whom, it is clear tendency that rich users are watched by other users due to the perceived usefulness of the bookmark collections, rather than social popularity. However, since this Citeulike dataset doesn't contain any record about the time when a user started to watch another user, it is hard to figure out the causal relations between the watching relations and the size of bookmarks. As explained, once a user declared a watching relation with another user, the library of the watched user is automatically updated to the watching user on his 'watchlist' page. Hence, it is unclear whether the watching users aggregated their rich collections by referring to their watched parties and copying some items from the watchlist or whether they actively collected information using their own strategies and the two users in watching relations got connected due to inherent similarities of their interests. Therefore, in the later section, I will explore whether watching relations are a convenient social mechanism to acquire useful items or a social network sharing not only the bookmarks but also inherent similarity of interests.

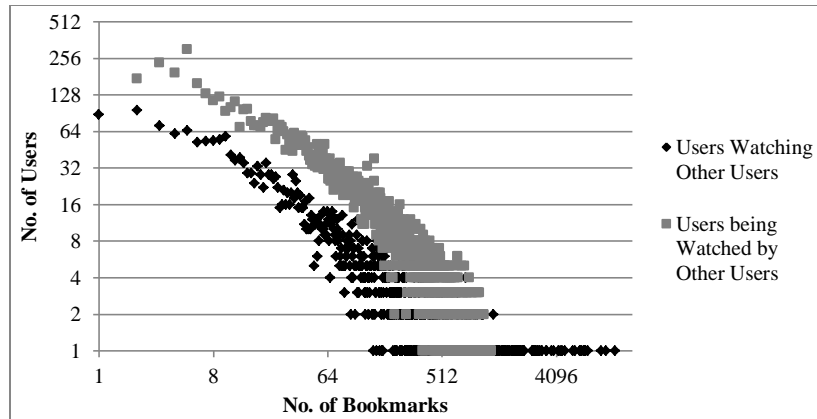


Figure 10. Bookmark distribution of users who are watching others, contrasted with the users who are being watched by others

5.2 FEASIBILITY OF WATCHING NETWORK AS A USEFUL INFORMATION

SOURCE

It is presumed that watching relations are highly centered on interesting objects, since the relations are usually initiated with users' perceived utility of the partners' information collection, and the major function of the relations is referrals of useful information. The above descriptive statistics also shows that richer users tend to be watched by more users. However, it is a vague presumption, and there is no clear evidence showing the feasibility of the social network as a useful source for recommendations. Even though it seems for users to get connected with a purpose to acquire information, are they actually influenced by their connections? Will information collected by a user be useful to his watching partners? Otherwise, are users racking up watching connections just for fun or out of curiosity as they do on various online social networking sites? Once the point that utility-based relations maintain high information similarity

is proved, which similarity measure is the most representable for the shared interests? In order to answer these questions, I will examine the following hypotheses.

H5.1. Information similarity between a pair of users who are in a watching relationship is higher than the information similarity between a pair of users who are not in a watching relationship.

H5.1.1. Information similarity between a pair of users who are in a direct watching relationship is higher than the information similarity between a pair of users in an indirect watching relationship.

H5.1.2. With the increase of social distance, the information similarity between two users is decreasing.

H5.2. The similarity between a pair of watching connections is comparable to the similarity of users' Top N peers which are chosen by traditional CF recommendation technology.

The target items of Citeulike dataset are scientific articles which contains various metadata. In order to determine the most representative similarity measure for watching network, the similarities based on the metadata will be compared with item-based similarities

H5.3. Metadata-based similarities are better than item-based similarities on demonstrating the similarities of watching connections.

5.2.1 Information Similarity Measures

In the Citeulike dataset, users' preferences were modeled by their bookmark history and expressed in unary ratings. Therefore, the similarity between a given users pair can start from counting how many items they co-bookmarked. In addition, the target items of the Citeulike data are scientific articles which contain various publication-related metadata from titles and abstracts to publication years and author names. When users bookmark the articles, the system enables them to add free-text tags, as well. Therefore, in this paper, the similarity of shared information was assessed using *item-based* and *metadata-based* approaches.

For the rest of this paper, we use the following notation. B is the user-item bookmark matrix, $B = \{B_{ui}\}_{L \times N}$ where L and N denote the number of users and items, respectively. $u \in \{u_1, \dots, u_l\}$ represents users, and $i \in \{i_1, \dots, i_n\}$ denotes items. B_i is the bookmark collection of item i and B_u is the bookmarks of user u .

5.2.1.1 Number of Co-bookmarks and Jaccard Coefficient

The very basic and traditional way to measure the information similarity using unary ratings is to count the number of items shared by two given users. In this study, as the item-based similarity measure, the number of common items found in both users' bookmark collections was counted. This refers to as '*the number of co-bookmarks*'. This is absolute number of common items, since we didn't consider the proportion of the common items to users' bookmark collections. However, sizes of bookmark collections varied dramatically from user to user. Let's assume that

there are two pairs of users – user a and b ; user c and d – have the same number of co-bookmarks, 10. The former pair is richer users, and the union collections of their bookmarks $|B_a \cup B_b|$ consist of 250 items. The union collections of the latter pair $|B_c \cup B_d|$ have 40 items. For the richer users (a and b), sharing 10 common items are rather trivial, but for the poorer users (user c and d), sharing 10 items means that most of their favorite items are common. We need to measure the number of co-bookmarks using *relative* (normalized) ratios, thereof. Based on this number of co-bookmarked items, there are various similarity measures, such as Jaccard, Dice, Cosine and Log-likelihood similarities. As the first normalized measure, we chose ‘*Jaccard coefficient*’. The Jaccard coefficient is the proportion of shared items in the union set of two users’ collections. When user a and user b are in consideration, we can compute the coefficient as the following.

$$\text{Jaccard Power } (u_a, u_b) = |B_a \cap B_b| / |B_a \cup B_b| \quad \text{Eq. 21}$$

5.2.1.2 Log-likelihood Similarity

The next similarity measure is ‘*the log-likelihood similarity*’, which is another typical similarity measure for the unary ratings. Log-likelihood similarity was initially proposed as a statistical analysis method for text processing. The strength of this similarity is that, unlike other similarity methods which take into account only the cases where an item was co-bookmarked, it takes into account not only the cases of co-bookmarking, but also the cases where an item was bookmarked by only one person and even the cases where an item was not bookmarked either of them. That is to say, this method measures how unlikely the overlap of interests was occurred just by chance [46, 75].

$$\Lambda(u_a, u_b) = \frac{p_{a,b}}{p_b} \times \frac{p_{\sim a, \sim b}}{p_b} = \left(\frac{n_{a,b}}{N_a} \times \frac{N_b}{n_a} \right) \times \left(\frac{n_{\sim a, \sim b}}{N_b} \times \frac{N_a}{n_b} \right) = \frac{n_{a,b} n_{\sim a, \sim b}}{n_a n_b} \quad \text{Eq.22}$$

where, $\ln(\Lambda(u_a, u_b))$ denotes the log-likelihood similarity between two user a and user b . $n_{a,b}$ is the number of items bookmarked by both user a and b . n_a and n_b is the number of items bookmarked by only user a or user b , respectively. $n_{\sim a, \sim b}$ is the number of items bookmarked by neither of them. A large positive similarity means that they are very similar and a large negative similarity means the dissimilarity. The similarity value of 0 indicates that two users in comparison are neither similar nor dissimilar [75].

5.2.1.3 Popularity Weight of Co-bookmarked Items

In measuring information similarity, due to the popularity bias (i.e. users are more likely to provide feedback on popular items than on rare items [166]), co-bookmarking very popular items doesn't reflect their shared interests as much as the co-bookmarks of rare items. For instance, in measuring movie tastes, the co-bookmarking of a very popular movie like 'Avartar' doesn't show the true shared interests between two users than the co-bookmarking of a unpopular independent film. Therefore, we incorporate the item popularity into information similarity using the average inverse item popularity as the following. We defined this similarity measure as '*popularity weight*' of co-bookmarked items.

$$IP_{ab} = \left(\log \sum_{i \in B_{ab}} \frac{N}{n_i} \right) / n_{ab} \quad \text{Eq.23}$$

where IP_{ab} is the average inverse popularity of co-bookmarks between user a and b . Item i is one of the items that the user a and b commonly bookmarked, and N is the total number of

items in the dataset. n_i is the number of bookmarks of the item i and n_{ab} is the number of co-bookmarks between the user a and b . Bookmarks containing less popular items are thus assigned a higher popularity weight value.

5.2.1.4 Metadata-based Similarity

Due to the irregular opportunistic nature of the bookmarking process, users with similar interests may not necessarily end up with very similar collections. In this case, similarity of interest might be more reliably measured on the level of item's meta-data than counting co-bookmarked items. Even though users' preferences were expressed in unary ratings, by taking into account the related metadata of bookmarked items, we are able to analyze why users bookmarked the corresponding items in more cognitive ways. For example, in *Citeulike* domain, bookmarking papers from the same authors or same journal (or conference) names can be considered as an indicator of similarity. When titles of two papers in comparison are constituent of the similar set of keywords, it is likely that they contain the similar contents and context. In addition, users added free-text social tags when they bookmarked items. Compared with item bookmarks simply showing users' interests on the corresponding items, social tags indicates users' more cognitive and descriptive interests on tagged items [71, 196]. With this reason, in this study using Citeulike dataset, we consider various kinds of metadata – paper titles, author names, and the social tags.

As the first metadata-based similarity, we computed '*title vector-based similarity*'. In order to compute the similarity of the textual metadata, we used the vector space model [156]. Before computing the title vector-based similarity, we built users' preference model from all titles of their favorite items. All terms in titles of a user's bookmarked papers were aggregated

into one bag of keywords. Then for effective comparison, we applied text processing techniques to the bag of title keywords. The terms were case-normalized to the lower-case letters, and all stop words⁹ were removed. We also applied Porter stemmer so as to reduce word variation to its stems or roots [107]. The processed bag of title keywords were transformed into title keyword vector consisting of keywords and the TF/IUF (Term Frequency/Inverse User Frequency) values. In this paper, we computed each user as a document. Therefore, instead of the document frequency value in the original TF/IDF formula, I used the user frequency indicating how many users bookmarked papers of which the titles contain the keywords. As Mori, et al. (2006) suggested, when words in two articles similar, the contextual representations of the article pair are also similar. In particular, titles of papers represent the topic of the paper succinctly. Therefore, this title vector-based similarity tries to compute how much two given users are interested in similar contents and contexts.

As the second metadata, we used authors' names. We named the similarity, as '*author name vector-based similarity*'. Because papers are published under the names of the authors, the topics of papers are directly connected with the authors, and the readers acknowledge the names as experts of a corresponding topic. When a user bookmarks several papers written by an author, it is easy to infer his interests lay on the author's research topics. Hence, author names are one of the important anchors to find interesting papers and indicative of interests on the related topics. Since the author names are proper nouns, none of the text-processing techniques were applied to this metadata. All authors' names of one user's bookmarked papers were aggregated as one bag of names and the TF/IUF values were computed. In here, in the same way of building title

⁹ <http://jmlr.csail.mit.edu/papers/volume5/lewis04a/a11-smart-stop-list/english.stop>

keyword vector, the user frequency value is about how many users bookmarked papers written by the same authors.

Along with the titles and author names, there are other kinds of metadata available to infer users' preferences such as abstracts, journal/conference names and publication year. However, I failed to find the abstracts and journal/conferences names for more than the half of the papers in this Citeulike dataset. In addition, I concluded that the publication years don't carry much meaning as a kind of metadata to infer users' information preference and similarity. Therefore, only two kind of metadata were considered in this dissertation.

5.2.1.5 Social Tag-based Similarity

The last metadata is users' social tags. Compared with titles and author names describing the contents and context of articles innately, social tags are about how each user approaches the articles from a personal point view. It is known that tags annotated by a user represent his conceptual understanding or categorization of a resource. Thus, users' social tags can serve as rich evidence about user interests. When two users have many common tags, we interpreted this result to mean that their interests are overlapped.

In particular, unlike other studies about social tag-based information similarity, we didn't count tag similarity based tri-partite (user-tag-resource) relations. Put differently, we didn't compute how two users annotate the same tags for the same articles. Instead, we computed how two users annotate the same tags regardless of the annotated articles. The reason is that our own previous studies showed that the tag similarity based on the tri-partite relations (i.e. how two users annotated the same items with the same tags) was too trivial (0.06% of the Jaccard efficient

for the Citeulike watching relations and 0.004% of the Jaccard efficient for the Citeulike group co-members) to suggest a meaningful similarity level between two users [96, 98].

Before we build social tag vector, due to the idiosyncratic nature of social tags, it is required to clean up the metadata. In Citeulike, users are able to add free-text tags of their choice on bookmarking items. However, when users don't assign any tags, Citeulike automatically assigns pseudo-tags, 'no-tag'. When users imported bibliographic information from a citation file such RIS, BibIX, ENL files, the system also automatically adds pseudo-tags like '*file-import-xx-xx-xx'. These pseudo-tags were added just for the system's sake and don't carry any meaning. Therefore, we eliminated the pseudo-tags for the comparison. Then, we built two kinds of tag vectors. For one tag vector, we didn't apply any text-processing techniques except case-normalization, and it is referred to as '*original tag vector-based similarity*.' For another kind of tag vector, the same text-processing techniques used for paper titles (i.e. case normalization, stop-word removal, and stemming) were applied to the tags, and it is referred to as '*processed tag vector-based similarity*.' These two tag vectors compute the degree of overlapped interests between two users. By comparing these two kinds of tag vectors to each other, we aim to test whether applying text processing techniques improve the effectiveness of the similarity measures. For both original tag vector and processed tag vector, we computed the TF/IUF values for each tag. The user frequency is about how many users annotated their bookmarks with a corresponding tag.

Once all metadata vectors – title vector, author name vector, original tag vector and processed tag vector – were built, the similarity between a pair of users was computed using the

cosine similarity for every metadata vector. In summary, the Figure 11 shows the similarity measures used in this study.

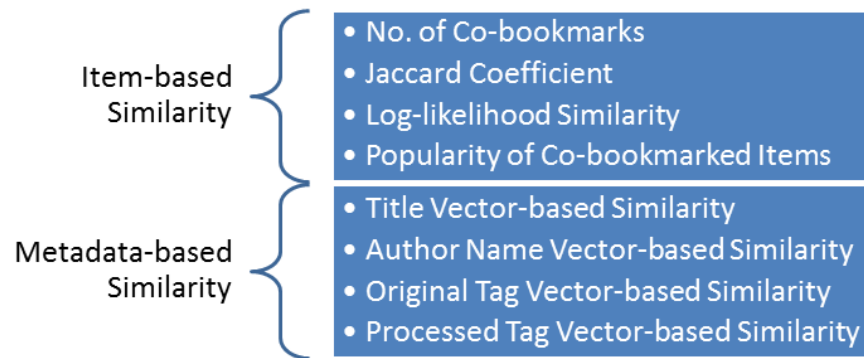


Figure 11. Summary of Similarity Measures

5.2.2 Information Similarity of User Pairs according to their Social Distances

This section is to test whether pairs of users in watching relationships share more similar information than users who are randomly paired and not socially associated to each other at all (H5.1). As the first step, the item-based information similarity was computed according to the users' social distances – direct, 1hop, 2hop and random pairs. In particular, the Citeulike dataset considered in this dissertation has two kinds of social networks. One kind is watching network which is explored in this chapter and another kind is group memberships which will be explored in the chapter 6. Thus, I classified two users neither in watching relations nor in group co-memberships as a random pair. Since they are not socially associated, and the social reachability is uncountable, their social distance was defined as infinite. The Table 4 shows the result.

Table 4. Comparison of Item-based Similarity Depending on the Distances of Watching Relations

	No. of Co-bookmarks	Jaccard	Log-Likelihood	Popularity
Direct	1.80	0.21%	.204	8.69
1Hop	.39	0.04%	.097	7.75
2Hop	.16	0.02%	.061	7.38
No Relation	.04	0.02%	.023	6.92
	$F = 5192.0,$ $p < .001$	$F = 10408.9,$ $p < .001$	$F = 7091.0,$ $p < .001$	$F = 2396.4,$ $p < .001$

For all kinds of item-based measures, the similarities were significantly different according to the social distances of user pairs (for number of co-bookmarks, $F = 5192.0$, $p < .001$; for the Jaccard Co-efficient, $F = 10409.0$, $p < .001$; for the log-likelihood similarity, $F = 7091.0$, $p < .001$; and for popularity, $F = 2396.4$, $p < .001$). For all kinds of the item-based measures, user pairs in direct watching relations produced the largest degree of similarities. On the contrary, the similarities of user pairs randomly coupled were the smallest for all measures. The differences were obvious, and the Scheffé post-hoc pairwise comparison also showed that the differences between direct watching connections and random pairs were significant. Hence, at least for the item-based similarities, the hypothesis 5.1 has met.

In addition, the similarities were decreased incrementally along with the increase of the social distance. The post hoc pairwise comparison displayed that the differences of all item-based similarities were significantly different among pairs in direct, 1hop and 2hop distance. The results proved that the hypotheses 5.1.1 and 5.1.2 are true for the item-based similarities. Therefore, we conclude that interests of direct watching connections are the most similar. According to the result of log-likelihood similarities, the co-shared items are based on inherently co-shared interests between two users, not by chance. The users in direct watching relationships also tend to share rare items, and user pairs in distant relationships or in no relation shared

popular or rather common items. We can identify this result as clear evidence about the homophily existing on the watching network.

We also compared the metadata-based similarity by the social distances and the following Table 5 is the result. The result of the metadata-based similarity is also the same with item-based similarity. For all kinds of metadata-based measures, the similarities were significantly different according to the social distances of user pairs (for title vector-based similarity, $F = 21296.7$, $p < .001$; for the author name-based similarity, $F = 8623.1$, $p < .001$; for original tag-based similarity, $F = 18692.5$, $p < .001$; and for processed tag-based similarity, $F = 19152.0$, $p < .001$). In the same patterns as the comparison of items-based similarities demonstrated, the similarities of users in direct watching relationships were the largest and the similarities of users in no relations were the smallest. The post-hoc pairwise comparison showed that the difference was significant. The hypothesis 5.1 hereby has the met completely. The similarities were decreased along with the increase of the social distance, as well. The post hoc pairwise comparison presented that the differences of all metadata-based similarities were significant among pairs in direct, 1hop and 2hop distance. Therefore, both hypotheses 5.1.1 and 5.1.2 are true for all kinds of similarities. Conclusively, users in direct relations shared items having the most similar topics or the most similar tag sets than user pairs in distant or infinite social distances.

Table 5. Comparison of Metadata-based Similarity Depending on the Distances of Watching Relations

	Title Vector	Author Name Vector	Original Tag Vector	Processed Tag Vector
Direct	.1440	.0149	.0362	.0505
1Hop	.0814	.0033	.0095	.0168
2Hop	.0626	.0020	.0062	.0114
No Relation	.0147	.0007	.0012	.0020
	$F = 21296.7,$ $p < .001$	$F = 8623.1,$ $p < .001$	$F = 18692.5,$ $p < .001$	$F = 19152.0,$ $p < .001$

5.2.3 Information Similarity vs. Social Features

In the above, we checked that watching relations share the most similar interests. As the close scrutiny of the information sharing patterns, I examined how the information similarities are correlated with the social features of watching relations. The selected social features are bibliographic coupling and edge betweenness centrality. First, bibliographic coupling of two vertices in a directed network (i.e. watching network) is the number of other users which both vertices watch. For instance, vertex i and j watches three of the same users and so have a bibliographic coupling of 3. In a directed network, if two users have social relationships with the same other people, the number of common neighbors can indicate how strongly they overlap socially [135, pp. 116-118]. In addition, this watching network is mainly to acquire useful information from other users whose information collections are interesting. Therefore, the number of the same partners in watching network can be a good indication that they deal with similar matters. When there is significantly positive correlation between the information similarities and this bibliographic coupling, it is clear proof that the presence of watching relations defines the similarity of the users' interests. Fortunately, the correlation test showed

that all eight kinds of similarities have significantly positive correlations with the values of bibliographic coupling (for number of co-bookmarks, $r = .06$, $p < .001$; for the Jaccard Coefficient, $r = .05$, $p < .001$; for the log-likelihood similarity, $r = .12$, $p < .001$; and for popularity, $r = .06$, $p < .001$; for title vector-based similarity, $r = .10$, $p < .001$; for the author name-based similarity, $r = .08$, $p < .001$; for original tag-based similarity, $r = 0.05$, $p < .001$; and for processed tag-based similarity, $r = .05$, $p < .001$).

Second, a betweenness centrality of an edge is the number of geodesic paths (i.e. the shortest path) passing through the edge. In order to compute this centrality value, geodesic paths between every pair of vertices in the network are computed, and the numbers of paths running along the corresponding edge are counted [135, p. 382]. Basically, this is a property for vertex; to measure how one vertex can act as a bridge to connect two vertices. In the similar sense, this edge betweenness is to consider influence of an edge within a network. In Citeulike watching network, let's assume that information flows from user to user, and each user has the same power of communication. The most effective way to deliver information from one vertex to another vertex is to take the shortest route. Therefore, it is more likely that the information will often pass down through user pairs having high edge betweenness than pairs with low edge betweenness. That is to say, edges with high tend to have important control over passing information within the network.

If there is a significant correlation between edge betweenness centralities and users' information similarities, the result could be interpreted to mean that the user pairs sharing more similar items tend to have more control over circulating information within the social network. When users formed their watching network around items of their interest, moreover, pairs of

users having common interests will share more coherent set of items and other users get around them so as to acquire the relevant information to their interests. Therefore, the user pairs whose interests are more largely overlapped may have stronger influential power. The Pearson correlation test found the significantly positive correlations of edge betweenness with the number of co-bookmarks, the Jaccard co-efficient and log-likelihood similarity (for number of co-bookmarks, $r = .10$, $p < .001$; for the Jaccard Co-efficient, $r = .06$, $p < .001$; for the log-likelihood similarity, $r = .15$, $p < .001$). However, this result raised one question. Isn't it likely that larger amount of information collections could lead to higher betweenness values because users may simply get around rich users to acquire information? According to correlation test, there was a significantly positive correlation between this edge betweenness and the number of items in union set of two users' collections ($r = .08$, $p < .001$). Then is it true that users just decided to get around rich users? In order to check this question more, it is tested how edge betweenness is correlated to item popularity and metadata-based similarities. I also found that significantly positive correlations for other remaining similarities (for popularity, $r = .10$, $p < .001$; for title vector-based similarity, $r = .11$, $p < .001$; for the author name-based similarity, $r = .09$, $p < .001$; for original tag-based similarity, $r = 0.04$, $p < .001$; and for processed tag-based similarity, $r = .05$, $p < .001$). This means that the user pairs having higher edge betweenness centrality tend to share rare items, and their interests described by the titles, author names, and social tags of their favorite articles are more similar than other pairs having lower edge betweenness. Users are not simply watching other rich users but watching other users who have coherent items matched to their interests.

5.2.4 Social Connections vs. Top N Anonymous Peers

The previous results demonstrated that users in watching networks do share their interests. Then is the watching network substitutable for anonymous peers selected by CF computations, in terms of the shared interests? This question will be answered in the next section about watching network-based recommendations (i.e. the Section 5.3) in a full scale. Prior to the Section 5.3, as a simple way to answer this question, the information similarities of watching connections were compared with the similarities of top N peers (H5.2). Since CF mainly used item-based similarity, top 20, 10, and 5 peers were selected using Jaccard co-efficient and log-likelihood similarity. The reason why these two item-based similarities are particularly considered is that these two similarity measures are widely used in recommendations using bookmark records.

For the comparison, users who watch other users were chosen as target users, and six groups of their anonymous peers (i.e. top 20, 10 and 5 peers based on Jaccard co-efficient and log-likelihood similarity, respectively) were selected for every target user. These peers could include the target users' watched partners since the selection of the peers were made by fully automated computations of the nearest neighbor approach, without any consideration of social information. In the same manner to the top peers, I classified target users' watched partners into eight groups (i.e. all of the watched partners and top 20, 10 and 5 partners based on Jaccard co-efficient and log-likelihood similarity, respectively). Then, I examined the differences on the information similarities between anonymous peers and social partners. The results demonstrated that target users' watching connections under-shared their bookmarks than anonymous peers.

According to a One-way within subjects ANOVA, for both Jaccard co-efficient and log-likelihood similarity, the differences were significant ($F = 610.0$, $p < .001$ for Jaccard co-efficient; $F = 972.3$, $p < .001$ for log-likelihood similarity). In the Bonferroni pairwise comparison to find the pattern of differences, for both the Jaccard and log-likelihood similarities, the values of anonymous peers were significantly higher than the ones of watching connections. That is, target users share more common items with their anonymous peers than the watching partners. This result suggests that, when taking into account the information similarity solely based on co-bookmarks, the recommendations based on this watching network might not be easy to beat traditional collaborative filtering.

If anonymous peers possess more common items of interests, are the watching connections less useful sources of information than anonymous peers? If anonymous peers are effective information sources, why did the target users connect to their watching partners, instead of anonymous peers? Is it because the target users simply couldn't find the anonymous peers or is there another reason for the target users to make the watching relationships with their current partners? In order to answer these questions, other aspects of information were explored – metadata. As explored above, metadata similarities could indicate overlapped interests between two users in more cognitive and descriptive ways. When anonymous peers share genuinely more common interests with our target users, it is expected that their metadata-based similarities will be also higher than the similarities of watching connections. The metadata-based similarities were compared for the same anonymous peers and watching partners. In particular, since Jaccard co-efficient and log-likelihood similarity picked heterogeneous groups consisting of different peers, the metadata-based similarities were computed and compared for both groups separately.

Figure 12 shows the results. The results demonstrate that watching connections have higher title and tag similarities than anonymous users, even though the similarities of author names were the opposite. The differences were all significant according to One-way within subjects ANOVA ($F = 1253.6, p < .001$ for title similarity; $F = 996.0, p < .001$ for author name similarity; $F = 542.3, p < .001$ for original tag similarity; $F = 606.1, p < .001$ for processed tag similarity). In case of title similarity, the results of Bonferroni pairwise comparison showed that title similarity of top 5 anonymous peers chosen by Jaccard co-efficient (i.e. Top 5 peers by Jaccard in the Figure 12) were insignificantly different with the title similarities of watching connections. However, the title similarities of the other peer groups were significantly lower than watching connections. In case of social tag similarity, our target users shared significantly more similar tags with their watching partners than peers, regardless whether the social tags were original or text-processed. Even though our target users have more common items with anonymous peers, their preferences expressed in a semantically rich format were not as much similar as the preferences of watching connections are. I interpret these results to mean that, due to the irregular and opportunistic nature of bookmarking process, item-level similarity may lag behind ‘true’ interests shared between pairs of watching connections. Therefore, notwithstanding the under-shared items, it is evident that users in a watching relation share similar interests and they could be a critical source for personalized recommendations to each other. However, author name-based similarities between pairs of watching connections were rather lower than the same similarities between target users and their anonymous peers. The differences were also significant. It hits that the author name is a feature closely related to the items per se. Therefore, the similarity based on author names might not be a good similarity measure to indicate watching connections’ shared

interests and further, to be used in watching network-based recommendations. To conclude, depending on how to measure the information similarities, our target users share more similar items with their anonymous peers and share more similar keywords and social tags with their watching connections. Therefore, the anonymous peers and watching connections have complementary ability to each other. Therefore, I suggested that H5.2 has met.

As shown up to now, it is important to decide how to measure the information similarity for watching network-based recommendations. Item-based measures were not sufficient to represent the similarities of watching connections. It seems that the shared interests of watching connections could be effectively expressed through titles of their favorite items and the social tags. In the next sub-section, in order to determine the best similarity measure for the watching network-based recommendations, the comparative effectiveness of each similarity measure will be assessed.

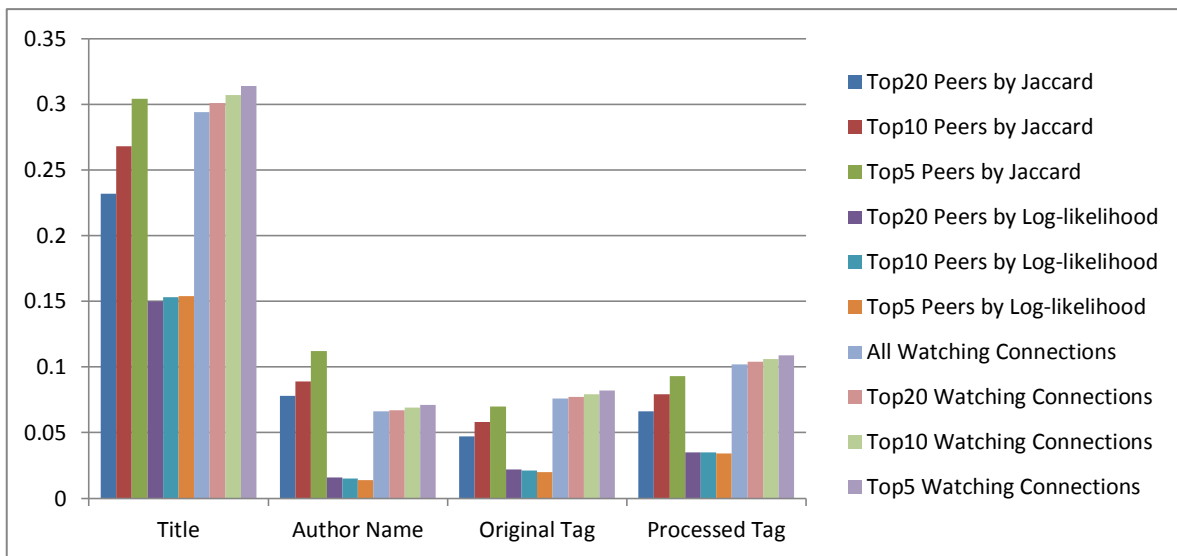


Figure 12. Metadata-based Similarity Comparison between Watching Connections vs. Top N Peers

5.2.5 The Best Similarity Measures for Watching Network

In the previous section, we verified that the watching relations enable Citeulike users to communicate knowledge and information to each other. The aforementioned results suggested that users in watching relationships possess higher information similarities than users in distant social distance or no connection. As explained in the section 5.2.4, nevertheless, some similarity measures are better than the others in representing the shared interests of watching relations, and the effectiveness of each similarity measure is still unclear. This section aims to test which measure indicates the similarity of the watching relations the most effectively. The aforementioned 8 similarity measures are assessed by seeing which measure predicts the existing relations the best. For this test, we chose the users who watch the other users as target users as I did on the section 5.2.4. Therefore, they are 3223 watching users. Then, we computed all eight similarity measures of the target users with every other user, regardless whether the two users are socially associated or not. Then, we picked the top N similar users of the target users and among the top N users, it is counted how many users were actually in social associations. The comparison was executed with increasing top N ranks such as top 50, top 20, top 10 and top 5. Under the precondition of homophily and social influence, we can reckon that the more similar interests two users commonly share, the more likely they are socially associated with each other. This comparison, thus, seeks to determine the similarity measure to meet this condition.

As the evaluation criteria of this prediction, traditional evaluation methods for information retrieval were used; precision and recall. The precision aims to measure how precise the predictions are and the recall aims to measure how complete the predictions are. More

specifically, precision at point N (precision@N) is the ratio of the number of correctly predicted items in the top-N list to N. Recall at point N (recall@N) is the ratio of the number of correctly predicted items in the top-N list to the total number of relevant items. As explained, we will compute the accuracy of the predictions according to the four different top N ranks. (N = 50, 20, 10, and 5).

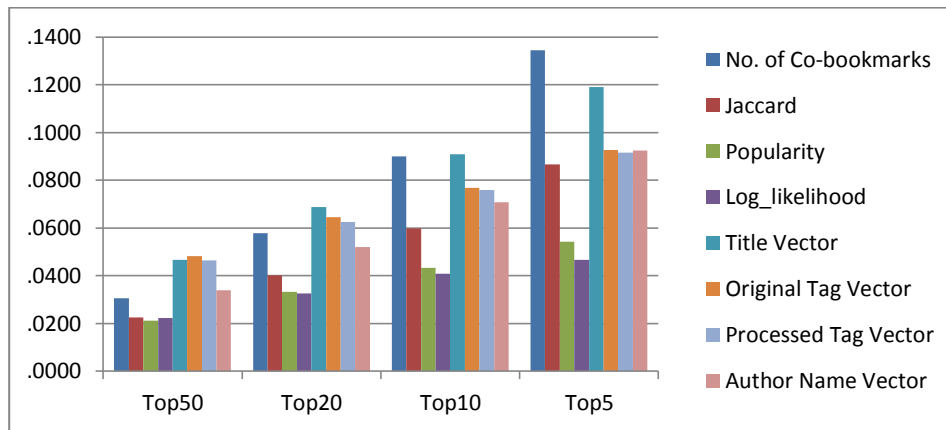
First, for 3223 target users who are watching users, we tested the effectiveness of individual measure by forecasting who they watched. Figure 13 shows the precision and recall of the predictions made by each similarity measure, respectively. According to the one-way ANOVA test, the differences in the precisions are significantly different depending on the kinds of measures for all ranks ($F = 106.7, p < .001$ for Top 50; $F = 92.0, p < .001$ for Top 20; $F = 100.2, p < .001$ for Top 10; and $F = 127.9, p < .001$ for Top 5). In lower ranks (i.e. top 50 and top 20), most metadata-based measures were significantly better than item-based similarity. However, in the higher ranks (i.e. top 10 and top 5), the number of co-bookmarks (i.e. absolute number of co-bookmarks) and title vector-based similarity produced significantly the most accurate predictions. Between these two measures, there was no significant difference. Two tag vector-based similarities and author name-based similarities were the second best measures, and among these measures, there were no significant differences in terms of precision. However, the Jaccard co-efficient (i.e. normalized and relative ratio of co-bookmarks) and the popularity weights made the worst predictions in the top N list.

Regarding the results of recall, like the results of precision, in higher N list (i.e. top 20, top 10 and top 5), the number of co-bookmarks produced the most complete predictions out of the target users' existing social relations. For top 20 and top 10, the title vector-based similarity

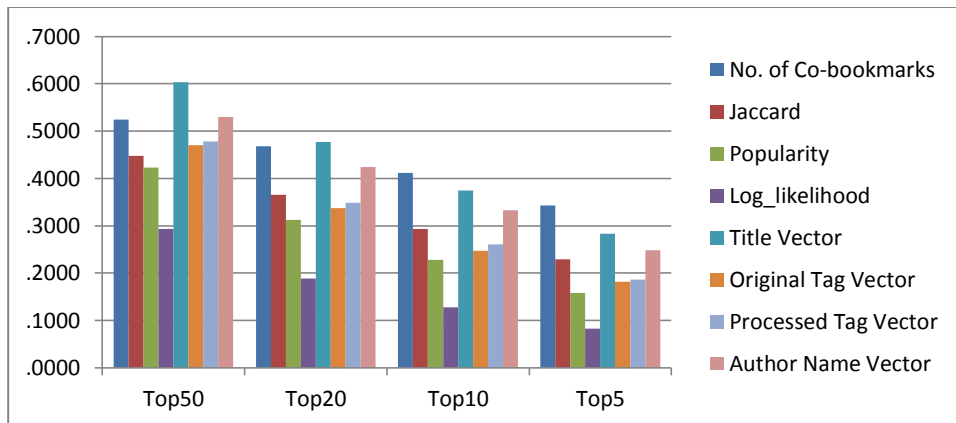
also produced the good results, and there was no significant difference between the absolute co-bookmark number and the title vector-based measure. However, in top 5 results, the number of co-bookmarks significantly outperformed the title vector-based measure. The pattern of difference regarding the recalls of the other similarities was not the same to the patterns of the precisions. In any top N ranks, log-likelihood measure did the least complete prediction. In addition, two tag-based similarities didn't performed well. For either kind of evaluation criteria, author name-based similarity was a good and effective measure, since it predicted the significantly second most correct and complete predictions in all ranks. Additionally, tag-based similarities are computed using term-frequency information. In particular, even though the original tags and the processed tags are rooted on the same set of tags, several text-processing techniques were applied to the later set in order to reduce word variation and to increase the chances of term matching. The Scheffé pairwise test result yielded that the two measures are not significantly different to each other regardless of the evaluation criteria and the N ranks. It demonstrates that text-processing techniques such as stop-word removal and stemming don't significantly contribute the effective comparison of tags.

Regardless of the evaluation criteria (either precision or recall), counting how many items a given pair of users co-bookmarked (i.e. absolute measure of co-bookmarks) is the most effective way to compute the information similarity of this utility-based watching networks. Title vector-based measure representing how much the topics of the bookmarked items are overlapped between two users is also effective. On the other hand, except the number of co-bookmarks, the item-based measures are less effective than the metadata-based measures. In spite of the good performance of absolute measure of co-bookmarks, the effectiveness of the normalized measures

(i.e. Jaccard co-efficient and log-likelihood) was disappointing. Another popularity weight was also below our expectation. Lastly, two tag-based measures performed quite the same level, since we failed to find any significant difference. Therefore, applying the text processing techniques didn't contribute distinctive improvement in representing shared tags between watching connections. Therefore, the hypothesis suggesting better performance of metadata-based similarities than item-based similarities (H5.3) is rejected.



(a) The results of the Precision



(b) The Results of the Recall

Figure 13. Results of the Test to Forecast Watching Relations

As the next step, we combined the individual measure to intensify the effectiveness and tested whether combining multiple effective measures together increase the performance or not. Each similarity measure has heterogeneous range of values. For instance, the values of co-bookmarks could range from 0 to 3,210,960 (i.e. total number of items in our dataset) and the values of Jaccard coefficients and all metadata-based similarities range from 0 to 1. Hence, before fusing similarity measures, it is required to normalize each similarity measure using the Standard Score (SS) like the following.

$$SS = \frac{x - \mu}{\sigma} \quad \text{Eq.24}$$

where x is one similarity value, and μ and σ is the mean and standard deviation of the corresponding similarity measure, respectively. Then the normalized similarity values were combined together, and existing watching relations were predicted using the combined similarities. The following Figure 14 and Figure 15 show the results of precision and recall of the combined measures along with the best two individual similarities (i.e. the number of co-bookmarks and title vector-based similarity).

We found that, regardless of the ranks, there are significant differences on the precisions ($F = 63.76, p < .001$ for top 50, $F = 57.54, p < .001$ for top 20, $F = 65.49, p < .001$ for top 10, $F = 83.84, p < .001$ for top 5). The Scheffé post-hoc pairwise test showed that the combinations of metadata-based measures produced significantly better results than the combinations of item-based measures, in lower ranks (top 50 and top 20). In higher ranks (i.e. top10 and top 5), on the other hand, the differences were blurred. The precisions were insignificantly different between the combinations of item-based measures and the combinations of metadata-based measures. However, when the number of co-bookmarks is fused with metadata-based similarities, it

produced the better precisions for both top 10 and top 5 results. Specifically, when the number of co-bookmarks, title vector-based similarity and processed tag vector-based similarity were combined together (i.e. ①+⑤+⑦ on the Figure 14), it significantly outperformed all the other similarities. Among the combinations, hence, combining the number of co-bookmarks, titles and processed tags is the most effective to accurately predict the existing watching relations and to represent the similarity of watching connections. However, interestingly, two best individual measures (i.e. the number of co-bookmarks and title vector-based similarity) also made almost equivalent level of predictions with the combinations even though the pairwise test concluded that the precisions of the two individual similarities are significantly lower than the best-performed combination.

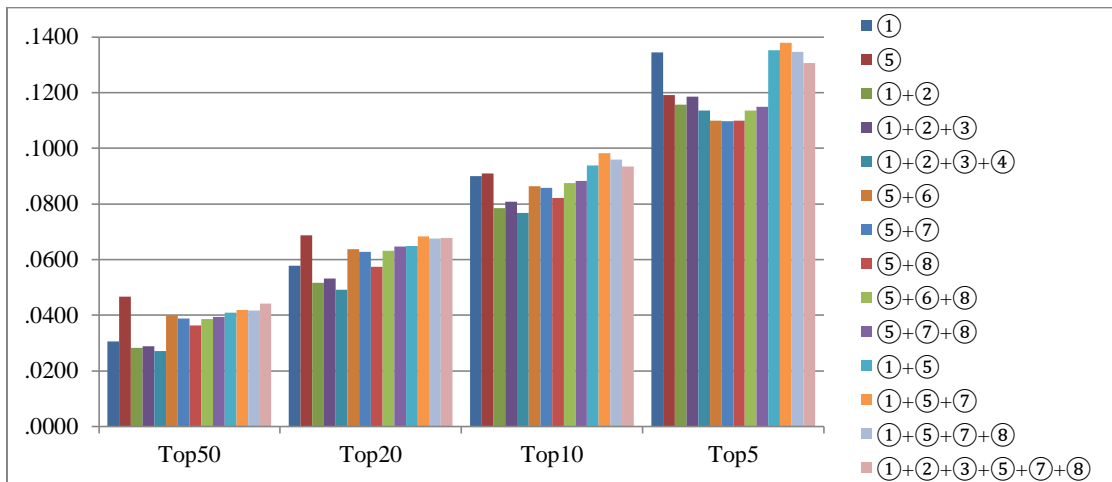


Figure 14. Precisions of Combined Measures for the Prediction of Watching Relations
 (①= No. of Co-bookmarks; ② = Jaccard; ③= Popularity; ④ = Log-likelihood; ⑤ = Title Vector; ⑥ = Original Tag Vector; ⑦ = Processed Tag Vector; ⑧ = Author Name Vector;
 The data series is displayed in the order of the legend)

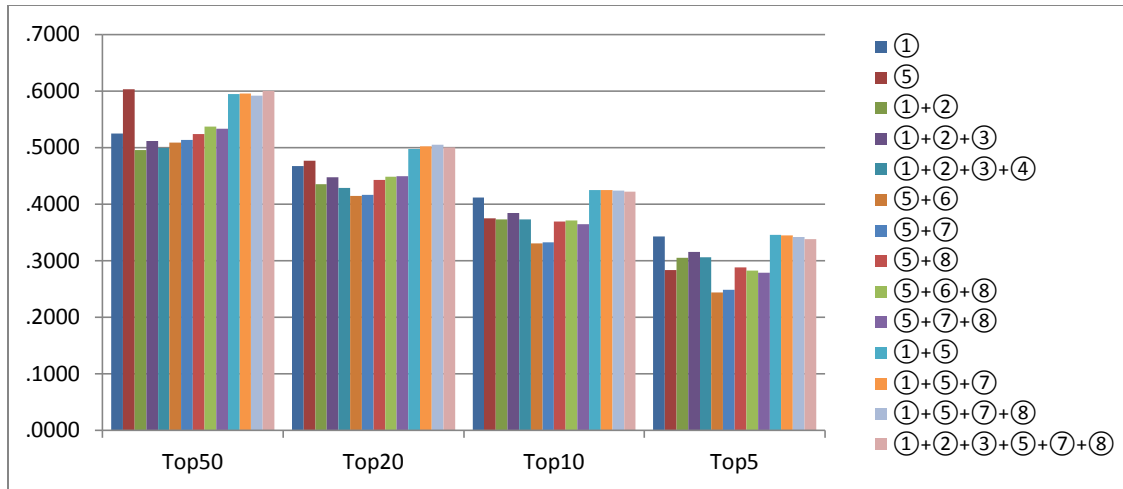


Figure 15. Recalls of Combined Measures for the Prediction of Watching Relations
 (①= No. of Co-bookmarks; ② = Jaccard; ③= Popularity; ④ = Log-likelihood; ⑤ = Title Vector; ⑥ = Original Tag Vector; ⑦ = Processed Tag Vector; ⑧ = Author Name Vector;
 The data series is displayed in the order of the legend)

One-way ANOVA test also showed that there are significant differences on the recalls across the similarity measures ($F = 81.61, p < .001$ for top 50, $F = 94.20, p < .001$ for top 20, $F = 101.04, p < .001$ for top 10, $F = 96.98, p < .001$ for top 5). The results indicate that, in lower ranks, it is failed to find any distinct pattern of differences between the combinations of metadata-based similarities and the combinations of item-based similarities. In higher ranks, the combinations of item-based similarities predicted more existing watching relations than the combinations of metadata-based similarities. In higher rank, as shown on the pattern of precisions, the combinations of the number of co-bookmarks with metadata-based similarities performed better. In particular, two kinds of combinations significantly outperformed – combinations of the number of co-bookmarks and title vector-based similarity (i.e. ①+⑤ on the Figure 15) and combinations of the number of co-bookmarks, title vector-based similarity

and processed tag vector-based similarity (i.e. ①+⑤+⑦ on the Figure 15). Between these two kinds of combinations, there was no significant difference. In addition, two best individual measures (i.e. the number of co-bookmarks and title vector-based similarity) also made good predictions, even though the recalls of the two individual similarities are significantly lower than the best-performed combinations.

To conclude, for both evaluation criteria, the combination of the number of co-bookmarks, title vector-based similarity and processed tag vector-based similarity explain shared interests of watching relations the most effectively. In order to predict and effectively show the sociality of watching relations, thus, the most critical information to consider is what they bookmarked, what the bookmarked items are about and which tags users annotate the bookmarked items with. In the following section, recommendations based on watching network using will be executed and assessed using various similarity measures including the best-performed combinations of similarities.

5.3 EVALUATION OF WATCHING NETWORK-BASED RECOMMENDATIONS

This section explores the potentials of watching network as an effective source of information by including the network in the personalized recommendations. Since the main focus of this part is on the benefit of the watching network as an alternative source to the anonymous peers, the watching network-based recommendations will be compared with typical CF recommendations. Therefore, the most critical hypothesis to be assessed in this part is the following.

H5.4. Watching network-based recommendations produce better recommendations than traditional CF approach.

Moreover, as mentioned, the final output of this recommendation is research articles which embed not only users’ bookmarks records about them but also various metadata such as titles, author names, publication journal or conference names, etc. and users’ annotated tags. In order to augment the recommendation quality, it is critical to determine the best way to take advantage of the watching network and the various metadata in the personalization process. Thus, in this watching network-based recommendation, we considered three aspects – users’ preferences on information items (i.e. research articles or books), watching network of users and content of items. Summarizes the overall design of recommendation approaches considered in this part.

		Computation of Prediction Probability		
		K-Nearest Neighbors		Matrix Factorization
		Bookmark	Titles & Social Tags	
Selection of Peers	Anonymous Peers	<ul style="list-style-type: none"> • CF • CFCW 		<ul style="list-style-type: none"> • CFSVD • CFSVD_CW
	Direct Watching	<ul style="list-style-type: none"> • Watch • Watch_CW 	<ul style="list-style-type: none"> • CBCW_Watch 	<ul style="list-style-type: none"> • WatchSVD • WatchSVD_CW
	Direct Watching + Watched	<ul style="list-style-type: none"> • ReciVarWatch • ReciVarWatch_CW 	/	<ul style="list-style-type: none"> • ReciVarWatchSVD • ReciVarWatchSVD_CW
	Direct Watching + Indirect Watching	<ul style="list-style-type: none"> • 1hopWatch • 1hopWatch_CW 		<ul style="list-style-type: none"> • 1hopWatchSVD • 1hopWatchSVD_CW
	Anonymous Peers + Direct Watching & Watched	<ul style="list-style-type: none"> • Hybrid • Hybrid_CW 		<ul style="list-style-type: none"> • HybridSVD • HybridSVD_CW

Figure 16. Overall Design of Watching Network-based Recommendation Algorithms

As mentioned on the section 4.2, CF recommendations are made in four steps – modeling user preferences, selection of peers, computation of prediction probability and suggestion of recommendations. In order to find better ways to generate the watching network-based recommendations, in every step, I will try out various ways of doing that.

Modeling Users Preferences

The recommendations based on K-nearest neighbors are generated using two different kinds of user preferences – bookmarks of items and metadata of items. The above section 5.2.5 shows that both users' bookmark records and metadata of items (in particular, titles and users' social tags) are workable similarity measures. In particular, when the bookmark records are fused with the metadata, it represents the shared interests of watching connections the best. Therefore, I modeled users' preferences in three ways – 1) based on users' bookmarks, 2) based on metadata of items (i.e. content boosted recommendations; that is, CBCW recommendations), and 3) based on the combinations of bookmarks and the metadata (i.e. recommendations using content similarity weight; that is CW).

Selection of peers

The results of the previous section indicated that users in direct watching connections tend to share more similar interests than indirect watching connections. Then is it better to include not only the connections to be watched by our target users but also the connections watching our target users? Even though the latter connections were not defined by our target users, the analysis result showed that they also have similar interests. In addition, when target users have very few watching connections, instead of relying on the limited number of social links, is it effective to include social links who are connected with distance (i.e. 1 hop)? Lastly, according to the results of the first part, the top N peers still have higher bookmark similarity than their watching

connections. Can hybrid recommendation fusing users' top N peers with their watching connection enhance the quality? Therefore, the point that I evaluate in this step (i.e. the selection of peers) is to determine the optimal scope of social network to be a component of the watching network-based recommendations through the following hypotheses.

H5.5. Recommendations based on direct watched partners are better than the recommendations based on extended scope of watching connections.

H5.5.1 Recommendations based on direct watched partners are better than the recommendations based on both direct and indirect watched partners.

H5.5.2 Recommendations based on direct watched connections is better than the recommendations based on both direct watching and watched connections.

H5.6 Pure social recommendations solely based on watching network is better than the hybrid recommendations based on both watching network and anonymous top N peer users.

Algorithms to compute the prediction probability

Once the better way to model users' preference is determined, in accordance with the resultant preferences, the better recommendation algorithm to compute the prediction probability will be chosen. In here, I test the K-Nearest Neighbor approach and matrix factorization algorithm.

Suggestion of recommendations

The recommendations mainly based on users' bookmark records are usually hard to make more precise suggestions than the recommendations based on numeric ratings. The bookmarks just represent users' simple interests on the corresponding items but don't reflect how much they are interested in or why the items interest them. Thus, for more accurate recommendations, we need to consider extra information. Citeulike is one of the typical social bookmarking systems and the

social bookmarking systems generally enable users to annotate social tags regarding what they think about the bookmarked items. As a social bookmarking system, of course, a part of the Citeulike data is the social tags. Additionally, other kinds of metadata such as title, abstract, author names, publication journal/conference names and publication years are also available. However, as explained, in the Citeulike data set considered in this dissertation, except title and author names, the other kinds of metadata are missing in many articles. Therefore, at the final step of the recommendations where the recommended items are selected by the prediction probability, users' extra preferences based on the metadata will be applied.

5.3.1 Recommendation Algorithms

5.3.1.1 Watching Network-based Recommendation and Hybrid Recommendations

The main focus of this study is to assess the substitutability of watching network for anonymous peers in recommendations. The watching network is a unilateral network defined by one party of the relations, not by mutual consents of two parties. According to the first part of this chapter, however regardless of the direction, the users in watching relations share common interests to each other. Therefore, even though the watched users didn't initiate the relations and don't know who are watching them, due to the shared interests, the information favored by their watching users could be a useful source. In addition, although the users connected with distance (i.e. 1hop) shared less similar interests than directly watching connections, their similarities were higher than the values of randomly paired users. Therefore, the users having a 1hop distance with our target users could be a valuable source, as well. In order to determine the optimal scope of

the watching connections to be a component of the watching network-based recommendations, I proposed three strategies in the selection of social connections – based on direct watching connections (‘Watch’), direct watching and watching connections (‘ReciWatch’) and direct and indirect watching connections (1hopWatch).

In the step to compute the prediction probability of recommendations, K-Nearest neighbors and matrix factorization algorithm were used alternatively. When the K-Nearest Neighbor approach is used, I only considered users’ watching networks. Let’s assume that the social recommendations are generated using the K-Nearest Neighbors based on direct watching connections. The recommendation considers only the bookmarks of our target users and the bookmarks of their partners being watching by our target users. In the execution of the matrix factorization approach, for every target user, I built a separate sub-matrix being made up of the target users’ bookmarks (in the training sets) and bookmarks of the direct watching connections. Due to the small number of bookmarks and much less sparse matrix, the recommendations could be efficiently generated. The data sparsity of the whole Citeulike dataset in our consideration is .9998 but the average data sparsity of watching network-based sub-matrices is 0.771 ($\sigma = 0.30$). The standard matrix factorization approach feeds the whole bookmark data onto the computation once. Therefore, the system requires a huge size of memory and often complains about the heap size error. When considering the whole dataset, the algorithm requires a large number of latent factors (e.g. > 100 for Netflix dataset [199]) and the large number is directly connected to the computational cost. On the other hand, watching connections are made up of relatively small number of links and focus on narrow scope of contents. Therefore, it doesn’t require computing a large number of factors. In addition, whenever a user updates his bookmark history or changes

his group-based social connections, the system is able to rebuild small sub-matrix only for the corresponding user and calculate the matrix factorization right away. In a situation where online users' participation is getting increased, the large-scale SVD has a serious scalability problem [23, 158, 195].

In the watching network-based recommendations based on the other two strategies to select social connections, the direct watching connections were altered to the corresponding social connections (i.e. watching and watched connections or direct and indirect watching connections) and the recommendations were produced in the same way like the above.

Lastly, as an extra option to consider, I also propose hybrid recommendations by combining target users' top N anonymous peers with their direct watching and watched connections (i.e. Hybrid, Hybrid_CW, HybridSVD, and HybridSVD_CW). In here, the variable N is determined as many anonymous peers as each target user's watching and watched connections are. According the experimental evaluation to decide the optimal setting of the recommendations, the CF recommendations using the K-nearest neighbors produced the best results when the considered anonymous peers were the equal number to the target users' watching network-based connections. A target user' top N anonymous peers were combined with his watching and watched connections. In particular, among various hybridization strategies, I used the mixed hybrid strategy [29]. Put differently, two sources of information – one source is target users' top N anonymous peers and another source is their watching network-based social connections – were mixed as one set, and then the mixed set was used as a foundation of the recommendations.

5.3.1.2 Content Similarity Weights

In the first part of this chapter, I found that in measuring information similarity among watching connections, metadata of items are equally critical as counting users' common bookmarks. Citeulike data consisting of social bookmarks represents what users are interested in, but not how much they are interested in the corresponding items or why the items interest them. In order to help users manage their favorite items more effectively and express their interests on the items in a more cognitive way, the Citeulike system enables the users to annotate free-text tags when they bookmark favorite items. However, the item-based similarities (such as Log-likelihood similarity, the number of co-bookmarks, and Jaccard coefficient) count just how many items they interested in but don't take the reasons for the interests into consideration. Therefore, the recommendations solely based on item-based similarity, it is hard to produce more elaborate and accurate suggestions. To enhance the recommendation accuracy and reflect users' cognitive understandings on the items and contents of items in recommendations, I utilize a content similarity weight.

In the above CF and social recommendations, at the final stage of recommendations, the final prediction probability of each candidate item (i.e. the value indicating how much the target user will like the candidate item) is computed by aggregating and averaging out the similarities of the peers who bookmarked the candidate item. Therefore, when some candidate items belong to the same set of peers, the probability values are exactly the same, even though the items differ in the contents and it is very likely that users added different tag sets. Using the content similarity weights, I try to add the different contents of items and social tags in the recommendations. The above CF and social recommendations are based on how much users'

like-minded anonymous peers or social connections liked the candidate items through the bookmark-based similarities. The recommendations using content similarity weight, along with to the like-minded anonymous peers or social connections' opinions, take into account how the topic of the items are matched with the target users' interests.

Once a list of candidate items were selected from CF recommendation and social recommendations, the system compares how the contents and social tags of these candidate items are similar to the target user's keyword vector-based and processed tag-based profiles. Although I already computed keyword vectors and processed tag-based vectors for the test about the information similarity (i.e. the test of the section 5.2), in this recommendation, users' data were split into a test set and training sets. Therefore, by excluding the test set of each test iteration, the target users' keyword-based and processed tag-based vectors were re-generated iteratively. Additionally, I generated the same kinds of vectors for each item, as well. Particularly, in processed tag vectors, I did not limit the scope of the social tags to the users' watching connections. I included all social tags annotate to the corresponding items, except target users' own tags once the item is one of their test set items.

The following equation denotes how the content similarity weights were applied in the recommendations.

$$CFCW_{u,i} = \frac{\sum_{v \in p_u} KNN_Prob_{u,v}}{V} \times \left(\left(\frac{k_{ui} - \mu_K}{\sigma_K} \right) + \left(\frac{t_{ui} - \mu_T}{\sigma_T} \right) \right) \quad \text{Eq. 25}$$

The above equation shows CF recommendations using content similarity weight (CFCW) of a candidate item i for a target user u . For every candidate item, the system calculated cosine similarity (k_{ui}) between user u 's keyword vector-based profile k_u and keyword vector-based profile of item i , k_i . It also calculated cosine similarity of (t_{ui}) of user u 's process tag-based

profile t_u and processed tag-based profile of item i , t_i . These two cosine similarities are combined as one content similarity weights, but in order to normalize the values, I applied the standard score using the mean μ and standard deviation σ of the corresponding similarity measure, respectively. This content similarity value is multiplied with the final prediction probability value of KNN recommendation (e.g., Jaccard similarity) for the item i . In here, v denotes one of the user u 's anonymous peers (p_u) and has the candidate item i in his bookmarks. V is the number of peers who have the candidate item i in their bookmarks. In the same way, the content similarity weights were applied to the resultant candidate items selected through matrix factorization algorithm as the following.

$$\text{CFSVD_CW}_{u,i} = \frac{\sum_{v \in p_u} \text{SVD_Prob}_{u,v}}{V} \times \left(\left(\frac{k_{ui} - \mu_K}{\sigma_K} \right) + \left(\frac{t_{ui} - \mu_T}{\sigma_T} \right) \right) \quad \text{Eq. 26}$$

5.3.1.3 Content Boosted Recommendations

The next content-boosted recommendations used hybrid recommendations fusing user preference with content information. The difference between this content-boosted recommendations and previous content similarity weight is when and how to apply the content information in recommendations. As explained, content similarity weights were applied at the final selection of recommended items. On the other hand, the content-boosted recommendations are to use users' content-based profile as their preferences instead of simple bookmark based ratings. This content-boosted approach is inspired by Melville, et al. [130]. The original content-boosted collaborative filtering (CBCF) was designed for sparse datasets where simple co-rating profiles can't produce sufficient number of peers. Briefly, the Melville's original algorithm links one vector consisting of actual user ratings with another vector consisting of item-to-item

similarities based on contents of the items. Therefore, even when two users didn't rate exactly the same items, if they rated the similar items in the similar manner, these two users could still be considered as 'peers' to each other. The Melville's CBCF approach was based on users' numeric ratings of items. Since this dataset doesn't have users' numeric ratings of items but only includes bookmarks, this study introduced a modified version of CBCF.

In the selection of peers, a target user's content profile was compared with other users' content profile in item-by-item basis. When two users bookmarked the same item, of course, it would be the perfect match, 1. For a target user's bookmarked item missing in his peer's bookmarks, among the peer users' bookmarks, the system picked the most similar item with a target user's bookmarked item. The similarity values of the most similar items was summed up and averaged out with the number of the target user's bookmarks (refer to equation 27).

$$CBSimilarity_{u,v} = \frac{\sum_{i=1}^n content_sim_{ij}}{n} \quad \text{Eq. 27}$$

u is a target user and v is his peer and the user u has totally n bookmarked items. j is user v 's the most similar items with every i . For instance, a target user u bookmarked talk #1 and #2 and his peer v bookmarked talk #1, #3 and #4. Because the talk #1 was bookmarked by both users, the content similarity of that item is the perfect match 1. However, the talk #2 was bookmarked by the user A but not by the user B. Then the system found that #3 is quite similar to #2 with the value of 0.9, and #4 is moderately similar with the value of 0.5. Then the content similarity between #2 and #4 was ignored and, instead, only the similarity with talk #3 was considered. The resultant similarity between users A and B ($S_{A,B}$) is $(1+0.9)/2$. During the final recommendation process to select recommendable items, among the candidate items, I applied the content similarity weights of the candidates in the same way with the section 5.3.1.2. Put

differently, this content similarity value of a corresponding item is multiplied to its final prediction probability value calculated by the content-boosted similarity.

5.3.2 Experimental Evaluation of Recommendations

For the evaluation of the proposed recommendation algorithms, 3223 users who were watching other user(s) and have at least one bookmarked item were selected as target users. Their bookmarks were randomly split into 10 almost equal-sized subsets so as to execute 10 cross validation. This strategy split items into N sets randomly. Hence, when target users have less than 10 bookmarks, I split the set of bookmarks into N sets consisting of one bookmark. Of course, the N is less than 10 for this case. For each iteration, one of the 10 subsets is used as a test set and the remaining 9 sets are used as a training set. Recommendations will be assessed by how many recommended items are actually in the test set as hits. This process is repeated with a different test set for 10 times and the hit numbers will be averaged out. I called this dataset as a ‘whole dataset.’ In addition, so as to determine various algorithmic parameters of the K-nearest neighbors and matrix factorization approaches, I needed a small random set. Among the 10 subsets, I used the first one set for this purpose and called this set as a ‘probe dataset.’

As the evaluation criteria, because Citeulike dataset doesn’t contain numeric ratings, traditional evaluation methods for information retrieval were used; precision and recall. The precision aims to measure how precise the recommendations are, and the recall aims to measure how complete the recommendations are. The accuracy of the recommendations will be computed according to the top 10, 5 and 2 recommendations ($N = 10, 5, \& 2$). The recommendations are

usually displayed in a ranked list. Users expected that items in a higher rank would be more important than the other items in lower ranks. Therefore, it is critical to place correct recommendations in a higher rank and the position of correct recommendations is a critical evaluation criterion of recommendations.

5.3.3 Social Network-based Recommendations using Various Settings

Before we start examining watching relation-based recommendations, it is required to find a better way to compute the similarity of the K-Nearest Neighbor algorithm and to optimize matrix factorization settings. In K-Nearest Neighbor algorithm, we need to determine a good similarity measure and an optimal value of K (i.e. how many like-minded anonymous peers or social peers will be considered). In addition, the optimal setting of matrix factorization is important criteria to affect the recommendation quality. Usually the good similarity measures of KNN and optimal settings of matrix factorization are determined by cross validation. However, due to the large dataset size, if I compute and compare the various algorithmic settings using the whole dataset, it is computationally too expensive. For this cross validation, thus, a smaller ‘probe dataset’ was used. The first of the users’ tenth dataset was withdrawn as a test set and using nine of the tenth datasets as training data, the recommendations were generated. Then by computing whether the items in the withdrawn test set were recommended or not, the F1 values were computed. The choice of the best similarity measure was determined by the highest F1 values in three ranks – top 10, top 5 and top 2. Then, once the good setting for each recommendation is decided, I will regenerate the recommendations with a bigger ‘whole dataset’ and compare the quality with other approaches. The more detailed explanations about the test

settings and results are described in the Appendix A.1. By following the best results, the best similarity measures of K-Nearest Neighbors and the optimal settings of the matrix factorization are as the following Table 6.

Table 6. The Optimal Setting of Each Recommendation Approach

Kinds of Recommendations		Optimal Settings
KNN-based	CF	Jaccard Coefficient with the same number of peers as much as the target users' social connections exist
	Watch	Log-likelihood Similarity with all social connections
	ReciWatch	Jaccard Coefficient with the top 50 similar social connections
	1hopWatch	Jaccard Coefficient with the top 20 similar social connections
Matrix Factorization-based	CFSVD	50 factors with $\lambda=0.15$
	WatchSVD	24 factors with $\lambda=0.15$
	ReciWatchSVD	20 factors with $\lambda=0.15$
	1hopWatchSVD	24 factors with $\lambda=0.15$

The next analysis is to check whether adding metadata information of items can enhance the recommendation quality or not. Table 7 and Table 8 display the precisions and recalls of top N results including the statistical significance. The results showed that even though some differences are not statistically significant, adding content property of items to the recommendations are consistently helpful to increase the accuracy and completeness of the recommendations without any exception.

Table 7. Differences of Recommendation Precision Depending on Metadata Properties of Items

	Top10	Top5	Top2	
CF	1.89%	2.18%	2.81%	t = -2.18, p = .029 for top 10; t = -2.66, p = .008 for top 5; t = -2.80, p = .005 for top 2
CF_CW	2.67%	3.53%	5.14%	
CFSVD	0.88%	1.13%	1.31%	
CFSVD_CW	1.93%*	2.53%*	3.10%*	t = -8.47, p < .001 for top 10; t = -7.55, p < .001 for top 5; t = -6.27, p < .001 for top 2
Watch	1.32%	1.78%	2.67%	t = -5.55, p < .001 for top 10; t = -6.27, p < .001 for top 5; t = -6.13, p < .001 for top 2
Watch_CW	2.30%*	3.44%*	5.57%*	
WatchSVD	1.00%	1.15%	1.39%	
WatchSVD_CW	2.50%*	3.11%*	3.98%*	t = -11.21, p < .001 for top 10; t = -10.87, p < .001 for top 5; t = -9.50, p < .001 for top 2
ReciWatch	1.32%	1.84%	2.84%	t = -7.68, p < .001 for top 10; t = -7.84, p < .001 for top 5; t = -6.64, p < .001 for top 2
ReciWatch_CW	2.46%*	3.60%*	5.62%*	
ReciWatchSVD	1.20%	1.37%	1.65%	
ReciWatchSVD_CW	2.41%*	3.05%*	3.79%*	t = -8.74, p < .001 for top 10; t = -8.98, p < .001 for top 5; t = -7.86, p < .001 for top 2
1hop	0.86%	1.10%	1.63%	t = -12.21, p < .001 for top 10; t = -12.73, p < .001 for top 5; t = -10.77, p < .001 for top 2
1hop_CW	2.07%*	2.94%*	4.54%*	
1hopWatchSVD	0.99%	1.13%	1.37%	
1hopWatchSVD_CW	2.16%*	2.74%*	3.43%*	t = -9.67, p < .001 for top 10; t = -9.72, p < .001 for top 5; t = -8.27, p < .001 for top 2
Hybrid	1.08%	1.33%	1.54%	t = -3.03, p = .002 for top 10; t = -3.15, p = .002 for top 5; t = -2.87, p = .004 for top 2
Hybrid_CW	2.44%	3.44%	4.87%	
HybridSVD	1.26%	1.48%	1.79%	
HybridSVD_CW	2.56%*	3.22%*	4.05%*	t = -7.97, p < .001 for top 10; t = -7.65, p < .001 for top 5; t = -6.19, p < .001 for top 2

Table 8. Differences of Recommendation Recall Depending on Metadata Properties of Items

	Top10	Top5	Top2	
CF	1.74%	0.97%	0.44%	t = -1.22, p = .224 for top 10; t = -1.53, p = .126 for top 5; t = -1.82, p = .070 for top 2
CF_CW	2.45%	1.68%	1.12%	
CFSVD	1.88%	1.19%	0.61%	
CFSVD_CW	2.93%*	2.05%*	1.15%*	t = -3.40, p < .001 for top 10; t = -3.42, p < .001 for top 5; t = -2.84, p < .001 for top 2
Watch	1.02%	0.72%	0.49%	t = -3.59, p < .001 for top 10; t = -3.89, p < .001 for top 5; t = -3.34, p < .001 for top 2
Watch_CW	1.80%*	1.48%*	0.98%*	
WatchSVD	1.14%	0.68%	0.37%	
WatchSVD_CW	3.12%*	1.98%*	1.00%*	t = -7.49, p < .001 for top 10; t = -6.64, p < .001 for top 5; t = -4.49, p < .001 for top 2
ReciWatch	0.96%	0.70%	0.48%	t = -4.33, p < .001 for top 10; t = -4.34, p < .001 for top 5; t = -3.42, p < .001 for top 2
ReciWatch_CW	1.73%*	1.40%*	0.90%*	

ReciWatchSVD	1.10%	0.69%	0.40%	t = -6.99, p < .001 for top 10; t = -6.42, p < .001 for top 5; t = -4.22, p < .001 for top 2
ReciWatchSVD_CW	2.61%*	1.82%*	0.96%*	
1hop	0.84%	0.56%	0.37%	t = -7.49, p < .001 for top 10; t = -7.04, p < .001 for top 5; t = -5.41, p < .001 for top 2
1hop_CW	1.89%*	1.44%*	0.98%*	
1hopWatchSVD	1.05%	0.60%	0.32%	t = -6.55, p < .001 for top 10; t = -6.34, p < .001 for top 5; t = -4.30, p < .001 for top 2
1hopWatchSVD_CW	2.44%*	1.71%*	0.86%*	
Hybrid	0.98%	0.83%	0.39%	t = -1.03, p = .299 for top 10; t = -.78, p = .432 for top 5; t = -1.03, p = .303 for top 2
Hybrid_CW	1.46%	1.16%	0.75%	
HybridSVD	1.16%	0.72%	0.43%	t = -5.78, p < .001 for top 10; t = -5.22, p < .001 for top 5; t = -3.22, p < .001 for top 2
HybridSVD_CW	2.77%*	1.93%*	0.98%*	

(* indicates that the differences are statistically significant)

The next analysis is to determine better algorithm to compute the recommendation probability between K-Nearest Neighbor and matrix factorization algorithms within each approach using the same kinds of peers. From this comparison, the whole dataset was used in the recommendations. Since the approaches with content similarity weights produced generally better results, the comparison was made on the center of those approaches. Table 9 and Table 10 show the results of precisions and recalls including the statistical significance levels, respectively. Unlike the distinct pattern of differences on the recommendation quality depending on the absence and presence of the content similarity weights, there is no clear winner between the K-Nearest Neighbor and matrix factorization. The differences were not significant either. The patterns of differences on precision and recall are also contrasting. In terms of precision, recommendations using the K-Nearest Neighbor were more accurate, but in terms of recall, recommendations using matrix factorization were more complete. In order to decide a better method to compute the prediction probability for each kind of recommendation approach, I calculated the extra F1 measures as shown on Table 11. The differences of F1 between K-Nearest Neighbors and matrix factorization are mostly insignificant and no distinct patterns of differences. Due to the insignificance of the differences, thereof, I chose the approach yielding

the higher F1 value in the highest rank (i.e. top 2). According to the choice of better recommendation algorithms for each recommendation approach, the originally proposed design of recommendation approaches (i.e. Figure 16) can be densely summarized as Figure 17. In the following analyses, the chosen algorithms will be used for the comparison.

		Bookmark-based / Metadata-based
Selection of Peers	Anonymous Peers	• CFCW
	Direct Watching	• Watch_CW • CB_Watch
	Direct Watching + Watched	• ReciWatch_CW
	Direct Watching + Indirect Watching	• 1hopWatch_CW
	Anonymous Peers + Direct Watching & Watched	• HybridSVD_CW

Figure 17. Revised Design of Watching Network-based Recommendation Algorithms

Table 9. Differences of Recommendation Precision between Different Methods to Compute Prediction Probability

	Top10	Top5	Top2	
CF_CW	2.67%	3.53%	5.14%	t = 2.90, p = .004 for top 10; t = 2.62, p = .009 for top 5; t = 3.30, p = .001 for top 2
CFSVD_CW	1.93%	2.53%	3.10%	
Watch_CW	2.30%	3.44%	5.57%	
WatchSVD_CW	2.50%	3.11%	3.98%	t = -.96, p = .337 for top 10; t = 1.10, p = .269 for top 5; t = 3.47, p = .001 for top 2
ReciWatch_CW	2.46%	3.60%	5.62%*	t = .29, p = .768 for top 10; t = 2.17, p = .030 for top 5; t = 4.63, p < .001 for top 2
ReciWatchSVD_CW	2.41%	3.05%	3.79%	
1hopWatch_CW	2.07%	2.94%	4.54%	t = -.55, p = .580 for top 10; t = .90, p = .364 for top 5; t = 3.14, p = .002 for top 2
1hopWatchSVD_CW	2.16%	2.74%	3.43%	
Hybrid_CW	2.44%	3.44%	4.87%	t = -.54, p = .589 for top 10; t = .65, p = .510 for top 5; t = 1.53, p = .126 for top 2
HybridSVD_CW	2.56%	3.22%	4.05%	

Table 10. Differences of Recommendation Recall between Different Methods to Compute Prediction Probability

	Top10	Top5	Top2	
CF_CW	2.45%	1.68%	1.12%	t = -.86, p = .387 for top 10; t = -.78, p = .434 for top 5; t = -.08, p = .930 for top 2
CFSVD_CW	2.93%	2.05%	1.15%	
Watch_CW	1.80%	1.48%	0.98%	t = -3.28, p = .001 for top 10; t = -1.63, p = .101 for top 5; t = -.08, p = .936 for top 2
WatchSVD_CW	3.12%	1.98%	1.00%	
ReciWatch_CW	1.73%	1.40%	0.90%	t = -3.211, p = .001 for top 10; t = -1.80, p = .071 for top 5; t = -.36, p = .718 for top 2
ReciWatchSVD_CW	2.61%	1.82%	0.96%	
1hopWatch_CW	1.89%	1.44%	0.98%	t = -2.08, p = .037 for top 10; t = -1.21, p = .224 for top 5; t = .71, p = .474 for top 2
1hopWatchSVD_CW	2.44%	1.71%	0.86%	
Hybrid_CW	1.46%	1.16%	0.75%	t = -3.67, p < .001 for top 10; t = -2.46, p = .014 for top 5; t = -1.05, p = .291 for top 2
HybridSVD_CW	2.77%	1.93%	0.98%	

Table 11. Differences of Recommendation F1 between Different Methods to Compute Prediction Probability

	Top10	Top5	Top2	
CF_CW	1.68%	1.54%	1.28%	t = .51, p = .608 for top 10; t = -.05, p = .956 for top 5; t = .19, p = .843 for top 2
CFSVD_CW	1.57%	1.56%	1.22%	
watch_cw	1.29%	1.37%	1.26%	t = -3.37, p < .001 for top 10; t = -1.46, p = .144 for top 5; t = .38, p = .701 for top 2
watchSVD_CW	1.81%	1.63%	1.19%	
reciWatch_CW	1.36%	1.40%	1.23%	t = -2.26, p = .023 for top 10; t = -.74, p = .457 for top 5; t = .61, p = .538 for top 2
ReciWatchSVD_CW	1.63%	1.50%	1.13%	
1hop_CW	1.26%	1.27%	1.18%	t = -2.08, p = .037 for top 10; t = -.86, p = .385 for top 5; t = 1.13, p = .256 for top 2
1hop_watchSVD_CW	1.48%	1.37%	1.01%	
Hybrid_CW	1.30%	1.28%	1.05%	t = -2.80, p = .005 for top 10; t = -1.69, p = .090 for top 5; t = -.50, p = .614 for top 2
HybridSVD_CW	1.73%	1.59%	1.15%	

5.3.4 Comparison of Various Recommendation Approaches

In previous section, the good way to compute the prediction probability was found for each recommendation approach along with the optimal algorithmic settings. Now the quality of

the watching network-based recommendations will be compared with the other recommendations including traditional CF recommendations. Figure 18 displays the comparison result of both precision and recall. According to the One-way ANOVA test, the differences of precision were significant but the differences of the recall were not, except the top 10 recall ($F = 20.10, p < .001$ for top 10 precision; $F = 22.92, p < .001$ for top 5 precision; $F = 25.95, p < .001$ for top 2 precision; $F = 4.18, p = .001$ for top 10 recall; $F = 1.71, p = .127$ for top 5 recall; $F = .63, p = .675$ for top 2 recall).

I checked the patterns of the differences through the pairwise test. First, I compared two approaches utilizing different methods to incorporate metadata information of items in recommendations – content similarity weight and content boosted recommendations. The approach adding content similarity weight outperformed the content boosted recommendation in precision. In the results based on recall, the differences were not significant but the quality of content boosted recommendations dropped along with the increase of the rank. Therefore, rather than taking into account the metadata properties of items in the step to compute the user-to-user similarity, it is more effective to apply the information at the final step to choose the recommended items.

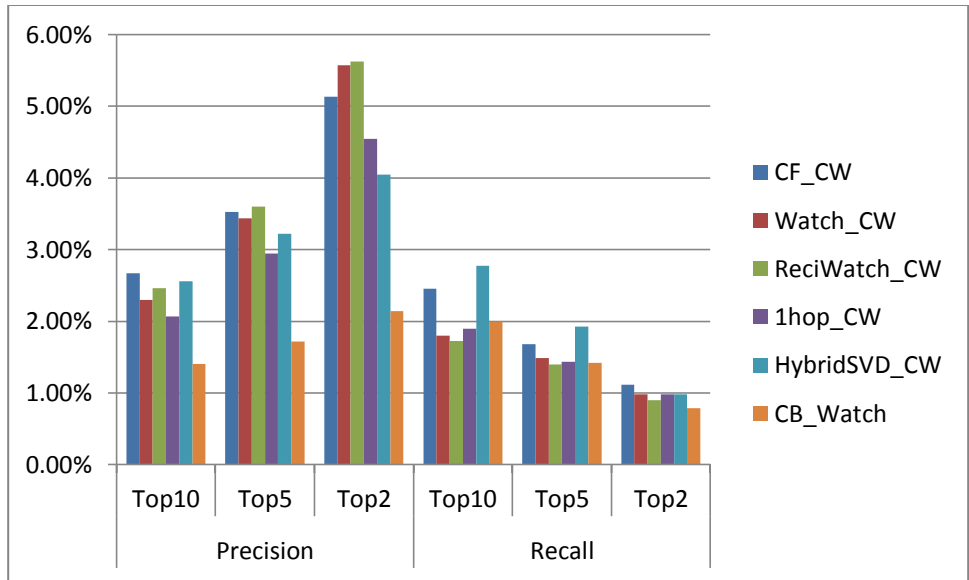


Figure 18. Differences of Recommendation Quality depending on the Kinds of Recommendation Approaches

The next analysis was to compare the recommendations based various scopes of watching network, so as to find out the best scope of watching network which can produce better social recommendations. As shown, this study considered three varying scopes of watching network – direct watching, direct watching and watched and direct and 1hop distance watching connections. The differences were not significant in precision. Despite of the insignificance, in higher ranks (i.e. top 5 and top 2), the quality of the recommendations based on watching and watched connections was the best among all proposed recommendations. The approach based on users’ direct and 1hop distanced watching connections was the worst. However, the results of the recall were not the same. The recall of the recommendations based on watching and watched connections were the poorest in all ranks and there is no consistently better approach among the three kinds of watching network-based recommendations. Therefore in the test of H5.5.1 which

aims to compare the direct watching connections with the connections combining users' direct and indirect social connections as viable information source, the hypothesis is rejected for both precision and recall. In the test of H5.5.2 for the comparison between the direct watching connections with both direct watching and watched connections, due to the statistical insignificance, in spite of the better performance of the latter social connection kind (i.e. both direct watching and watched connections), the hypothesis is also rejected. As the relevant hypothesis, the H5.5 is also rejected. All scopes of watching network considered in this chapter have equivalent values as information sources.

Then did the watching network-based recommendation perform better than traditional CF recommendations, which is a baseline of this comparison? According to the pairwise test, the quality of CF recommendations was not significantly different with watching network-based recommendations. That means that watching network-based recommendations didn't performed better than CF recommendations, but can produce the suggestions as good as the CF recommendations produce. Therefore, the relevant hypothesis H5.4 is rejected. If the target users' anonymous peers and their watching network-based connections are equally good information source, when we combine these two groups of peers together, can the recommendations based on the combined peers generate better recommendations? As explained, I generated hybrid recommendations by taking into account the preferences of both anonymous peers and watching connections. In the comparison centered on precision, in the lowest rank, it performed the second best following CF recommendations. Nonetheless, as the rank is getting higher, the quality dropped. Therefore, the relevant hypothesis H5.6 is partially rejected when we just count the results of precision. In terms of precision, there is no value to combine different

peer groups. On the other hand, the comparisons based on recall demonstrated different patterns. In the lower ranks (i.e. top 10 and top 5), the recall of the hybrid recommendations performed significantly better than the other recommendations even though, in the highest rank, the recall of the hybrid recommendations didn't show significant improvement. When there are some enhancements regarding the completeness of the recommendations, combining anonymous peers and watching network could be meaningful. By following the results, the relevant hypothesis H5.6 is partially accepted when solely counting the results of recall.

5.3.5 Watching Network-based Recommendation for Cold-start Users

The descriptive statistics of this Citeulike data source (introduced in the section 5.1) explained that users participating in the watching network are more familiar with the Citeulike system and use the system more often than the other users who don't participate in the network. In the former user group (i.e. members of the Citeulike watching network), however, there are still 316 users (9.80% of our target users) having less than 5 bookmarks. Usually, in recommendations, users having this amount of bookmarks or ratings are hard to receive reasonably good recommendations since the numbers of bookmarks are insufficient to present what they like properly. Therefore, these users are called as cold-start users. One weakness of the CF recommendations is this cold-start user problem. Since the recommendations rely on the overlapped interests or tastes among users, when users have insufficient information about interests or tastes, it is difficult to guess presumably favorable items. As explained in the chapter 2, some researchers suggest that social network-based recommendations could be a good solution to solve this cold start user problem. Hence, in this subsection, I will test whether the

best recommendation approach differs according to target users' bookmark numbers and whether watching network-based recommendation is a good option for cold-start users. For this test, I classified users into three clusters according to their number of bookmarks – cold-start users ($n < 5$), user having medium level of bookmarks ($5 \leq n < 200$) and users having large number of bookmarks ($n \geq 200$). There are 316 users, 2,080 users and 827 users in each cluster, respectively. I evaluated which recommendation approach is the best for each cluster of users.

Figure 19 is the result of precision and Figure 20 is the result of recall. In order to find out the statistical significance of the difference, I executed Two-way ANOVA test. The test demonstrated that the differences on the precision and recall according to different recommendation approaches were statistically significant among the user clusters ($F = 3.32, p < .001$ for top 10 precision; $F = 2.83, p = .002$ for top 5 precision; $F = 3.00, p < .001$ for top 2 precision; $F = 5.55, p < .001$ for top 10 recall; $F = 6.40, p < .001$ for top 5 recall; $F = 8.64, p < .001$ for top 2 recall). As the next analysis, I find the patterns of difference on the various recommendation approaches among three clusters of users.

First, for users having large information collection, the differences of the various recommendation approaches are significant ($F = 9.02, p < .001$ for top 10 precision; $F = 9.52, p < .001$ for top 5 precision; $F = 11.37, p < .001$ for top 2 precision; $F = 6.83, p < .001$ for top 10 recall; $F = 7.21, p < .001$ for top 5 recall; $F = 8.74, p < .001$ for top 2 recall). For this cluster of users, CF recommendation is always the best option in terms of both precision and recall. However, I failed to find any watching network-based recommendation which is working as consistently well as the CF recommendation is.

Second, in case of the cluster of users having medium level of information collections, the precision differences of the various recommendation approaches are significant but the differences of recall are not significant ($F = 5.11, p < .001$ for top 10 precision; $F = 5.53, p < .001$ for top 5 precision; $F = 7.73, p < .001$ for top 2 precision; $F = 2.03, p = .070$ for top 10 recall; $F = 1.06, p = .378$ for top 5 recall; $F = 1.82, p = .105$ for top 2 recall). The both precision and recall demonstrated that watching network-based recommendation was the best option for this cluster of users having medium level of information collection. In terms of precision, the watching network-based recommendations performed significantly better than the other recommendations including the CF recommendations. Even though the differences were insignificant, the watching network-based recommendations yielded the highest recall values than the others, as well.

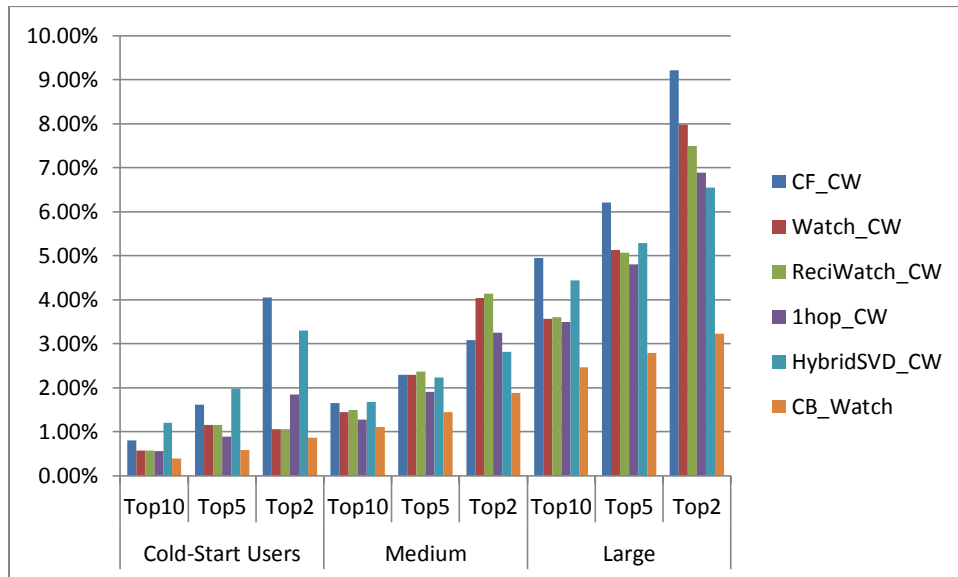


Figure 19. Differences of Precision Depending on Users' Number of Bookmarks

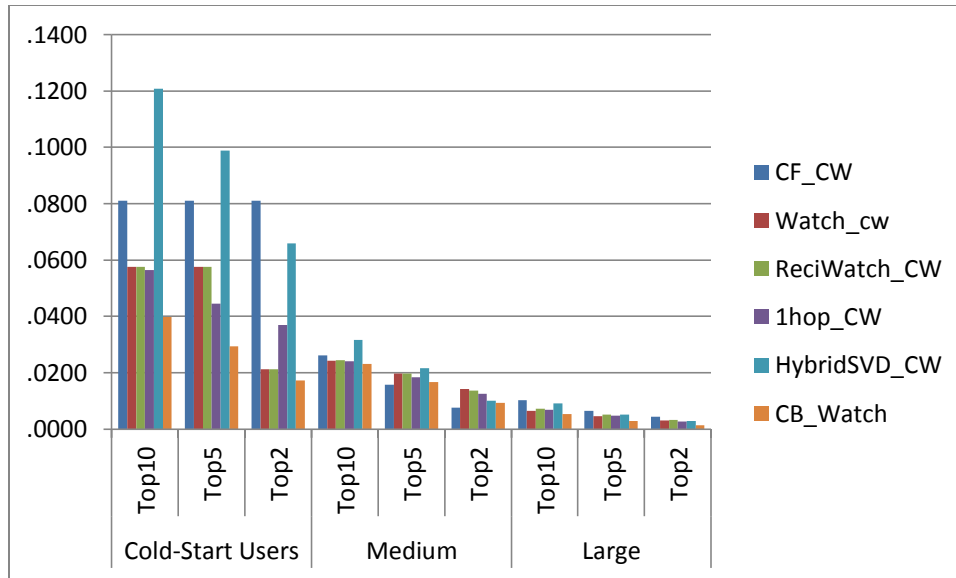


Figure 20. Differences of Precision Depending on Users' Number of Bookmarks

Lastly, for the cold-start users, the differences of the various recommendation approaches are not significant for both precision and recall ($F = 2.06$, $p = .068$ for top 10 precision; $F = 2.00$, $p = .076$ for top 5 precision; $F = 2.06$, $p = .068$ for top 2 precision; $F = 2.06$, $p = .068$ for top 10 recall; $F = 2.00$, $p = .076$ for top 5 recall; $F = 2.06$, $p = .068$ for top 2 recall). Although the differences are insignificant, the results displayed on Figure 19 and Figure 20 indicated that hybrid recommendations and CF recommendations are good options to cold-start users. In lower ranks, the hybrid recommendations performed better than the other approaches in both precision and recall. However, in the highest rank, the CF recommendations were the best approach. Therefore, utilizing cold-start users' social connections as a part of their recommendations is helpful to increase the quality of the suggestions. However, the CF recommendation is still the best approach that generates the most accurate suggestions.

5.4 CONCLUSION

In this chapter, I examined the feasibility of watching network as a useful information source through 1) the examination of shared interests among watching connections and 2) the empirical evaluation of watching network-based recommendations.

In this study based on Citeulike data set, as the first part of this chapter, the changes of information similarities according to the social distances of user pairs were explored. The direct watching connections shared the most similar information. In particular, they favored not only the exact same items, but also the items containing similar metadata, such as papers consisting of a set of similar keywords, written by similar authors and annotated by similar social tags. The users' similarity decreased along with the increase of their social distances. Since I found the shared interests existing on the watching network, I tested how the social connections are valuable information sources in more details by utilizing them as a part of the personalized recommendations.

The second part of this chapter of which the focus was the evaluation of the recommendations based on watching network showed that watching network-based recommendations were comparable in the quality to traditional CF recommendations. In terms of how the suggestions are accurate and complete, the watching network-based recommendations have equivalent quality to the CF recommendations. In addition, when combining users' watching connections with their anonymous peers, the recommendations significantly improved at least regarding the evaluation based on recall. I interpreted these results to mean that Citeulike users are connected with their watching partners via their perceived utility of the partners'

information. Conclusively, users' watching connections are useful information sources and a feasible foundation for their personalized recommendations.

Lastly, according to the descriptive statistics, compared with the other users who are outside of the watching network, the users engaging in the watching network are more familiar with the Citeulike system and more active in using the system. However, there are still cold-start users whose bookmark collections are insufficient to represent what they are interested in and further, due to the insufficient preferences, it is hard for them to receive reasonably good recommendations. For those cold-start users, including users' social connections as a part of recommendations enhance the quality but it is not enough to use the social connections alone in the recommendations. It is more effective to take into account the watching connections as a supplementary component of the CF recommendations for cold-start users.

6 GROUP MEMBERSHIP-BASED RECOMMENDATIONS

This chapter aims to recommend favorable items to individual users based on their online group memberships. Group activities are usually centered on one recognizable topic, to follow a community of interest or practice. The social dynamics in groups also target to distribute or contribute topic-relevant information. When a user participates in a group, this membership shows his interests or relevancy to the corresponding topic of the group. The first part of this chapter substantiates this suggestion through Citeulike group membership data. In particular, as I did for watching networks, information similarities of members of the same group will be compared in various ways. Regarding users' group-based sociality, Citeulike data provides us two references to consider – co-members of the same group and groups per se. Citeulike allows for users not only to manage their individual personal collections but also to participate in their group's communal library. Therefore, this part assesses whether either of them is really a good information source and which reference is a better information source. The objective of the first part is to check whether online group activities, which are targeted at sharing topic-relevant information, are really good indicators of users' interests on the relevant topic and plausible sources to provide users useful information.

Most of recent studies under the name of 'group recommendations', however, aggregate tastes of a group of people into one set, and suggest recommendations for that group of people,

not an individual user. The second part of this chapter attempts to recommend items to individual users using their group membership information. The quality of various group-based recommendations is compared with traditional CF approach. In addition, there are some users having more than one kind of social networks on Citeulike; both watching connections and co-members. In this section, the quality differences of SN recommendations generated by each social network and both social networks will also be compared.

6.1 DATA SOURCE FOR GROUP MEMBERSHIP: CITEULIKE

The data source of group membership is the same Citeulike dataset used for watching network, which was made by the system administrators on May 15, 2011. The dataset contains the list of whole groups and the members of the Citeulike, as of the time when this dataset was made. In Citeulike, when users find interesting information, they are able to send the information to their group spaces, as well as to save the information in their personal repositories (refer to the Table 28). Hence, the dataset also provides both bookmark histories of all group members and bookmark collections of each group which is jointly contributed by members of the group. When a member adds information to their group spaces, all other members can see the newly added information and the annotated tags, on the page titled ‘group library’. To be short, the Citeulike data set regarding group memberships contains group ids, all members of each group, group library, and group members’ personal repositories. The group libraries and members’ repositories are constituted of article ids, social tags and time stamps of bookmarks. Because metadata of items (i.e. titles, authors, publication journals, abstracts and publication years) was

missing in the data set, I collected the metadata separately. Hence, to summary both group libraries and members' personal repositories consist of bookmarks of items along with the time-stamps of bookmarks and the metadata such as, titles, authors, publication years, publication journals/conferences, abstracts and social tags of the bookmarked items.

In this dissertation, I excluded group members having no items in their personal collections and groups having a single member. In case of groups having no bookmark on their group library, once the members have bookmark(s) in their personal collections, the groups and the members were included in the study. As the result, I found 1,870 groups and 8,009 members which are 8.5% of the whole Citeulike users. The following Table 12 is the descriptive statistics. Figure 21 and Figure 22 show the distributions of groups and group members according to the number of bookmarks. These groups and the members have about 1.4 million bookmarks being made up of 1,044,393 articles. 22.9% of groups and 35.8% of members have at most 5 bookmarks. In addition, 77.4% of the articles ($n = 808,797$) were bookmarked by only one user. The bookmark sparsity of this Citeulike data for group memberships is 0.9998. As you can see on Table 12, groups have larger collection of information than the members. It seems natural because group libraries are collaboratively composed by members.

Table 12. Descriptive Statistics about the Dataset of Citeulike Group Memberships

Groups & the Memberships	No. of Groups	1,870
	No. of Users having at least one group membership	8,009
	No. of Group Membership	11,863
	Avg. No. of Memberships per Group	6.34 ($\sigma = 16.3$)
	Avg. No. of Memberships per Member	1.48 ($\sigma = 1.7$)
Information Collections	No. of Distinct Articles in Group Library or Members' Repositories	1,044,393
	No. of Distinct Articles in Group Library	260,567
	No. of Distinct Articles in Members' Repositories	982,298
	No. of Bookmarks in Group Library or Members' Repositories	1,446,607
	Total No. of Bookmarks in Group Library	324,486
	Total No. of Bookmarks in Members' repositories	1,122,121
	Avg. No. of Bookmarks Per Group	173.52 ($\sigma = 784.1$)
	Avg. No. of Bookmarks Per Member	140.11 ($\sigma = 1133.8$)
	Avg. No. of Tags Per Group	704.98 ($\sigma = 7421.7$)
	Avg. No. of Tags Per Member	533.65 ($\sigma = 6784.5$)
	Avg. No. of Bookmarks Per Articles	1.39 ($\sigma = 1.3$)
	Avg. No. of Tags Per Article	5.51 ($\sigma = 8.8$)

Figure 23 is the distribution of groups according to the number of members. The groups have 6.3 members on average and the number of members per group ranges from two members to 588 members. Groups usually have about 174 bookmarks in the group libraries and about 22.9% of groups have at most five items in the library ($n = 428$). Figure 24 shows the distribution of users according to the number of their group memberships. Users are members of 1.5 groups, on average. The group members have usually about 140 bookmarks in their personal repositories and 21.8% of them have only one item ($n = 1,742$). Before addressing an in-depth study about information sharing patterns of group members, I will examine whether users have heterogeneous information usage patterns according their social status.

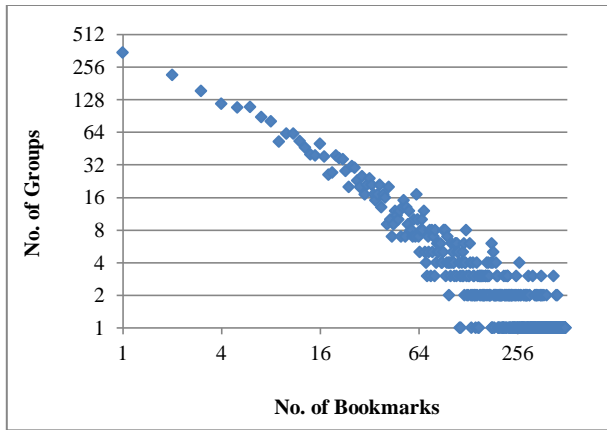


Figure 21. Distribution of Groups according to the Bookmark Number of their Libraries

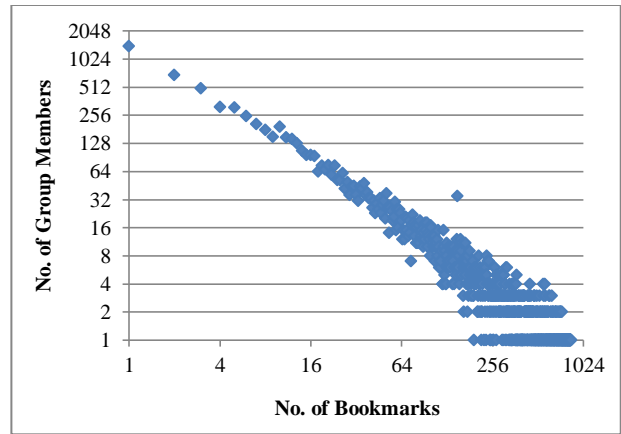


Figure 22. Distribution of Group Members according to the Bookmark Number of their Personal Repositories

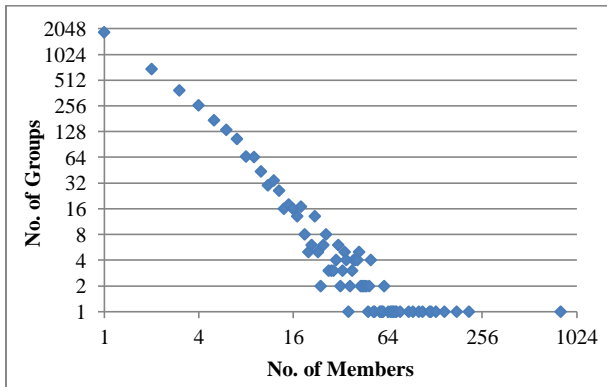


Figure 23. Distribution of Groups according to the Number of Members

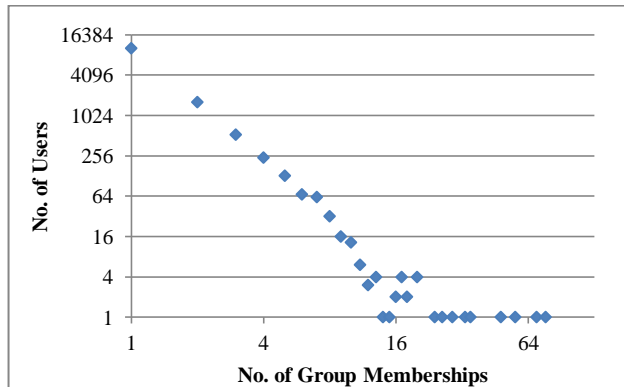


Figure 24. Distribution of Users according to the Number of Group Memberships

Information Usage Patterns of Users according to Their Social Status

In the descriptive statistics about Citeulike watching network presented at the chapter 5, we checked that users in watching relations tend to join the Citeulike system earlier, to use it often and for longer period of time than users not in watching relations. The users in watching relations also are richer users in terms of their bookmark collections than users outside of the watching relations. Then, what about users participating in groups and the others who are not a part of group activities? Is there any difference on the information usage patterns between users in watching relations and the other users in group memberships? Is there any notable and discrete pattern of users who are participating in both watching network and group membership?

First, information usage patterns of group members were compared with the others who in neither group membership or watching network. The users participating in groups joined the system earlier ($M = 900.6$ days ago from the time when this data set was made, $\sigma = 598.3$ days ago), had used it often ($M = 22.19$ times on different days, $\sigma = 63.4$ times on different days) and for longer period of time ($M = 312.6$ days, $\sigma = 486.2$ days) than the other users not engaging in any social network (they joined the system 801.4 days ago ($\sigma = 567.7$ days ago), had used it 2.6 times on different days ($\sigma = 6.7$ times) and for 52.6 days ($\sigma = 180.2$ days)) on average. The users in group memberships also possess more than 4 times larger bookmark collections ($M = 140.1$ bookmarks, $\sigma = 1133.8$) than the other users outside of Citeulike social networks ($M = 13.8$ bookmarks, $\sigma = 183.0$). Because all differences of the values are so obvious, I didn't run a separate statistical test for the differences. Hence, group members were more actively using the Citeulike, and it is likely that they knew the system functionalities better than users outside of the

group membership network. Figure 25 displays the bookmark distribution of group members contrasted with the distribution of users outside of group membership network.

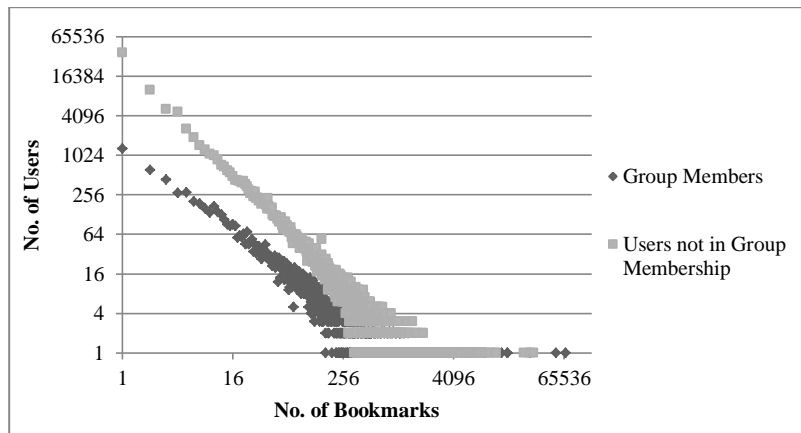


Figure 25. Bookmark distribution of users who involve in group membership contrasted with the one of users not involved

The results of the section 5.1 and the above results demonstrate that, users engaging in either watching network or group membership were more actively using Citeulike and more familiar with the system than the other users. Then is there any difference between users in the watching network and group members, in terms of their information usage pattern? In the comparison between group members and users in watching networks, 1,118 members were watching other users at the same time. These 1,118 users are involved in both social networks with their choice. If the users being watched by someone else are counted, the number will increase. However, the watched users didn't know the fact that they were a part of watching network. That is, the watching network is not their self-selected social network, and they became a part of the network by someone else's choice. That is the reason why they were excluded in target recipients of watching network-based recommendations, even though they were actively

used as a component of their watching partners' recommendations. Therefore, in this comparison, I compared three clusters of users – group members who didn't watch others (n = 6,891), users engaging in watching network (both watching users and watched users) and not participating groups (n = 8,865), and the others participating in groups and watching others (n = 1,118). Interestingly, group members joined the system later, had used it less often and for shorter period of time than the other two clusters of users, on average. The bookmark collections of group members were smaller than the other clusters of users. In addition, the users in both group membership and watching network the most actively used the system (joined the system the earliest, used it the most often for the longest period of time and possessed the largest bookmark collections) than the other two clusters of users. According to One-way ANOVA and the Schaffé pairwise test, these differences are significant in the 0.001 significance level (refer to Table 13). Figure 26 shows the bookmark distributions according to the clusters of users. In summary, the more users are active and familiar with the system, the more likely they participate in social networks. Additionally, the more users are active and familiar with the system, the more social networks they were engaged in. Actually, it is still unclear about the cause-and-effect relationship. Since there is no information about the time when each social link was made, we don't know whether users' engagements of various social networks make it possible for users to collect many objects or not and furthermore whether having social networks are helpful for users to acquire information yet. The users may collect larger collections of bookmarks, since they used the system for a longer period of time. However, it is clear that users in Citeulike social networks know how to use the system and more enjoyed using it than the other users. In particular, it seems that users associated with watching networks are more active than group

members. Therefore, in following analysis, along with the information similarity between group co-members, I execute the comparison of information between a pair of watching connections and a pair of group co-members.

Table 13. Summary of Differences according to Users' Clusters
 (* indicates the differences are significant)

	Time to join Citeulike	The frequency of system usage (by different date of bookmarks)	Period of time to use Citeulike	No. of Bookmarks
Group Members	838.3 days ago	107.03	236.4 days	14.03
Users in Watching Network	934.9 days ago	189.70	392.4 days	27.85
Users in both Networks	1284.5 days ago	343.99	782.9 days	72.47
	$F = 42.8^*$	$F = 694.7^*$	$F = 263.9^*$	$F = 482.0^*$

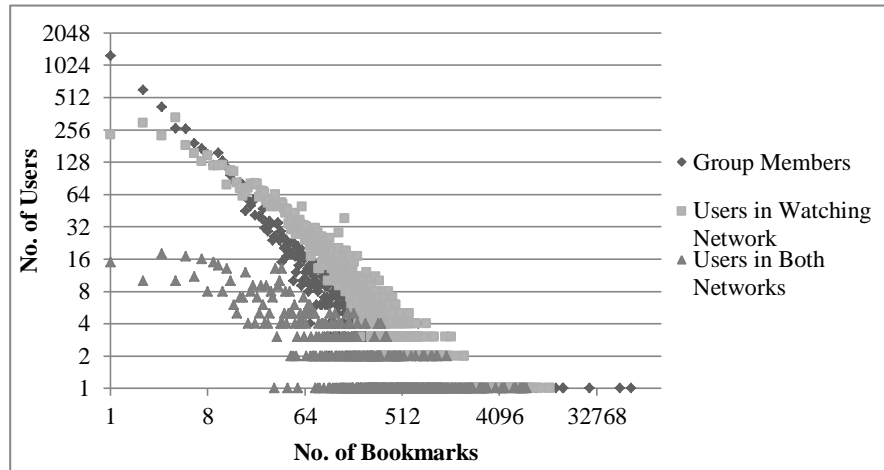


Figure 26. Bookmark distributions of three user clusters

Lastly, I tested whether the number of groups which a user participates in is correlated with the size of his bookmark collections. We can reckon that the more groups a user participates

in, the larger his bookmark collection is, since group activities mainly intend to acquire information relevant to a certain topic. The result of Pearson correlation test showed that this guess is right. The number of users' group memberships is significantly correlated with the size of their bookmark collections ($r = .08, p < .001$). As the next test, the correlation between the number of group members and the size of group library is tested. If group members collaborate to compose their group library, we can expect a positive correlation, as the result. According to Pearson correlation test, there is a significantly positive correlation between the number of group members and the size of group library ($r = .12, p < .001$). That is, the more members a group has, the larger the group's library is. The more detailed understanding about how group members contribute to their group library will be addressed in the section 6.2.1.

6.2 FEASIBILITY OF GROUP MEMBERS AS A USEFUL INFORMATION

SOURCE

In this section, we will explore how co-members of the same group share similar information and whether being a member of a group is beneficial to obtain useful information or not. Between group library and co-members' personal repositories, which one can group members more rely on in order to obtain useful information? Do group members refer to both group library and co-members' repositories? Compared with lurkers who don't contribute any item to their group library, do users who contribute a lot of items on their group library tend to share more similar items with co-members? In the situation where users express their interests in simpler way (such as unary rating, presence/absence of bookmark) than numeric ratings (such as

5 star or 10 star ratings) as Citeulike system does, what is a proper measurement to express shared interests among group members: counts of co-bookmarked items, titles consisting of similar sets of keywords, or co-annotated social tags? So as to answer these questions, I define the following hypotheses. The feasibility of group memberships as useful information source is a critical point we need to investigate before proceeding to the group membership-based recommendations, because the recommendations are basically based on the similarity of preferences between a target user and his reference peers. As ways to check how group memberships are viable to be useful information sources, the following hypotheses were tested.

H6.1. Information similarity between a pair of co-members is higher than the similarity between two users who are randomly paired.

H6.2 Information similarity between a member and his group is higher than the similarity between two co-members of the same group.

H6.3. The similarity between a user and his co-members of the same group is comparable to the similarity between him and his top N anonymous peers who are chosen by traditional CF recommendation technology.

As explained, the data source of this group membership has the same data structure of the Citeulike data used for watching network-based recommendations (refer to chapter 5), containing the same types of users preferences on information objects – bookmarking interesting items and annotating social tags to the items. The metadata of the information objects (i.e., scientific articles) is also the same – titles, abstracts, publication journal/conference names, authors and social tags. However, as described in the section 5.2.1.4, in the Citeulike data source, the abstracts and journal/conference names are missing for many articles. Therefore, based on users’

bookmarking history and metadata information available for the majority of objects, I computed item-based similarity measures and metadata-based similarity measures in the exact same kinds as I did for watching network (refer to the section 5.2.1). Figure 27 displays the eight kinds of information similarity measures to be explored in this part.

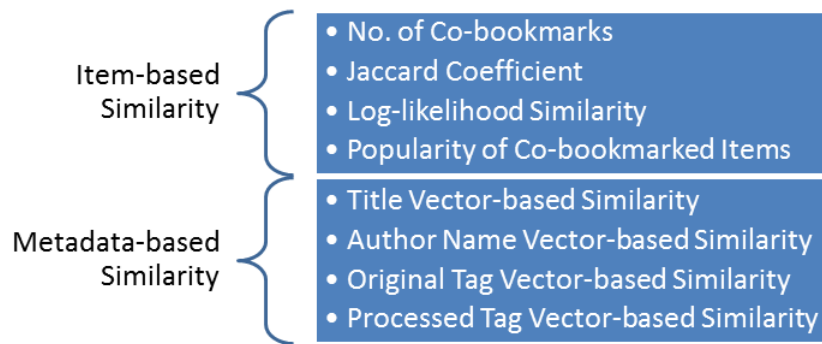


Figure 27. Summary of Similarity Measures

6.2.1 Members' Contributions on Group Library

Before examining information sharing patterns in online groups, we need to consider one important point. Citeulike provide communal space of every group, to which all members of that corresponding group can contribute. That is to say, it is very likely that a part of a group library is posted by a member. The items contributed by the member are literally his bookmarks. In the information comparison, thus, the corresponding member's contributions need to be excluded. Otherwise, members' contributed items will be accidentally counted as matches. In addition, if the contributed items are included as a part of candidate items, and recommendations are generated from the candidate list, it will turn out to be a target leak, which is an unintentional bug to provide ground truth in the training data. As a way to pick each user's own contributions

on his group library, I used the time stamp of each bookmark. When a group member finds an interesting item, he is able to post the item to his personal repository and his group library. Particularly, the Citeulike interface enables users to add their favorite items into both their personal repositories and their group libraries at the same time (refer to Figure 28). When one item is posted on both a member's personal repository and his group library with less than 60 second interval, we can assume that the item was posted by that member. If one member copied the same item from his group library, there would be amount of time interval larger than 60 seconds, because he may not refer to the updates of his group library all the time. Therefore, the interval of 60 seconds is reasonable time for one user to add interesting items to both personal repository and group library. Figure 29 shows the ratio of members' contributions on group libraries. Put differently, this graph shows that how much percent of group library was contributed by each member. If a member's contributions reached around 80% or 100%, this means that 80 or 100 percent of the group library was contributed by that corresponding member. Therefore, we can state that the members whose contributions on their group library are 100% totally dominate the libraries. They are the only one who composes the group libraries. I suggest that the members' contributions in terms of group library show the degree of control over their groups and their roles in the group. When a member contributes his group library a lot, he is prone to lead the group activities and may be an important member of that group.

On the other hand, Figure 30 shows members' contribution ratio in terms of their personal repository. In this graph, if one member's contribution ratio is 50%, half of his bookmarks were posted to that corresponding group's library at the same time. Put differently, his contributions on the group came from the half of his personal bookmark collection regardless

how much his items occupy the group library. When a user joined multiple groups, in this analysis, his contribution to each group was counted separately. I suggest that the contribution ratio to members' personal repository shows that users' perceived importance of their groups. If a user posted items on both his repository and group library in many times, it is likely that every time he finds an interesting item, he posted it not only on his personal collection but also his group library. Therefore, we can interpret this to mean that he has a definite desire to share interesting items with their group co-members.

Copy article: where would you like to file it?
Let us know where, and how, you want this filed.

Copy Article

Title: Adapting neighborhood and matrix factorization models for context aware recommendation
Published: 2010
Abstract: ▶ In this paper, we describe our solutions to the weekly recommendation track and social network track of tl recommendation trac...
Authors: Liu NN, Cao B, Zhao M, Yang Q

Post to: Your library Adaptive-Web cafesmu social_navigation Social Web **A**

(The article already exists in any libraries highlighted in **bold**. It will not be copied again and you will not be form without selecting a new library. Unselecting these libraries will not remove the article, but will remove t the available tags list below.)

Tags: **Show all tags** **B**

Priority:

Privacy: Don't let other people see that I've posted this article
Author: I am one of the authors of this article (This is only applied to articles in your own library, not groups.)

▶ **Notes:**

BibTeX Key:

Copy Article **C**

Figure 28. Citeulike interface enabling users to post favorite items to both their personal repositories and group libraries simultaneously; For instance, this user is posting an item to her personal repository (“your library”) and her two groups (“Adaptive-Web” and “Social Web”) according to the A area. Social tags annotated to this item by this user (B area) will be shown to both the repository and the group libraries.

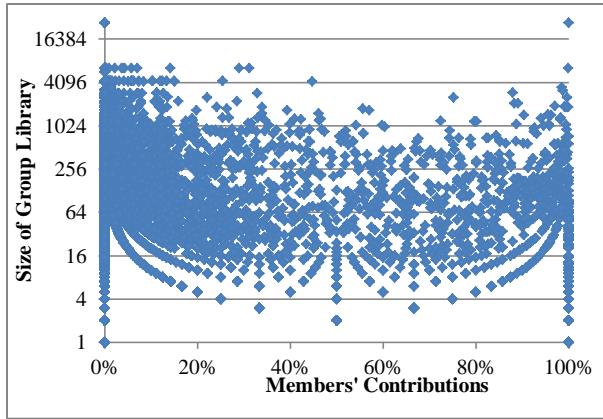


Figure 29. Ratio of Member’s Contributions in terms of the Size of Group Library

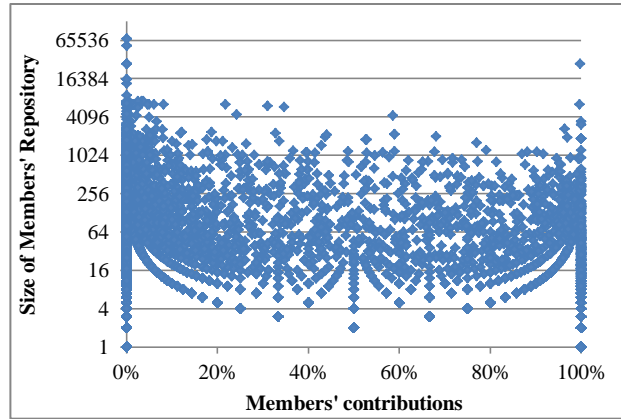


Figure 30. Ratio of Members’ Contributions in terms of the size of their Personal Repository

When combining these two kinds of contributions, the following Figure 31 shows the intertwined distributions. Even though the graph doesn’t represent any interpretable pattern, I suggest that there is a group of users whose group-based recommendations are in question. For instance, members in the area A on the Figure 31 contributed their group libraries a lot (more than 80%). The contributions are a great part of their personal repositories, as well. Group-based recommendations are made from the list of candidate items, which shown on group library and the members’ repositories but not discovered by target users yet. For these users who dominate their group libraries, thus, their group libraries hardly provide a list of candidate items, since they are mostly made up of the users’ own items. In addition, unless they share interests with their co-members of the same group, the group-based recommendation is obsolete to them.

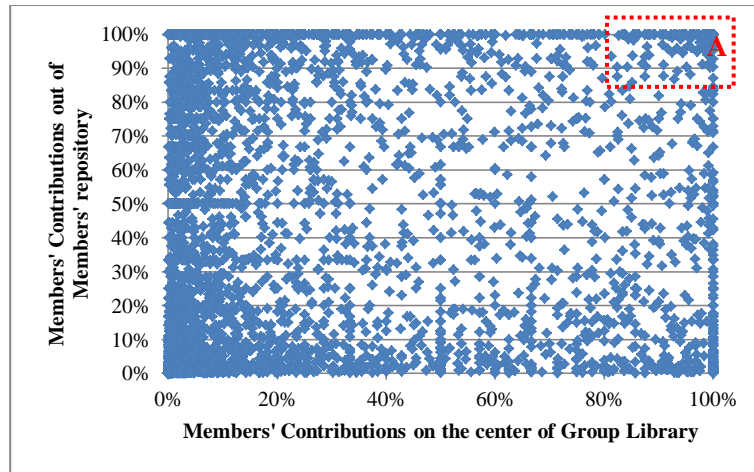


Figure 31. Contrast of Members' Contributions in terms of both Group Library and the Members' Personal Repository

As the next step, I tested how the members' contributions are correlated with the information similarity with their co-members (in here, the Jaccard similarity of co-members). I computed the information similarity of each member with his co-members and average out the similarities. Then I calculated Pearson correlation between members' contributions and the similarities of co-members. There was a significantly small correlation between members' contributions in terms of group library and the similarity of the co-members ($\gamma = .08, p < .01$). However, there was no significant correlation between members' contribution ratio to members' repositories and the similarity of the co-members. In order to examine the correlation of members' group contributions with the similarity of co-members in more details, I classed members into 10 clusters depending on the members' contributions to their group library with 10% unit increases (i.e. 0%, 10%, 20% ... 100%). I ran the one-way ANOVA and, like the results of the Pearson correlations, there was a significant difference on the similarity of co-members among the members' clusters ($F = 14.97, p < .001$). In particular, post-hoc pairwise

comparison with Schaffe adjustment examined the pattern of differences on the similarity of co-members among the clusters, in more details. Co-member similarity of two extremes (i.e. members' contributions $\leq 20\%$ and $\geq 90\%$) are significantly lower than other clusters having moderate levels of member contributions ($20\% < \text{members' contributions} < 90\%$). The Figure 32 is the result.

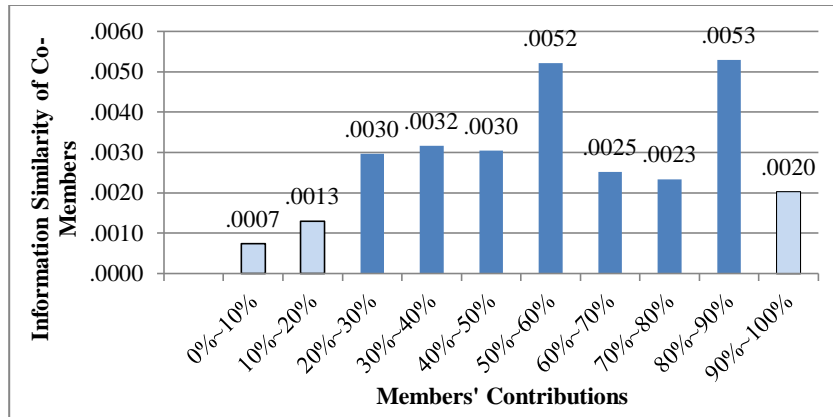


Figure 32. Members' Contributions vs. Information Similarity of Co-Members

The graph demonstrates members, who dominate their group activities or who barely contribute their group activities (clusters on two extremes), tend not to share similar interests with their co-members. Said differently, these dominators and lurkers are not interested in other co-members' personal repositories. If so, are they interested in items contributed by other co-members? In order to test this question, after excluding each member's own contributions on his group library, I computed how much his remaining items in the personal library are similar to the library. I ran one-way ANOVA on the similarities of all members using the same classification of 10 clusters depending on members' contributions. There was also a significant difference on the similarity between a member and his group after excluding his contributions among the

clusters of contributions. According the Schaffe post-hoc pairwise comparison, like the results of the similarity between two co-members, the similarities of two extremes are significantly lower than the similarities of other clusters in moderate levels of members' contributions. The Figure 33 is the result.

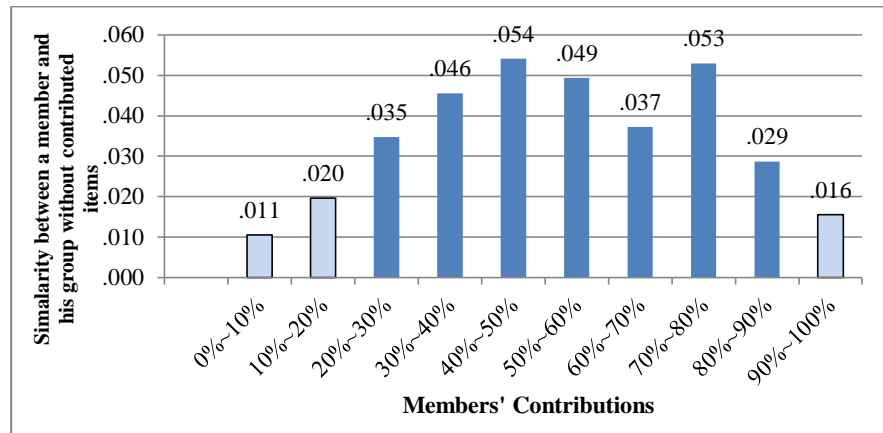


Figure 33. Members' Contributions vs. the Similarity between a Member and his Group without his Own Contributed Items

The graph demonstrates members, who dominate their group activities or who rarely contribute their group activities, tend not to share similar interests with their group library. In case of dominators, it is natural to have very low similarity with group library, since their items mostly form the group library, and this similarity was measured after getting rid of their own items. However, lurkers are uninterested in co-members' repository or other members' contributions on the group. Through these results, we can speculate that group membership-based recommendations might not work for some users like lurkers and dictators. Therefore, in the second part of this chapter, I will examine how the recommendations are useful depending on the different degrees of user contributions.

6.2.2 Information Similarity of Group Members

As mentioned, the main objective of this first part is to examine whether engaging in group activities is helpful for users to acquire useful information. To achieve this objective, the very basic step is to test whether users participating in the same groups are sharing similar interests. In this section, I compared whether two co-members of the same group share more similar information than other pairs of users who are not socially associated with (H6.1). For this comparison, users who are not in any Citeulike social networks (i.e. watching networks and group memberships) were chosen. Then, they were randomly paired up, and their information similarities were computed for all eight kinds of similarity measures.

As the first analysis, the item-based similarities were compared between co-member pairs and the random pairs. For all kinds of item-based measures, the similarities of co-member pairs were significantly different than the similarities of the random pairs ($t = 32.22, p < .001$ for the number of co-bookmarks; $t = 58.62, p < .001$ for the Jaccard coefficient; $t = 32.24, p < .001$ for the log-likelihood similarity; $t = 54.54, p < .001$ for the popularity). Table 14 is the summary of the similarity differences. The results partially proved that the hypothesis H6.1 has met for item-based similarities.

Table 14. Item-based Similarity of Group Co-members

	No. of Co-bookmarks	Jaccard	Log-Likelihood	Popularity
Group Co-Members	.26	1.59%	.050	8.00
Users Outside of Citeulike Social network	.04	0.02%	.023	6.92

I also compared metadata-based similarities between two kinds of user pairs, and Table 15 is the differences. The results of the metadata-based similarities showed the same pattern with the item-based similarities. For all kinds of metadata-based measures, the similarities of group co-members were significantly higher than the similarities of random pairs ($t = 53.69, p < .001$ for the title vector-based similarity; $t = 322.47, p < .001$ for author name vector-based similarity; $t = 178.46, p < .001$ for the original tag vector-based similarity; $t = 165.71, p < .001$ for processed tag vector-based similarity). The hypothesis 6.1 is completely accepted for all kinds of similarity measure. Apparently, group co-members shared more similar items and their topics of interests are also similar than the user pairs without any social connection.

Table 15. Metadata-based Similarity of Group Co-members

	Title	Author Name	Original Tag	Processed Tag
Group Co-Members	.1117	.0222	.0492	.0595
Users Outside of Citeulike Social network	.0147	.0007	.0012	.0020

6.2.3 Group Co-members vs. Top N CF Peers

According to the previous results, users participating in the same groups do share their interests. Then is the group network substitutable for anonymous peers selected by CF recommendations? As a simple way to answer this question, in this section, the information similarities of co-members are compared to the similarities of top N peers with the hypothesis H6.3. When our target users share similar interests with their co-members as much as they did with the top N peers, it will provide us a good reason to consider the co-members as a component of recommendations. For this comparison, using the CF recommendation algorithms

based on users' bookmark records, top 20, 10 and 5 peers were selected through Jaccard coefficient. That means that these top N peers were chosen just because they bookmarked many common items with our target users. Since CF recommendation algorithm picks the peers by fully automate computations of the nearest neighbor approach, without any consideration of the group membership information, the anonymous peers could include the target users' co-members. In the same manner to choose the top N anonymous peers, I chose the three kinds of top N social connections (i.e. top 20, 10 and 5 co-members based on Jaccard co-efficient).

First, I compared the differences on the Jaccard-based similarity between peers and co-members. Because the top peers were literally the users who bookmarked the most common items with our target users, naturally, the Jaccard-based similarities of all three kinds of peers ($M = .047$ for top 20, $M = .054$ for top 10, $M = .063$ for top 5) were significantly higher than the similarities of the co-members ($M = .030$ for top 20, $M = .042$ for top 10, $M = .053$ for top 5; $F = 408.3$, $p < .001$). Then is the group network unable to substitute for anonymous peers? As mentioned, these peers were like-minded peers when we counted the co-bookmarked items. As we explored in the chapter 5, bookmarking an item is rather opportunistic. Even though there is a favorable item, it is easy for users to miss it. Therefore, in order to check whether target users share similar interests with not only their anonymous peers but also their co-members, we need to consider other aspects such as contents and authors of their favorite items or their annotated tag sets.

When the popularity of co-bookmarked items is counted, target users tend to share rare items with their co-members ($M = 15.53$ for top 20, $M = 19.26$ for top 10, $M = 25.68$ for top 5) than their peers ($M = 9.12$ for top 20, $M = 9.19$ for top 10, $M = 9.27$ for top 5). The differences

on the popularity of co-bookmarked items were significant. It seems that co-members could be a good source to acquire useful and sometimes whimsical information.

However, the comparisons of other kinds of metadata, unfortunately, demonstrated that anonymous peers could be better information sources than co-members. For all four kinds of information similarities, the similarities of top N peers were higher than the similarities of co-members, regardless of the top N ranks. Users even annotated more similar tags with their peers than co-members. The Table 16 shows the differences and statistical significance. Consequently, although co-members share similar items and interests to each other, co-members might not be able to substitute for the anonymous peers and might not be better information sources than anonymous peers. The hypothesis H6.3 is rejected and the similarity between a user and his co-members of the same group is lower than the similarity between him and his top N anonymous peers.

Table 16. Metadata-based Similarity Comparison between Co-members vs. Top N Peers

	Title	Author Name	Original Tag	Processed Tag
Top20 Co-members	.1978	.0617	.0551	.0710
Top10 Co-members	.2192	.0772	.0597	.0769
Top5 Co-members	.2441	.0950	.0658	.0844
Top20 Peers	.2890	.1268	.0545	.0723
Top10 Peers	.3111	.1439	.0606	.0796
Top5 Peers	.3370	.1649	.0688	.0892
	<i>F=2325.10, p < .001</i>	<i>F=2079.31, p < .001</i>	<i>F=75.60, p < .001</i>	<i>F=86.58, p < .001</i>

6.2.4 Group Library vs. Co-members' Personal Repository

The above results showed that users share similar items and interests with their co-members, but not as much as they did with their anonymous peers. Then what about users' group

library? Do group members have more similar information with their group library than co-members' personal repositories? Do group members have more similar information with their group library than top N peers? As the next step, I will answer these questions based on the hypothesis H6.2. Throughout the following comparisons, when computing the information similarity between members and their group libraries, I used the group library data where the members' own contributions were eliminated.

On the H6.2 assumption ("Information similarity between a member and his group is higher than the similarity between two co-members of the same group."), more weights were placed on the similarity with group libraries than group co-members. It is because group library is a collection of information collaboratively aggregated by group members who are interested in or relevant to the corresponding topics of the group. Additionally, group members can have diverse interests and their personal repository reflect their various interests. Therefore, I assume that group libraries may be coherent collections which are more relevant to the groups' topic than members' personal collections. According to the results shown on Table 17, this assumption is right except the result of the Jaccard coefficient. When I counted the number of co-bookmarked items, the results of the independent t-test demonstrated that users shared significantly more items with their groups ($M = 2.63$) than co-members ($M = .26$, $t = 60.39$, $p < .001$). Users have more similar item collections and share more rare items with their group libraries than their co-members' personal repositories ($t = 103.34$, $p < .001$ for log-likelihood similarity; $t = 42.44$, $p < .001$ for popularity weight of co-bookmarked items). However, the result of Jaccard coefficient showed that the values of the group library are lower than the values of co-members, even though there was no significant difference ($t = -1.34$, $p = 0.18$).

In case of metadata-based similarities, for all kinds, users have more common interests with their group libraries than co-members. Metadata about the item properties like titles and author names, group libraries ($M = .198$ for titles; $M = .056$ for author names) were significantly more similar than co-members ($M = .112$, $t = 56.78$, $p < .001$ for titles; $M = .022$, $t = 31.13$, $p < .001$ for author names). On the other hand, the metadata assigned by users themselves, the similarity with groups ($M = .082$ for original tags; $M = .098$ for stemmed tags) was significantly smaller than the similarity with co-members ($M = .049$, $t = 23.38$, $p < .001$ for original tags; $M = 0.60$, $t = 26.02$, $p < .001$ for the processed tags). Conclusively, the H6.2 is true for all kinds of similarity measures except the Jaccard coefficient. Thus, it is certain that users have common interests with their group libraries and their co-members of the same group, compared with random pairs. In addition, the group libraries, communal spaces of groups, are better information source than group co-members' personal repositories.

Table 17. Comparison of Information Similarity between Group vs. Co-members

	Group	Co-member
No. of Co-bookmarks	2.63	.26
Jaccard	1.46%	1.59%
Log-likelihood	26.87%	4.98%
Popularity	28.10	8.00
Title	19.83%	11.17%
Author Name	5.58%	2.22%
Original Tag	8.16%	4.92%
Processed Tag	9.82%	5.95%

As the next analysis, I also compared the information similarities of the group libraries with the top N anonymous peers. Through this analysis, I check whether users' any social reference related to their group activities can substitute for their anonymous peers. The results

indicated that, in terms of shared items and properties of items, the top N peers are still better sources than group libraries. Interestingly, however, users have more similar tag sets with their group libraries than the anonymous peers. In this case, of course, when I got rid of users' own contributions, the related tags were also eliminated before the comparison. Therefore, I suggest that group library could be a plausible source to provide users items of their interests, because some items consisting of the group library correspond to users' interests. In addition, due to the strong similarities of top N peers, we need to consider both group library and anonymous peers at the same time. Therefore, in the second part of this chapter, I will generate the hybrid recommendations by fusing these different kinds of peers as one reference group.

Table 18. Metadata-based Similarity Comparison between Co-members vs. Top N Peers

	Group Library	Top20 Peers	Top10 Peers	Top5 Peers	
Title	.198	.289	.311	.337	$F= 785.9, p < .001$
Author Name	.056	.127	.144	.165	$F= 681.9, p < .001$
Original Tag	.082	.055	.061	.069	$F= 74.0, p < .001$
Processed Tag	.098	.072	.080	.089	$F= 54.1, p < .001$

6.2.5 Two Social Networks in Citeulike: Watching Network vs. Group Membership

In chapter 5, we discovered that the watching network has been built based on the shared interests. In this chapter, we also found that group co-memberships are rooted on common interests. When Citeulike users have two options for users to socially associate to each and the purposes of both options are to share and acquire useful information, which option would be more helpful for them to collect favorable items? In this section, this question will be answered. First, for the 10,114 users who have either group membership or watching relations, the

information similarities between group co-members and the similarity between members and their groups were compared with the similarities of watching connections.

Table 19 shows the differences of item-based similarities between group-based relations and watching relations. In the situation where all differences are significant, group library scored the winner. However, among the rest relations, it is hard to clearly say that which relationship is a better source, since the patterns of differences are quite heterogeneous according to the similarity measures. In terms of the number of co-bookmarks, co-members shared lower number of co-bookmarks than the indirect watching relations with 1 hop distance and this relationship is the worst in terms of the log-likelihood similarity. However, other two similarity measures depict different pictures. Consequently, when users have two options, group activities are more beneficial to collect favorable items and, in particular rather than the information collections of the co-members in the same group, group library is a better source.

Table 19. Comparison of Item-based Similarity between Watching Relations and Group-based Relations

	No. of Co-bookmarks	Jaccard	Popularity	Log-likelihood
Direct Watch	1.80	0.21%	1.80	20.42%
1hop Watch	.39	0.04%	0.76	9.71%
2hop Watch	.16	0.02%	0.45	6.06%
Co-members	.26	1.01%	8	4.98%
Group-Member	2.63	1.46%	28.10	26.87%
	$F = 3552.6, p < .001$	$F = 6845.4, p < .001$	$F = 1933.7, p < .001$	$F = 6151.2, p < .001$

However, the results of the metadata-based comparison show more consistent patterns of differences as Table 20 displays. For these metadata-based similarities, the group library is a definite winner. In addition, co-members also shared more similar interests, specifically such as similar authors and tags, than watching relations. In case of watching relations, they shared more

similar keyword sets than co-members but the similarity is lower than group library. Therefore, the results explained that users share more resembling interests with their group-based social connections than their watching connections.

Table 20. Comparison of Metadata-based Similarity between Watching Relations and Group-based Relations

	Title	Author Name	Original Tag	Processed Tag
Direct Watch	14.40%	1.49%	3.62%	5.05%
1hop Watch	8.14%	0.33%	0.95%	1.68%
2hop Watch	6.26%	0.20%	0.62%	1.14%
Co-members	11.17%	2.22%	4.93%	5.95%
Group-Member	19.83%	5.58%	8.16%	9.82%
	$F = 17143.6, p < .001$	$F = 7937.9, p < .001$	$F = 17894.9, p < .001$	$F = 18535.7, p < .001$

In spite of the differences between watching relations and group co-members, both connections have higher similarities in every measure that we consider. Beyond the item-based similarities, the higher values of metadata-based similarities represent that they bookmarked items of which the topics are similar and the authors are similar. They also have similar conceptual and structural knowledge, in accordance with the higher similarity of their tag sets. Therefore, we suggest that users in utility-based social networks are based on similar interests and actually influence each other. The connections serve as a useful source of information.

However, these results about the differences between watching relations and group co-members are rather general, since I computed the similarities for the whole user population. We cannot decide yet which utility-based relations are more effective aids to acquire useful information between two social networks. Therefore, for the users who participate in both watching network and group activities, I will examine which connections are more similar to them. As explained, in our Citeulike dataset, there are 1118 users who are watching users and, at

the same time, group members. Their similarities with their group-based relations and watching relations will be compared. Then, between their watching connections and co-members in the same group, which connections have the higher similarity was explored. On average, each target user has 20.4 watched users and 40.9 co-members. According to the results on the Table 21, group library is the best information source and their watching relations are the second best information source for all kinds of information similarity. On the other hand, they share the lowest similarities with their co-members. Specifically, the users share more common but relatively rare items with their group library. Additionally, they tend to bookmark items containing similar contents and authors, and similar tags with their group library. I interpret this result to mean that, as a source to acquire useful information, group activities are more beneficial than watching relations. Therefore, rather than collections built by individuals, communal spaces jointly contributed by people with underlying common interests are more valuable. Since the users didn't actively initiate the relations like watching relations, and most of the social interactions occur on the group space, users are inclined to pay less attentions on the co-members' library. However, this conclusion is only true when a system provides groups' communal spaces on which members are able to collaboratively build the group collections.

Table 21. Similarity Comparison between Watching Connections & Group Co-members for the Same Users

	Watching	Co-members	Group-Members	
No. of Bookmarks	2.85	1.29	5.33	$F= 188.7, P < .001$
Jaccard	.0026	.0018	.0122	$F= 567.9, P < .001$
Popularity	26.4251	11.3726	48.6086	$F= 2196.2, P < .001$
Log-likelihood	.2688	.1848	.4525	$F= 639.9, P < .001$
Title	.1648	.1203	.2369	$F= 1017.4, P < .001$
Author Name	.0171	.0111	.0537	$F= 1087.4, P < .001$
Original Tag	.0370	.0229	.0869	$F= 970.6, P < .001$
Processed Tag	.0529	.0337	.1077	$F= 956.9, P < .001$

Up to now, we provided evidences that utility-based social networks, both watching relation and group co-member, embed critical social properties – homophily and social influence. Since there is no timestamp when each user started his watching relations and when he joined a group in our dataset, it is hard to draw a clear line between homophily and social influence. However, it is evident that users didn't collect their utility-based connections just for fun or curiosity. They actually referred to the social connections and shared the similar items and interests. Between watching relations and co-members, even though there are some similarity differences, both social networks are important references to acquire favorable items.

6.2.6 The Best Similarity Measures for Watching Network

Once we determine that the homophily and social influence exist on group membership, it is required to consider how to exploit the network on accessing useful information. In this section, I will decide the most effective similarity measures to indicate the object-centered sociality. As a way to assess the effectiveness, I will check which measure predicts the existing group memberships the best.

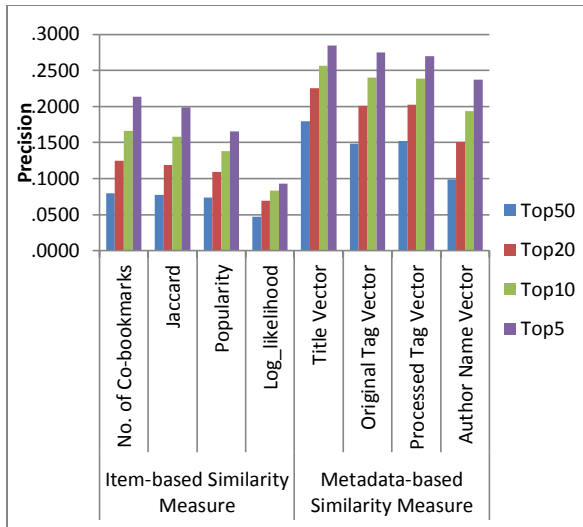
For this comparison, we chose the users who participate in group activities as our target users – 8009 group members. I examine the comparison with increasing top N ranks such as top 50, top 20, top 10 and top 5. As the evaluation criteria of this prediction, precision and recall are used as I did at the study about watching network. Thus, precision at point N (precision@N) will measure the ratio of the number of correctly predicted items in the top-N recommendation list to N and recall at point N (recall@N) will measure is the ratio of the number of correctly predicted

items in the top-N recommendation list to the total number of relevant items. For the equations and the detailed explanation, refer to the section 5.2.5. The accuracy of the predictions will be evaluated according to the four different top N ranks. ($N = 50, 20, 10,$ and 5).

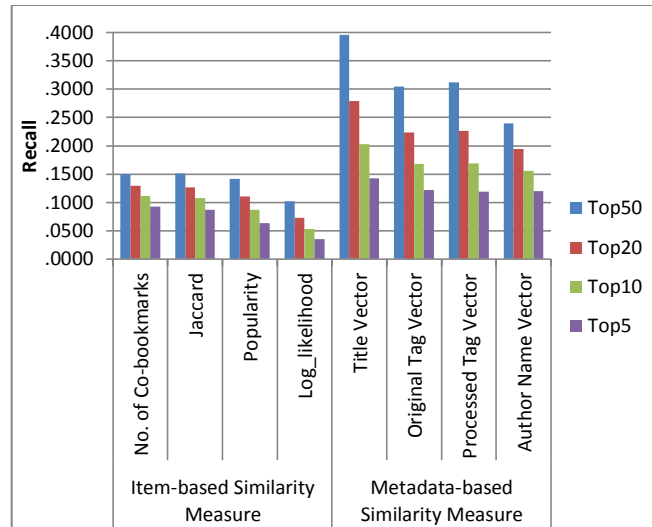
First, for 8009 target users who are members of group(s), we predicted who co-members of the same groups are. Figure 34 show the precisions and recall of the predictions made by each similarity measure. In all top N lists and for both evaluation criteria, all kinds of metadata-based measures significantly outperformed the item-based measures ($F = 353.4, p < .001$ for precision of the top 50; $F = 329.9, p < .001$ for precision of the top 20; $F = 305.9, p < .001$ for precision of the top 10; $F = 289.5, p < .001$ for precision of the top 5; $F = 897.4, p < .001$ for recall of the top 50; $F = 472.2, p < .001$ for recall of the top 20; $F = 279.4, p < .001$ for recall of the top 10; $F = 174.5, p < .001$ for recall of the top 5). None of the item-based similarities could defeat the metadata-based measures. The group co-members share common interests and common conceptual and knowledge structure, rather than exactly same information items. Therefore, all metadata-based similarities could predict the existing co-members much better than item-based similarities. The number of co-bookmarks, which was the best predictor of the watching relations, performed worse than any of the metadata-based similarities for the group co-member relations, even though it was the best measures among item-based similarities. Among the metadata-based similarities, title vector-based similarities predicted the most correct co-members and two tag-based similarities are the second best. Between two tag-based similarities, there was no significant difference.

We also predicted the existing group co-member relations using the combined measures. Figure 35 and Figure 36 show the results of precisions and recalls along with the three best

individual measures, respectively. Regarding the precisions, as we can expect easily, the combinations of metadata-based similarities performed significantly better than any other measures combined with item-based similarities ($F = 205.7, p < .001$ for precision of the top 50; $F = 136.9, p < .001$ for precision of the top 20; $F = 116.1, p < .001$ for precision of the top 10; $F = 70.3, p < .001$ for precision of the top 5; $F = 701.9, p < .001$ for recall of the top 50; $F = 283.0, p < .001$ for recall of the top 20; $F = 156.9, p < .001$ for recall of the top 10; $F = 67.2, p < .001$ for recall of the top 5). In particular, fusing the title vector-based similarity and tag vector-based similarities was an effective measure in higher ranks. However, the most notable result is that the best individual measure, title vector-based similarity, predicted as many correct relations as the best combinations or much more correlations than the other kinds of combined measures. This best performance of the title vector-based similarity was the same for the recall. Among various combinations, the combinations of all metadata-based similarities (i.e. combining titles, tags and author names) performed the best. However, title vector-based similarity predicted similarly complete suggestions in terms of recall. According to the Scheffé pairwise test, there was no significant difference between the individual title vector-based similarity and the combinations of the metadata-based similarities for both evaluation criteria.



(a) The Results of The Precision



(b) The Results of The Recall

Figure 34. Results of the Test to Forecast Group Co-members

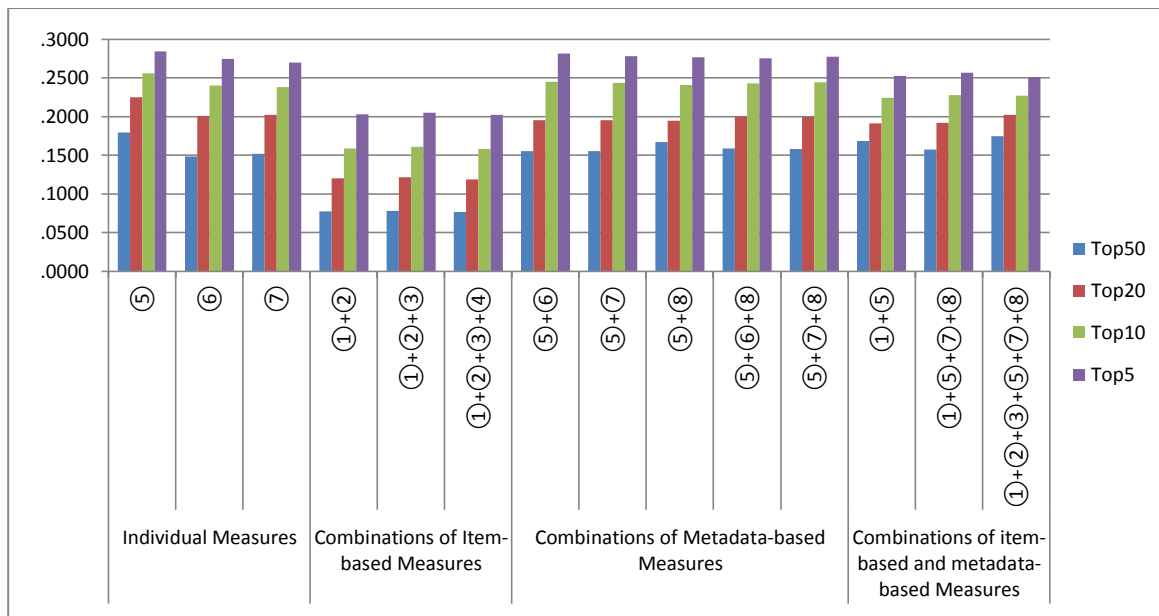


Figure 35. Precisions of Combined Measures for the Prediction of Group Co-members

(①= No. of Co-bookmarks; ② = Jaccard; ③= Popularity; ④ = Log-likelihood; ⑤ = Title Vector; ⑥ = Original Tag Vector; ⑦ = Processed Tag Vector; ⑧ = Author Name Vector)

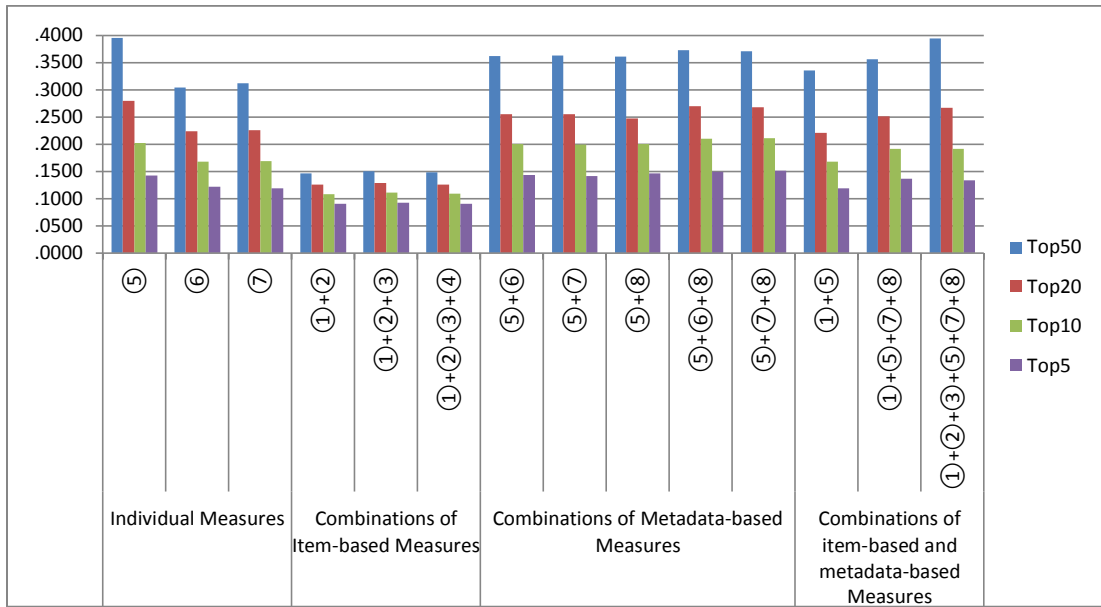


Figure 36. Recalls of Combined Measures for the Prediction of Group Co-members
 (①= No. of Co-bookmarks; ② = Jaccard; ③= Popularity; ④ = Log-likelihood; ⑤ = Title Vector; ⑥ = Original Tag Vector; ⑦ = Processed Tag Vector; ⑧ = Author Name Vector)

6.3 EVALUATION OF GROUP MEMBERSHIP-BASED RECOMMENDATIONS

The second part of this chapter aims to check how feasible users' group-based social connections are as a useful information source in a full scale, through the recommendations generated by the co-members' preferences. As the previous part demonstrated, users in the same group frequently focus on sharing topic-relevant information. Users' group-based activities and the social associations are one of the typical examples of object-centered sociality, which explained that information objects are social interaction triggers and anchors of communications. However, interestingly, there have been few recommendation studies considering this sociality.

Most of the attempts to utilize users' group activities in the recommendations were to suggest items which are seemingly favorable to the whole group of people, instead of the suggestions for individual members. Therefore, the main objective of this part is to check whether the group-based sociality is an effective information source for individual group members. Furthermore, I will check whether users' group co-members possess a comparable quality to their anonymous peers as information source by substituting the peers with group co-members in recommendation process. The most critical hypothesis to be explored in this part is the following.

H6.4 Group Membership-based recommendations are better than traditional CF recommendation approach.

In addition, since Citeulike system offers users two kinds of object-centered sociability – watching network and group membership network, for the users who are participating in both social networks, it will be tested which of the two kinds performed better as a component of personalized recommendations. The overall design of the proposed recommendation algorithms are displayed on Figure 37.

		Computation of Prediction Probability				
		Non-Personalized	K-Nearest Neighbors		Matrix Factorization	
			Bookmark	Titles & Social Tags		
Selection of Peers	Anonymous Peers	<ul style="list-style-type: none"> Community Community_CW 	<ul style="list-style-type: none"> CF CFCW 		<ul style="list-style-type: none"> CFSVD CFSVD_CW 	
	Group Members		<ul style="list-style-type: none"> GMem GMem_CW 	/		<ul style="list-style-type: none"> GMemSVD GMemSVD_CW
	Group Members + Group		<ul style="list-style-type: none"> Group Group_CW 			<ul style="list-style-type: none"> CP_Group
	Group Members + Watching Connections		<ul style="list-style-type: none"> GMemWatch GMemWatch_CW 	/		<ul style="list-style-type: none"> GMemWatchSVD GMemWatchSVD_CW
	Anonymous Peers + Group Members + Group		<ul style="list-style-type: none"> Hybrid Hybrid_CW 			<ul style="list-style-type: none"> HybridSVD HybridSVD_CW

Figure 37. Overall Design of Group Membership-based Recommendation Algorithms

The experimental comparison of the group-based recommendations will be performed by following three steps of the CF recommendations – selection of peers, computation of prediction probability and suggestion of recommendations.

Selection of Peers

The results of the previous section showed that group co-members and group libraries have different degrees of information similarities. It seems that, rather than group co-members, group libraries are more reliable sources of information to the members. In this part, by generating recommendations based on group members or (and) group library information, I will test this hypothesis is right or not.

H6.5 Recommendations based on both groups and co-members’ information collections are better than the recommendations solely based on co-members’ collections.

In Citeulike system, along with this group membership network, there is watching network which is explored in previous chapter 5. Both of the networks aim to share and distribute interesting information, and there are a certain number of users who engage in both networks. Therefore, I will compare the quality of the recommendations made by each of the networks. Additionally, by aggregating two social networks into one set of social links, I will generate the recommendation based on the unified set of social links. Lastly, when I compared various kinds of information similarities between users' group co-members and their top N peers, the top N peers have mostly higher similarities (hence, H6.3 was rejected). In here, it is necessary to fuse uses' self-defined social networks with their top N anonymous peers and to test whether the hybrid recommendations enhance the quality than the situation where recommendations take into account each kind of peer separately. For the test, I will examine the following hypotheses.

H6.6 Recommendations based on one social network are better than recommendations based on two social networks.

H6.7 Hybrid recommendations based on users' group-based social connection and their top N anonymous peers are better than traditional CF recommendation approach.

Algorithms to compute the prediction probability

In this step, I will choose better recommendation algorithm to compute the prediction probability. K-Nearest Neighbor approach and matrix factorization algorithms will be tested. By comparing the performance, better algorithm will be chosen.

Algorithms to compute the prediction probability

The results of the section 6.2.6 indicated that metadata information is more important similarity measures for group co-members than item-based similarities. In addition, when we consider users' bookmarks as preferences, since the bookmarks simply represent users' interests on the items without any further information such as the degree of preferences, candidate items from the same peer have the exactly same prediction probability values. However, each item contains different contents and it is very often that users' annotated tags vary item-by-item. In this study, thus, as one way to incorporate the metadata information in recommendations, when the system selects the final list of recommended items, I add content properties of items to the recommendation prediction probabilities of the items.

6.3.1 Recommendation Algorithms

6.3.1.1 Group-based Recommendations

My proposed social network-based recommendation algorithms mainly aim that, a target user's self-defined social connections play a role of the reference peers. The social connections substitute for his anonymous peers picked by CF recommendation algorithms, at the step to choose the target user's reference peers for his recommendation. The group-based recommendations were generated in this way. The top N anonymous peers defined by automated CF computations were replaced by group and (or) co-members. When considering users' group memberships, I proposed two strategies: 1) considering co-members' personal repositories (i.e. GMem) and 2) considering both co-members' personal repositories and group library (i.e., Group). The former recommendations are intended to utilize so-called 'social comparison' or

‘constructuralism.’ These theories suggest that people use the others who are similar to them as a reference group and compare themselves with the references to get information or make a decision. As I suggested in the previous section, co-members of the same group share the reasonably similar interests. Therefore, they could be good references to each other. The later recommendation approach is intended not only to utilize co-members’ repositories but also to utilize mutually aggregated information collection of groups, as the theory of ‘communal sharing’ suggests. Since group members collaboratively have built the group library, the information on the library could be communal assets to the members. In here, the recommendations solely based on group library (not considering group co-members) weren’t considered. The results of the first part (i.e. the section 6.2.1) demonstrated that some users devoted themselves so much on composing their group library. Accordingly, they dominated the group library and I named the type of users as ‘dictators’. When the dictator members’ contributions on the group library were excluded for the recommendations, there is very few or no item left on the library. That means that, for this dictator type of members, the recommendations only based on the group library are literally impossible.

In the step to compute the prediction probabilities of recommendations, K-Nearest neighbors and matrix factorization approaches were used alternatively. When I use the K-nearest neighbor approach, I only considered the information of users’ group-based connections (i.e. GMem and Group). In the execution of the matrix factorization approaches (i.e. GMemSVD and GroupSVD), for every target user, I built a separate sub-matrix only being made up of the target user’s bookmarks and his social connections’ bookmarks. Then the matrix factorization approach was applied on that sub-matrix. Due to the small number of bookmarks and less sparse matrix

(average sparsity of the matrices), the recommendations could be efficiently generated. The standard matrix factorization approach feeds the whole bookmark data onto the computation at a time. Therefore, the system requires a huge size of memory and often complains about the heap size error. When considering the whole dataset, the algorithm requires a large number of latent factors (e.g. > 100 for Netflix dataset [199]) and the large factors are directly connected to the computational cost. As explained, because members of the group tend to have a narrow focus on a certain topic, it doesn't need to compute the large number of factors. In a situation where online users' participation is getting increase, the large-scale based SVD has a serious scalability problem [23, 158, 195]. Therefore, whenever there are modifications on bookmarks of a group's social connections, we can update the small sub-matrix for them and applied the matrix factorization method.

6.3.1.2 Recommendations based on Two Social Networks and Hybrid Recommendations

In order to assess the quality of various online social networks as information sources, as mentioned, the recommendations based on different social networks will be compared. As Citeulike provides users opportunities to participate in watching network and group memberships, for the Citeulike users who associate both social networks, I generated the recommendations based on both networks (i.e. GMemWatch, GMemWatch_CW, GMemWatchSVD, and GMemWatchSVD_CW) and compared the quality with the other recommendations using each of them separately. In the same way to produce the group membership-based recommendations, the recommendations consider the bookmarks of users'

group co-members and watching networks. When there are overlapped social links between two social networks, I counted them only once.

I also produced hybrid recommendations by combining users' top N peers and their group-based social connections (i.e. Hybrid, Hybrid_CW, HybridSVD, and HybridSVD_CW). In here, the variable N is determined as many as each target user's group-based social connections are. According the experimental evaluation, when peers equaling to the number of the target users' group-based connections were used in the K-nearest neighbors, the CF recommendations produced the best results. A target user' top N anonymous peers were combined with his co-members of the same group and his group library. In particular, among various hybridization strategies, the mixed hybrid strategy was chosen. Put differently, two sources of information – one is target users' top N anonymous peers and another is their group-based social connections – were mixed as one set, and then the mixed set was used as a foundation of the recommendations.

6.3.1.3 Adding Metadata Information of Items in Recommendations

In the first part of this chapter, the similarity of watching connections showed that, rather than counting users' bookmark records, metadata of items are more critical to represent the shared interests. In this group-based recommendation, two strategies were proposed to incorporate the metadata information of items in recommendations – content similarity weights and metadata-based preferences. I already explained the content similarity weights on the 5.3.1.2. Briefly, the content similarity weight is to reflect users' cognitive understandings on the items and contents of items in recommendations at the final selection of recommended items.

Particularly, in this content weight similarity, I did not limit the scope of social tags to the users' group co-members. Put differently, regardless the annotators are socially associated with our target users or not, I considered all social tags annotated to items. In the comparison of shared interests between group co-members and top N anonymous peers (i.e. the section 6.2.3), our target users shared more common metadata with the top N anonymous peers than their co-members. I interpreted this result to mean that, despite the fact that they are socially disconnected in general, like-minded users' opinions are still not negligible. In addition, according to my own preliminary study [95], content similarity weights performed better than content boosted approach or content-based recommendations.

In previous recommendations with content similarity weight, first it is computed the similarity about how much peers in consideration (whether they are anonymous peers or group-based social connections) are alike to our target users. Then the level of how much candidate items are favorable to our target users in terms of the metadata was added. In the recommendations based on metadata-based preferences (CP_Group), I computed the suggestions in the other way. It is computed first how much our target users' bookmarked items are matched with their overall tastes and then the degree about how much the peers would like their favorable items was applied. As explained, bookmark-based preference is unary rating – a user like this item. Therefore, it is hard to generate more elaborate and accurate recommendations. In here, the metadata-based preferences represent what target users like derived from their keywords and social tags. As I did for content similarity weight, I built target users' keyword vector-based profiles and processed tag vector-based profiles. In this case, of course, the items in users' test set were excluded. For the comparison, I also built the keyword vector-based profiles and

processed tag vector-based profiles of the target users' bookmarked items. Then, the cosine similarities indicating how each bookmarked items are similar to the corresponding users' whole metadata-based profiles were calculated. The cosine similarities altered the bookmark-based unary ratings. That is, the cosine similarity between users' metadata vectors and other metadata vectors of their favorite items is now the numeric ratings showing that how much our target users liked the corresponding items. The equation to compute the metadata-based preferences is the following. In this equation, the item i is one of the items in user u 's bookmark history.

$$CP_{u,i} = \left(\frac{k_{ui} - \mu_K}{\sigma_K} \right) + \left(\frac{t_{ui} - \mu_T}{\sigma_T} \right) \quad \text{Eq. 28}$$

Once the metadata-based preferences are computed, I computed the K-Nearest Neighbor recommendations using the Pearson correlation. Since now the metadata-based preference have numeric values, the Pearson correlation is applicable. The matrix factorization recommendations were also executed using the metadata-based numeric preferences. In this study, the main target of the social recommendations is the group-based recommendations. Therefore, I applied the metadata-based preferences solely to the recommendations based on group co-members and group library. According to the preliminary analysis, this recommendation approach performed the best. This metadata-based preference is an attempt to increase the effectiveness of the group-based recommendations.

6.3.1.4 Community vote-based Recommendations

This approach (i.e. Community) is the simplest approach among all proposed algorithms. For each target user, I picked the most popular items among his group libraries and his co-members' repositories. Once items do not appear on target user's collection, I counted the bookmark

frequency of the items in both group libraries and co-members' collections. According to the descending order of the frequencies, the items were sorted and suggested as community vote-based recommendations. Because group members have interests on a certain topic, they gathered on the relevant online place (i.e. their group library) to share the topic relevant information. Therefore, the popular items among the people whose interests are similar could be of a value to be a part of the recommendations.

6.3.2 Experimental Evaluation of Recommendations

For the evaluation of the proposed recommendation algorithms, the targets were 8009 users who are a member of groups. Their bookmark sets were split into 10 equal sized subsets so as to do 10 cross validation. For each iteration, one set of the 10 sets were used as a test set and other 9 sets were used as a training set. This process was repeated with a different test set for 10 times. I called this data set as 'whole set.' In addition, so as to decide various settings or parameters to optimize the recommendation algorithms (such as the similarity measures for the N-nearest approach and the number of features and value of SVD ALSWR factorizer), I used the first iteration of the above 10 subsets and called this small data set as a 'probe set.' When I use this 'probe set', first subset will be the test set and remaining 9 subsets will be the training data without any iteration. As the evaluation criteria, because Citeulike dataset doesn't contain numeric ratings, traditional evaluation methods for information retrieval were used; precision and recall.

6.3.3 Social Network-based Recommendations using Various Settings

As the evaluation, through H6.4, I plan to compare the quality of group membership-based recommendations with baseline CF recommendations. Additionally, as depicted on Figure 37, depending on the scope of social links, I considered five types of peers including peers with anonymity and social associations. Through the various scopes of social links (i.e. group co-members only, group co-member and group library, group co-members and watching connections, and anonymous peers and group-based social links), it is designed to assess the quality of social recommendations across the various scopes. However, before the comparisons, it is required to find a good setting for each kind of social recommendations. In K-Nearest Neighbor based recommendations, as demonstrated on the first part of this chapter, there are several similarity measures and top N numbers to consider. We need to decide a good similarity measure and N number produces the best results, thereof. In matrix factorization, we also need to determine the optimal number of latent factors and λ values to the weighted regularization. Therefore, as a pre-analysis before the full-scale comparison, I try to find the best algorithmic settings within each kind of social recommendations.

As the first step of the analysis, I tested the good similarity measures of K-Nearest Neighbor approaches and the optimal settings for matrix factorization approaches. In this analysis, the smaller probe data set was used and the F1 was the evaluation method. For K-Nearest Neighbor approach, three sorts of similarity measures – the number of co-bookmarks, Jaccard co-efficient, and log-likelihood similarity – and four numbers of N – top 100, 50, 20 and N equaling to the number of target users' social connections – were tested. The detailed results

were explained on the Appendix A.2.1. In case of matrix factorization, several numbers of latent factors and multiple kinds of λ regularization value were tested. The detailed test settings and results were explained on the Appendix A.2.2. The following Table 22 shows the final results. Since there are no significant differences, even though Jaccard coefficients were generally selected, it is hard to conclude that one item-based similarity outperformed the others.

Table 22. The Optimal Setting of Each Recommendation Approach

Kinds of Recommendations		Optimal Settings
KNN-based	CF	Jaccard Coefficient with the same number of peers as much as the target users' social connections exist
	GMem	Jaccard Coefficient with the top 50 similar social connections
	Group	Jaccard Coefficient with the top 50 similar social connections
	GMemWatch	Loglikelihood similarity with the top 20 similar social connections
Matrix Factorization-based	CFSVD	50 factors with $\lambda=0.15$
	GMemSVD	16 factors with $\lambda=0.15$
	GroupSVD	16 factors with $\lambda=0.15$
	GMemWatchSVD	20 factors with $\lambda=0.15$

The next analysis is to check whether adding metadata information of items can enhance the recommendation quality or not. Table 23 and Table 24 display the precisions and recalls of top N results including the statistical significance. The results showed that even though some differences are not statistically significant, adding content property of items to the recommendations are generally helpful to increase the accuracy and completeness of the recommendations, except one case – community vote-based recommendations. I have one more strategy to utilize the metadata information of items in recommendations – metadata-based preferences. In the next section, I will compare the recommendation quality between two strategies –content similarity weight and metadata-based preferences. Due to the improved performance, in the next analysis, I considered the all recommendations with content similarity

weight except the community vote-based recommendations. Both community and community_CW will be considered in the next analysis.

Table 23. Differences of Recommendation Precision Depending on Metadata Properties of Items

	Top10	Top5	Top2	
CF	1.10%	2.21%	1.72%	t = -8.42, p < .001 for top 10; t = -2.95, p = .003 for top 5; t = -7.96, p < .001 for top 2
CF_CW	2.00%*	2.75%*	3.95%*	
CFSVD	0.91%	1.17%	1.58%	t = -8.63, p < .001 for top 10; t = -7.91, p < .001 for top 5; t = -6.42, p < .001 for top 2
CFSVD_CW	1.62%*	2.14%*	2.92%*	
GMem	1.20%	1.63%	2.53%	t = -7.92, p < .001 for top 10; t = -8.03, p < .001 for top 5; t = -6.94, p < .001 for top 2
GMem_CW	1.95%*	2.80%*	4.55%*	
GMemSVD	1.29%	1.64%	2.18%	t = -6.44, p < .001 for top 10; t = -6.31, p < .001 for top 5; t = -5.61, p < .001 for top 2
GMemSVD_CW	1.86%*	2.44%*	3.30%*	
Group	0.94%	1.20%	1.72%	t = -9.33, p < .001 for top 10; t = -9.76, p < .001 for top 5; t = -8.72, p < .001 for top 2
Group_CW	1.70%*	2.42%*	3.67%*	
GroupSVD	1.61%	2.18%	3.18%	t = -4.45, p < .001 for top 10; t = -4.36, p < .001 for top 5; t = -3.83, p < .001 for top 2
GroupSVD_CW	2.08%*	2.85%*	4.09%*	
GroupWatch	1.15%	1.28%	1.47%	t = -6.12, p < .001 for top 10; t = -6.91, p < .001 for top 5; t = -7.13, p < .001 for top 2
GroupWatch_CW	2.64%*	3.52%*	4.98%*	
GroupWatchSVD	1.98%	2.57%	3.46%	t = -4.54, p < .001 for top 10; t = -3.07, p < .001 for top 5; t = -2.26, p = .024 for top 2
GroupWatchSVD_CW	3.48%*	3.98%*	5.11%	
Hybrid	0.76%	1.00%	1.49%	t = -8.82, p < .001 for top 10; t = -8.66, p < .001 for top 5; t = -7.15, p < .001 for top 2
Hybrid_CW	1.48%*	2.15%*	3.19%*	
HybridSVD	0.74%	0.91%	1.39%	t = -12.67, p < .001 for top 10; t = -13.02, p < .001 for top 5; t = -10.73, p < .001 for top 2
HybridSVD_CW	2.08%*	2.90%*	4.18%*	
Community	0.86%	1.24%	1.69%	t = -1.92, p = .054 for top 10; t = -2.20, p = .027 for top 5; t = -3.60, p < .001 for top 2
Community_CW	0.95%	1.39%	2.13%*	

(* indicates that the differences are statistically significant)

Table 24. Differences of Recommendation Recall Depending on Metadata Properties of Items

	Top10	Top5	Top2	
CF	3.01%	3.01%	1.17%	t = -3.81, p < .001 for top 10; t = -0.99, p = .31 for top 5; t = -4.47, p < .001 for top 2
CF_CW	4.43%*	3.35%	2.36%	
CFSVD	3.53%	2.48%	1.41%	t = -2.32, p = .02 for top 10; t = -1.49, p = .13 for top 5; t = -1.75, p = .07 for top 2
CFSVD_CW	4.27%	2.87%	1.77%	
GMem	4.07%	2.81%	1.80%	t = -2.96, p = .003 for top 10; t = -3.33, p < .001 for top 5; t = -3.35, p < .001 for top 2
GMem_CW	5.44%	4.13%*	2.96%*	
GMemSVD	3.78%	2.68%	1.70%	t = -2.63, p = .008 for top 10; t = -1.77, p = .075 for top 5; t = -0.81, p = .41 for top 2
GMemSVD_CW	4.58%	3.15%	1.87%	
Group	3.63%	2.35%	1.30%	t = -4.63, p < .001 for top 10; t = -5.20, p < .001 for top 5; t = -4.98, p < .001 for top 2
Group_CW	5.50%*	4.14%*	2.64%*	
GroupSVD	4.78%	3.47%	2.34%	t = -2.52, p = .012 for top 10; t = -2.47, p = .013 for top 5; t = -1.38, p = .175 for top 2
GroupSVD_CW	5.65%	4.21%	2.69%	
GroupWatch	1.03%	0.67%	0.27%	t = -2.74, p = .006 for top 10; t = -2.94, p = .003 for top 5; t = -4.02, p < .001 for top 2
GroupWatch_CW	1.89%	1.42%	0.87%*	
GroupWatchSVD	1.88%	1.27%	0.81%	t = -1.89, p = .058 for top 10; t = -0.98, p = .324 for top 5; t = -0.13, p = .889 for top 2
GroupWatchSVD_CW	2.77%	1.60%	0.85%	
Hybrid	2.60%	1.81%	1.09%	t = -4.87, p < .001 for top 10; t = -5.07, p < .001 for top 5; t = -4.04, p < .001 for top 2
Hybrid_CW	4.30%*	3.37%*	2.05%*	
HybridSVD	1.81%	1.26%	0.91%	t = -11.17, p < .001 for top 10; t = -10.18, p < .001 for top 5; t = -7.85, p < .001 for top 2
HybridSVD_CW	5.63%*	4.28%*	2.84%*	
Community	4.79%*	3.61%*	1.91%	t = 6.04, p < .001 for top 10; t = 4.87, p < .001 for top 5; t = 1.54, p = .123 for top 2
Community_CW	3.17%	2.47%	1.64%	

(* indicates that the differences are statistically significant)

The next analysis is to determine a better algorithm to compute the recommendation probability between K-Nearest Neighbor and matrix factorization algorithms within each recommendation approach. Since the approaches with content similarity weights produced generally better results, the comparison was made on the center of those approaches.

Table 25 and Table 26 show the results of precisions and recalls respectively. Unlike the distinct patterns of differences on the recommendations depending on the absence and presence of the content similarity weights, there is no clear winner. In the recommendations based on anonymous peers and group co-members, the quality of K-Nearest Neighbors approach was more accurate and complete. In particular, for the recommendations using metadata-based preferences, the K-Nearest Neighbor approach was better than matrix factorization. According to the better performance of each proposed recommendation algorithms, the overall design of the recommendation approaches shown on Figure 37 can be revised in a summarized form as the following Figure 38.

		Computation of Prediction Probability	
		Non-Personalized	Bookmark-based or Metadata-based
Selection of Peers	Anonymous Peers	<ul style="list-style-type: none"> • Community • Community_CW 	<ul style="list-style-type: none"> • CFCW
	Group Members		<ul style="list-style-type: none"> • GMem_CW
	Group Members + Group		<ul style="list-style-type: none"> • GroupSVD_CW • CP_Group
	Group Members + Watching Connections		<ul style="list-style-type: none"> • GMemWatchSVD_CW
	Anonymous Peers + Group Members + Group		<ul style="list-style-type: none"> • HybridSVD_CW

Figure 38. Revised Design of Group Membership-base Recommendation Algorithms

Table 25. Differences of Recommendation Precision depending on Probability Computation Algorithms

	Top10	Top5	Top2	
CF_CW	2.00%*	2.75%*	3.95%*	t = 3.56, p < .001 for top 10; t = 3.66, p < .001 for top 5; t = 3.66, p < .001 for top 2
CFSVD_CW	1.62%	2.14%	2.92%	
GMem_CW	1.95%	2.80%	4.55%*	t = 0.77, p = .441 for top 10; t = 2.29, p = .022 for top 5; t = 4.76, p < .001 for top 2
GMemSVD_CW	1.86%	2.44%	3.30%	
Group_CW	1.70%	2.42%	3.67%	t = -3.57, p < .001 for top 10; t = -2.76, p = .005 for top 5; t = -1.64, p = .100 for top 2
GroupSVD_CW	2.08%*	2.85%	4.09%	
GroupWatch_CW	2.64%	3.52%	4.98%	t = -2.59, p = .009 for top 10; t = -1.04, p = .296 for top 5; t = -0.18, p = .853 for top 2
GroupWatchSVD_CW	3.48%	3.98%	5.11%	
Hybrid_CW	1.48%	2.15%	3.19%	t = -4.78, p < .001 for top 10; t = -4.00, p < .001 for top 5; t = -3.11, p = .001 for top 2
HybridSVD	2.08%*	2.90%*	4.18%	
CP_Group	1.48%*	2.05%*	2.55%	t = 6.07, p < .001 for top 10; t = 5.79, p < .001 for top 5; t = 2.92, p = .003 for top 2
CP_GroupSVD	0.76%	1.03%	1.61%	

Table 26. Differences of Recommendation Recall depending on Probability Computation Algorithms

	Top10	Top5	Top2	
CF_CW	4.43%	3.35%	2.36%	t = 0.46, p = .64 for top 10; t = 1.57, p = .114 for top 5; t = 2.32, p = .020 for top 2
CFSVD_CW	4.27%	2.87%	1.77%	
GMem_CW	5.44%	4.13%	2.96%*	t = 2.18, p = .028 for top 10; t = 2.95, p = .003 for top 5; t = 3.85, p < .001 for top 2
GMemSVD_CW	4.58%	3.15%	1.87%	
Group_CW	5.50%	4.14%	2.64%	t = -0.38, p = .696 for top 10; t = -0.21, p = .831 for top 5; t = -0.15, p = .878 for top 2
GroupSVD_CW	5.65%	4.21%	2.69%	
GroupWatch_CW	1.89%	1.42%	0.87%	t = -2.07, p = .038 for top 10; t = -0.62, p = .529 for top 5; t = 0.09, p = .922 for top 2
GroupWatchSVD_CW	2.77%	1.60%	0.85%	
Hybrid_CW	4.30%	3.37%	2.05%	t = -3.10, p = .001 for top 10; t = -2.39, p = .016 for top 5; t = -2.58, p = .009 for top 2
HybridSVD	5.63%	4.28%	2.84%	
CP_Group	3.83%	2.97%*	1.41%	t = 3.19, p = .001 for top 10; t = 3.84, p < .001 for top 5; t = 1.14, p = .254 for top 2
CP_GroupSVD	2.43%	1.59%	1.09%	

6.3.4 Comparison of Various Recommendation Approaches

Once the optimal setting for each recommendation approach is decided, it is the stage to compute the quality of social recommendations among the various recommendation approaches including the CF recommendations. Figure 39 and Figure 40 display the precision and recall results, respectively. According to the One-way ANOVA, the quality of these recommendation approaches are significantly different ($F = 122.79$, $p < .001$ for top 10 precision; $F = 90.24$; $p < .001$ for top 5 precision; $F = 63.55$, $p < .001$ for top 2 precision; $F = 29.32$, $p < .001$ for top 10 recall; $F = 23.55$; $p < .001$ for top 5 recall; $F = 15.41$, $p < .001$ for top 2 recall). For the specific patterns of the differences, I executed the Schaffé pairwise comparison.

First, I compared two approaches utilizing different methods to incorporate metadata information of items in recommendations – content similarity weight (i.e. GroupSVD_CW) and recommendations utilizing metadata-based preferences (i.e. CP_group). In the tests base on precision and recall, the approach adding content similarity weight outperformed the recommendation utilizing metadata-based preferences. Therefore, rather than taking into account the metadata properties of items in modeling users' preferences, it is more effective to apply the information at the final step to choose the recommended items.

Second, I compared all recommendations utilizing group membership network (i.e. GMem_CW, GroupSVD_CW, Community and CP_Group) in order to discover proper scope of group-based social connections as a foundation of users' personalized recommendations. Among them, the matrix factorization recommendations based on co-members and group library (GroupSVD_CW) and another recommendation based on group co-members (GMem_CW) was

better than the others in the perspective of both precision and recall. Between the two approaches, there was no significant difference. Therefore the relevant hypothesis 6.5 is rejected. On the other hand, the community vote-based recommendations performed the poorest among the four proposed approaches in the respect of precision and recall. Even though group co-members share similar interests to each other, the detailed preferences of the individual members were different to each other and require more sophisticated personalization tactics like group membership-based recommendations.

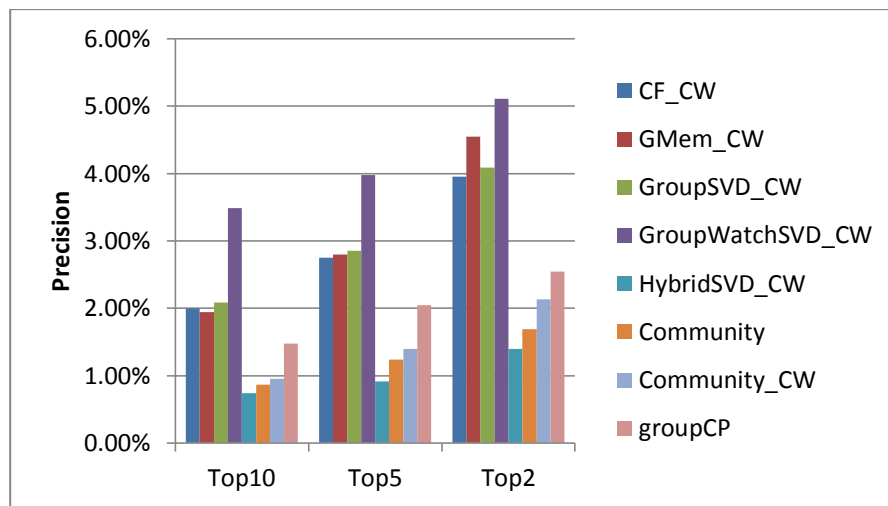


Figure 39. Differences of Recommendation Precision depending on the Approaches

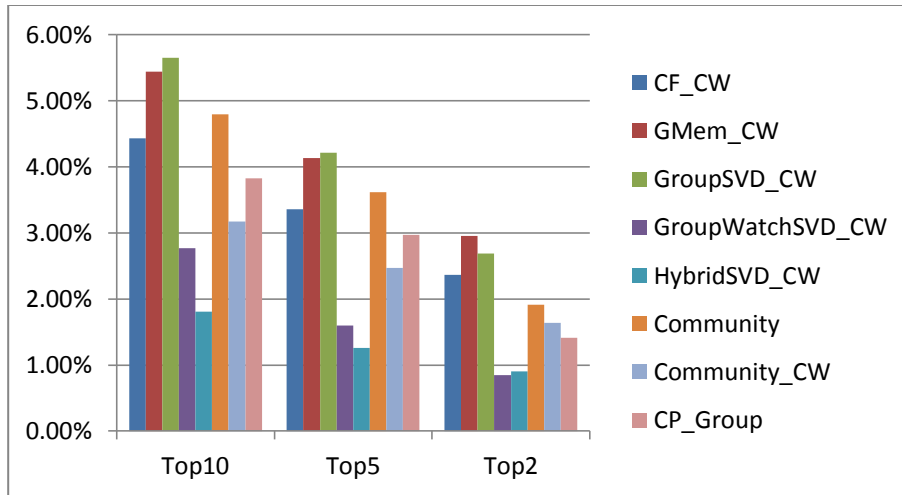


Figure 40. Differences of Recommendation Recall depending on the Approaches

The next analysis is to compare recommendations the group membership-based recommendations with traditional CF recommendations. In the results of precision, the quality of recommendations based users' group membership-based connections was not significantly different with the traditional CF recommendations in all ranks. However, the differences of the recall demonstrated a different pattern. In all ranks, GMem_CW and GroupSVD_CW significantly outperformed the other approaches, while the difference between these two kinds of recommendations was insignificant. Put differently, the recommendations based on group membership-based social links can produce as accurate suggestion as the traditional CF recommendations can. However, the recommendations are more complete than the CF recommendations. While the hypothesis H6.4 cannot be accepted solely by the results of precision, when counting the results of recall, the hypothesis H6.4 is accepted partially.

If the both target users' anonymous peers and their group membership-based connections are good information source, when we combine these two groups of peers together, can the

recommendations based on the combined peers generate better recommendations? I generated hybrid recommendations by taking into account the preferences of both anonymous peers and group membership-based connections. Regardless of the evaluation criteria (whether it is precision or recall), the hybrid recommendations were always the worst in all ranks. This personalized recommendation approach was worse than community vote-based recommendations which are non-personalization approach. Therefore, the hypothesis 6.7 is rejected. Rather than mixing up different groups of peers, it is better to take into account individual kinds of peers separately.

6.3.5 Comparison of Group Membership-based Recommendations with other Social Recommendation Approaches

In the previous section, one approach on the Figure 39 and Figure 40 – the recommendation considering two kinds of Citeulike social networks into one set of peers (i.e. GMemWatchSVD_CW) – wasn't mentioned. The results showed the recommendation approach outperformed in the perspective of precision but not so good in the results of the recall. However, not every target user was able to receive this recommendation. In Citeulike, there are two social networks – watching network introduced in the chapter 5 and group membership explored in this chapter. With their own choice, Citeulike users are able to participate in both networks. In the Citeulike dataset, I found 1118 users who are engaging in both networks. For these 1118 users, the recommendations considering two kinds of social networks were available. In this section, between two possible socialities of these target users, I assessed which sociality is a better information source. These 1118 target users can receive three kinds of social recommendations –

group membership-based recommendations (i.e. GroupSVD_CW), watching network-based recommendations (i.e. ReciWatch_CW) and both network-based recommendations (i.e. GMemWatchSVD_CW). Within each recommendation approach, I selected the algorithm yielded the best results. Along with CF recommendations for these target users, the quality of the three social recommendations was compared.

Figure 41 shows the comparison results in terms of precision, recall and F1 measure. The differences of precision and recall were insignificant ($F = 1.17, p = .310$ for top 10 precision; $F = .64; p = .526$ for top 5 precision; $F = 2.63, p = .072$ for top 2 precision; $F = 3.09, p = .046$ for top 10 recall; $F = .38; p = .678$ for top 5 recall; $F = .28, p = .753$ for top 2 recall). Despite of the insignificance, there are contrasting patterns between precision and recall. Watching network-based recommendations produced better recommendations than the other two approaches including regarding the precision. However the result of recall displayed that the performance of watching network-based recommendations was worse than the others. In order to decide the best performing social recommendations, F1 measures were computed. The F1 Measure also showed no significant difference ($F = 2.71, p = .067$ for top 10 F1 value; $F = 23.55; p = .001$ for top 5 F1 value; $F = .123, p = .885$ for top 2 F1 value). In addition, I failed to find any consistent patterns of differences in all ranks. Therefore, at least for 1118 Citeulike users, combining multiple social networks didn't improve the quality, compared with the recommendations using individual social networks separately. Additionally, for this 1118 target users, social network-based recommendations are as good as the traditional CF recommendations. Therefore, the hypothesis 6.6 is rejected. All three kinds of social recommendations and CF recommendations performed equally.

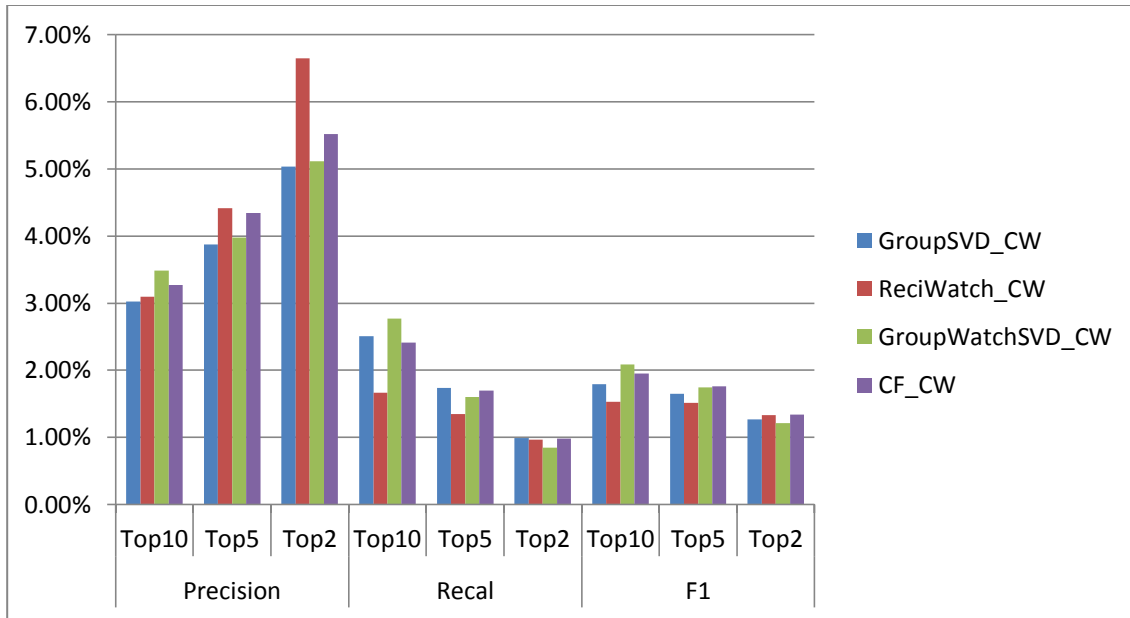


Figure 41. Comparison of Various Citeulike Social Recommendations

6.3.6 Recommendation Quality according to the Ratio of Group Members' Contributions

In the first part of this study (i.e. the section 6.2.1), I explained various patterns about how group members contribute on their group library. According to their patterns, the users were classified into three clusters – lurkers, moderators and dictators. Again, lurkers didn't contribute the communal spaces of their groups and are liable not to share any interests with their group co-members. On the other hand, dictators dominate their group libraries. Their contributions occupied more than 90% of the group library and they seldom share interests with their co-members. The moderators are the users who contributed their group to a moderate level and share similar interests with their co-members. Since users belonging to each cluster have distinct patterns in distributing and sharing information within their social network, I examined whether

group-based recommendations produce the different levels of quality for each cluster or not. We can reckon that the most desirable user cluster is the moderator type of users. However, for the users in other two clusters, the group network-based recommendations are in question. For this analysis, first I classified group members into the three clusters. Since users are able to participate in multiple groups and have different degree of contributions per group, the classification was made by the median value of the corresponding member's contributions over his all groups. The contributions usually have skewed distributions. Hence, the median value is a more suitable 'representative value' than mean value. By following the same classification schema used in the section 6.2.1, when a member's median contribution is less than 20%, he is a lurker. When another members' median contribution is more than 90%, he is a dictators. The remaining users whose median values are more than 20% and less than 90% are moderators. Among 8009 target users, the majority was the lurkers (84.8%, $n = 6788$ users) and their number of groups were 1.47 ($\sigma = 1.58$) on average. 216 users were the dictators (2.7%) and their groups were 1.62 ($\sigma = 3.80$). 1005 users (12.6%) were moderators and joined 1.51 groups ($\sigma = 1.32$).

Figure 42 and Figure 43 display the results of precision and recall. There are significant differences on the quality of various recommendations depending on the user clusters ($F = 16.61$, $p < .001$ for top 10 precision; $F = 12.81$, $p < .001$ for top 5 precision; $F = 8.41$, $p < .001$ for top 2 precision) in terms of precision. However, in the results of recall, the quality of various recommendations was not significantly different according to the user clusters ($F = 2.09$, $p = .009$ for top 10 recall; $F = 1.37$, $p = .157$ for top 5 recall; $F = .33$, $p = .990$ for top 2 recall). As the next analysis, I find the patterns of difference on the various recommendation approaches within each cluster of users.

First, for lurkers, the differences of the various recommendation approaches are significant ($F = 81.07, p < .001$ for top 10 precision; $F = 62.66, p < .001$ for top 5 precision; $F = 44.89, p < .001$ for top 2 precision; $F = 27.59, p < .001$ for top 10 recall; $F = 21.98, p < .001$ for top 5 recall; $F = 13.51, p < .001$ for top 2 recall). In particular, for this cluster of users, the recommendations combining their group membership-based social connections with their watching partners yielded significantly higher precisions than the other approaches. Group membership-based recommendations (i.e. GMem_CW and GroupSVD_CW) were the second best approaches followed by CF recommendations. Among the group membership-based recommendations and CF recommendations, there is no significant difference in the precision. In terms of recall, two group membership-based recommendations produced significantly higher recall than the others. Due to the rare contributions on the group library and uncommon interests with their co-members, the lurkers are regarded to be uninterested in their group related social activities. However, the recommendation results indicated that their group libraries and co-members are valuable sources of information and they are participating group activities with genuine interests on the corresponding topic and for acquiring information of their interests.

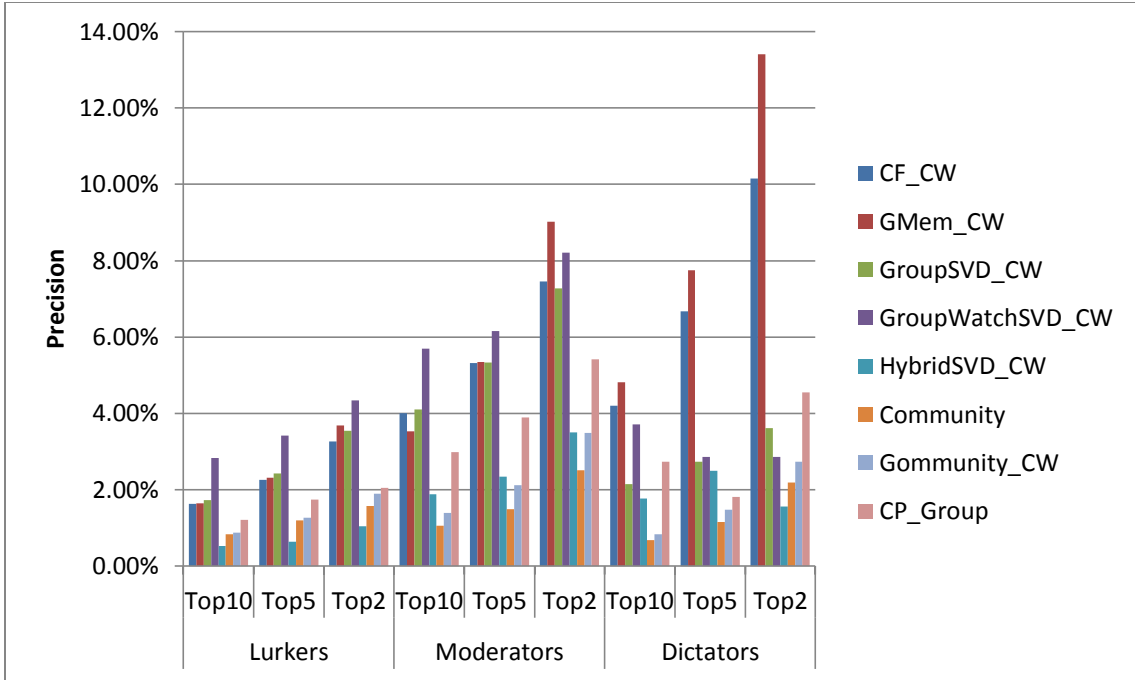


Figure 42. Differences of Precision according to Users' Contribution Patterns

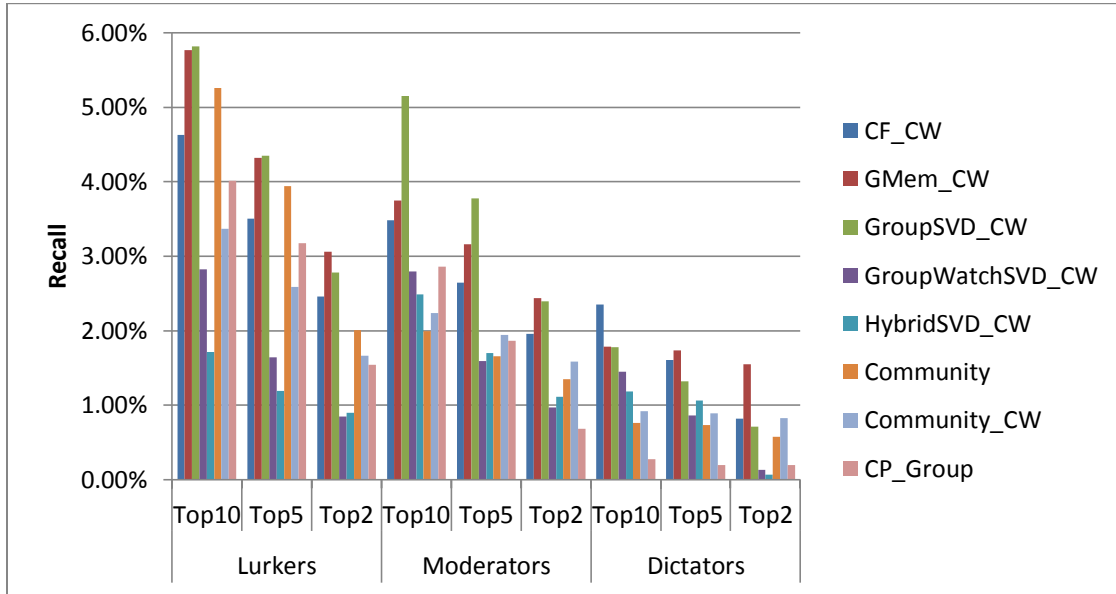


Figure 43. Differences of Recall according to Users' Contribution Patterns

Next, I assessed the best recommendation algorithms for moderators. The precision differences of the various recommendation approaches are significant ($F = 36.47, p < .001$ for top 10 precision; $F = 26.11, p < .001$ for top 5 precision; $F = 15.77, p < .001$ for top 2 precision). However, the differences of recall are insignificant except the recall of the lowest rank ($F = 5.32, p < .001$ for top 10 recall; $F = 3.40, p = .001$ for top 5 recall; $F = 2.00, p = .051$ for top 2 recall). In the test based on precision, like the results of lurkers, in lower ranks, the recommendations combining users' group-based connections with watching connections are significantly better than the others. However, in the highest rank where the most accurate predictions are supposed to be place, the recommendations based on group co-members (i.e. GMem_CW) was significantly the best. In the test based on recall, even though there are no statistical significances, group membership-based recommendations (i.e. GMem_CW and GroupSVD_CW) always produced the highest recall values in all ranks.

Lastly, for the dictators, it is hard to expect the good performance of group membership-based recommendations. In their group library, after excluding their contributions, very little information has left. They are seldom alike to their co-members. Like the result of the lurkers where group membership-based recommendations performed well despite of the uncommon interests with the co-members, can we expect the same improvement of the group membership-based recommendations for these dictators? While the precision differences of the various recommendation approaches are significant ($F = 5.13, p < .001$ for top 10 precision; $F = 5.12, p < .001$ for top 5 precision; $F = 5.51, p < .001$ for top 2 precision), the differences of recall are insignificant ($F = .91, p = .491$ for top 10 recall; $F = .41, p = .89$ for top 5 recall; $F = 1.34, p = .22$ for top 2 recall). According to the test based on precision, the recommendations based on their

group co-members were the best recommendation algorithms than CF recommendations. However, the recommendations utilizing both group co-members and group libraries performed poorly. In the test based on recall, the results of the lowest rank indicated that CF recommendations were better than the others. However, as the rank is getting higher, the recommendations based on their group co-members became the best recommendation algorithms than CF recommendations. Interestingly, the second best recommendations in the highest rank is community vote based recommendations, which consist of items popularly favored by their co-members. I interpreted this result to mean that, even though the dictators didn't pay attention to their co-members, people whose interests are relevant to the topics of their groups gathered together in the dictators' groups. Therefore, the group co-members are now a beneficial source of information to the dictator users.

6.3.7 Group Membership-based Recommendation for Cold-start Users

In this dataset, 32.33% of our target users (2,589 users) have less than 5 bookmarks. Usually, in recommendations, users having this amount of bookmarks or ratings are hard to receive reasonably good recommendations since the numbers of bookmarks are insufficient to present what they like properly. Therefore, these users are called as cold-start users. One weakness of the CF recommendations is this cold-start user problem. Since the recommendations rely on the overlapped interests or tastes among users, when users have insufficient information about interests or tastes, it is difficult to suggest presumably favorable items. As explained in the chapter 2, some researchers suggest that social network-based recommendations could be a good solution to solve this cold start user problem. Hence, in this subsection, I test whether the best

recommendation approach differ according to target users' bookmark numbers and whether group network-based recommendation is a good approach for cold-start users. For this test, I classified users into three clusters according to their number of bookmarks – cold-start users ($n < 5$), user having medium level of bookmarks ($5 \leq n < 200$) and users having large number of bookmarks ($n \geq 200$). There are 2,589 users, 4408 users and 1012 users in each cluster, respectively. For the users belonging to each cluster, I compared the recommendation algorithms.

Figure 44 and Figure 45 display the results of precision and recall. The differences on the quality of various recommendations are significant depending on the user clusters ($F = 16.87, p < .001$ for top 10 precision; $F = 10.25, p < .001$ for top 5 precision; $F = 4.66, p < .001$ for top 2 precision; $F = 16.37, p < .001$ for top 10 recall; $F = 12.74, p < .001$ for top 5 recall; $F = 10.27, p < .001$ for top 2 recall).

As the next analysis, I investigated the patterns of difference on the various recommendation approaches within each cluster of users. First, for users having large information collection, the differences of the various recommendation approaches are significant ($F = 9.02, p < .001$ for top 10 precision; $F = 9.52, p < .001$ for top 5 precision; $F = 11.37, p < .001$ for top 2 precision; $F = 6.83, p < .001$ for top 10 recall; $F = 7.21, p < .001$ for top 5 recall; $F = 8.74, p < .001$ for top 2 recall). In both tests based on precision and recall, the recommendations based on group co-members and watching connections and another kind of recommendations based on group co-members generated the better recommendations than CF recommendations.

Second, for users having medium number of bookmarks, the differences of the various recommendation approaches are significant ($F = 64.01$, $p < .001$ for top 10 precision; $F = 55.40$, $p < .001$ for top 5 precision; $F = 38.63$, $p < .001$ for top 2 precision; $F = 40.15$, $p < .001$ for top 10 recall; $F = 31.39$, $p < .001$ for top 5 recall; $F = 19.80$, $p < .001$ for top 2 recall). In both tests based on precision and recall, the group membership-based recommendations based on group co-members and group library (i.e. GroupSVD_CW) generated improved suggestions than CF recommendations for all ranks. In addition, the recommendations based on group co-members and watching connections were also good way to suggest items.

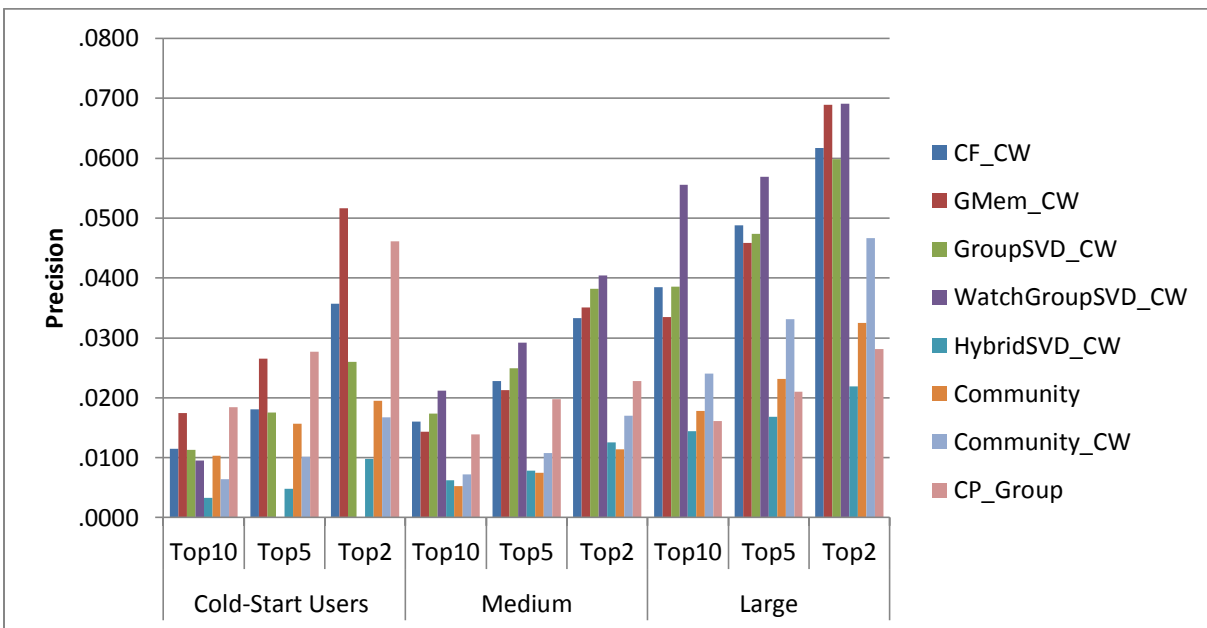


Figure 44. Differences of Precision according to Users' No. of Bookmarks

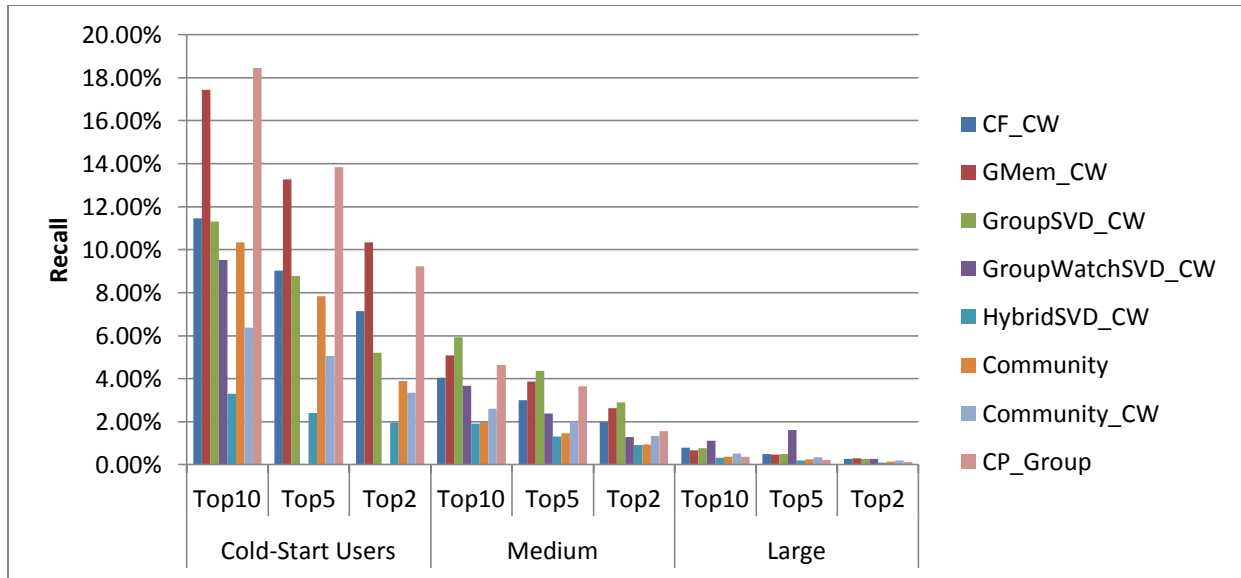


Figure 45. Differences of Recall according to Users' No. of Bookmarks

The last analysis is to find the best recommendation algorithm for cold-start users. The differences of the various recommendation approaches are significant ($F = 14.21, p < .001$ for top 10 precision; $F = 10.88, p < .001$ for top 5 precision; $F = 9.56, p < .001$ for top 2 precision; $F = 14.21, p < .001$ for top 10 recall; $F = 10.88, p < .001$ for top 5 recall; $F = 9.56, p < .001$ for top 2 recall). In both tests based on precision and recall, the recommendations based on group membership-based recommendations (i.e. GMem_CW and CP_Group) consistently generated the better recommendations than CF recommendations in all ranks. Interestingly, for the cold-start users, the metadata-based preference was a workable option to generate the recommendations. I suggest that in the situation where users' preference is insufficient, it is wiser to consider the metadata information of items.

6.4 CONCLUSION

In this chapter, I examined the feasibility of users' group based social connections as a useful information source through 1) the examination of shared interests among co-members of the same group and the overlapped interests between a member and his group library and 2) the empirical evaluation of group membership-based recommendations.

In the first part of this study, the information similarities of co-member pairs were significantly higher than the values of the random pairs. Rather than the co-members of the same group, however, users share more common interests with their group library. I suggest that group library is a collection of information collaboratively aggregated by group members who are interested in or relevant to the corresponding topics of the group. Therefore, group libraries are coherent collections which are more relevant to the groups' topic than members' personal collections. The group libraries are also better information source than watching networks. Regardless whether the similarity was measure by shared items or shared metadata, the similarity between group members and their group libraries was higher than the similarities of two group co-members or two watching connections. In addition, between group co-members and watching connections, watching connections is a better information source than group co-members in terms of both item-based similarities and metadata-based similarities.

The comparison of recommendation quality among various recommendation approaches demonstrated that the social recommendations based on group membership network can generate quality suggestions equivalent to or sometimes better than CF recommendations. In group membership-based recommendations, group co-members and group's communal spaces are

equally important information sources. However, the hybrid recommendations combining users' group membership network with their anonymous peers are ineffective and harm the quality of group-based social recommendations. In addition, for the users participating in both the watching network and group membership network on the Citeulike system, the best social recommendation approach was tested. The result showed by combining two object-centered sociality wasn't effective to enhance the recommendation quality. Regardless whether the recommendations were generated by one kind of social network or two kinds of social networks, the predictions were equivalent level.

In the analysis about users' contributions on their group libraries, I found that the group membership-based recommendation might not work well for every group member. Some users didn't contribute their group activities at all nor have common interests with their co-members. Some other users contribute their group activities too much but didn't pay attentions to the other co-members. Therefore, I tested whether we need to consider different kinds of recommendation approach depending on users' patterns about distributing and sharing information with group co-members. The results showed that we don't need to. For all group members, regardless of their social interaction patterns, the group membership-based recommendations were better than the other recommendations.

Finally, in order to solve the cold-start problem, I tested which recommendation method is the most suitable for the users whose bookmarks are insufficient. For them, group membership-based recommendations are significantly the best than the other recommendation options. In addition, unlike other users having sufficient number of bookmarks, the recommendations utilizing the metadata-based preferences worked well. Therefore, for the users

don't have enough data to represent their interests or tastes, it is wiser to utilize the metadata information of items.

7 **COLLABORATION NETWORK-BASED RECOMMENDATIONS**

This chapter aims to explore the feasibility of recommendations based on users' research collaboration network. Compared with the era when people collaborated mainly offline, with the development of a wide variety of computer-supported cooperative work applications, we are able to work together online without any limitation on time and regions. In addition, in respect of the major cost savings through computer-supported cooperative works, when people find other experts whose interests are matched with theirs, they can more easily cooperate with them. Even though the current collaboration networks are prosperous in online spaces and are based on shared interests and expertise, there are very few attempts to utilize the collaboration network as a component of users' personalized recommendations. Therefore, I will investigate patterns about how research collaborators share information and how to generate users' recommendations based on their collaborators' preferences.

In particular, the target system of this study is a conference navigation system. Academic conferences provide exciting opportunities for researchers. They can attend many interesting presentations and panels to learn about works of others and to place their own works in a broader context. Frequently the ultimate goal of attending talks is to meet other researchers who are doing similar or relevant research, to share opinions about their works, and in many cases, to look for a chance of interesting research collaborations. However, finding interesting talks to

attend is a real challenge. In conferences, several sessions are typically held at a time, and there are a lot of research-related activities, such as tutorials, industrial discussion, social events, keynote speech, etc., for a short period of time. Therefore, there is little time for attendees to analyze all alternatives and decide where to go. It is easy for them to miss important talks and further to miss important opportunities of future collaborations. A recommender system that suggests attendees which talks are worthy to attend could be of real help in this context. This chapter aims to examine patterns about how conference attendees share interesting research talks with their collaborators and propose recommendation algorithms to suggest interesting talks for conference attendees. Particularly, this domain requires harmonizing information items with social context in recommendations. Research collaborators share overlapped research interests and topics. Hence, when researchers attend conferences with their collaborators, talks caught their attentions could also interest their collaborators. This is the first attempt to generate personalized recommendation for conference talks.

As mentioned, collaborators tend to share similar research interests. In order to generate recommendations of research talks based on the shared research interests, what is the best way to represent their similar interests? When there is another kind of social network existing on our target system, which social network is better information source? In the first part of this chapter, these questions will be answered. Therefore, the objectives of this first part are to understand information sharing patterns among research collaborators in the context of attending research conferences and the comparison of information similarity between two social networks existing on our target system – one is a collaboration network and another is a personal acquaintance network. In particular, in this system, since the collaboration network was built based on users’

own publication records and their co-authorship history, the collaboration network is co-authorship network. In this chapter, collaboration network and co-authorship network will be used interchangeably.

However, the conference context is not typical for recommender systems. The majority of recommendation algorithms were created for a large-scale and long-term context where user actions (such as ratings) are collected from many users over a long period of time. On the other hand, at conferences, the time to collect data and the number of users are severely limited. Users usually pay attentions on papers presented on conference right before and during the conferences. The users are also from a specific research community of the corresponding topic. In this unusual context, what kind of recommendation approaches could work best? To answer this question, the second part of this study will study various recommendation algorithms to suggest interesting talks.

In this chapter, I will check whether this suggestion is right or not; collaboration network is a feasible information source to acquire useful research talks in the first part and a viable foundation for users' personalized recommendations of conference talks in the second part.

7.1 DATA SOURCE FOR COLLABORATION NETWORK: CONFERENCE

NAVIGATOR 3

This study about the research collaboration network is based on a social adaptive navigation system. The system is 'Conference Navigator' which aims to support conference

attendees' navigations and effective scheduling. The system was introduced in 2006 by the Personalize Adaptive Web Systems Lab, the University of Pittsburgh. Since then, three versions of the system were used at 16 conferences up to July, 2012. The most recent third version of the system (that is, Conference Navigator 3, CN3 in abbreviation) is available at <http://halley.exp.sis.pitt.edu/cn3>. In its essence, as Figure 46 shows, CN3 displays all talks in a conference with navigation supports. It also provides various information access methods from navigations via titles, presentation types, and authors' names to advanced search and personalized recommendations. Whenever users find interesting talks, they are able to bookmark them, and then the bookmarked items are added to their schedule automatically. For more detailed information of this Conference Navigator, please refer to [27, 188]. In conferences, talks are usually to present conferences papers. Hence, hereafter, talks and conference papers will be used interchangeably.

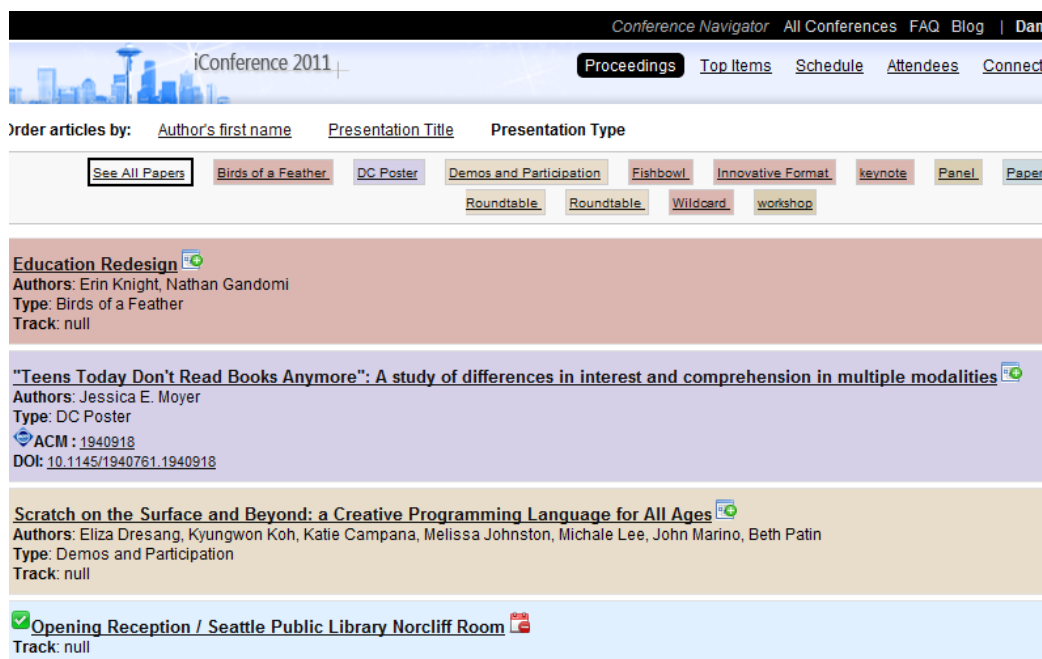


Figure 46. Main Page of Conference Navigator

For this study, with my own privilege to access the system, the system was backed up on November 30, 2011. This dataset covers bookmark history from July, 2008 to November, 2011 of 13 conferences ('Hypertext 2008, 2009, 2011', 'Adaptive Hypermedia 2008', 'UMAP 2009, 2010, 2011', 'EC-TEL 2009, 2010, 2011', 'ASIS&T 2010 Annual Meeting', 'iConference 2011' and 'TPRC 2011'). The data in this source contains user information (i.e. name, email address, affiliation and job title, and user id) and users' bookmarks (i.e. the time when each bookmark was created), conference names, venues, conference talks (including detailed talk information such as titles, abstracts, and authors of the talks), and the date and times when each talk was held. This dataset originally has about 2,200 users, 1,878 talks and 6658 bookmarks of 13 conferences. However, as described this data is for academic conferences and in conferences, the organizers usually offer several social events, such as social dinner and lunch, coffee break, opening/welcoming reception, etc. In particular, I classified keynote speeches also as social events. It is because most of conference attendees attend the keynote speeches and the bookmarking activities really don't represent users' real interests. I got rid of all records about the social events and the bookmarks. In addition, only 24.3% of whole user population ($n = 532$) bookmarked at least one talk. Hence, I excluded all other users who didn't bookmark any talk or who did bookmark only social events. I also removed talks which were not bookmarked at all. Lastly, I excluded administrative users, developers and test users. The final dataset has now 454 users, 1,000 talks and 5,094 bookmarks. About 37% of the users ($n = 171$) have at most 5 bookmarks and 30% of the talks ($n = 298$) were bookmarked by only one user. The bookmark sparsity of this dataset is 0.9886.

CN3 system enables the users to get socially connected with other users on the system. There are two kinds of CN3 social networks; following and connections. The following network is characteristic of unilateral relation, like watching relations introduced in the section 3.1 and chapter 5 and doesn't require obtaining the followed party's consent. The latter network is a reciprocally agreed-upon relation between two social parties, which is equivalent to the online friendships on Facebook or mySpace. Among 454 users in this dataset, 21 friendships (consisting of 30 users) and 123 follower relations (consisting of 48 distinct followers and 80 distinct followees) were found. In spite of the way to initiate the relations and the different degree of reciprocity, from interface point of view, there isn't any marked difference between two social networks. Both social networks aim to help users acquire useful talk information through their social partners. From users' point of view, the system shows the same social navigations and provides the social interaction functions for both kinds. There are 16 identical pairs (being made up of 25 users) between two social networks. That means 25 users in 16 friendships are also associated with the same social partners as following relations and it remains only 5 unique friendships in CN3. It seems that separating friendships and following relations is meaningless. Therefore, in this paper, these networks are consolidated as one kind of directed social network, as a 'CN3 social network.' Following network is originally a directed network but friendship network is originally an undirected network. Therefore, the links were counted twice with different directions. For instance, when user A is following user B, there is one CN3 social network from user A to user B. When user A is befriending with user C, there are two CN3 social networks from user A to user C and vice versa. In this way, I found 143 CN3 social links consisting of 105 users (some users followed each other reciprocally, even though they were

friends on CN3, which are reciprocal relationships). After excluding followees, 57 users are actually self-defined their CN3 social connections.

The users selected their CN3 social partners explicitly, but the relations are not necessarily based research collaborations. On the other hand, this study aims to exploit shared interests among users' research collaborators as a component of users' recommendations. Basically, the CN3 users are attendees of conferences. They are mostly research scientists or students and have records of their research publications to represent what their expertise and interests are. Furthermore, their publication records indicate who their colleagues are. When a pair of users has worked together, their research interests are largely overlapped or highly co-related. Hence, when a user bookmarked a conference talk, it is plausible to expect that the talk also interests his collaborators. In addition, users' publication records will be another important source to deduce users' interests. Accordingly it would be able to generate better recommendations. In order to build the users' collaboration network and to infer users' research interests, thus, their publication records were collected. Users' publication history and the resultant collaboration network were gathered not only from the conference talks and the authorship information existing on CN3 system, but also from two external data sets; DBLP (Digital Bibliography & Library Project) and publications of iSchool faculty members and graduate students. The first data source, CN3 dataset has accumulated publication records of 13 conferences. We expect that CN3 users attended the conferences so as to present their works. Therefore, titles, abstracts and author names of CN3 conferences talks were used. As the second data source for publication history, the DBLP is a bibliography Website for computer science and information science discipline. This site, operated by Universität Trier, Germany, provides

bibliographic information of numerous conference papers, journal articles and books of several publishers (e.g. ACM digital library, IEEE computer society digital library, computer science-related conference and journal list on Wikipedia, and so forth). Since this Website covers wide range of topics, the data contains huge amount of publications (more than 1.3 million articles in January, 2010). Since all conferences on CN3 are about information science, the DBLP is one of the best sources to acquire users' publication history [175, 187]. The developer team of CN3 system obtained the DBLP backup dataset directly from the Website. It contains 2,103,480 articles and the authorship information. It also has brief metadata of each article such as title, journal/book/conference name, publication year, DOI (digital object identifier) and publisher. As the final source for users' publication history, all publication information of iSchool faculty members and graduate students (For more information about iSchool, refer to <http://www.ischools.org/site/about/>) is collected by Dr. Jungsun Oh, who is an assistant professor of the University of Pittsburgh. She first collected the names of all faculty members and graduate students at every iSchool (registered in iSchools organization, <http://www.ischools.org/>) and crawled their publications. This dataset contains 9315 people and their 5962 publications.

In order to build CN3 users' collaboration network, first, the users' real names were manually disambiguated on the center of the last names. Hence, when their last names are too common to disambiguate the corresponding user with the other users such as Kim, Johnes, Lee, Liu, Lim, Scott, Wu, Xie, etc., the users were excluded. As the next step, duplicated publications in three sources of publications (i.e. CN3 conference talks, DBLP articles, and publications of iSchool faculty and graduate students) were eliminated. Through this process, among 454 users in this CN3 dataset, I found 362 users' publications. Out of these author users, I found 189 co-

authorship pairs being made up of 163 CN3 users, which is 35.7% of the CN3 user population. Among 362 users having publication record(s), however, I found only 7 CN3 social links. Table 27 is the descriptive statistics. Figure 47 and Figure 48 are the distributions of users and articles according to the number of bookmarks, respectively.

Table 27. Descriptive Statistics of Conference Navigator Dataset

Social Networks	Collaboration Network	No. of users engaging in this network	163
		No. of social connections	189
		Avg. No. of social connection per User	2.32 ($\sigma = 2.1$)
	CN3 Social Network	No. of users engaging in this network	105
		No. of users after excluding followees	57
		No. of social connections	143
Avg. No. of social connection per User		2.42 ($\sigma = 2.7$)	
Information Collections	No. of distinct conference talks	1,000	
	No. of users who have bookmark(s)	454	
	No. of users who have at least one own publication	362	
	No. of bookmarks	5,094	
	Avg. No. of bookmarks per user	11.72 ($\sigma = 10.9$)	
	Avg. No. of bookmarks per talk	5.04 ($\sigma = 6.2$)	
	Avg. No. of publications per user found on three publication sources	16.45 ($\sigma = 28.0$)	

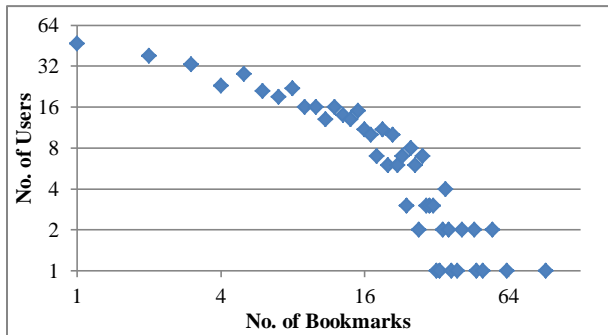


Figure 47. Distribution of Users according to the Number of their Bookmarks

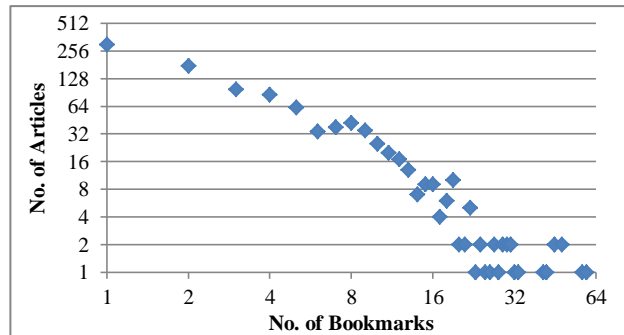


Figure 48. Distribution of Articles according to the Number of their Bookmarks

Figure 49 shows the distribution of collaboration networks. 50.31% of them have one collaborator. Actually, this collaboration network only counted the existing connections among CN3 users. These users also have other collaborators who are not CN3 users. Figure 49 shows the contrast of users' collaborators in CN3 with their total numbers of collaborators. On average, collaboration connections existing on CN3 is 13.1% of their total collaborators ($\sigma = 11.8\%$). Figure 50 is the distribution of publication numbers per user. Among 362 users who have publication records, 50% of them have 5 or less publications.

I also examined the time points when users usually bookmarked items. In a research conference context, there are three possible temporal divisions: before a conference begins, during a conference is held and after a conference ends. If users collected similar amount of bookmarks in each temporal division, it would be worthy to consider this temporal information in recommendations. Among 454 users whose average number of bookmarks is 11.7, 345 users made 10.8 bookmarks on average before conferences began ($\sigma = 9.6$). 160 users added 8.1

bookmarks ($\sigma = 7.6$) while conferences were being held. After conferences ended, 27 users added a few bookmarks ($n = 3.4$, $\sigma = 3.7$). This result means that users used the conference navigation system mainly before they attend conferences and prepared the schedule of research talks to attend. While they were attending conferences, users also bookmarked some talks, but not as much as they did once the conferences ended. Thus, it seems that recommending interesting talks is of a real help before conferences are started.

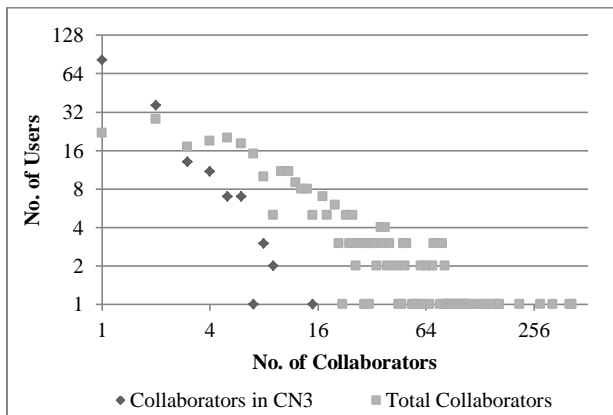


Figure 49. Distribution of Users' Collaborators

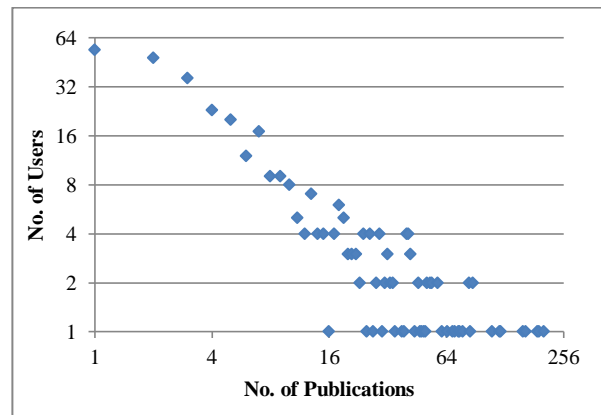


Figure 50. Distribution of Users' Publications

Lastly, I performed a textual clustering method in order to discover how many topics CN3 data consists of. Even though all conferences in CN3 are about information science discipline, specific focus of each conference is quite different. When meaningful clusters are discovered, it might possible to compute item-to-user similarity based on the results of this topic clustering. In addition, when a user has bookmarks of two or more conferences and topics of the conferences are distinctively different, it might require taking the bookmarks of different conferences individually. I ran K-means clustering from 2 to 5 of variable k with canopy

centroids and Latent Dirichlet Allocation (LDA) from 2 to 5 topics [19]. However, the results of both tests produced only one topical cluster. Although specific focuses of the conferences were defined differently, contents of the papers are similar to one another.

7.2 FEASIBILITY OF RESEARCH COLLABORATORS AS A USEFUL INFORMATION SOURCE

Since research collaborators have worked together, they are familiar with each other's expertise and research interests. In addition, their collaborations are usually based on shared interests or highly relevant expertise. Therefore, when we attend a conference with our colleagues, I suggest that, because research talks of conferences are highly correlated with the attendees' research interests and expertise. In this part, I will prove that research collaborators sharing similar research interests are a viable information source to acquire interesting research talks. However, as explained in the above 7.1 section, CN3 data has another social network – CN3 social network. According to the descriptive statistics of their co-authorship records, most of them are not associated with each other as collaborators. As the system intended, however, users chose their social connections as one kind of information source of talks to attend (when a user bookmarks a certain talk, CN3 system showed the bookmark record to his social partners). Collaboration network is based on users' shared interests and expertise but doesn't necessarily aim to share interesting talks in CN3. On the other hand, CN3 social network is mainly to share talk information in CN3. With the purpose to find interesting and favorable talks, which social network is a better information source? When the collaboration network bears a high degree of

interest similarity, does CN3 social network share comparably similar interests? Are the collaborators really good sources for talk information? Among various types of data expressing users' interests and preferences, which one does express the shared interests among collaborators and CN3 social connections better? In the first part of this chapter, I will answer these questions through in-depth study focusing on users' information sharing. The second part of this chapter will answer these questions through the evaluation of recommendations based on these social networks. The hypotheses to explore regarding these questions are the following.

H7.1. Information similarity between a pair of collaborators is higher than the similarity of a randomly coupled pair who is not in the collaboration network or in the CN3 social network.

H7.2. Information similarity between a pair of collaborators is higher than the similarity of a user pair who is in a CN3 social relation.

H7.3. Metadata-based similarity is better than item-based similarity to represent collaborators' shared interests.

7.2.1 Data Analysis Methods

CN3 data set has quite similar data structure to the Citeulike watching network and group membership dataset. The information objects are conference papers, and users expressed their preferences through bookmarking the objects. However, one difference is that users' tags are unavailable. Actually, while the CN3 system enables users to add their tags when they bookmark an article, users are not quite active to use the function. The system doesn't have enough tags to be used in personalized recommendations. Therefore, I excluded users' tags in this study.

Because of the resembling data structure, the same types of similarity measures are used. For the detailed description of the similarity measures, refer to the section 5.2.1.

First, as item-based similarity measures, I consider the number of co-bookmarks, Jaccard coefficient and log-likelihood similarity. Since the number of items and the number of users are too small to take into account of the item popularity, popularity weight of co-bookmarked items was not considered.

Second, as metadata-based similarity measures, I consider keyword vector-based similarity and author name-based similarity, of which both are using the vector space model. For keyword vector-based similarity, in this study using CN3 data, both titles and abstracts of items were taken into account, since all of the items have the information. As I did for the Citeulike datasets, all words in titles and abstracts of each user's favorite articles were aggregated together into a bag of keywords. That is, each bag of keywords is made up of all keywords appearing in one user's favorite items. For an effective comparison, text processing methods were applied to the bag of keywords – case normalization, stop words removal and stemming. Then the processed bag of keywords was transformed into keyword vector consisting of keywords and the TF/IUF (Term Frequency/Inverse User Frequency) values. Therefore, each keyword vector corresponds to each user. In this case, one document represents one user's keyword set, thus the user frequency shows how many users bookmarked papers of which the titles or abstracts contain the keywords. As the final comparison, I computed how two given keyword vectors are far from each other using the Cosine similarity.

In case of author name-based similarity, since they are proper nouns, I didn't apply any text processing method. In the same way with the keyword vector-based similarity, all names of

authors who wrote each user's favorite items were aggregated into one bag of names. The bag of names was transformed into an author name vector. Then, I computed how much two users favored items written by the same authors using the same Cosine similarity.

Along with these two preferences, one extra similarity was considered for CN3 data. Unlike Citeulike datasets, the CN3 dataset contains additional information about users' preferences; users' publication records. Users' own publications represent their research expertise and topics of interests. The items of interests in CN3 dataset are also scientific publications. Therefore, it is legitimate, as one way to measure the similarity of users' preferences, not only to count how many bookmarks two collaborators share, but also to compare how similar their expertise and research interests are through their own publications. It is to know how much two collaborators' works are correlated with their favorite items and further to conjecture whether users' professional records are useful to be used as a part of their preferences for recommendations. In order to measure users' research interests, '*research profile-based similarity*' was compared. As users' research profiles, keywords in titles and, if possible, abstracts of one user's publications were aggregated, text-processed and transformed into a research profile vector. In particular, each user's research profile was built in two ways – one profile is based on his whole publications and another profile is based on his recent publications which are published since 2008. The similarity based on the former profile is the '*research profile-based similarity*' and the similarity based on latter profile is '*recent research profile-based similarity*.' Recent publications could represent their current expertise and research interests better than the whole list, especially if users have a long list of publications. In addition, all

conferences for which CN3 system was used were recently held. Therefore, it is expected that the profiles based on users' recent publications would be more effective.

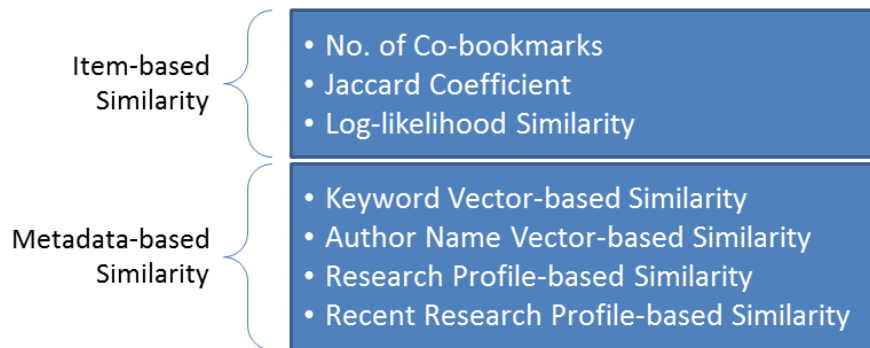
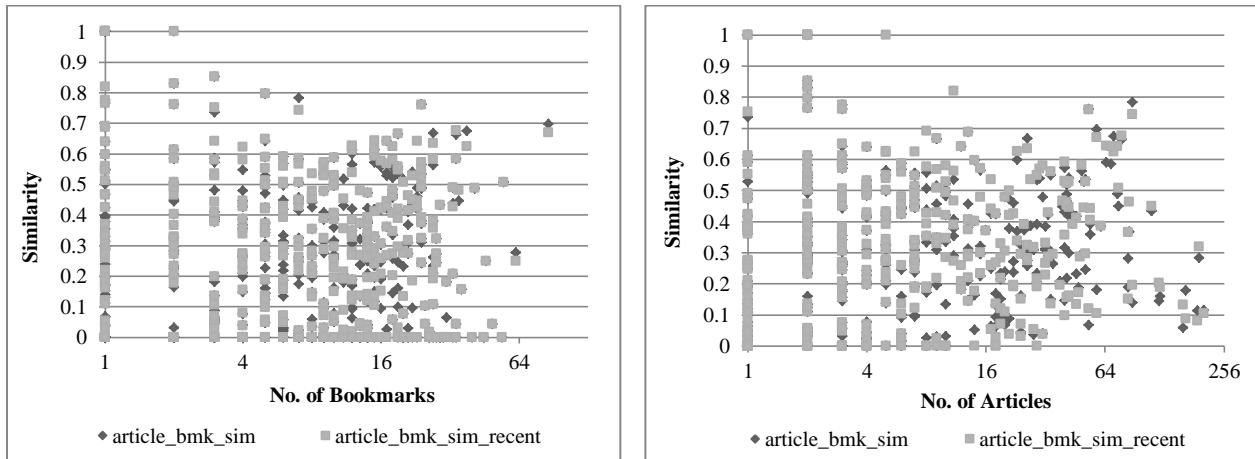


Figure 51. Summary of Similarity Measures

7.2.2 Relevance of Users' Research Profiles with Contents of their Favorite Talks

Before the comparison of users' information similarity, I examined how a user's research profile representing his expertise and interests is semantically alike with his favorite items. If users' research profiles are similar to their bookmarks, then we may be able to utilize the profiles in their recommendations as a part of user preferences. In addition, if we discover a high level of relevance between a user's research profile and his bookmarks, and pairs of collaborators share highly overlapped research profiles, we will be able to positively utilize collaborator's bookmark information in the recommendations for their colleagues. As explained above, users' publication records and bookmarks were already aggregated and transformed into research profiles and keyword vectors, respectively. Using the Cosine similarity, thus, I compared how one vector corresponding to a user's research profile is close to another vector corresponding to his bookmark keywords. In particular, since there are two kinds of research profiles – one profile

made up of users' whole publications (i.e. research profile) and another profile made up of their recent publications (i.e. recent research profile), the comparison was performed using both kinds of research profiles. Users' research profiles are alike to their keyword vectors to a moderate level (similarity = 0.31 on average). When I calculated how users' recent research profiles are similar to their keyword vectors, the similarity was also a moderate level (similarity = 0.34 on average). It was expected that users' research profiles (whether they are based on users' whole publications or recent publications) would be highly similar to their keyword vectors. The profiles are similar to their keywords, but not as highly as expected. Even though recent publications reflected the users' expertise and interests of the time when the conferences were held, the recent research profiles were alike to their keyword vectors to a moderate level. Figure 52 shows the distribution of the similarities according to their number of bookmarks and articles.



(a) The distribution of similarity according to users' number of bookmarks

(b) The distribution of similarity according to users' number of articles

Figure 52. Similarity between Users' Research Profiles and their Keyword Vectors

7.2.3 Information Similarity of User Pairs according to the Kinds of their Social Links

This section is to test whether pairs of users in collaboration relations share more similar information than the other pairs of users in CN3 social network or the remaining users who are not socially engaged at all (H7.1 and H7.2). Both research collaborators' network and CN3 social network are quite small networks, because one is consisting of less than 200 edges and another is consisting of less than 150 edges. Therefore, it is unreasonable to compare information similarity with direct social connections and indirect connections having 1 hop or more hop distances, as I did in chapter 5. Instead, the information similarities of different social networks in the CN3 system are compared.

For this test, the number of co-bookmarked items between every user pair was computed. The results of the one-way ANOVA results represent that information similarities of the three kinds of social links (collaborators, CN3 social network, and random pairs) are significantly different ($F = 405.7$, $p < .001$). According to the Scheffe post-hoc pairwise test, collaborators shared significantly larger co-bookmarked items than random pairs. However, the co-bookmarks of collaborators are significantly smaller than the ones of CN3 social connections.

The comparison of Jaccard efficient and log-likelihood similarity also showed the same results. There are significant differences on the Jaccard coefficient ($F = 423.7$, $p < .001$) and log-likelihood similarity ($F = 391.7$, $p < .001$) depending on the types of social links. In particular, both similarities of collaborator pairs were significantly higher than random pairs, but significantly lower than pairs of CN3 social links. Figure 53 displays the comparison of similarities and the average similarity values.

These results show that the first hypothesis, H7.1 is accepted for item-based similarities. All item-based information similarities of the collaborators are higher than random pairs. On the other hand, the above result showed that H7.2 hypothesis is rejected for item-based similarities. All item-based information similarities of the collaborators are lower than the CN3 social relations.

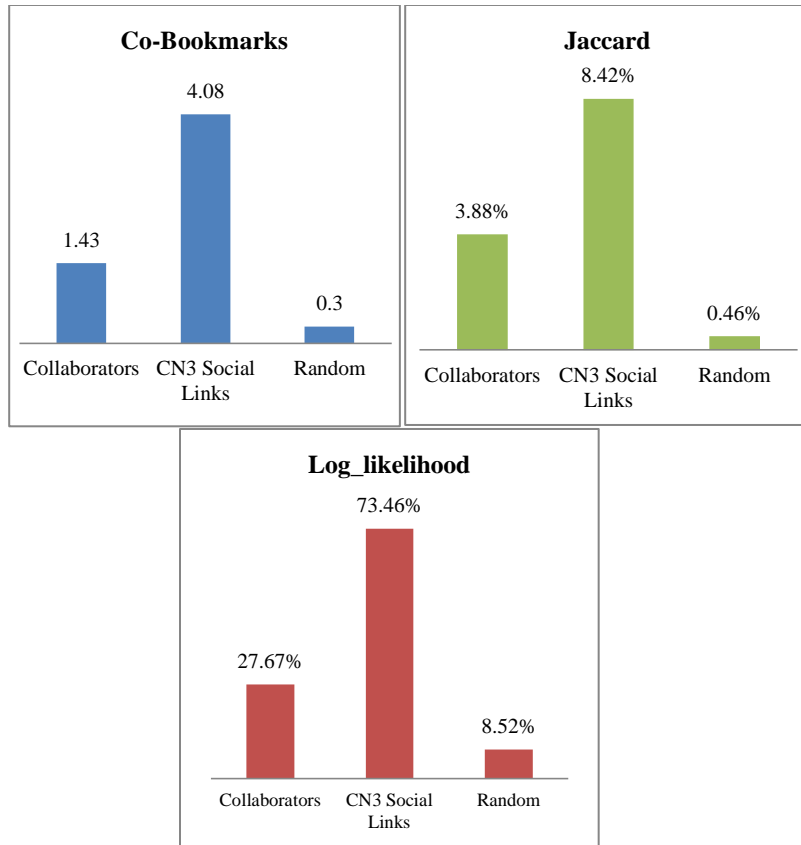


Figure 53. Comparison of Item-based Similarities according to the Kinds of Social Links

The next analysis performs the comparison of metadata-based similarities according to the social status of a given user pair. In this study, as explained, four kinds of metadata-based

similarities were explored – keyword vector-based, author name vector-based, research profile-based and recent research profile-based.

First, the comparison of keyword vector-based similarity displayed that users share significantly different level of similarities according to how they are socially associated to other users ($F = 359.0, p < .001$). According to the post-hoc pairwise test, the keyword-based similarity of CN3 social network was significantly higher than the similarity of co-authors. Like the results of item-based similarities, the keyword-based similarity of random pairs was the significantly lowest.

The results of author name vector-based similarity showed the exactly same pattern. The differences of author name-based similarities were significant across the three types of social status ($F = 830.7, p < .001$). The similarity of CN3 social network was the highest, and the value of co-authors was the second highest. The similarity of random pairs was the lowest. The differences of these pairwise comparisons were also significant.

The results of the similarity comparisons based on research profiles, nevertheless, displayed different patterns. As expected, the co-authors have more similar research profiles than pairs of CN3 social connections or randomly coupled pairs regardless whether the profiles are based on users' whole publications or recent publications. These differences are statistically significant ($F = 84.1, p < .001$ for research profiles; $F = 74.2, p < .001$ for recent research profiles). Figure 54 shows the comparison of similarities depending on how they are socially associated.

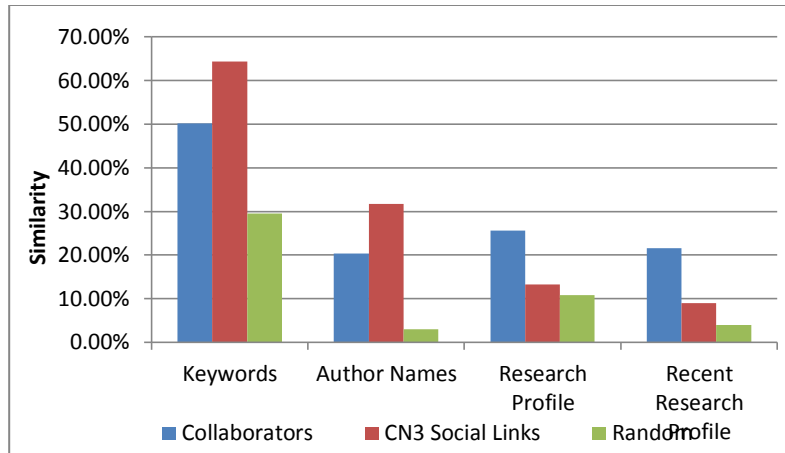


Figure 54. Comparison of Metadata-based Similarities according to the Kinds of Social Links

Therefore, for all item-based similarities and all metadata-based similarities, the first hypothesis H7.1 is accepted. The similarities of collaborators are larger than randomly coupled users. However, in case of the hypothesis H7.2, depending on the types of similarities, the results are different. For all item-based similarities, keyword vector-based similarities and author name vector-based similarities, the hypothesis is rejected. However, for the similarities based on research profiles, the hypothesis is accepted. The relatively higher similarity of collaborators' research profiles seems natural, since users' collaboration network has been built on the center of their publications.

To summary, collaborators share similar expertise and research interests. Compared with the other users who are not socially engaged at all, they also favored similar talks. However, CN3 social network of which the main purpose is to share useful information on CN3 is better source than collaborators. Therefore, in the second part of this chapter, I will consider both kinds of social networks as components of the CN3 recommendations.

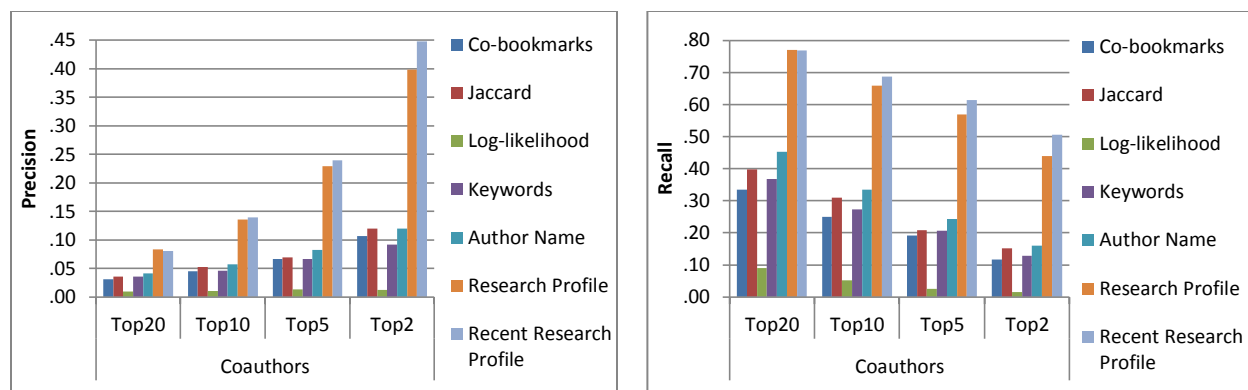
7.2.4 The Best Similarity Measures for Collaboration Network

In above section, various kinds of information similarities were compared according to the social status between two users. The results showed that, depending on the social status, the best similarity measure reflecting the shared interests was different. For instance, recent research profile-based similarity is better than the others for collaborators. However, the best similarity measure for CN3 social network is still unclear, since most of the measures performed the highest for this social network. Moreover, the individual measures can be combined into another kind and could represent users' common interests better. Therefore, this section aims to test which measure indicate the similarity of collaborators and the similarity of CN3 social connections the most effectively. The aforementioned 7 similarity measures are assessed by comparing which measure predicts the existing relations the best. For collaboration network, I chose 163 target users who have colleagues in CN3 system. I computed all seven similarity measures of a target user with every other users, regardless whether they are collaborators or not. Then the top N similar peers of the target user were chosen and among them, it is counted how many peers were actually collaborators. In the same way, I computed combined similarity measures for the target users. The comparison was executed with increasing top N ranks such as top 20, top 10, top 5 and top 2.

As the evaluation criteria of this prediction, precision and recall were used. Precision at point N is the ratio of the number of correctly predicted users in the top N predictions to the rank N. Recall at point N is the ratio of the number of correctly predicted users in the top N

predictions to the total number of real social connections. For more detailed way of calculation and the equations, refer to the section 5.2.5.

Figure 55 is the result of collaborator predictions using individual similarity measures. According to the one-way ANOVA test, the differences of the precisions were significantly different depending on the kinds of similarity ($F = 54.3, p < .001$ for top 20, $F = 64.6, p < .001$ for top 10, $F = 68.9, p < .001$ for top 5, $F = 71.1, p < .001$ for top 2). The post-hoc pairwise test showed that research profile-based similarities are significantly better than other measures. However, between two research profiles-based similarities, there is no significant difference, even though recent research profile-based similarity produced higher precision. Overall, metadata-based similarities were better than item-based similarities. Among the item-based similarities, despite of the statistical insignificance, the Jaccard coefficient was better than the other two, and log-likelihood similarity was the worst among all individual measures. The results of recall were also the same. The differences of the recalls were significantly different depending on the kinds of similarity ($F = 68.4, p < .001$ for top 20, $F = 61.4, p < .001$ for top 10, $F = 59.8, p < .001$ for top 5, $F = 49.4, p < .001$ for top 2). In terms of recall, the research profile-based similarities were also significantly better than the others. The recalls also displayed that most of the metadata-based similarities are better than item-based similarities. Among item-based similarities, the Jaccard coefficient was the best to depict the common interests.



(a) The Precision Results

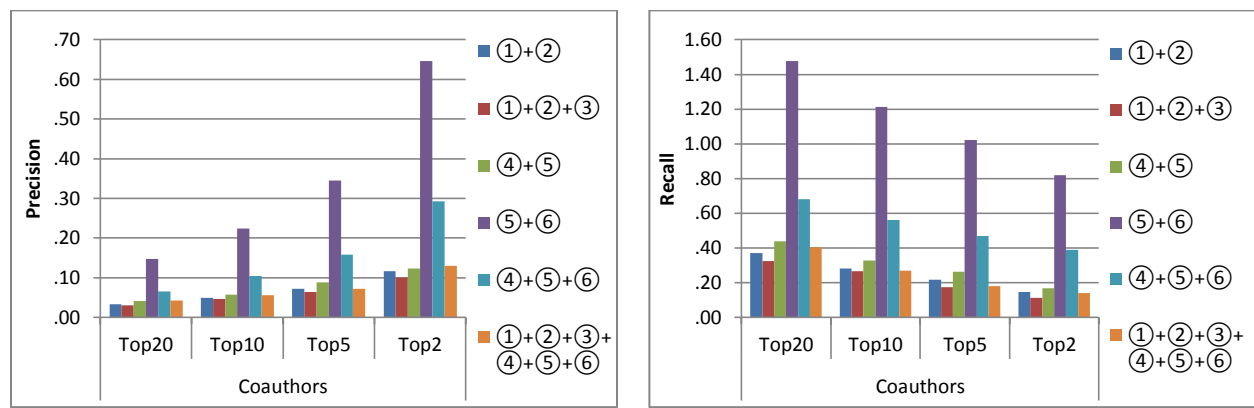
(b) The Recall Results

Figure 55. Results of the Test to Predict Collaborator Relations using Individual Similarity Measures

Figure 56 shows the results collaborator predictions using combined similarity measures. Since recent research profile-based similarity was insignificantly different to another kind of research profile-based similarity, the former was only considered in this collaborator prediction. According to the one-way ANOVA test, the differences of the precisions were significantly different depending on the kinds of similarity ($F = 100.8, p < .001$ for top 20, $F = 94.3, p < .001$ for top 10, $F = 81.0, p < .001$ for top 5, $F = 60.0, p < .001$ for top 2). The post-hoc pairwise test showed that, when research profile-based similarity was fused with author name vector-based similarity, the precision was significantly the highest. The similarity combining all three metadata-based similarities yielded the second best precision. Whether it is individual measure or combined measure, item-based similarities didn't represent the shared interest of collaborators very well. The results of recall were also the same. The differences of the recalls were significantly different depending on the kinds of similarity ($F = 140.0, p < .001$ for top 20, $F = 92.4, p < .001$ for top 10, $F = 72.7, p < .001$ for top 5, $F = 53.5, p < .001$ for top 2). The

similarity combining research profiles with author name vector-based similarity produced the highest recall, and the similarity combining all three metadata-based similarities produced the second highest recall, as well. The recalls also displayed that most of the metadata-based similarities are better than item-based similarities.

Consequently, in order to present collaborators' common interests, the similarity combining metadata-based similarities is better than individual measures or the similarity combining item-based similarity. In particular, the research profile-based similarity fused with users' favorite author information is the best similarity measure. Therefore, the relevant hypothesis H7.3 is accepted.



(a) The Precision Results

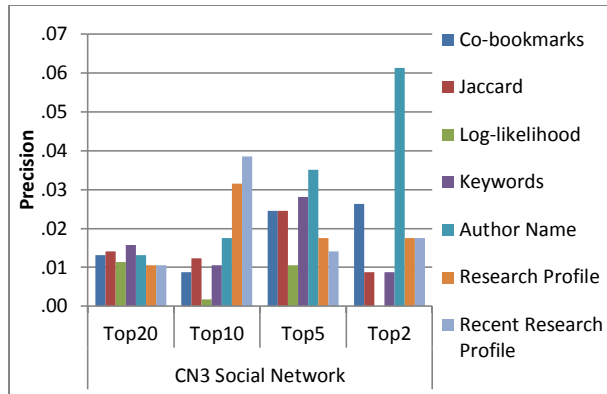
(b) The Recall Results

Figure 56. Results of the Test to Predict Collaborator Relations using Combined Measures (①= No. of Co-bookmarks; ② = Jaccard; ③= Log-likelihood; ④ = Keyword Vector; ⑤ = Author Name Vector; ⑥ = Recent Research Profile. The data series is displayed in the order of the legend)

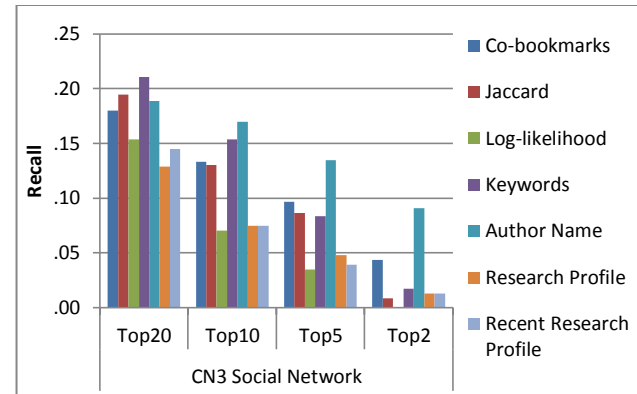
As the next analysis, I predicted CN3 social links for 57 target users. I computed the seven individual similarity measures and six combined similarity measures of the target users

with every other users and, in the same way as I did for the predictions of collaborators, I chose the top N similar peers of the target users. Among them, it is counted how many peers were actually CN3 social links. The comparison was also executed with increasing top N ranks such as top 20, top 10, top 5 and top 2.

The results of the predictions based on individual similarity measure are shown on Figure 57. The differences of the precisions were not significantly different depending on the kinds of similarity ($F = 0.39, p = .88$ for top 20, $F = 3.37, p = .003$ for top 10, $F = 1.05, p = .38$ for top 5, $F = 2.43, p = .03$ for top 2). In spite of the statistical insignificance, top5 and top 2 results indicated that author name vector-based similarity is better than the other individual measures. In this prediction of CN3 social networks, the differences between item-based similarities and metadata-based similarity weren't distinct. While item-based similarities performed worse than metadata-based similarities, in the highest rank, the number of co-bookmarks was the second best. The differences of recalls were not significantly different, either ($F = 0.42, p = .87$ for top 20, $F = 1.14, p = .34$ for top 10, $F = 1.32, p = .25$ for top 5, $F = 2.72, p = .013$ for top 2). However, top 10, top5 and top 2 results of recalls indicated that author name vector-based similarity performed the best. In this recall, the number of co-bookmarks performed the best among item-based similarities.



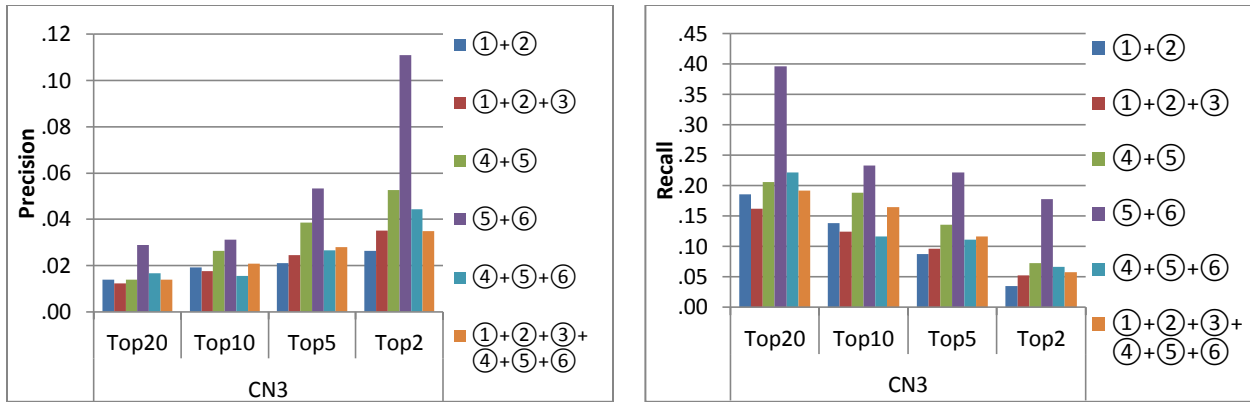
(a) The Precision Results



(c) The Recall Results

Figure 57. Results of the Test to Predict CN3 Relations using Individual Similarity Measures

Figure 58 shows the results collaborator predictions using combined similarity measures. The differences of the precisions were not significantly different depending on the kinds of similarity ($F = 2.22, p = .05$ for top 20, $F = .71, p = .614$ for top 10, $F = 1.0, p = .413$ for top 5, $F = 2.43, p = .19$ for top 2). In spite of the statistical insignificance, the results of all ranks indicated that the similarity combining research profile-based similarity and author name vector-based similarity is the best. The differences of recalls were not significantly different, either ($F = 1.81, p = .11$ for top 20, $F = .63, p = .34$ for top 10, $F = .90, p = .48$ for top 5, $F = 1.52, p = .18$ for top 2). However, top 10, top5 and top 2 results of recalls indicated that author name vector-based similarity performed the best. For CN3 social network, the combined similarities fusing item-based similarities were not good measure. For this reason, in order to present CN3 social connections' common interests, the similarity combining the research profile-based similarity with users' favorite author information is also the best similarity measure.



(a) The Precision Results

(b) The Recall Results

Figure 58. Results of the Test to Forecast CN3 Relations using Combined Similarity Measures (①= No. of Co-bookmarks; ② = Jaccard; ③= Log-likelihood; ④ = Keyword Vector; ⑤ = Author Name Vector; ⑥ = Recent Research Profile. The data series is displayed in the order of the legend)

7.3 EVALUATION OF COLLABORATION NETWORK-BASED RECOMMENDATIONS

This section aims to check how viable users' collaborators are as a useful information source in a full scale, through the recommendations generated by the collaborator's preferences. As mentioned, recommending conference talks is a special domain. Even though the final outputs of this recommendation are simply talks, the talks embed implicit metadata such as authorship, abstract, title, etc. These items can be extended to social context, as well. They are core objects which users' social interactions in a conference are usually initiated with and centered on. Put differently, conference attendees usually start their conversations with other attendees regarding

the papers they presented at the conference and share the related research interests through the papers. The users also have important metadata such as their publication records. Therefore, there are various aspects we need to consider in recommendations.

I considered three aspects – users’ preferences on information items, social network of users and content of items – and generated recommendations using one of these aspects or hybrid approaches fusing all of them. Figure 59 summarizes the space of the recommendation approaches considered in this part.

		Computation of Prediction Probability				
		Non-personalized	K-Nearest Neighbors		Matrix Factorization	Content-based
			Bookmark	Profile & Author Name		
Selection of Peers	Anonymous Peers	• Community	• CF • CFCW		• CFSVD • CFSVD_CW	• Profile • Content
	Collaborators		• SN_Coll • SNCW_Coll	• SNP_Coll • SNPCW_Coll	• SNSVD_Coll • SNSVD_Coll_CW	
	CN3 Social Links		• SN_CN3 • SNCW_CN3	• SNP_CN3 • SNPCW_CN3	• SNSVD_CN3 • SNSVD_CN3_CW	
	Collaborators + CN3 Social Links		• SN_All • SNCW_All	• SNP_All • SNPCW_All	• SNSVD_All • SNSVD_All_CW	

Figure 59. Overall Design of CN3 Talk Recommendation Algorithms

As explained in previous chapters, collaborative filtering (CF) systems typically perform in three steps – selection of peers, computation of prediction probability and suggestion of recommendations. In this part, I specified one more step – expression of users’ preferences. Therefore, the recommendations will be tested in the following step.

- Modeling users’ preferences
- Selection of peers

- Algorithms to compute prediction probability of each candidate item
- Selection of recommendations among candidate items

Modeling Users' Preferences

The recommendations based on K-Nearest Neighbors are generated using two different users' preferences. The results of the section 7.2.4 displayed that rather than item-based similarities, metadata-based similarities can be better similarity measure for both social networks. In particular, the combined similarity fusing users' recent research profiles and their favorite author information could be the best similarity measure for both kinds of social networks. Therefore, this comparison will check which one is a better preference between bookmark-based similarity (i.e. Jaccard coefficient) and the combined similarity.

Selection of peers

In previous section, we found that users not only share similar research interests but also favored similar talks with their collaborators. It seems that collaborators are useful sources to provide favorable talk information to our target users. In CN3 system, along with the collaboration network, there is CN3 social network of which main purpose is to share interesting talk information. The results of the above section showed that users' interests about talks are closer to their CN3 social connections than their collaborators, in spite of the collaborators' highly relevant research expertise and interests. Therefore, I will produce two kinds of recommendations based on each kind of social network and compare the performance. In addition, when users have both kinds of connections in CN3 system, do the recommendations based on the information of both kinds of social links perform better than the recommendations based one kind? In order to answer this question, the following hypotheses will be tested.

H7.4. Recommendations based on collaboration network perform better than the recommendations based CN3 social network.

H7.5. Recommendations based on both kinds of social network perform better than the recommendations based one kind of social network.

Algorithms to compute the prediction probability

Once the better method to model users' preference is decided, based on the resultant preference, I will choose the better algorithm to compute the prediction probability. As proposed, I tested the K-Nearest Neighbor approach and matrix factorization algorithm. By comparing the performance, better algorithm will be chosen.

Suggestion of recommendations

The recommendations of talks mainly based on bookmark records which reflect users' unary rating. When a user bookmarks a certain item, the action of bookmarking showed his interest on the item. When the user didn't bookmark another item, it doesn't necessarily represent he dislike it or has no interest on it. Due to the simple unary ratings, when some candidate items belong to the same set of peers, the recommendation probabilities of the items are exactly same, even though they are different items and contain different contents. It is because, in the choice of recommendable items, the recommendation probability is computed by aggregating the similarities of the peers who have the corresponding items. As one solution of this problem, this study added extra information in the final step of the recommendations. When choosing candidate information, content properties of each candidate item were considered.

As the final and the most critical hypotheses, I will test the following.

H7.6. Recommendations based on users' own social networks perform better than the traditional CF recommendations.

7.3.1 Recommendation Algorithms

7.3.1.1 Basic Recommendation Approaches

The first three recommendations take advantage of the three information sources, respectively – preferences of conference attendees, contents of conference talks and the attendee's social context information. While 'CF recommendations' are based on preferences of the whole population of the CN3 dataset regardless whether the users are socially associated with our target users or not, 'SN recommendations' are solely based on users' own social connections (i.e. their research collaborators and their CN3). 'Content-based recommendations' are based on the content properties of talks and the descriptions of users' interests.

As the first recommendation algorithm and the base line recommendations of this study, CF recommendations are based on all preferences of the whole population in the CN3 dataset. The CN3 system doesn't provide numeric rating mechanism when users bookmark conference talks. However, the system enables the users to express their interests on certain talks by bookmarking them. Hence, users' preferences of talks were encoded as binary ratings; 0 (i.e. no interest) or 1 (i.e. interest). For CF recommendations, I used the exactly same algorithm to the Citeulike recommendations. For the detailed recommendation algorithm, refer to the chapter 4 and the section 5.3.1. For CF recommendations using K-Nearest Neighbors approach (CF, hereafter), the Jaccard coefficient was used to compute bookmark similarities among users [99].

Based on this Jaccard coefficient, we picked the most like-minded 20 users who have the highest bookmark similarity with our target users. The like-minded users are called as top 20 anonymous peers. That is to say, the CF recommendations take into account the preferences of 20 the most similar anonymous users. After computing the similarities, we chose candidate items in peer users' bookmarks, which are not bookmarked by our target user. Then in order to select the most presumably favorable items, we aggregated the similarities of the peers to whom each candidate item belongs. In case of CF recommendations using matrix factorization (i.e. CFSVD; for detailed description of matrix factorization algorithm, refer to the chapter 4), I don't need to limit the top peers. Instead, it is required to determine the optimal number of factors.

As another baseline recommendation algorithms, content-based recommendation aims to recommend items which are similar to those users liked in the past in terms of contents of target items. Hence, the first stage of the recommendation is for the systems to comprehend what users liked in the past through users' bookmarked items. In the empirical evaluation of the recommendation, I used cross-validation evaluation strategy. That means users' partial bookmarks will be taken into account as a set of test items. By excluding the set of test items of the corresponding round, thus, I iteratively built keyword vector of each user's all bookmarks using the titles and abstracts for 10 times. All terms in the titles and abstracts of the bookmarked papers were aggregated into a bag of words. Then, after case normalization and stop words removal, the bag of words was pre-processed through Porter stemmer. As the next step, the bag of words was transformed into a keyword vector consisting of all terms and the TF-IDF values, as a user's keyword vector-based preference. In the same way, I also built keyword vector of each item (i.e. conference talks), having all terms appearing in the title and the abstract and the

TF-IDF values. Then, the similarity between the user profile and each item vector was computed using the cosine similarity. Using the resultant similarity, the system made a ranked list of candidate items (Lops et al. 2011). In the final selection of recommendation items, of course, the users' bookmarked items were excluded.

7.3.1.2 Collaboration Network-based Recommendations

The SN recommendations consider the attendee's social context information. For SN recommendations, how to generate the recommendations is almost the same with CF recommendations. However, one distinguishing feature of the proposed approach is that users' top anonymous peers were substituted by their social connections. In collaboration network-based recommendations (i.e. SN_Coll, SNCW_Coll, SNP_Coll, SNPCW_Coll, SNSVD_Coll and SNSVD_Coll_CW), the preferences of the target users' collaborators were only considered. After computing the similarities between our target users and their collaborators, candidate items were chosen among the collaborators' bookmarks, which are not bookmarked by the target user. When we suggest CN3 social network-based recommendations (i.e. SN_CN3, SNCW_CN3, SNP_CN3, SNPCW_CN3, SNSVD_CN3 and SNSVD_CN3_CW), the recommendations took into account only the bookmarks of the target users' all CN3 social networks. For the users who have both kinds of social networks, the recommendation taking into account both social links were produced as well. In this recommendation algorithm, their duplicated social links between two social networks were eliminated.

In the SN recommendations using K-Nearest Neighbors, it is critical to use a good similarity measure to pick the best set of the neighbor neighbors. According to the results of the

previous section 7.2, two preferences were tested for collaboration network-based recommendations – the bookmark-based preferences (i.e. Jaccard coefficient) and the combined similarity fusing users’ recent research profiles and their favorite author information. The similarity based on the bookmark-based preferences (i.e. SN, and SNCW) is to compute how much two given users (a target user and one of his colleagues) bookmarked the same items. The similarity based on the combined preferences (i.e. SNP and SNPCW) is to suggest how much two given users’ research profiles are alike and they favored the same authors’ talks. When I combine the recent research profile-based similarity with the author name vector-based similarity, the similarity values were normalized using the Standard Score (SS) like the following.

$$SS = \frac{x - \mu}{\sigma} \quad \text{Eq. 29}$$

where x is one similarity value, and μ and σ is the mean and standard deviation of the corresponding similarity measure, respectively. These two kinds of similarities have the same range of values, but the distributions of the similarities are different. Therefore, the normalized similarity values were summed together. Along with the K-Nearest Neighbor approach, I also calculated the collaboration network-based recommendations using the matrix factorization approach (i.e. SNSVD and SNSVD_CW). When the matrix factorization approach is used, for every target user, I built a separate sub-matrix only being made up of the target users’ bookmarks and his collaborators. In the same way as I did for collaboration network-based recommendations, the recommendations based on CN3 social network or both kinds of social networks are calculated.

7.3.1.3 Content Similarity Weights

In above CF recommendations, the recommender system computed the Jaccard similarity or combined similarity between a target user and his peers. In the choice of recommendable items, the similarities of the peers who bookmarked the corresponding items were aggregated and averaged out as the recommendation probability. When some candidate items belong to the same set of peers, the recommendation probabilities of the items are exactly same, even though they are different items and contain different contents. The same problem occurs in the SN recommendations. In order to resolve this problem, another source of information - content weights (i.e. CW) - was added at the final stage of recommendation (i.e. in the selection of recommended items) for CF and SN recommendations.

Once candidate items were selected from anonymous peers' or social connections' bookmarks, the system compared how the contents of these candidates are similar to a target user's keyword vector-based profile, which is built in the content-based recommendations. The content similarity between a target user and the candidate item was computed using the same way with the content-based recommendation. Each candidate paper's keyword-based vector was compared with user's keyword-based vector using the cosine similarity.

$$CFCW_{u,i} = \frac{\sum_{v \in p_u} Jaccard_{u,v}}{V} \times \frac{U \cdot I}{\|U\| \|I\|} \quad \text{Eq. 30}$$

Equation 30 shows CF recommendation using K-Nearest Neighbor and content similarity weight (i.e. CFCW) of items i for a target user u . For every candidate item i , the system calculated cosine similarity of user u 's keyword profile U and keyword vector of item I . This content similarity value is multiplied with the Jaccard similarity of user u with peer v , who is one

of user u 's top peers (p_u) and has the candidate item i in his bookmark. V is the number of peers who bookmarked the candidate item i .

$$CFSVD_CW_{u,i} = \frac{\sum_{v \in p_u} SVD_Prob_{u,v}}{V} \times \frac{U \cdot I}{\|U\| \|I\|} \quad \text{Eq. 31}$$

Equation 31 shows CF recommendations using matrix factorization algorithm and content similarity weight (i.e. CFSVD_CW) of item i for a target user u . In here, as the prediction probability, the resultant values of matrix factorization algorithm were used. SN-based recommendations using content similarity weight (i.e. SNCW and SNSVD_CW) are also similar. At the last stage of the recommendation, content of candidate items which were picked from the target user's social connections were compared with the target user's content profile.

7.3.1.4 Profile-based Recommendations

The results of the section 7.2.2 showed that users' research profiles are similar to their favorite items to a moderate degree. When their research profiles indicating the expertise and research interests are similar to their bookmarked items to a certain degree, is it useful to use the profiles as users' preferences for recommendations? In order to answer this question, profile-based recommendations (i.e. SNP and SNPCW) were computed. First, I selected all research talks presented at conference(s) which a target user attended, as candidate items. Then, I computed how the contents of these candidates (i.e. keyword-based vectors of the items which are computed in the content-based recommendations) are similar to the target user's recent research profile using the cosine similarity. The strength of this recommendation is that, once the system collects a target user's publication records, it is possible to generate the profile-based recommendations, without any users' input data.

7.3.1.5 Community Vote Recommendations

The last recommendation approach simply takes into account item popularity among conference attendees, so-called ‘community vote (i.e. Community)’. Conferences focus on one research topic, area or discipline. Hence, conference attendees tend to form a community around the topic or research area. If many attendees bookmarked a certain item, that item could have reasonably good quality appreciated by many researchers. Therefore, rather than considering individual user’s personal preferences, the system recommended top 10, 5 and 2 popular items per conference. Particularly, for each test user, his own bookmarked items were excluded, and the top N popular items for the conference he attended were suggested. The reason why these three specific N numbers were selected is that, for other aforementioned recommendation algorithms, the same N recommendations were generated.

7.3.2 Experimental Evaluation of Recommendations

For the evaluation of recommendation algorithms, I selected 203 target users who have either their colleagues or CN3 social connections in the system. 163 target users and 57 target users have colleagues and CN3 social connections, respectively. Among them, 17 users participate in both social networks. These target users 12.7 bookmarks on average ($\sigma = 12.4$).

In order to assess the proposed recommendations, this paper used the N cross-validation strategy (in this paper, the value N is 10). Among the whole collection of each user’s bookmarks, this strategy splits items into N sets randomly. It uses one set per iteration as a test set and the remaining sets as training set and generates recommendations. For the uses of which the

bookmark number is less than 10, I split their data into equal-sized set, even though the N is less than 10. To assess the quality of the recommendation, it was checked whether the test items appeared in the suggested recommendation or not. If recommendations contained test items, the recommendations was counted as hits (value of 1) and otherwise, counted as no-hit (value of zero). This process was iterated with a different test set for N times, and the recommendation quality was average out. In order to compute the accuracy of the recommendations, the number of hits scored from the lists of top 10, 5 and 2 recommendations was counted respectively.

Note that each recommendation approach generates a ranked list. Items with the higher ranks are expected to be more important than items in lower ranks. Therefore, it is critical to place correct recommendations in a higher rank. The position of correct recommendations is a critical evaluation criterion of recommendations. Therefore, precision at point N (precision@N) and Recall at point N (recall@N) are used as the evaluation criteria in this party.

7.3.3 Social Network-based Recommendations using Various Settings

Before testing whether collaboration network-based recommendations produce better quality than baseline CF recommendations, we need to find good settings for each recommendation approach. As explained, the recommendations can be generated through different strategies in every step of the way. They can be based on different kinds of user preferences, be generated by different groups of peers and different algorithms to compute the prediction probability and consider extra information such as content properties of each items at the final step to decide recommendable items. Therefore, before comparing the recommendation performance with other

kinds of recommendations, we need to determine good settings for the social recommendation algorithm.

First, how to compute the information similarity between our target users and their social partners is a critical determinant of the recommendation quality, especially when the recommendation algorithm is the K-Nearest Neighbors. The results of the chapter 7.2 showed that metadata-based similarity, specifically when users' recent research profiles are fused with the information about their favorite authors, is better to represent the shared interests between collaborators and CN3 social links than the preferences solely based on users' bookmark records. Therefore, before comparing the performance of collaboration network-based recommendations with typical CF recommendations, I test whether users' recent research profiles and their favorite author's information are really a good way to present users' shared interests and further to improve the recommendation quality. Therefore, as shown on Figure 60, the KNN algorithms are using two kinds of user preferences – one is based on the similarity of users' bookmark records and another is based on the similarity of users' recent research profiles and list of their favorite authors.

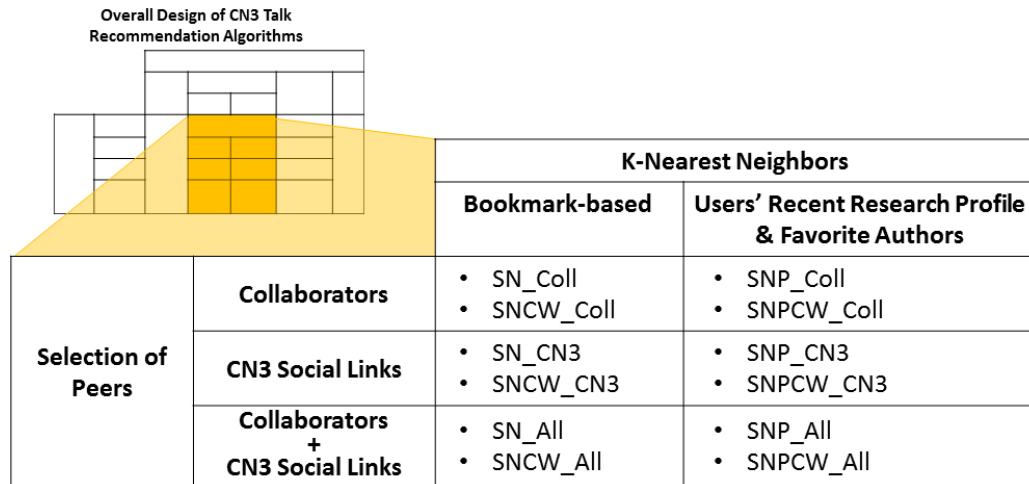


Figure 60. KNN Recommendations using Different Kinds of User Preferences

The statistical test to see the differences of the recommendation quality according to the kind of preferences showed that there is no consistently predominant preference. The Table 28 and Table 29 show the results and none of the differences were statistically significant. Therefore, I just determined the preferences yielding better precisions and recalls. According to the results, Figure 60 was revised as the following Figure 61.

		K-Nearest Neighbors
Selection of Peers	Collaborators	<ul style="list-style-type: none"> • SNP_Coll • SNCW_Coll
	CN3 Social Links	<ul style="list-style-type: none"> • SN_CN3 • SNCW_CN3
	Collaborators + CN3 Social Links	<ul style="list-style-type: none"> • SNP_All • SNCW_All

Figure 61. KNN Recommendations using Better User Preferences

Second, in order to increase the accuracy of the suggestions, recommendations can consider content property of each item at the final stage of the recommendation. For every kind

of recommendations, except non-personalized community and content-based approaches, I multiplied the content weight to the recommendation probability of every candidate items. As the results, adding content property is definitely helpful to enhance the recommendation quality. According to the results of Table 30 and Table 31, the differences of many recommendations are statistically significant regardless whether the quality was evaluated by precision or recall. Even for the recommendations of which the differences are not statistically significant, the suggestions with the content weights yielded higher mean precision or mean recall values.

Table 28. Differences of Recommendation Precision according to the Kind of Preferences

		Top10	Top5	Top2	
Collaborator-based	SN_Coll	3.08%	4.04%	6.11%	t = -1.05, p = .29 for top 10; t = -1.26, p = .209 for top 5; t = -0.18, p = .85 for top 2
	SNP_Coll	3.54%	5.00%	6.40%	
	SNCW_Coll	5.39%	8.40%	14.89%	t = 1.40, p = .16 for top 10; t = 1.00, p = .32 for top 5; t = 1.01, p = .31 for top 2
	SNPCW_Coll	4.59%	7.36%	12.53%	
CN3 Social Network-based	SN_CN3	3.63%	3.97%	4.00%	t = .32, p = .75 for top 10; t = .94, p = .35 for top 5; t = .39, p = .70 for top 2
	SNP_CN3	3.44%	3.33%	3.57%	
	SNCW_CN3	5.96%	8.83%	12.61%	t = .63, p = .53 for top 10; t = .44, p = .66 for top 5; t = .46, p = .65 for top 2
	SNPCW_CN3	5.43%	8.22%	11.27%	
Both Network-based	SN_All	3.09%	4.14%	5.28%	t = -1.26, p = .21 for top 10; t = -0.87, p = .38 for top 5; t = -0.40, p = .69 for top 2
	SNP_All	3.55%	4.68%	5.80%	
	SNCW_All	5.48%	8.22%	13.31%	t = 1.45, p = .15 for top 10; t = .96, p = .34 for top 5; t = .75, p = .45 for top 2
	SNPCW_All	4.78%	7.41%	11.88%	

Table 29. Differences of Recommendation Recall according to the Kind of Preferences

		<i>Top10</i>	<i>Top5</i>	<i>Top2</i>	
Collaborator-based	SN_Coll	23.81%	16.30%	10.22%	$t = -1.35, p = .18$ for top 10; $t = -1.41, p = .16$ for top 5; $t = -0.37, p = .71$ for top 2
	SNP_Coll	29.47%	21.64%	11.43%	
	SNCW_Coll	39.12%	31.03%	22.65%	
	SNPCW_Coll	35.23%	28.48%	19.27%	$t = .83, p = .41$ for top 10; $t = .58, p = .56$ for top 5; $t = .84, p = .40$ for top 2
CN3 Social Network-based	SN_CN3	21.25%	12.07%	5.71%	$t = .14, p = .89$ for top 10; $t = .54, p = .59$ for top 5; $t = .37, p = .71$ for top 2
	SNP_CN3	20.60%	10.52%	4.94%	
	SNCW_CN3	33.34%	26.83%	18.37%	$t = .38, p = .70$ for top 10; $t = .24, p = .81$ for top 5; $t = .36, p = .72$ for top 2
	SNPCW_CN3	31.32%	25.55%	16.40%	
Both Network-based	SN_All	22.35%	15.76%	8.35%	$t = -1.76, p = .08$ for top 10; $t = -1.29, p = .20$ for top 5; $t = -0.73, p = .47$ for top 2
	SNP_All	28.43%	19.78%	10.21%	
	SNCW_All	37.00%	28.90%	20.37%	
	SNPCW_All	34.82%	27.73%	18.37%	$t = .57, p = .57$ for top 10; $t = .33, p = .74$ for top 5; $t = .60, p = .55$ for top 2

Table 30. Differences of Recommendation Precision Depending on Content Properties of Items

	Top10	Top5	Top2	
CF	2.44%	2.55%	3.71%	$t = -5.90, p < .001$ for top 10; $t = -5.90, p < .001$ for top 5; $t = -4.14, p < .001$ for top 2
CFCW	4.64%*	5.89%*	8.78%*	
CFSVD	1.63%	1.66%	2.08%	$t = -2.20, p = .028$ for top 10; $t = -3.45, p = .001$ for top 5; $t = -3.21, p = .001$ for top 2
CFSVD_CW	2.28%	3.16%*	4.39%*	
SNP_Coll	3.54%	5.00%	6.40%	$t = -3.63, p < .001$ for top 10; $t = -3.72, p < .001$ for top 5; $t = -4.21, p < .001$ for top 2
SNCW_Coll	5.39%*	8.40%*	14.89%*	
SNSVD_Coll	4.71%	7.12%	10.75%	$t = -0.71, p = .47$ for top 10; $t = -1.13, p = .25$ for top 5; $t = -1.80, p = .07$ for top 2
SNSVD_Coll_CW	5.13%	8.31%	15.11%	
SN_CN3	3.63%	3.97%	4.00%	$t = -3.50, p = .001$ for top 10; $t = -4.61, p < .001$ for top 5; $t = -4.15, p < .001$ for top 2
SNCW_CN3	5.96%*	8.83%*	12.61%*	
SNSVD_CN3	4.92%	5.74%	7.69%	$t = -0.43, p = .66$ for top 10; $t = -2.15, p = .03$ for top 5; $t = -2.39, p = .019$ for top 2
SNSVD_CN3_CW	5.25%	8.39%	14.02%	
SNP_All	3.55%	4.68%	5.80%	$t = -4.45, p < .001$ for top 10; $t = -4.68, p < .001$ for top 5; $t = -4.59, p < .001$ for top 2
SNCW_All	5.48%*	8.22%*	13.31%*	
SNSVD_All	4.81%	6.21%	9.47%	$t = -0.69, p = .49$ for top 10; $t = -2.48, p = .014$ for top 5; $t = -2.57, p = .010$ for top 2
SNSVD_All_CW	5.15%	8.24%	14.32%	

(* indicates that the differences are statistically significant)

Table 31. Differences of Recommendation Recall Depending on Content Properties of Items

	Top10	Top5	Top2	
CF	17.42%	9.45%	6.06%	t =-5.23, p < .001 for top 10; t =-4.84, p < .001 for top 5; t =-3.35, p < .001 for top 2
CFCW	32.55%*	22.06%*	14.07%*	
CFSVD	11.16%	5.46%	2.84%	t =-2.14, p = .033 for top 10; t =-3.55, p < .001 for top 5; t =-2.95, p =.003 for top 2
CFSVD_CW	16.10%	11.75%*	6.37%	
SNP_Coll	29.47%	21.64%	11.43%	t =-2.10, p < .001 for top 10; t =-2.21, p < .001 for top 5; t =-3.01, p = .001 for top 2
SNCW_Coll	39.12%*	31.03%*	22.65%*	
SNSVD_Coll	33.52%	26.31%	16.26%	t =-0.78, p =.43 for top 10; t =-0.99, p = .33 for top 5; t =-1.58, p = .116 for top 2
SNSVD_Coll_CW	37.01%	30.43%	22.62%	
SN_CN3	21.25%	12.07%	5.71%	t =-2.61, p = .010 for top 10; t =-3.60, p < .001 for top 5; t =-3.24, p = .002 for top 2
SNCW_CN3	33.34%	26.83%*	18.37%	
SNSVD_CN3	28.92%	17.61%	10.09%	t =-0.13, p = .89 for top 10; t =-1.49, p = .14 for top 5; t =-1.97, p = .052 for top 2
SNSVD_CN3_CW	29.56%	24.75%	18.89%	
SNP_All	28.43%	19.78%	10.21%	t =-2.27, p = .024 for top 10; t =-2.62, p = .009 for top 5; t =-3.32, p = .001 for top 2
SNCW_All	37.00%	28.90%	20.37%*	
SNSVD_All	32.93%	22.49%	13.96%	t =-0.53, p = .59 for top 10; t =-1.85, p =.066 for top 5; t =-2.29, p =.022 for top 2
SNSVD_All_CW	34.84%	28.67%	21.26%	

(* indicates that the differences are statistically significant)

Lastly, I compared the recommendations based on K-Nearest Neighbors with the recommendations generated by matrix factorization algorithm. Since the approaches with content weights produced always better results, the recommendation quality was compared on the center of those approaches. Before this comparison, the matrix factorization with an optimal setting was selected (refer to Appendix A.3).

Table 32 and Table 33 show the results. The t-test statistical test showed that K-Nearest Neighbor worked significantly better than matrix factorization approach for CF recommendations. However, for other social recommendations, the differences were not significant and distinct in precision and recall. For instance, for the recommendations based on

collaborators, the precision result showed that matrix factorization (SNSVD_Coll_CW) is better than K-Nearest Neighbor in the highest rank. However, the recall result showed the opposite. Therefore, additionally, I computed F-1 measure to determine the better approach. For all three kinds of social recommendations, recommendations generated using matrix factorization algorithm performed better than K-Nearest Neighbor approach. By following from the results of the best settings up to now, the overall design of cn3 talk recommendation algorithms (i.e. Figure 59) was modified as Figure 62.

		Computation of Prediction Probability		
		Non-personalized	Bookmark-based	Content-based
Selection of Peers	Anonymous Peers	• Community	• CFCW	• Profile • Content
	Collaborators		• SNSVD_Coll_CW	
	CN3 Social Links		• SNSVD_CN3_CW	
	Collaborators + CN3 Social Links		• SNSVD_All_CW	

Figure 62. Modified Design of CN3 Tal Recommendation Algorithms

Table 32. Differences of Recommendation Precision Depending on Probability Computation Algorithms

	Top10	Top5	Top2	
CFCW	4.64%*	5.89%*	8.78%*	t =6.67, p < .001 for top 10; t =4.77, p < .001 for top 5; t =3.85, p < .001 for top 2
CFSVD_CW	2.28%	3.16%	4.39%	
SNCW_Coll	5.39%	8.40%	14.89%	t =.43, p =.67 for top 10; t =.08, p = .93 for top 5; t =-0.08, p= .93 for top 2
SNSVD_Coll_CW	5.13%	8.31%	15.11%	
SNCW_CN3	5.96%	8.83%	12.61%	t =.95, p =.34 for top 10; t =.34, p = .74 for top 5; t =-.50, p= .62 for top 2
SNSVD_CN3_CW	5.25%	8.39%	14.02%	
SNCW_All	5.48%	8.22%	13.31%	t =.70, p =.49 for top 10; t =-.025, p = .98 for top 5; t =-.52, p= .60 for top 2
SNSVD_All_CW	5.15%	8.24%	14.32%	

(* indicates that the differences are statistically significant)

Table 33. Differences of Recommendation Recall Depending on Probability Computation Algorithms

	Top10	Top5	Top2	
CFCW	32.55%	22.06%	14.07%	t =5.83, p < .001 for top 10; t =4.08, p < .001 for top 5; t =3.62, p < .001 for top 2
CFSVD_CW	16.10%	11.75%	6.37%	
SNCW_Coll	39.12%	31.03%	22.65%	t =.46, p = .65 for top 10; t =.14, p = .89 for top 5; t =.007, p = .994 for top 2
SNSVD_Coll_CW	37.01%	30.43%	22.62%	
SNCW_CN3	33.34%	26.83%	18.37%	t =.79, p = .43 for top 10; t =.42, p = .67 for top 5; t =-0.10, p = .92 for top 2
SNSVD_CN3_CW	29.56%	24.75%	18.89%	
SNCW_All	37.00%	28.90%	20.37%	t =.60, p = .55 for top 10; t =.07, p = .95 for top 5; t =-.26, p = .79 for top 2
SNSVD_All_CW	34.84%	28.67%	21.26%	

7.3.4 Comparison of Social Recommendations with Typical Approaches

Once the optimal setting for each recommendation approach is decided, it is the stage to compare the quality of social recommendations with typical approaches including CF recommendations and content-based recommendations. First, Figure 63 and Figure 64 display the precision and recall results, respectively. The quality of these seven recommendation approaches are significantly different ($F = 8.65, p < .001$ for top 10 precision; $F = 8.37, p < .001$ for top 5 precision; $F = 6.10, p < .001$ for top 2 precision; $F = 8.65, p < .001$ for top 10 recall; $F = 8.37, p < .001$ for top 5 precision; $F = 6.10, p < .001$ for top 2 precision).

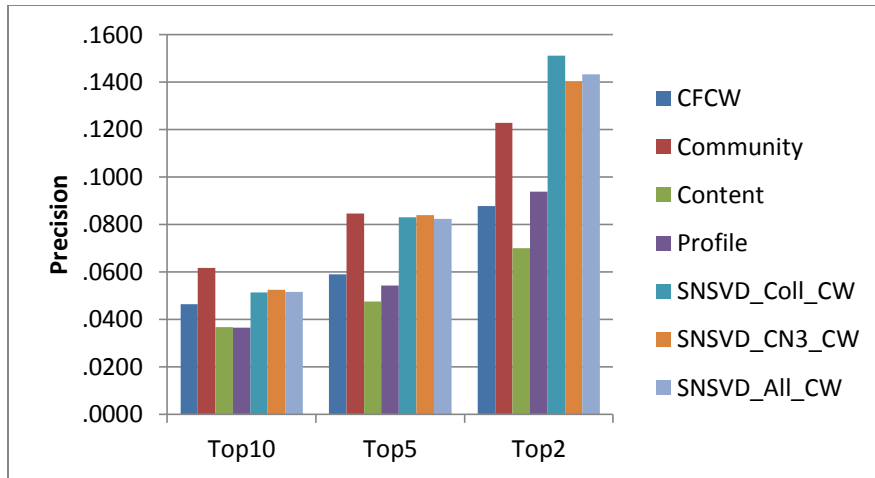


Figure 63. Differences of Recommendation Precision Depending on the Approaches

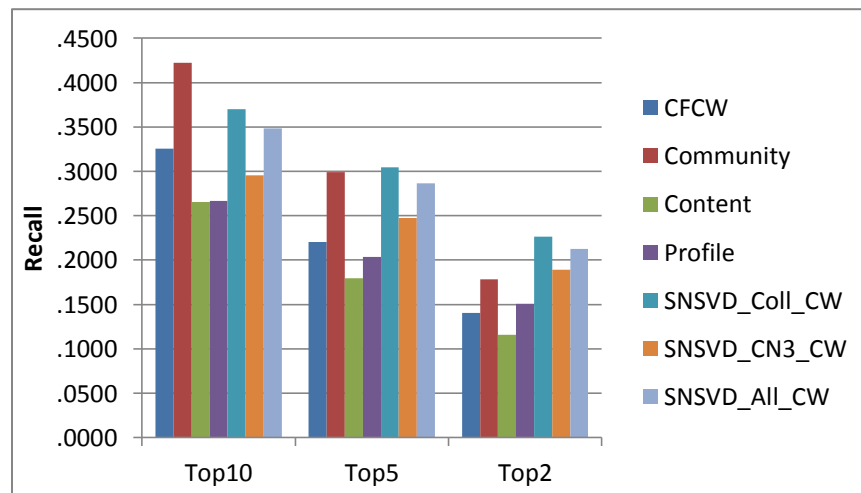


Figure 64. Differences of Recommendation Recall Depending on the Approaches

For the specific patterns of the differences, I executed the Schaff é pairwise comparison. According to the results, CF and content-based recommendations, which are considered as baseline recommendations performed the worst. Interestingly, in lower rank (i.e. top 10), the quality of the recommendations utilizing users' recent research profiles were as bad as the

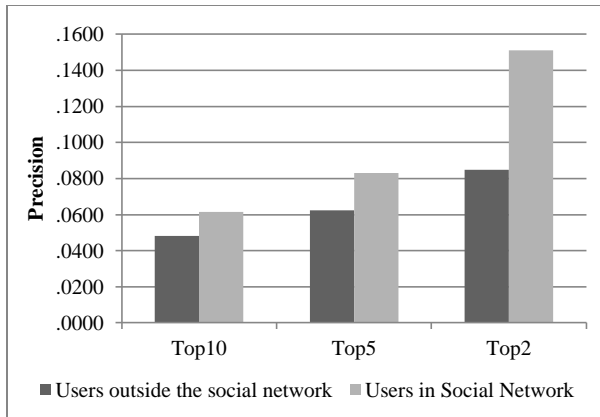
baseline recommendations. However, as the rank is getting higher, the quality became better and better. Thus, in the top 2 results, the profile-based recommendation performed significantly better than baselines. However, other two kinds of recommendations – community and social recommendations – were significantly better than these baseline approaches and the profile-based approach. Therefore, the hypothesis 7.6 is accepted. The recommendations based on users' own social networks performed better than the traditional CF recommendations.

In social recommendations, whatever kinds of social networks the recommendations utilized, there was no significant difference. Put differently, whether the recommendations are based on the collaboration network (i.e., SNSVD_Coll_CW), CN3 social network (i.e., SNSVD_CN3_CW) or both of them (i.e., SNSVD_All_CW), the qualities were similar level. Therefore, the relevant two hypotheses H7.4 and H7.5 are rejected. Despite of the insignificant difference, according to the F-1 measure, for all ranks, the recommendations based on the collaboration network (F-1 value of the top 10 is 8.71%; F-1 value of the top 5 is 12.50%; F-1 value of the top 2 is 17.29%) tend to yield better suggestions than the recommendations based on CN3 social network (F-1 value of the top 10 is 8.54%; F-1 value of the top 5 is 11.82%; F-1 value of the top 2 is 15.13%) or both social networks (F-1 value of the top 10 is 8.63%; F-1 value of the top 5 is 12.16%; F-1 value of the top 2 is 16.26%). Another interesting result is that in the lower rank, the non-personalized recommendation – community vote-based recommendations – performed the best (for top 10) or as good as the social recommendations (for top5). On the other hand, in the top 2 result where the most accurate recommendations should be placed, all three kinds of social recommendations significantly outperformed the community vote-based recommendations. This pattern is identical for both precision and recall. That means that in order

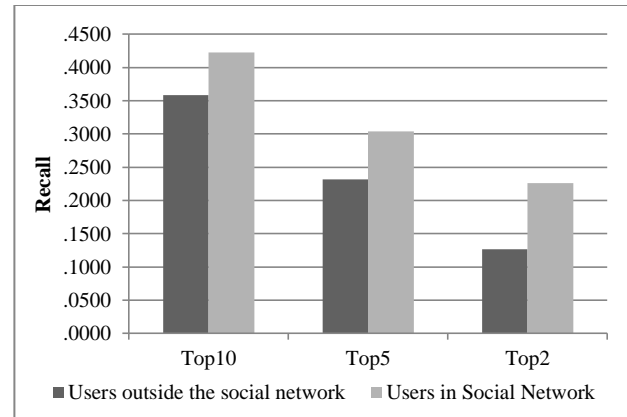
to generate recommendations for conference talks, the suggestions based on users' own social connections (whether the social connections are collaborators or CN3 social connections) is the best option to consider. In addition, unexpectedly good performance of community vote-based recommendations is also impressive. This result can be interpreted to mean that conference attendees understand conference talks in the similar manner. Because they usually have background knowledge about the conference topics or areas, they have insights to discover interesting or good papers. Thus their insights are helpful to other attendees when their opinions were aggregated as a form of collective intelligence (i.e. community vote-based recommendations).

The quality of social recommendations was tested for the 203 users who are participating at least one kind of social network. Then, is being a part of the social networks beneficial to get better recommendations? For this comparison, for the remaining 251 users who don't engage in any social network, I computed four kinds of recommendations – CF with content weight (i.e. CFCW), community vote-based, content-based and their profile-based recommendations. Between the two groups of users, according to the t-test with a 0.001 significance level, the numbers of bookmarks are not significantly different ($t = .013$, $p < .001$). Users engaging in social network have 12.63 bookmarks ($\sigma = 12.40$) and users not in engaging in any social network have 10.06 ($\sigma = 8.77$) on average. Among various recommendation approaches, different kinds of recommendation were the best for these two groups of users. The above showed that, for lower rank, the community vote-based recommendations were the best for the users in social networks. Notwithstanding this result, in higher ranks, social recommendations were better than community-based approaches. For the users outside of the social network, as expected,

community vote-based recommendation approach is always significantly the best for the precision and recall-based evaluations ($F = 8.26, p < .001$ for top 10 precision; $F = 5.14, p = .002$ for top 5 precision; $F = 2.98, p = .031$ for top 2 precision; $F = 7.21, p < .001$ for top 10 recall; $F = 3.98, p = .008$ for top 5 precision; $F = 1.61, p = .185$ for top 2 precision). Then the best approach for users in social network was compared with the community vote-based recommendations of the users outside of the social networks. In case of the top 10 results, since the community vote-based recommendations was the best option for both groups of users, naturally, there is no significant difference between two users ($t = -3.26, p = .002$ for the precision; $t = -2.39, p = .017$ for the recall). For higher ranks (i.e. top 5 and top 2), as explained, the social recommendation approach based on users' collaborators was the best option for the users having social connections and the community vote-based recommendation was still the best top option for the users not participating any social network. Hence, I compared the quality differences of two recommendation approaches according to their social status. Figure 65 shows the results. In case of the top 5 result, there is no significant difference between collaborator-based recommendations for the users in social network and community vote-based recommendations for the users outside of the social networks ($t = -2.66, p = .008$ for the precision; $t = -2.45, p = .015$ for the recall). However, in the highest rank, the recommendations of the users in social network yielded significantly better suggestions ($t = -4.11, p < .001$ for the precision; $t = -3.82, p < .001$ for the recall). We can interpret this result to mean that being engaging in social networks is more beneficial to receive more accurate and better information recommendations.



(a) Comparison of Precision



(b) Comparison of Recall

Figure 65. Comparison of Recommendation Quality according to Users' Social Status

7.3.5 Collaboration Network-based Recommendations for Cold-start Users

Lastly, in spite of the good popularity, one well-known shortcoming of CF is cold-start user problem. Until users rated or bookmarked sufficient number of items, it is hard for them to get any CF recommendations or to receive reasonable quality of recommendations. We can think about several solutions about this problem. First available solution is to utilize their research profile. Research profile is built from users' own publication records, thus without any users' input, the system already know what they may be interested in. Second, when the information objects have various content or metadata information, content-based recommendations are available even with a small number of bookmarks or ratings. The last solution is to take advantage of users' social connections. A series of my own studies [93, 99, 95] have provided some empirical evidences that adding online friends is more helpful for cold start users to receive good recommendations rather than adding a few more bookmarks or ratings. Therefore, to cope

with this cold-start user problem, several recommendations will be assessed. I compared the precisions and recalls of each recommendation depending on how many items each user bookmarked. First users were divided into three groups: cold start users ($n \leq 4$), users having moderate number of items ($n \leq 15$) and users having relatively larger number of items ($n > 15$). 63, 73 and 67 users belong to each group, respectively.

Figure 66 and Figure 67 displays the differences of recommendation precisions and recalls according to the user groups. Two way ANOVA test demonstrated that the differences on the recommendations precisions according to different recommendation approaches were not statistically significant among the user groups ($F = 2.49, p = .003$ for the top10 precision; $F = 2.13, p = .013$ for the top 5 precision; $F = 2.45, p = .004$ for the top2 precision; $F = 1.89, p = .032$ for the top10 recall; $F = 2.12, p = .014$ for the top 5 recall; $F = 2.51, p = .003$ for the top2 recall). However, within each group of users, the differences of recommendation precisions were significantly different according to the kinds of recommendation approaches ($F = 9.056, p < .001$ for the top10 precision; $F = 8.43, p < .001$ for the top 5 precision; $F = 7.62, p < .001$ for the top2 precision; $F = 7.73, p < .001$ for the top10 recall; $F = 7.06, p < .001$ for the top 5 recall; $F = 6.61, p < .001$ for the top2 recall). All social recommendations, regardless of the social network kinds, outperformed other kinds of recommendations in terms of both precision and recall. Interestingly, the next best-performing approach for these cold-start users is CF recommendations. However, other content-based, profile-based and community vote-based recommendations didn't perform well. In particular, for cold-start users who have insufficient bookmarks, the community votes, which are made up of collective intelligence of the conference attendees and produced generally good results for other two groups of users, were not useful.

Instead, cold-start users tend to follow their social partners' opinions when looking for interesting papers. For other two groups of users, nonetheless, community vote-based recommendations produced the best or second best suggestions depending on the ranks. Therefore, for richer users who have sufficient number of bookmarks, the opinions of the majority are important source of information. For these users, the social recommendations were the second best.

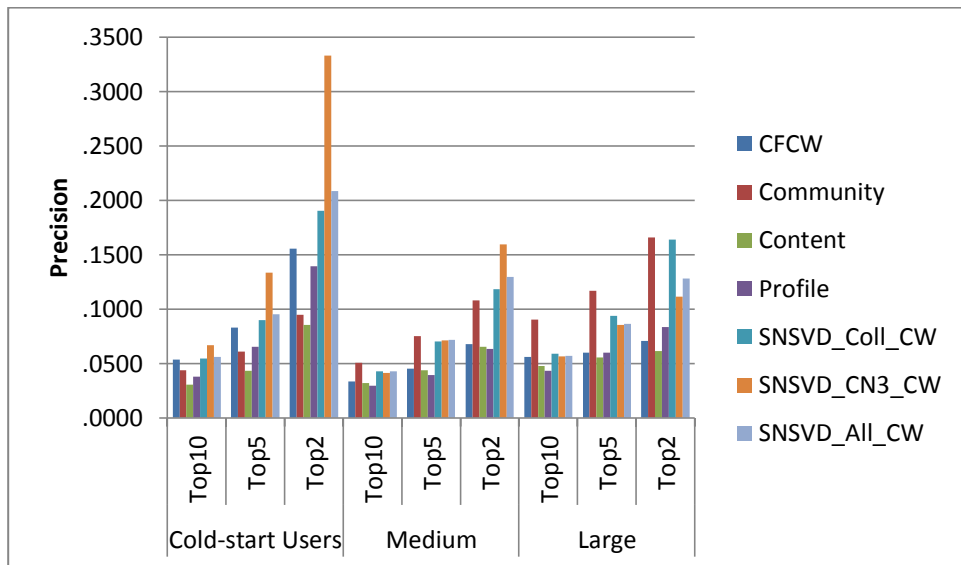


Figure 66. Differences of Precision Depending on Users' Number of Bookmarks

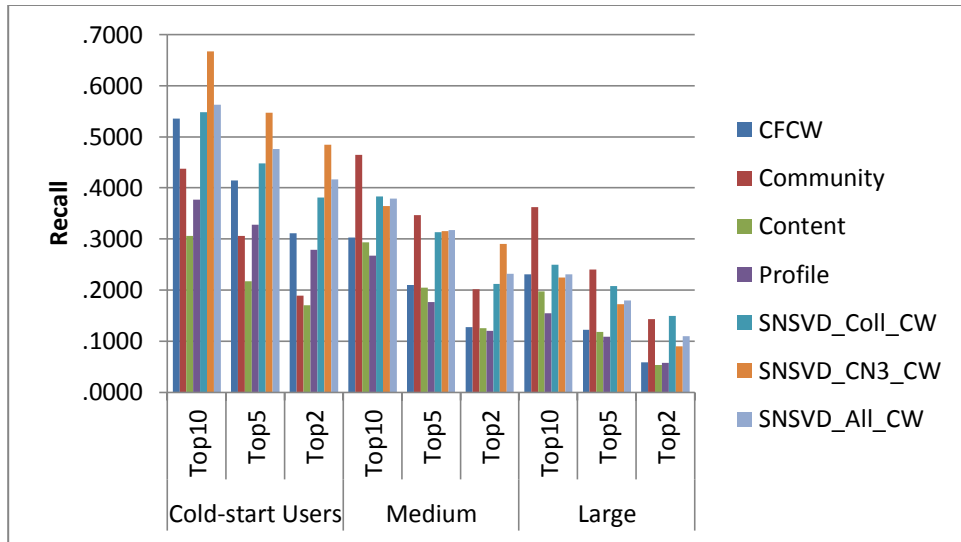


Figure 67. Differences of Recall Depending on Users' Number of Bookmarks

7.4 CONCLUSION

This chapter was to explore patterns about how research collaborators share information and how to generate users' recommendations based on their collaborators' preferences. This chapter consists of two parts – 1) the examination of shared interests among collaborators and 2) the empirical evaluation of collaboration network-based recommendations. The target information objects of this study were conference talks. Compared with other product recommendations, at conferences, the time to collect data and the number of users are severely limited. In order to cope with this unusual context, it requires harmonizing information items with social context in recommendations. Research collaborators share overlapped research interests and topics. Hence, when researchers attend conferences with their collaborators, talks caught their attentions could

also interest their collaborators. To our best knowledge, this is the first study related to the problem of recommending conference talks.

In the first part to examine shared interests among collaborators, the information similarity between collaborators is higher than other pairs who are not socially associated. However, the similarity was lower than another kind of social network, CN3 social network. Because the collaborators are social connections built around their research achievement, the collaborators' similarities based on their publications were higher than the other users. Except this similarity about users' own publications, users shared less common information with their collaborators than their CN social connections. The analysis to find the optimal measurement to show the collaborators' relevant interests demonstrated that users' research profiles consisting of their own publications are the most effective. This research profile-based similarity is also an effective way to represent the common interests of CN3 connections.

In the second part, the usefulness of the collaboration network as an information source was tested by utilizing the social connections as a component of the recommendations. The evaluation using the 10 cross-validation test, I found that social network-based approach always performed the best regardless whether the evaluation was done by precision or recall. While the social network-based recommendations performed better than traditional CF recommendations, the recommendations based various users' social networks generated equally good suggestions. Put differently, whether the base of the recommendations is collaborators, CN3 social connections or a combined set of both social networks, the social recommendations were equivalently good. In addition, community vote-based recommendations were the second best. Even though the community vote-based recommendations couldn't produce as accurate and

complete as the social recommendations are, I interpreted the good performance of the community vote-based approach to mean that majority of conference attendees understand conference talks/papers in the similar manner as a research community. For the cold-start users who have insufficient number of bookmarks to receive reasonable recommendations, the social network-based recommendations were always the most effective way to personalize the information. On the other hand, the community vote-based approach wasn't helpful for the cold-start users. Cold-start users don't have enough bookmarks to follow and form the votes of the majority. Conclusively, when generating personalized recommendations of conference talks, the most critical source to rely on is users' own social connections.

8 CONCLUSION AND DISCUSSION

This dissertation aims to check the value of users' various online social connections as information sources and to explore how to include the social connections as a foundation for personalized recommendations. In spite of the widely available and diversified online sociability, most recent social network-based recommendations have concentrated on limited kinds of online sociality (i.e., trust-based networks and online friendships). Thus, this study tried to prove the expandability of social network-based recommendations to more diverse and less explored social networks. The online social networks considered in this dissertation include: 1) a watching network, 2) a group membership network, and 3) an academic collaboration network (i.e. co-authorship network). The preceding three chapters described the findings regarding these three kinds of online social networks, respectively. Particularly, each chapter was made up of two parts – the feasibility of social networks as a useful information source and the evaluation of social network-based recommendations.

The first part was to understand how users' social network structure is correlated to their information collections, to precede an investigation regarding personalized recommendations based on the social networks. This first part was designed to find answers to the research questions proposed in the introduction, which were the following:

- **Q1.** Compared with pairs of users who are not socially associated, do socially connected users share more common interests?
- **Q2.** What is a proper measurement for information similarity between two socially connected people?
- **Q3.** Do users have an equivalent degree of common interests with their social connections as they do with their top N anonymous peers, which are chosen by conventional CF recommendations?

The second part then goes on to develop and evaluate recommendation algorithms using various social networks. The viability of multiple online social networks was demonstrated as a means to gather useful information and as having potential for developing more diverse social network-based recommendations. Thus, this part was designed to find the answers to the following research questions proposed in the introduction:

- **Q4.** Do social network-based recommendations produce better suggestions than traditional CF recommendations?
- **Q5.** For the given three types of social network, are all of them useful sources of information?
- **Q6.** Can SN recommendations solve the cold-start user problem?

The research results were summarized and concluded as responding to the research questions to address how each question was answered and what were the major findings of this study.

8.1 FINDINGS OF RESEARCH QUESTIONS

8.1.1 Findings of the First Part: Feasibility of Social Networks as Useful Information

Source

Q1. Compared with pairs of users who are not socially associated, is it true that socially connected users share more common interests?

As explained, three types of online social networks were investigated in this dissertation. For all types, our target users shared common interests with their self-defined social partners. In particular, the information similarity of direct watching connections was the largest and the similarities decreased incrementally along with an increase in social distance. The similarity between two co-members of the same group was also larger than random pairs. Group members, however, share more resembling interests with their group library. The group library is a communal space collaboratively composed by group members. That means that the group library is a more coherent collection of information relevant to the topic of the corresponding group. Therefore, group members who are interested in the group's topic tend to refer to their group library more than the individual spaces of group co-members. In addition, when a user participates in both watching network and group membership network, he tends to share the most similar interests with his group library followed by his watching connections. The user didn't share common interests with his co-members of the same group, as much as with the group library and watching connections.

Information similarities between two collaborators are higher than randomly coupled pairs as well. However, the similarity was lower than another kind of social network that exists

on the CN3 system (i.e., CN3 social network), that primarily aims to share useful information. Because the collaborator relations were built around their achievements, i.e., their own publications, similarity in research profiles was higher than with another CN3 social network. However, from the perspective of shared bookmarks and common metadata, the similarities between collaborators were lower than the similarities between the CN3 social connections.

Q2. What is a proper measurement for information similarity of two socially connected people?

In this study, two data sources were used. For the watching network and group membership, Citeulike data source was used. For the collaboration network, CN3 data source was used. Both Citeulike and CN3 data consist of users' bookmark records and metadata information of items. Since the target objects of both data sources were academic papers, the available metadata information consisted of titles, author names, abstracts, publication journal/conference names, and the most importantly, users' social tags. When calculating information similarity solely based on bookmark records, the calculation of the similarity between our target users and their peers (whether the peers are anonymous or social connections) simply relies on what they like. Therefore, in order to model users' preferences in a more sophisticated format and to better understand the social interactions around users' information spaces, it requires to consider more varying aspects, such as why users are interested in certain items, what are the contents of users' favorite items and authors of the favorite items. In the analysis based on the Citeulike and the CN3 data source, the effectiveness of the various similarity measures to indicate the object-centered sociality was investigated.

For the watching connections based on Citeulike data source, the number of common items and metadata information of items such as keywords derived from titles and users' social

tags are good ways to represent the shared interests between two social connections. Rather than those individual similarity measures, the combination of these well-working measures was the best way to represent the common interests between watching connections. In case of the group membership network based on the same Citeulike data source, metadata-based similarity was significantly better than all kinds of item-based similarity. Like the results of the watching network, the combinations of the best individual measures (such as all four kinds of metadata-based similarities) performed best. However, the performance of the best individual measure – the keyword vector-based similarity – was as good as the combination. The analysis on the collaboration network also showed a better performance of metadata-based similarity than the similarity using bookmark records. In particular, collaborator’ research profiles and author names of their favorite items were the most well-working individual measures, and the combination of these two measures was the best measure to present the shared interests of collaborators, and also the common interests of CN3 social connections.

To conclude, in the data sources consisting of users’ bookmark records, along with the bookmark records, the metadata of items is critical, or sometimes better, information to measure the shared interests of online social networks.

Q3. Do users have an equivalent degree of common interests with their social connections as they do with their top N anonymous peers, which are chosen by conventional CF recommendations?

The main purpose of this dissertation is to test whether users’ various online social networks are substitutable for their top N anonymous peers. Actually, in the second part of this dissertation, this is addressed in a comprehensive manner. However, in the first part, as a preceding and

simpler test, the information similarity of social networks was compared to the similarity of top N anonymous peers. The results showed that the similarity of social networks was lower than the similarity of top N anonymous peers. When counting the similarity based on metadata, users in a watching relation have more common interests than their anonymous peers. However, this result is only unique for this watching network. For the other social networks, users' shared interests with their social connections were less common than top N anonymous peers. However, this analysis is based on a simple comparison of information similarities. A more in-depth and sophisticated examination about how users' social networks are substitutable for the top N anonymous peers was explored in the second part of this dissertation.

8.1.2 Findings of the Second Part: Evaluation of Social Network-based Recommendations

Q4. Do SN recommendations produce better suggestions than traditional CF recommendations?

The main purpose of this dissertation is to test whether users' self-defined social networks have a comparable value to their anonymous peers as a useful information source and whether the social connections are substitutable for their anonymous peers picked by the fully-automated CF-based computation. For this test, the recommendations based on users' social networks were compared with CF recommendations.

The results showed that it is worthwhile to utilize users' self-defined social networks as an alternative or a complementary mechanism to anonymous peers. According to the test of the recommendations based on various social networks, the social networks have better or sometimes equivalent quality as a useful information source. That means, the recommendations

based on online social networks performed better than CF recommendations, and according to the domain, the performance of the social network-based recommendations was as good as CF recommendations.

In terms of how the suggestions are accurate (i.e. precision) and complete (i.e. recall), the watching network-based recommendations have equivalent quality to the CF recommendations. In addition, when combining users' watching connections with their anonymous peers, the hybrid recommendations significantly improved at least regarding the evaluation base on the completeness (i.e. recall). The social recommendations based on group membership network generated quality suggestions equivalent to or sometimes better than the CF recommendations. Specifically, in the results of precision, the quality of group membership-base recommendations was not significantly different with the traditional CF recommendations. However, the differences of the recall demonstrated that group membership-based recommendations significantly outperformed the CF recommendations. That is, the recommendations based on group membership-based social links can produce as accurate suggestion as CF recommendations can and more complete suggestions than CF approach. Finally, collaboration network-based recommendations always performed best regardless whether the evaluation was done by precision or recall in all ranks.

Q5. For the given three social networks types, are all of them useful sources of information?

The answer to this question is yes. All three types of social networks are useful sources of information and when we take into account the social networks as a part of the recommendations, the quality of recommendations was better than or as good as conventional CF recommendations. In particular, watching network-based recommendations are comparable to the CF

recommendations. In addition, when users' watching connections were fused with their top N anonymous peers, the recommendations were significantly improved. Therefore, users' watching connection alone has value as a useful information source and further in hybrid recommendations, it has a complementary ability to CF recommendations. Next, group-based social connections are also useful sources of information. The social recommendations based on group membership network generated quality suggestions equivalent to or sometimes better than the CF recommendations. In group membership-based recommendations, group co-members and group's communal spaces are equally important information sources. However these group membership-based connections don't have complementary power to the CF recommendations like the watching connections. The hybrid recommendations combining users' group membership network with their anonymous peers were ineffective. Finally, collaboration network is also important information source. The collaboration network-based recommendations always outperformed the CF recommendations. Particularly, the results of the social recommendations indicated that collaborators, CN3 social connections or a combined set of both social networks were equally useful information sources.

Since each social network is based on different data source and the recommendations were generated for different target users, it was impossible to compare the quality of the social recommendations network-by-network. In studies such as those based on Citeulike, however, since the system provided two kinds of object-centered sociality (i.e. watching network and group membership), I was able to compare the quality of recommendations based on the two kinds for the same target users. The result showed that combining two kinds of object-centered sociality wasn't effective in enhancing the recommendation quality. Regardless whether the

recommendations were generated by one kind of social network or two kinds of social networks, the predictions were at an equivalent level. Therefore, the watching connections and group-based social connections are equally useful information sources.

Q6. Can social network-based recommendations solve cold-start user problem?

First and foremost, the answer is yes. For all social network types and data sources considered in this study, the recommendation algorithm that worked best for the cold-start users was the social network-based recommendations. The social network-based recommendations were always the most effective for personalizing the cold-start users' information. Therefore, we can conclude that in order for cold-start users to receive reasonably good recommendations, it is more effective to be socially associated with other users, rather than collecting a few more items.

8.1.3 Conclusion

With the proliferation of online social networking applications and the abundant Web-based sociability, this study explored the expandability of social network-based recommendations to more diverse kinds of online social networks, especially having high degree of object-centered sociality and greatly information centered. This study investigated three different types of online sociality and two kinds of data sources. The results showed that the online social connections were built based on shared interests and thus have a significant value as a legitimate source for users to acquire interesting information. This study demonstrated that different social networks of a novelty value can improve conventional personalized technology. Watching network-based recommendations are as good as the CF recommendations. The group membership-based recommendations and collaboration network-based recommendations enhanced the quality of the

personalization of users' information spaces over the traditional collaborative filtering approach. This dissertation contributed to the body of knowledge by comprehensively identifying the nature of various social networks relating to information similarity among members of the network.

It also demonstrated how these less focused and information-centered social networks are viable as a means for gathering useful information and proved that there is potential for the application of social network-based recommendations on various collective intelligence applications. In particular, I suggest that collective intelligence applications, focusing on one solid kind of information items and of which the main social dynamics are to share information items, are beneficial to use these information-centered social networks in personalized recommendations. As we can witness in our daily lives, we cannot agree with and be influenced by our social connections in every perspective. We maintain our preferences to every social connection in a difference aspect. For example, in order to buy a car, we can ask an opinion to Bob who is a mechanic. However, his opinions about movies may not necessarily be valuable as much as his car-related advice. On Web 2.0 applications mainly focusing on users' social networking (i.e. social networking sites, in short, SNS), users tend to share all possible types of information in inconsistent manners. On these SNSs, we cannot guarantee that shared interests among social connections on one kind of information are analogous with the other kinds of information. For instance, it is hard to say that a friend pair sharing similar movie tastes always has common preferences about restaurants. In addition, for recommendations, it is not easy to compute users' preference models for the wide spectrum of information and utilize the models in recommendations. Recommender system cannot generate reasonably good recommendations of

restaurants based on users' book preferences [159]. Therefore, rather than SNSs which mostly promote online social networking and encourage sharing a variety of information about users' daily lives such as Facebook, Twitter, LinkedIn, MySpace, Google Plus, Cyworld, etc., collective intelligence systems to manage and share one solid kind of information such as Flixster, Delicious, Flickr, Youtube, Yelp, SlideShare, etc. are better system to consider social network-based recommendations.

8.2 LIMITATIONS OF THE STUDY AND FURTHER RESEARCH DIRECTION

In the design of this dissertation, some limitations were encountered. First, even though I pointed that existing studies did not use diverse data sources, in this study, only two data sources were used. For watching network and group membership network, I even used the same data source. According to domain properties and unique characteristics of sociability existing on various online applications, although the sociability is the same kind, it is hard to guarantee that the results of this study will be universally identical for all applications. That means that it is immature to generalize the results of this study for other applications yet. Therefore, in future, the study will be executed using more diverse datasets. For instance, in order to explore watching network-based recommendations, a Delicious data set has been collected. Further, to generate collaboration network-based recommendations, publication records of biomedical researchers were collected.

Another limitation of this study is the lack of user study. Even though the evaluation is based on users' bookmarks, the points of evaluation were how much the predictions were

accurate and complete. However, in the recommendation discipline, there are various points of evaluations such as novelty, diversity, satisfaction, perceived usefulness of the recommended items, trustworthiness of the recommender system, etc. Initially I designed these social network-based recommendations as one way to let users get involved in their own recommendations and give users some controls over the personalization of their information spaces. With the cross validation test, however, it was impossible to assess whether the recommendations based on users' self-defined social networks can give users some feelings of control. In order to fill this gap, as the future direction, I plan to do a user study about the perceived quality of the social network-based recommendations proposed in this study. I've already developed a cultural event recommender system with a purpose of allowing users to share interesting cultural events with their online social links. The user study will be executed via this application.

While studying the recommendations based on online social networks, I found that one problem of the social recommendation is lower coverage ratio than the CF recommendations. On the center of the top 10 precision results, recommendation based on the collaboration network produced at least one correct prediction just for 25.3% of the whole user population. However, the CF recommendations produced at least one correct prediction for 70.7% of the whole user population. One big assumption of this social network-based recommendation is that users have to have their social partner(s). On CN3 system which was the data source for collaboration networks, for instance, the number of the users having at least one social connection was 44.7% of the whole user population. For the remaining 55.3% users, that means the recommender system cannot suggest any social network-based recommendations. In case of the other two recommendations (i.e. watching network-based recommendations and group membership-based

recommendations), I didn't compare the recommendation quality between the users having social connection(s) and the other users not having any social connection, as I did for CN3 system. Therefore, I cannot compare the hit ratio between social network-based recommendations and CF recommendations. However, as mentioned, for watching network-based recommendations, users who cannot receive any social recommendations, since they don't have any social link, were 96.6% of the whole user population and for group network-based recommendations, they were 91.5% of the whole population. In order to address this problem, in future, I plan to explore the solution. I will explore recommendations based on implicit social networks. As explained in the section 2, there are various studies about how to build implicit social networks mainly by utilizing users' rating/bookmarking patterns.

Throughout this dissertation, I executed two parts of studies. The first part was to find information sharing patterns to prove the significance of the social networks. As the results, I found that users shared significantly overlapped interests with their social connections. The second part was to generate and evaluate the personalized recommendation based on the online social networks and I discovered that these social connections could be feasible sources of useful information. My original intension was to utilize the information sharing patterns of the first part in the second part as one component of recommendation process. However, I admit that I fail to connect these two parts together. In addition, there could be some disagreements on the results of the first part with the second part results. For instance, in the first part of the chapter 5 which was to explore watching connections, the results showed that watching connections have significantly higher information similarities than the other random pairs who are not socially associated, at all. Hence we can reckon that the watching network-based recommendations performed better than

CF recommendations, since watching network-based recommendation actively utilizes users' social links but CF recommendations don't count users' social links. In the second part of the chapter 5, nonetheless, the results showed that recommendations based on watching connections and CF recommendations based on anonymous peers equally good. There could be two reasons. One reason is that the anonymous peers could include target users' watching connections. The CF recommendation algorithm picks the most like-minded peers solely based on the degree of information similarity regardless of the social connections between anonymous peers and the target users. Another reason is that the random pairs used in the first parts targeted on the whole user population but the CF recommendations targeted on a limited set of like-minded peers. Conclusively, in future, I will concentrate on how I can adopt users' information sharing patterns with their social partners and social networking properties as a component of the social recommendations so as to improve the quality of social network-based recommendations.

Lastly, a weakness of this study is that the quality of overall social recommendations was not evidently better than other social network-based recommendation studies. The main focus of this study is to pick the best performing approaches among social and non-social recommendations. Hence, how to increase the performance of social network-based recommendation per se was out of the focus. Therefore, the next step of the social network-based recommendations will be how to take into account users' information more effectively and how to generate more improved social recommendations.

Appendix A

FINDING GOOD SETTINGS FOR RECOMMENDATIONS

A.1 FINDING GOOD SETTINGS FOR WATCHING NETWORK-BASED RECOMMENDATIONS

A.1.1 Good Similarity Measures for K-Nearest Neighbor Recommendations

Similarity Measure for KNN Recommendations based on Anonymous Peers (CF)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on anonymous peers (i.e. CF). In this experiment, I computed the KNN recommendations with varying number of peers (top 100, top 50, top 20 and the equal number of target users' social connections) and three different kinds of similarity measures (i.e. log-likelihood similarity, the number of co-bookmarks and Jaccard co-efficient). First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood, the suggests considering the number of peers equal to each target user's watching connections (i.e. Loglikelihood_All) significantly performed better than others based on top N peers ($F = 31.1, p < .001$ for top 10; $F = 17.9, p < .001$ for top 5; $F = 4.0, p = .007$ for top 2). For the recommendations based on the number of co-bookmarks also showed that the recommendations based on the number of anonymous peers equal to the target users' watching connections (i.e. Cobmk_All) performed significantly better than the others ($F = 24.7, p < .001$ for top 10; $F = 12.5, p < .001$ for top 5; $F = 2.9, p = .033$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results. The recommendations based on the number of anonymous peers equal to the target users' watching connections (i.e. Jaccard_All) performed significantly better than the others ($F = 62.4, p < .001$ for top 10; $F = 41.7, p < .001$ for top 5; $F = 15.6, p < .001$ for top 2). Lastly, the differences on the

recommendation quality according to the similarity measures were compared. Among three kinds of similarities, the recommendations based on Jaccard co-efficient produced significantly better suggestion than the other two kinds of recommendations ($F = 11.9, p < .001$ for top 10; $F = 8.3, p < .001$ for top 5; $F = 2.6, p = .08$ for top 2). Figure 68 shows the results of KNN recommendations based on anonymous peers. Thus, as KNN recommendations using anonymous peers, I selected KNN recommendation using the Jaccard co-efficient based on as many anonymous peers as the target user’s watching relations exist (i.e. **Jaccard_All**).

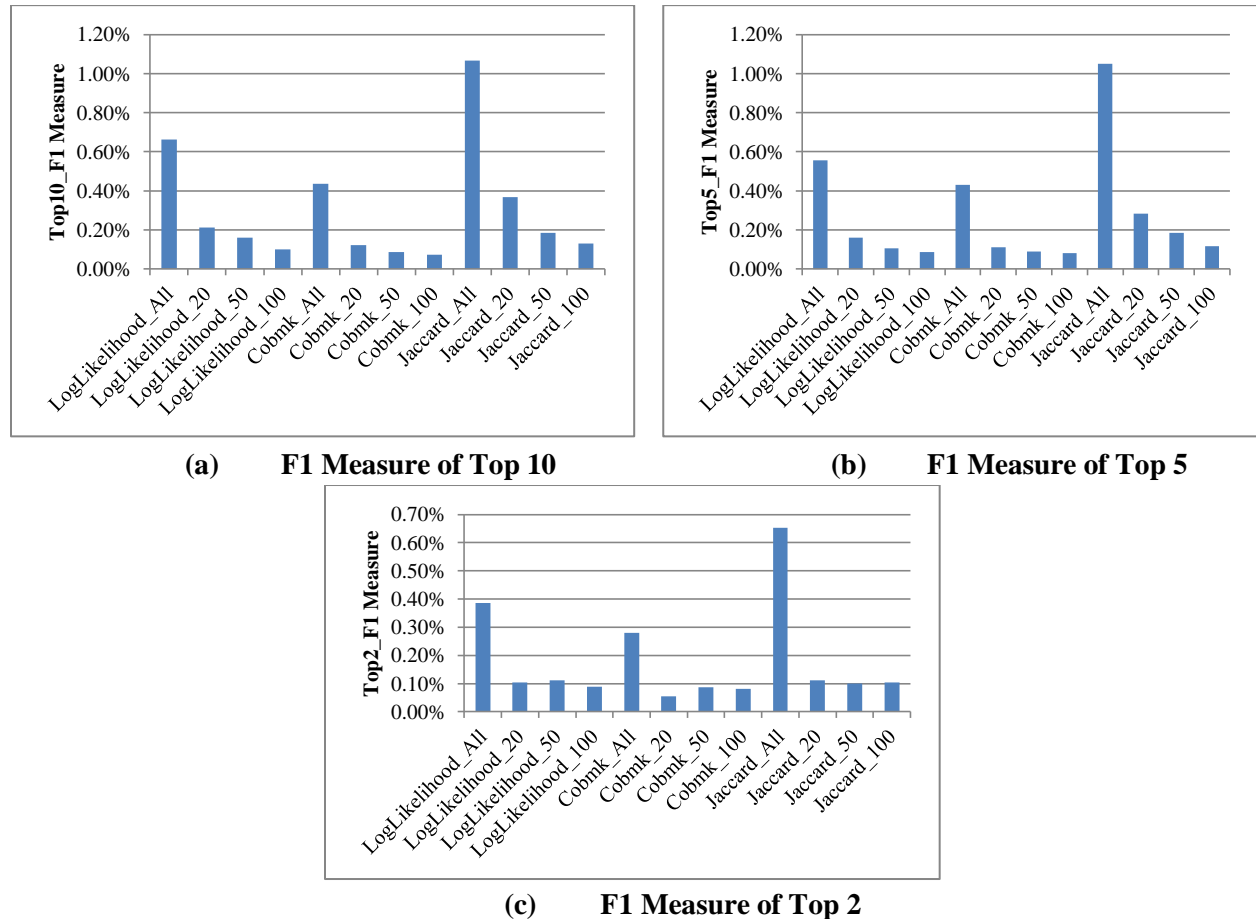
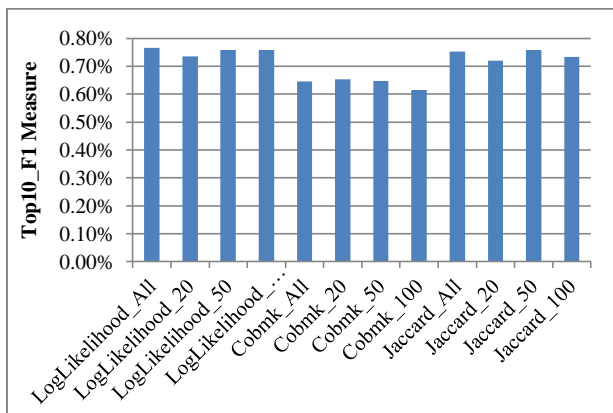


Figure 68. Comparison of KNN Recommendations based on Anonymous Peers (CF)

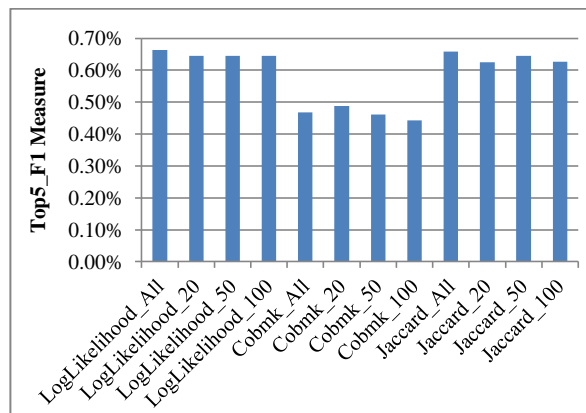
Similarity Measure for KNN Recommendations based on Direct Watching Relations (Watch)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users’ direct watching connections (i.e. Watch). In this experiment, I computed the KNN recommendations with varying number of connections (top 100, top 50, top 20 and each target user’s all social connections) and three different kinds of similarity measures (i.e. log-likelihood similarity, the number of co-bookmarks and Jaccard co-efficient). First, the recommendation

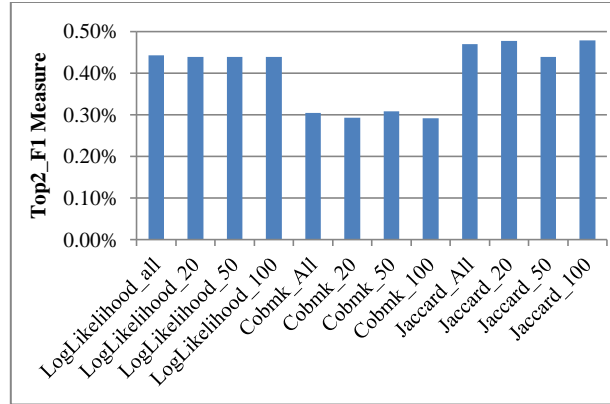
quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .02, p = .99$ for top 10; $F = .01, p = .99$ for top 5; $F = .00, p = 1.00$ for top 2). However, in lower ranks evaluations (i.e. top 10 and top 5), the recommendations generated by each target user's all watching users produced slightly better results. The recommendations based on the number of co-bookmarks also demonstrated that there is no significant difference on the quality according to the different numbers of connections ($F = .04, p = .99$ for top 10; $F = .06, p = .99$ for top 5; $F = .012, p = .99$ for top 2). The result of recommendations based on Jaccard coefficient also showed the same results ($F = .03, p = .99$ for top 10; $F = .02, p = .99$ for top 5; $F = .02, p = .99$ for top 2). Due to the insignificance, I chose the recommendations based on the all connections of target users and compared the quality according to the similarity measures. Among three kinds of similarities, there was no significant difference on the performance of recommendations according to the similarity measures ($F = .03, p = .99$ for top 10; $F = .02, p = .99$ for top 5; $F = .02, p = .99$ for top 2). Figure 69 shows the results of KNN recommendations based on users' direct watching connections. Due to the insignificant difference, the highest values on the graph were selected. For all ranks, the similarity measure yielded the highest F1 value was the log-likelihood similarity. Thus, as KNN recommendations using users' direct watching connections, I selected KNN recommendation using the log-likelihood similarity based on the target user's all watching relations (i.e. **Loglikelihood_All**).



(a) F1 Measure of Top 10



(b) F1 Measure of Top 5



(c) F1 Measure of Top 2

Figure 69. Comparison of KNN Recommendations based on Direct Watching Relations (Watch)

Similarity Measure for KNN Recommendations based on Relations in both Directions (reciWatch)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' both watching and watched relations (i.e. reciWatch). In this experiment, I computed the KNN recommendations with the same setting to the above KNN recommendations based on direct watching connections. First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .007, p = .99$ for top 10; $F = .001, p = .99$ for top 5; $F = .001, p = .99$ for top 2). The recommendations based on the number of co-bookmarks also demonstrated that there is no significant difference on the quality according to the different numbers of connections ($F = .04, p = .99$ for top 10; $F = .105, p = .96$ for top 5; $F = .001, p = .99$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results ($F = .007, p = .99$ for top 10; $F = .006, p = .99$ for top 5; $F = .001, p = 1.00$ for top 2). Due to the insignificance, I compared the quality of recommendations including four types of social connection numbers according to the similarity measures. Among three kinds of similarities, there was no significant difference on the performance of recommendations according to the similarity measures ($F = 1.4, p = .26$ for top 10; $F = 2.3, p = .10$ for top 5; $F = 1.4, p = .25$ for top 2). Figure 70 shows the results of KNN recommendations based on target users' both watching and watched relations. Due to the insignificant difference, the highest values on the graph were selected. For all ranks, the recommendations based on the Jaccard co-efficient with the 50 nearest connections yielded slightly higher F1 value. Thus, as KNN recommendations using users' both watching and watched relations, I selected KNN recommendation using the Jaccard co-efficient based on the target user's top 50 similar connections (i.e. **Jaccard_50**).

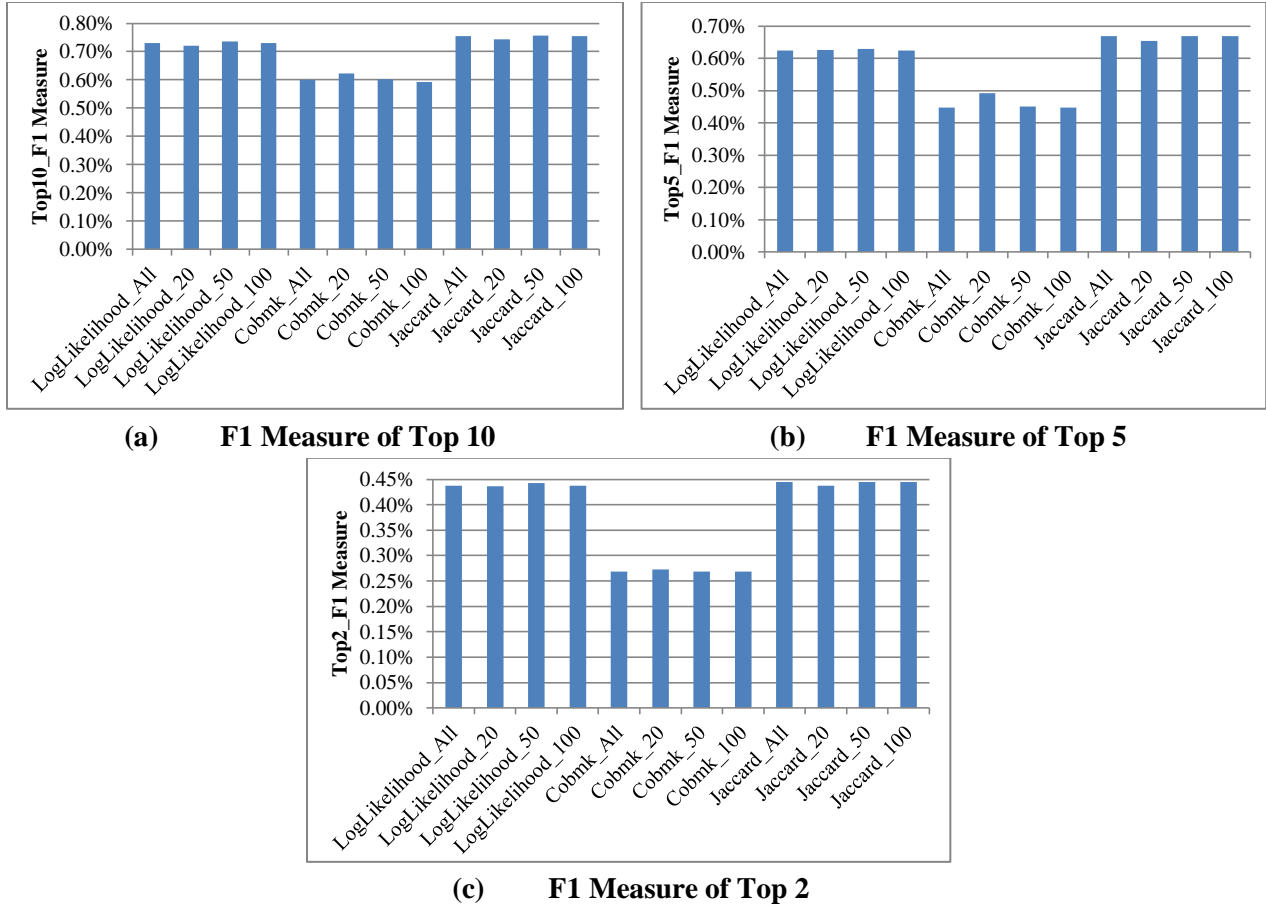


Figure 70. Comparison of KNN Recommendations based on Watching and Watched Relations (ReciWatch)

Similarity Measure for KNN Recommendations based on Direct and Indirect Relations (IhopWatch)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' direct and indirect relations (i.e. IhopWatch). In this experiment, I computed the KNN recommendations with the same setting to the above KNN recommendations based on direct watching connections. First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .02, p = .99$ for top 10; $F = .02, p = .99$ for top 5; $F = .03, p = .99$ for top 2). The recommendations based on the number of co-bookmarks also demonstrated that there is no significant difference on the quality according to the different numbers of connections ($F = .01, p = .99$ for top 10; $F = .04, p = .99$ for top 5; $F = .02, p = .99$ for top 2). The result of recommendations based on Jaccard coefficient also showed the same results ($F = .09, p = .96$ for top 10; $F = .08, p = .97$ for top 5; $F = .01, p = .99$ for top 2).

=.99 for top 2). Among three kinds of similarities, there was no significant difference on the performance of recommendations according to the similarity measures ($F = 2.70, p = .07$ for top 10; $F = .71, p = .49$ for top 5; $F = .53, p = .59$ for top 2). Figure 71 shows the results of KNN recommendations based on target users' direct and indirect watching relations. Due to the insignificant difference, the highest values on the graph were selected. For all ranks, the recommendations based on the Jaccard co-efficient with the 20 nearest connections yielded slightly higher F1 value. Thus, as KNN recommendations using users' direct and indirect watching relations, I selected KNN recommendation using the Jaccard co-efficient based on the target user's top 20 similar connections (i.e. **Jaccard_20**).

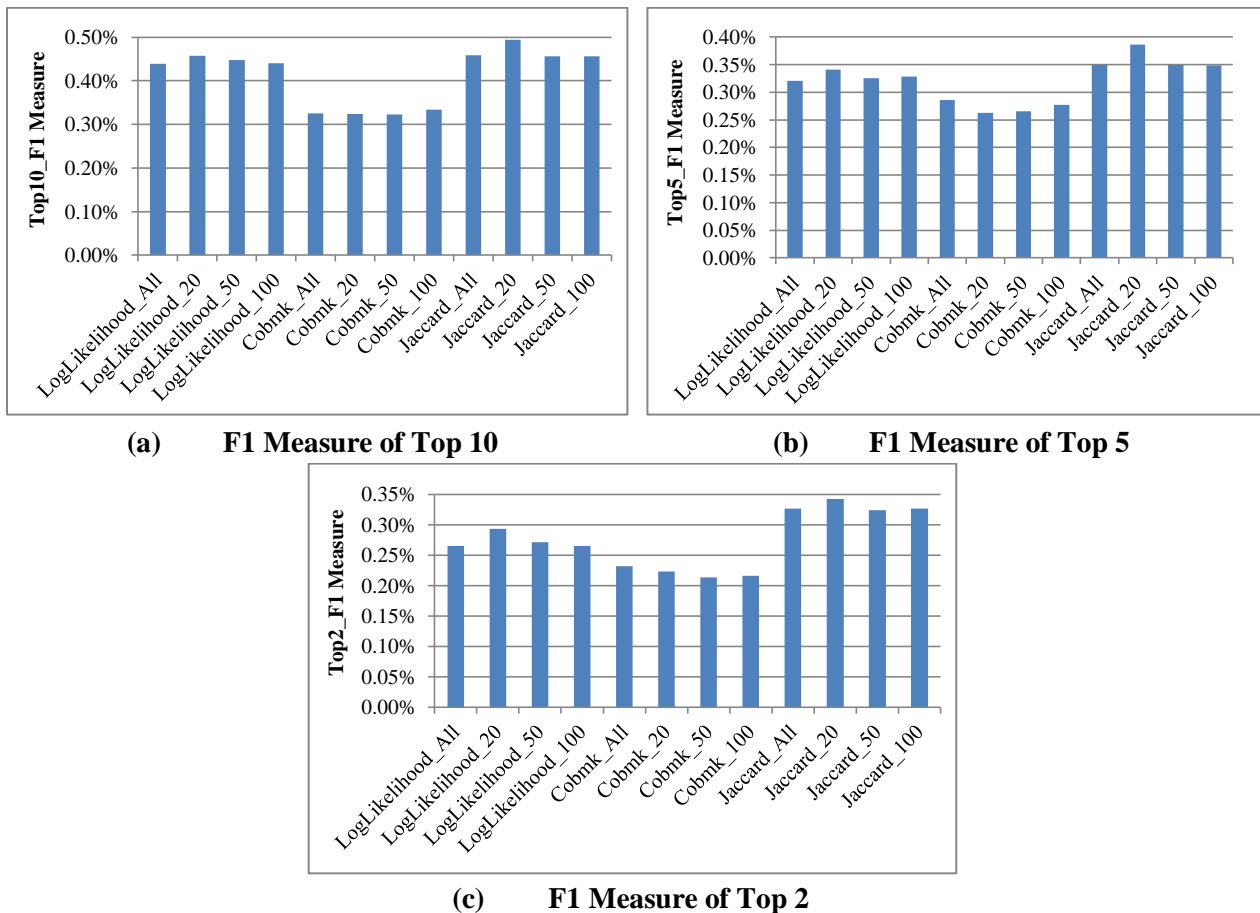


Figure 71. Comparison of KNN Recommendations based on Direct & Indirect Watching Relations (1hopWatch)

A.1.2 Finding Optimal Setting for Matrix Factorization Recommendations

Settings for SVD Recommendations based on Anonymous Peers (CFSVD)

Figure 72 shows the results of matrix factorization recommendations based on users' anonymous users (i.e. CFSVD). According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there was no significant difference in the F1 measure across all settings ($F = 1.34, p = .23$ for top 10; $F = .86, p = .54$ for top 5; $F = .43, p = .88$ for top 2). Due to the statistical insignificance, as the graph shows, the recommendations having 50 factors with $\lambda=0.15$ which yielded the highest F1 value were selected.

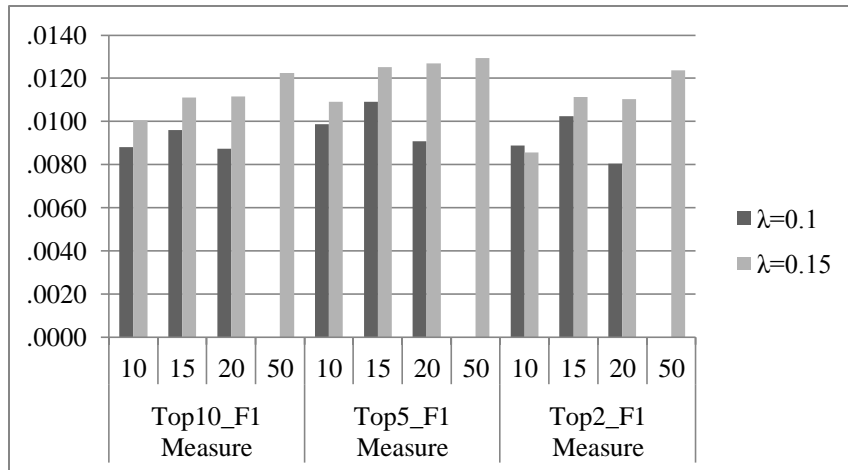
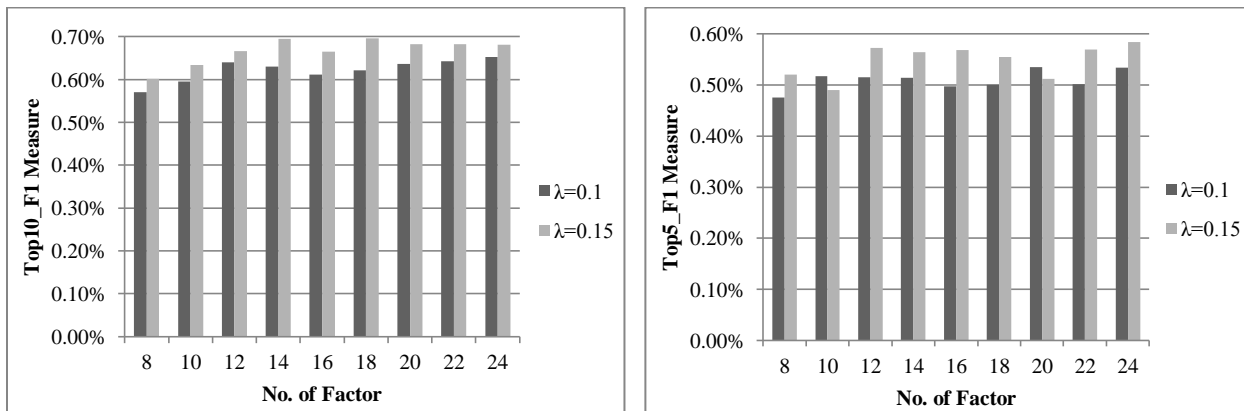


Figure 72. Comparison of Matrix Factorization Recommendations based on Users' Anonymous Peers (CFSVD)

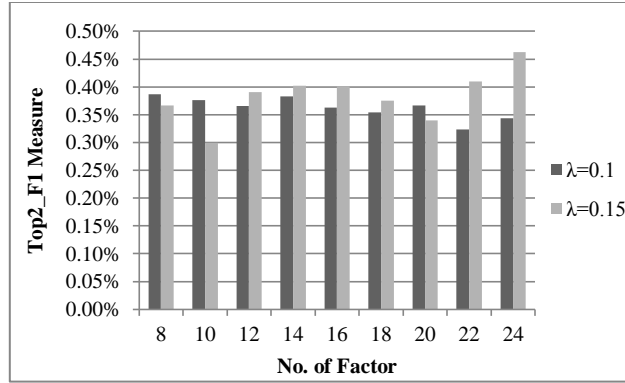
Settings for SVD Recommendations based on Direct Watching Relations (WatchSVD)

Figure 73 shows the results of matrix factorization recommendations based on direct watching relations (i.e. watchSVD). According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .37, p = .99$ for top 10; $F = .24, p = .99$ for top 5; $F = .22, p = 1.0$ for top 2). Due to the statistical insignificance, as the graph shows, the recommendations having 24 factors with $\lambda=0.15$ which yielded the highest F1 value were selected.



(a) F1 Measure of Top 10

(b) F1 Measure of Top 5

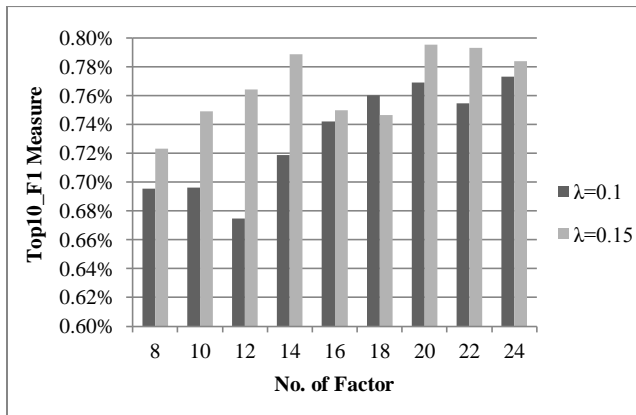


(c) F1 Measure of Top 2

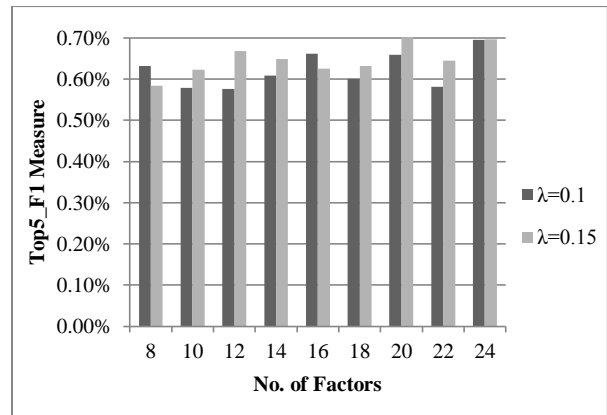
Figure 73. Comparison of Matrix Factorization Recommendations based on Direct Watching Relations (WatchSVD)

Settings for SVD Recommendations based on Relations in Watching and Watched Connections (reciWatchSVD)

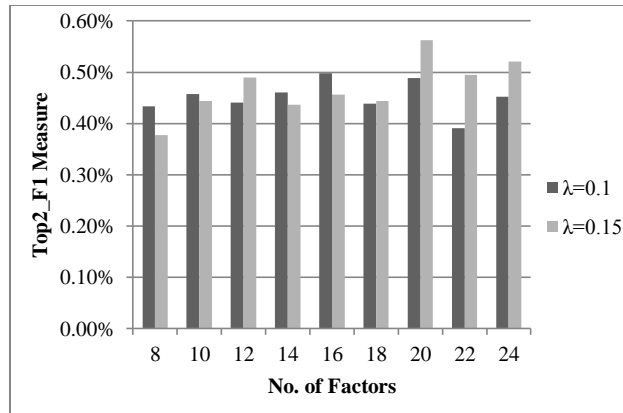
Figure 74 shows the results of matrix factorization recommendations based on watching relations in both directions (i.e. reciWatchSVD). According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .32, p = 1.00$ for top 10; $F = .35, p = .99$ for top 5; $F = .30, p = .99$ for top 2). Due to the statistical insignificance, as the graph shows, the recommendations having 20 factors with $\lambda=0.15$ which yielded the highest F1 value were selected.



(a) F1 Measure of Top 10



(b) F1 Measure of Top 5

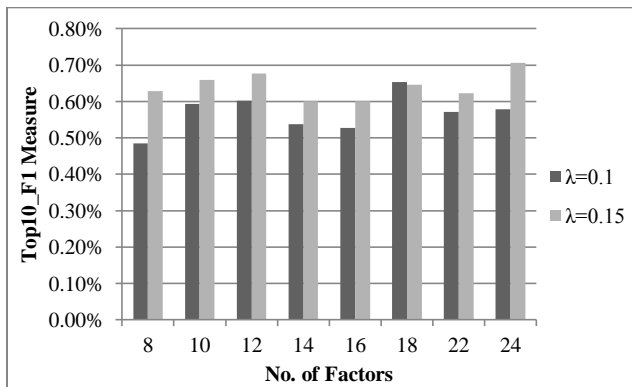


(c) F1 Measure of Top 2

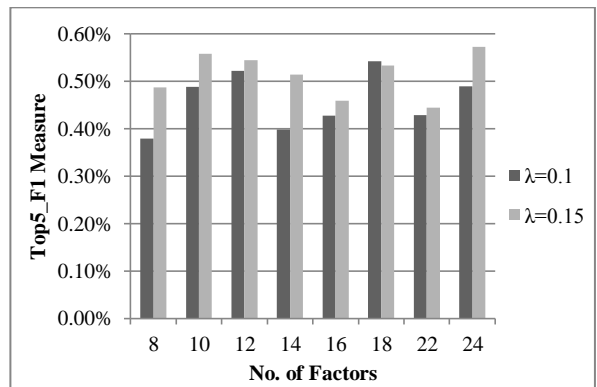
Figure 74. Comparison of Matrix Factorization Recommendations based on Relations in both Directions (reciWatchSVD)

Settings for SVD Recommendations based on Direct and Indirect Relations (1hopWatchSVD)

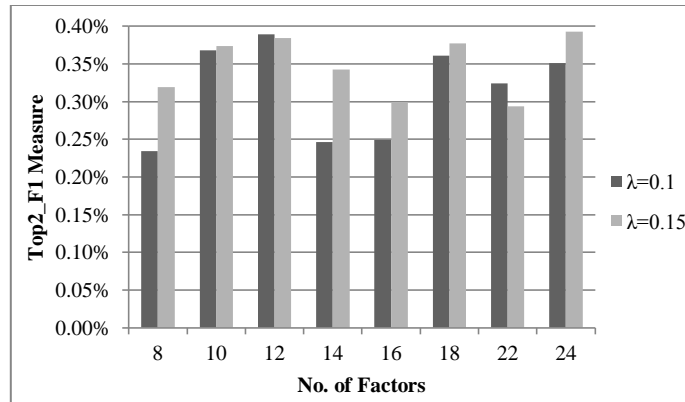
Figure 75 shows the results of matrix factorization recommendations based on watching relations in both directions (i.e. 1hopWatchSVD). According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = 1.05, p = .40$ for top 10; $F = .90, p = .57$ for top 5; $F = .60, p = .88$ for top 2). Due to the statistical insignificance, as the graph shows, the recommendations having 24 factors with $\lambda=0.15$ which yielded the highest F1 value were selected.



(a) F1 Measure of Top 10



(b) F1 Measure of Top 5



(c) F1 Measure of Top 2

Figure 75. Comparison of Matrix Factorization Recommendations based on Relations in Direct and Indirect Social Connections (1hopWatchSVD)

A.2 FINDING GOOD SETTINGS FOR GROUP MEMBERSHIP-BASED RECOMMENDATIONS

A.2.1 Good Similarity Measures for K-Nearest Neighbor Recommendations

Similarity Measure for KNN Recommendations based on Anonymous Peers (CF)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' anonymous peer (i.e. CF). Since the recommendations based on watching connections and group co-members targeted on different sets of users, I tested the traditional CF recommendations, separately. These CF recommendation settings were tested for the situation where the target users are group members. In this experiment, I computed the KNN recommendations with varying number of peers (top 100, top 50, top 20 and the equal number of target users' social connections) and three different kinds of similarity measures (i.e. log-likelihood similarity, the number of co-bookmarks and Jaccard co-efficient). First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood, the suggestions considering the number of peers equal to each target user's group co-members (i.e. Loglikelihood_All) significantly performed better than the others based on top N peers in lower ranks ($F = 12.72, p < .001$ for top 10; $F = 8.24, p < .001$ for top 5; $F = 2.97, p = .03$ for top 2). For the recommendations based on the number of co-bookmarks also showed that the recommendations based on the number of anonymous

peers equal to the target users' group co-members (i.e. Cobmk_All) performed significantly better than the others ($F = 10.7, p < .001$ for top 10; $F = 5.40, p = .001$ for top 5; $F = 1.23, p = .30$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results. The recommendations based on the number of anonymous peers equal to the target users' group co-members (i.e. Jaccard_All) performed significantly better than the others ($F = 25.38, p < .001$ for top 10; $F = 15.08, p < .001$ for top 5; $F = 5.47, p = .001$ for top 2). Lastly, the differences on the recommendation quality according to the similarity measures were compared. Among three kinds of similarities, the recommendations based on Jaccard co-efficient produced significantly better suggestion than the other two kinds of recommendations ($F = 11.05, p < .001$ for top 10; $F = 6.76, p = .001$ for top 5; $F = 3.3, p = .04$ for top 2). Figure 76 shows the results of KNN recommendations based on anonymous peers. Thus, as KNN recommendations using anonymous peers, I selected KNN recommendation using the Jaccard co-efficient based on as many anonymous peers as the target user's group co-members exist (i.e. **Jaccard_All**).

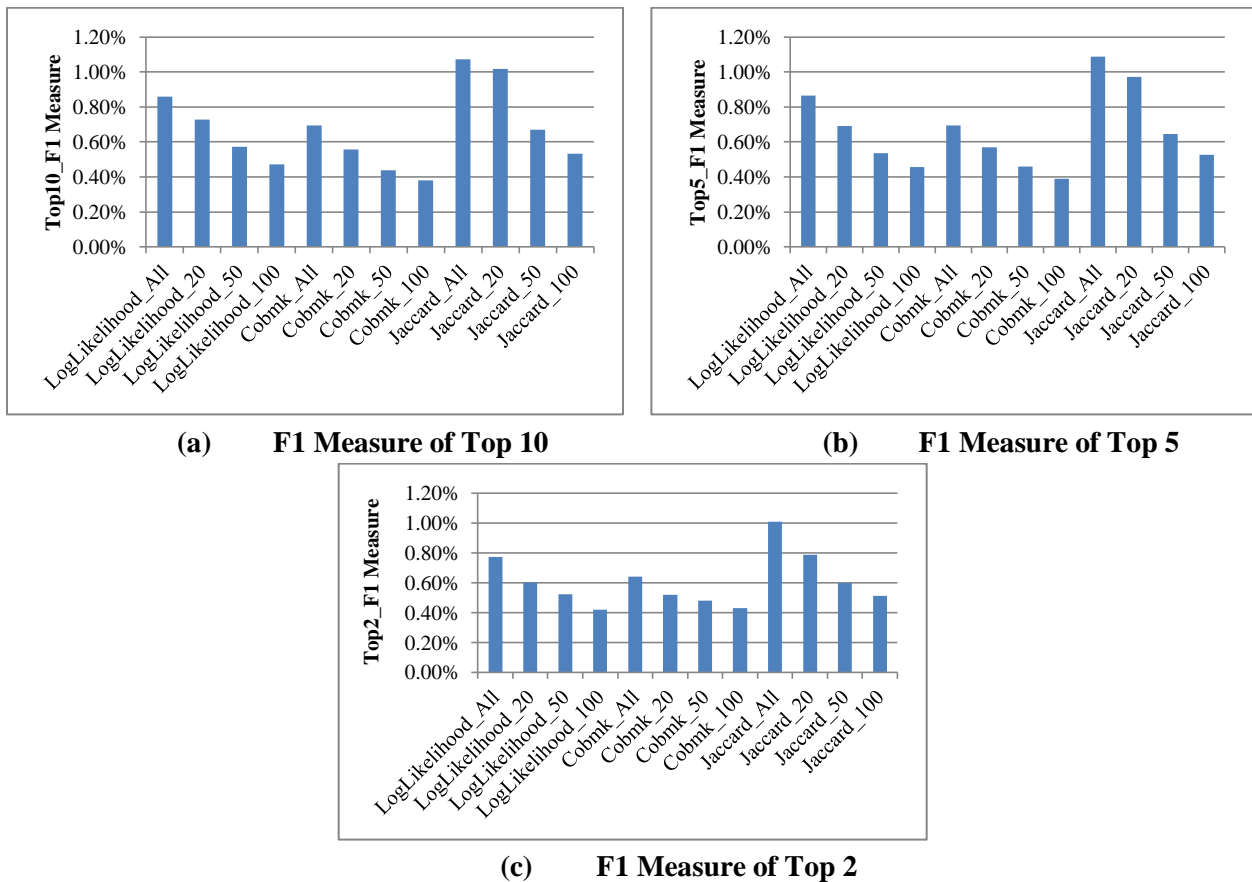
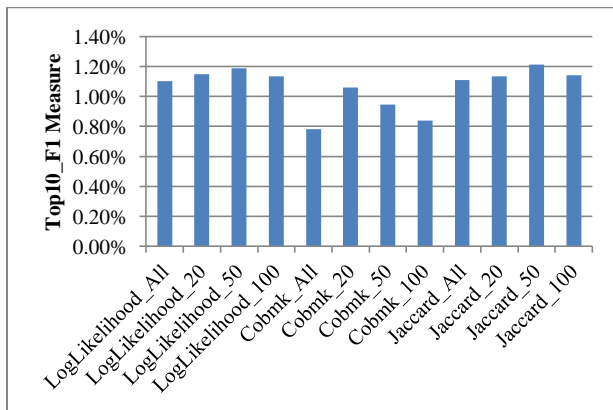


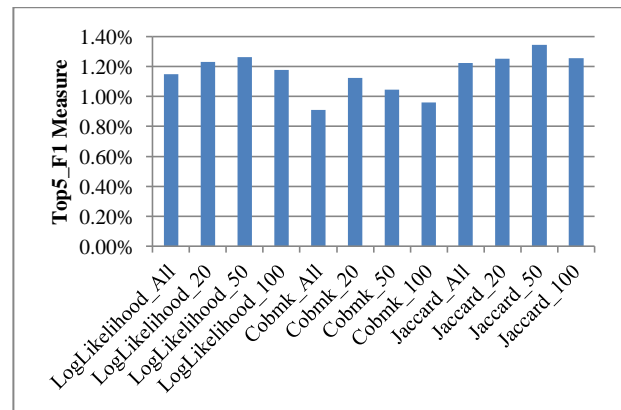
Figure 76. Comparison of KNN Recommendations based on Users' Anonymous Users (CF)

Similarity Measure for KNN Recommendations based on Group Co-members (GMem)

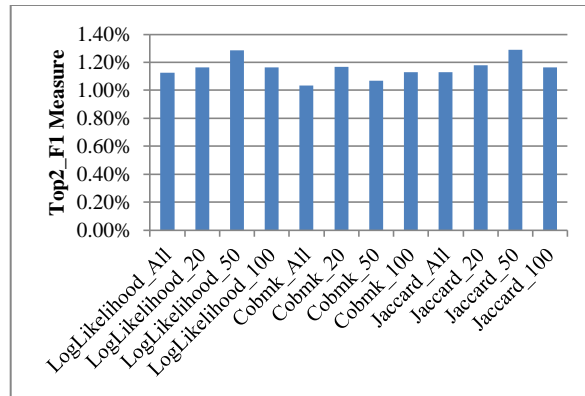
This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' co-members of the same group (i.e. GMem). In this experiment, I computed the KNN recommendations with varying number of peers (top 100, top 50, top 20 and the equal number of target users' social connections) and three kinds of similarity measures (i.e. log-likelihood similarity, the number of co-bookmarks and Jaccard co-efficient). First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .16, p = .92$ for top 10; $F = .18, p = .91$ for top 5; $F = .16, p = .92$ for top 2). The recommendations based on the number of co-bookmarks also demonstrated that there is no significant difference on the quality according to the different numbers of connections ($F = 3.3, p = .02$ for top 10; $F = .96, p = .41$ for top 5; $F = .18, p = .91$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results ($F = .25, p = .86$ for top 10; $F = .17, p = .92$ for top 5; $F = .17, p = .92$ for top 2). Due to the insignificance of the quality, I compared the quality according to the similarity measures including all numbers of N peers. Among three kinds of similarities, there was no significant difference on the performance of recommendations according to the similarity measures ($F = 1.06, p = .35$ for top 10; $F = .96, p = .38$ for top 5; $F = .20, p = .82$ for top 2). Figure 77 shows the results of KNN recommendations based on target users' co-members. Due to the insignificant difference, the highest values on the graph were selected. For all ranks, the recommendations based on the Jaccard co-efficient with the 50 nearest connections yielded slightly higher F1 value. Thus, as KNN recommendations using users' group co-members, I selected KNN recommendation using the Jaccard co-efficient based on the target user's top 50 similar connections (i.e. **Jaccard_50**).



(d) F1 Measure of Top 10



(e) F1 Measure of Top 5



(f) F1 Measure of Top 2

Figure 77. Comparison of KNN Recommendations based on Users' Group Co-members (GMem)

Similarity Measure for KNN Recommendations based on Users' Co-members and their Group Library (Group)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' co-members of the same group and the group library (i.e. Group). In this experiment, I computed the KNN recommendations with the same setting to the above KNN recommendations based on users' co-members. First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .04, p = .99$ for top 10; $F = .22, p = .89$ for top 5; $F = .15, p = .93$ for top 2). The recommendations based on the number of co-bookmarks also demonstrated no significant difference on the quality according to the different numbers of connections ($F = .86, p = .46$ for top 10; $F = .26, p = .86$ for top 5; $F = .04, p = .99$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results ($F = .15, p = .96$ for top 10; $F = 2.96, p = .02$ for top 5; $F = 2.62, p = .03$ for top 2). Due to the insignificance of the quality regardless of the number of social connections, I compared the quality according to the similarity measures including all numbers of N peers. Among three kinds of similarities, however, there were significant differences on the performance of recommendations according to the similarity measures, especially in lower ranks ($F = 9.25, p < .001$ for top 10; $F = 7.22, p = .001$ for top 5; $F = 6.33, p = .002$ for top 2). According to the pairwise test, the recommendations based on Jaccard co-efficient and log-likelihood similarity performed significantly better than the recommendations based on the number of co-bookmarks, even though there was no difference between the Jaccard co-efficient and log-likelihood similarity. Figure 78 shows the results of KNN recommendations based on target users' co-members and their group library. Due to the insignificant difference between the Jaccard and log-likelihood similarity, as KNN recommendations

using users' co-members and group library, I selected KNN recommendation using the Jaccard coefficient based on the target user's top 50 similar connections (i.e. **Jaccard_50**), which produced relatively higher F-1 value.

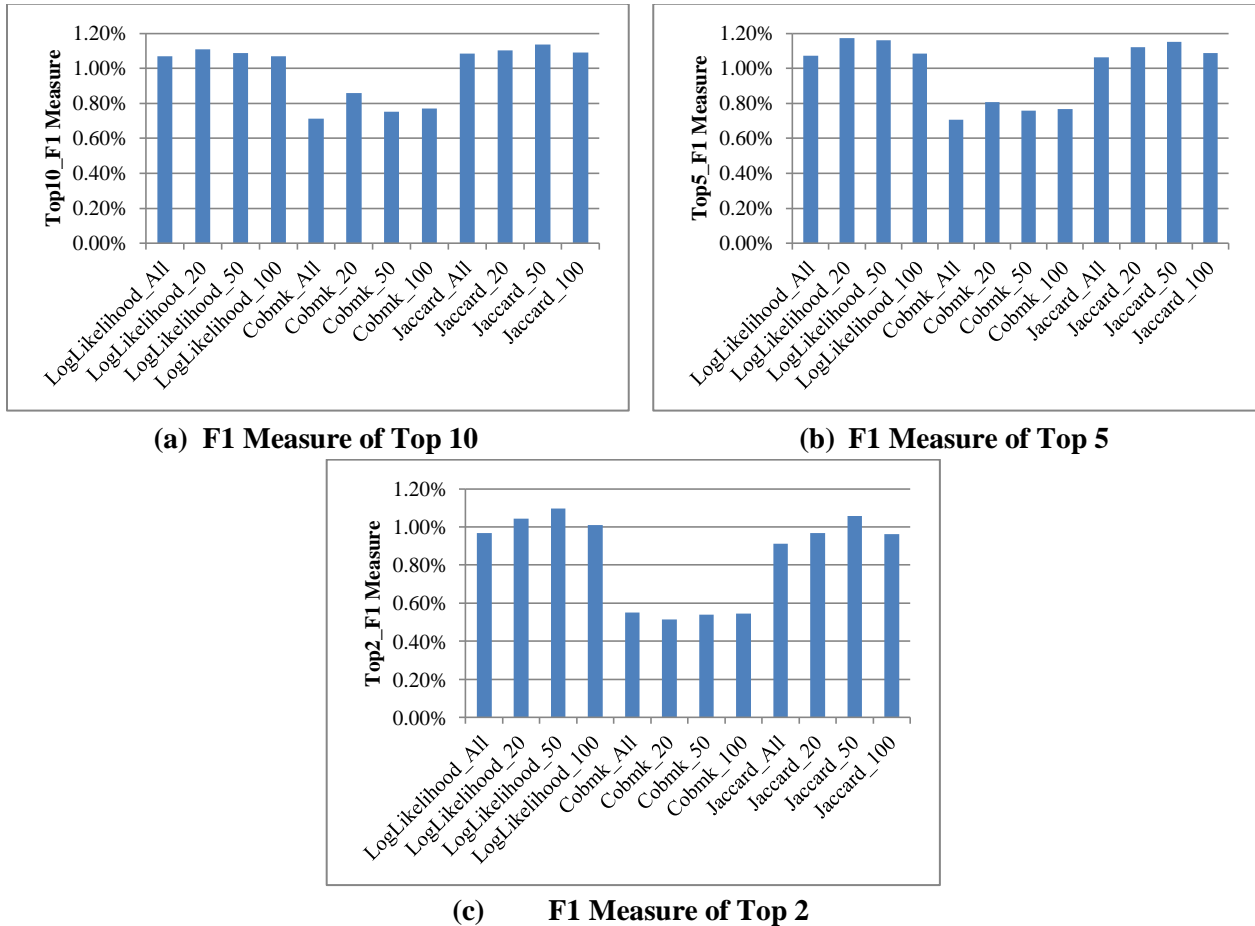
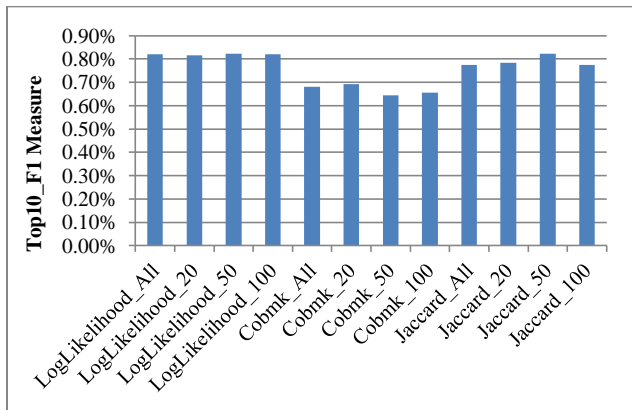


Figure 78. Comparison of KNN Recommendations based on Co-members and Group Memberships (Gmem)

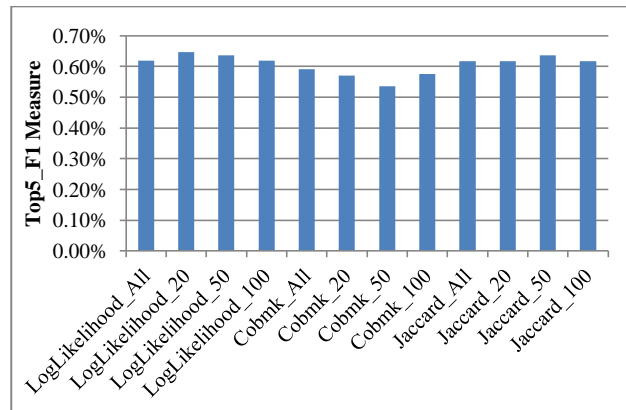
Similarity Measure for KNN Recommendations based on Watching Networks and Group Memberships (GMemWatch)

This section is to find a good similarity measure and optimal number of peers for KNN recommendations based on target users' direct watching connections and their co-members of the same group (i.e. GMemWatch). In this experiment, I computed the KNN recommendations with the same setting to the above KNN recommendations based on users' co-members. First, the recommendation quality was compared according to the varying number of peers. The results of the recommendations based on Log-likelihood showed that, regardless of how many social connections the recommendation algorithm considered, the qualities were not significantly different ($F = .00, p = 1.00$ for top 10; $F = .009,$

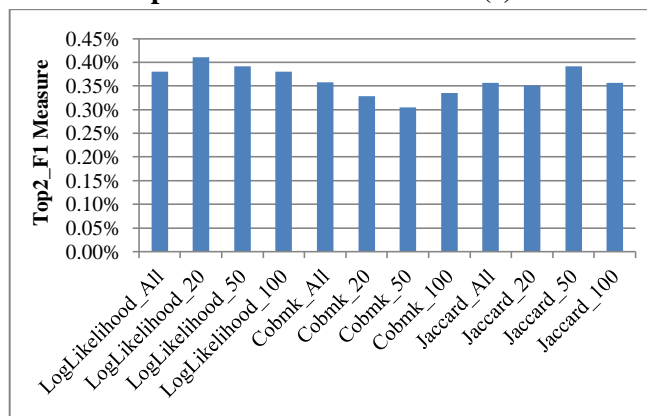
$p = .99$ for top 5; $F = .018$, $p = .99$ for top 2). The recommendations based on the number of co-bookmarks also demonstrated that there is no significant difference on the quality according to the different numbers of connections ($F = .03$, $p = .99$ for top 10; $F = .04$, $p = .99$ for top 5; $F = .06$, $p = .98$ for top 2). The result of recommendations based on Jaccard co-efficient also showed the same results ($F = .02$, $p = .99$ for top 10; $F = .01$, $p = .99$ for top 5; $F = .03$, $p = .99$ for top 2). Due to the insignificance of the quality, I compared the quality according to the similarity measures including all numbers of N peers. Among three kinds of similarities, there was no significant difference on the performance of recommendations according to the similarity measures ($F = .29$, $p = .75$ for top 10; $F = .05$, $p = .95$ for top 5; $F = .07$, $p = .93$ for top 2). Figure 79 shows the results of KNN recommendations based on target users' watching connections and their co-members. Due to the insignificant difference, the highest values on the graph were selected. For all ranks, the recommendations based on the log-likelihood similarity with the 20 nearest connections (i.e.) yielded slightly higher F1 value. Thus, as KNN recommendations using users' both watching relations and group co-members, I selected KNN recommendation using the log-likelihood similarity based on the target user's top 20 similar connections (i.e. **loglikelihood_20**).



(d) F1 Measure of Top 10



(e) F1 Measure of Top 5



(f) F1 Measure of Top 2

Figure 79. Comparison of KNN Recommendations based on Watching Relations and Group Memberships (GMemWatch)

A.2.2 Finding Optimal Setting for Matrix Factorization Recommendations

Settings for SVD Recommendations based on Anonymous Users (CFSVD)

Figure 80 shows the results of matrix factorization recommendations based on users’ anonymous users (i.e. CFSVD). As explained in the A.2.1, this test targeted on the target users consisting of group members. According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there was no significant difference in the F1 measure across all settings ($F = 2.89, p = .008$ for top 10; $F = 2.40, p = .025$ for top 5; $F = 1.74, p = .018$ for top 2). Due to the statistical insignificance, as the graph shows, the recommendations having 50 factors with $\lambda=0.15$ which yielded the highest F1 value were decided.

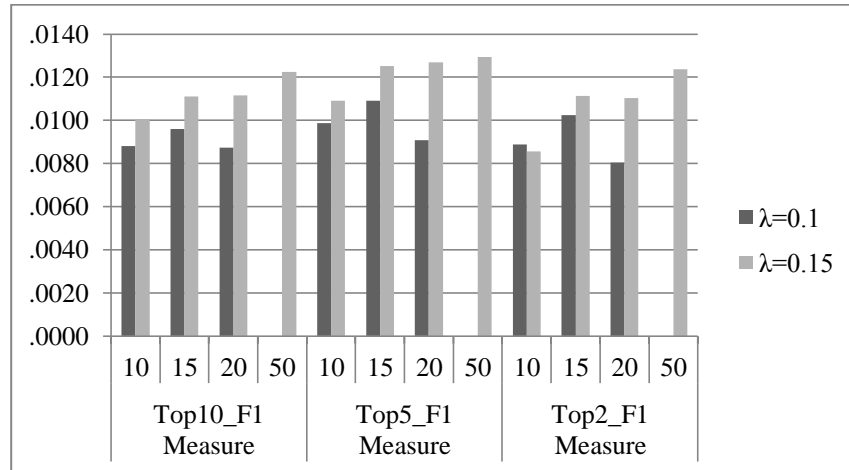


Figure 80. Comparison of Matrix Factorization Recommendations based on Users’ Anonymous Peers (CFSVD)

Settings for SVD Recommendations based on Group Co-members (GMemSVD)

Figure 81 shows the results of matrix factorization recommendations based on users’ co-members of the same group (i.e. GMemSVD). According to the One-way ANOVA, there were significant differences in the F1 measure across all settings ($F = 5.24, p < .001$ for top 10; $F = 3.23, p < .001$ for top 5; $F = 1.61, p = .04$ for top 2). In spite of the overall statistical significance, according to the pairwise tests, there were no significant differences on factor by factor and depending on the different λ values. Thus, as the graph shows, the recommendations with 16 factors and $\lambda=0.15$ which yielded the highest F1 value were decided.

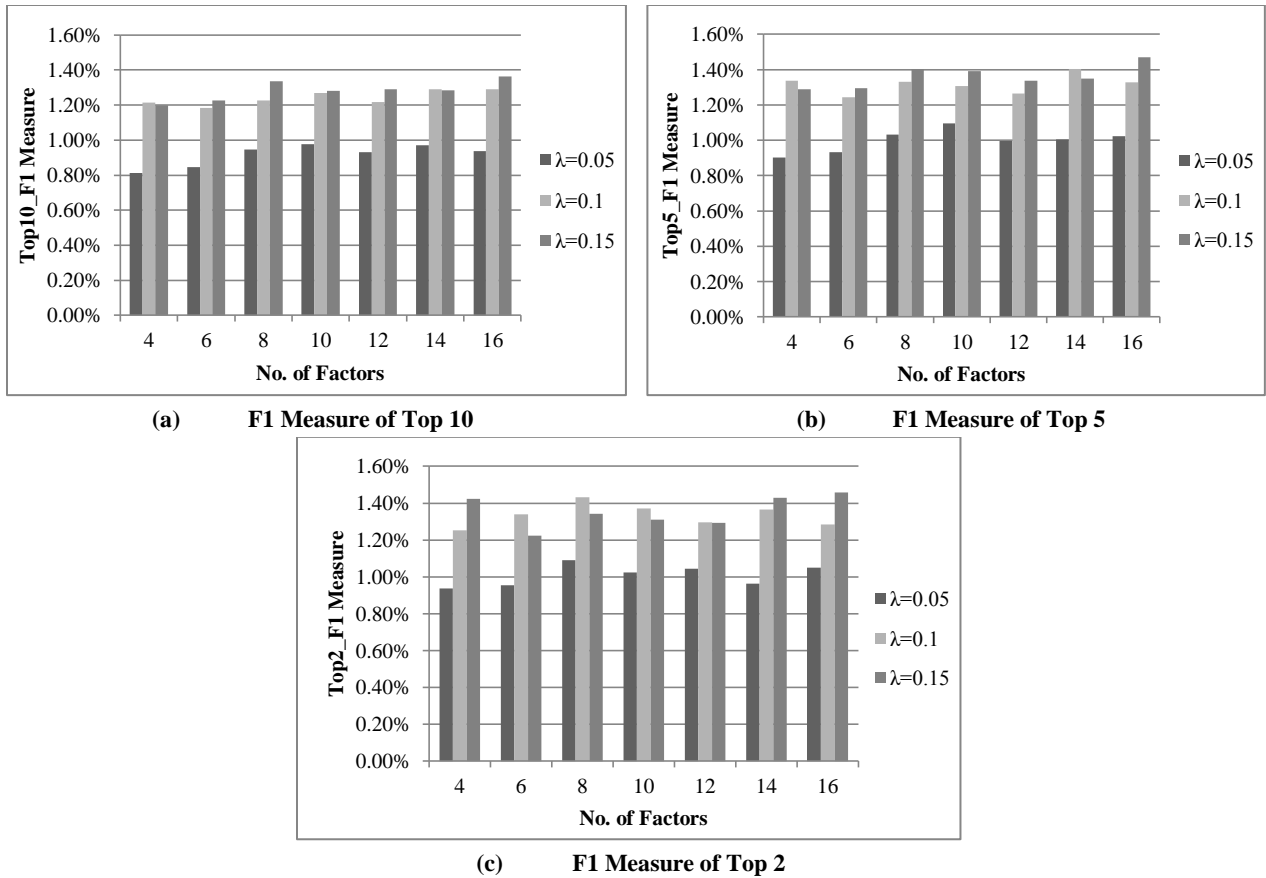


Figure 81. Comparison of Matrix Factorization Recommendations based on Users' Co-members (GMemSVD)

Settings for SVD Recommendations based on Users' Co-members and their Group Library (GroupSVD)

Figure 82 shows the results of matrix factorization recommendations based on users' co-members and their group library (i.e. GroupSVD). According to the One-way ANOVA, there were significant differences in the F1 measure across all settings ($F = 7.50, p < .001$ for top 10; $F = 5.87, p < .001$ for top 5; $F = 4.11, p < .001$ for top 2). In spite of the overall statistical significance, according to the pairwise tests, there were no significant differences on factor by factor and depending on the different λ values. Thus, as the graph shows, the recommendations with 16 factors and $\lambda=0.15$ which yielded the highest F1 value were decided.

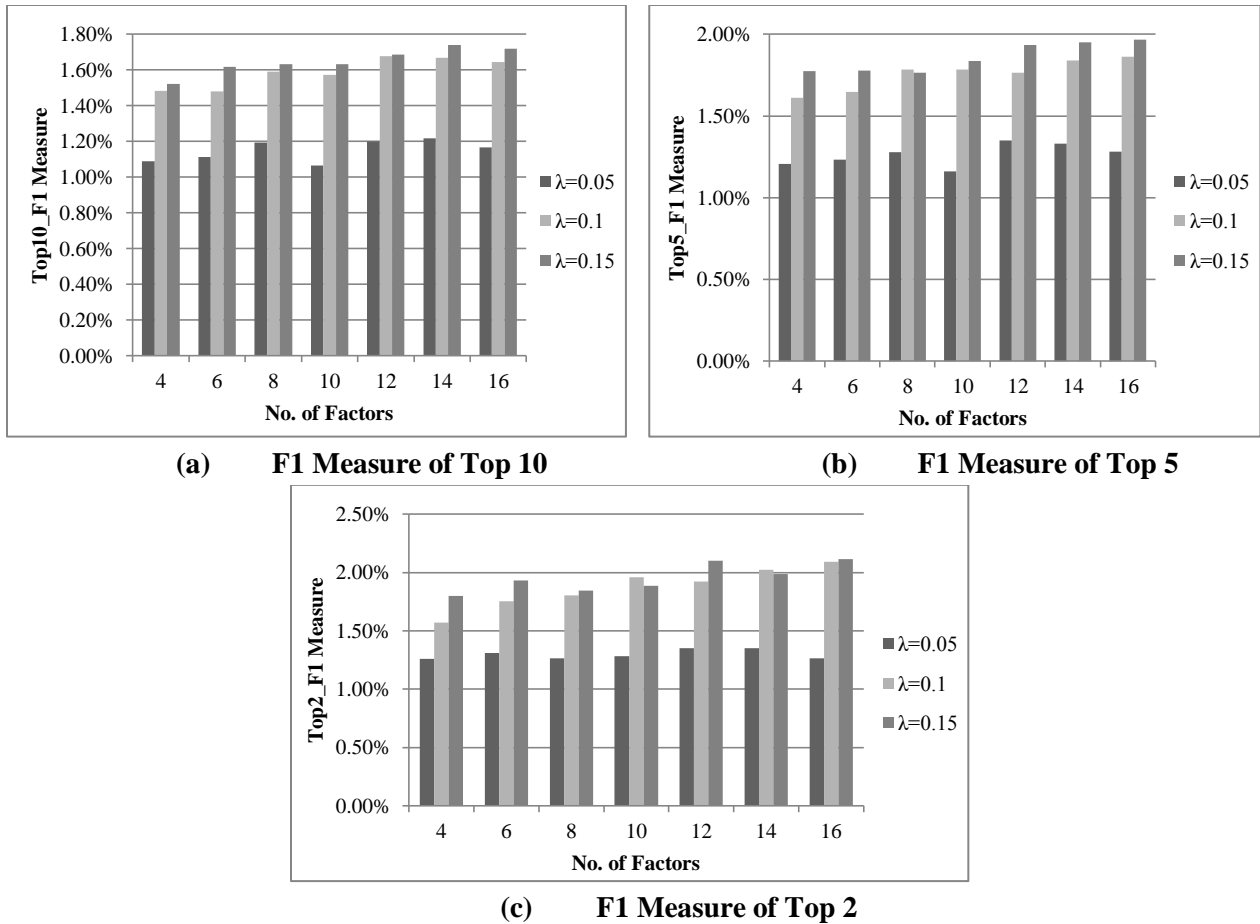


Figure 82. Comparison of Matrix Factorization Recommendations based on Users' Co-members and their library (GroupSVD)

Settings for SVD Recommendations based on Watching Networks and Group Memberships (GMemWatchSVD)

Figure 83 shows the results of matrix factorization recommendations based on watching relations and group co-members (i.e. GMemWatchSVD). According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .30, p = .997$ for top 10; $F = .39, p = .99$ for top 5; $F = .19, p = 1.0$ for top 2). In spite of the statistical insignificance, the graph shows that the F1 values of 20 factors with $\lambda=0.15$ were getting higher on higher ranks.

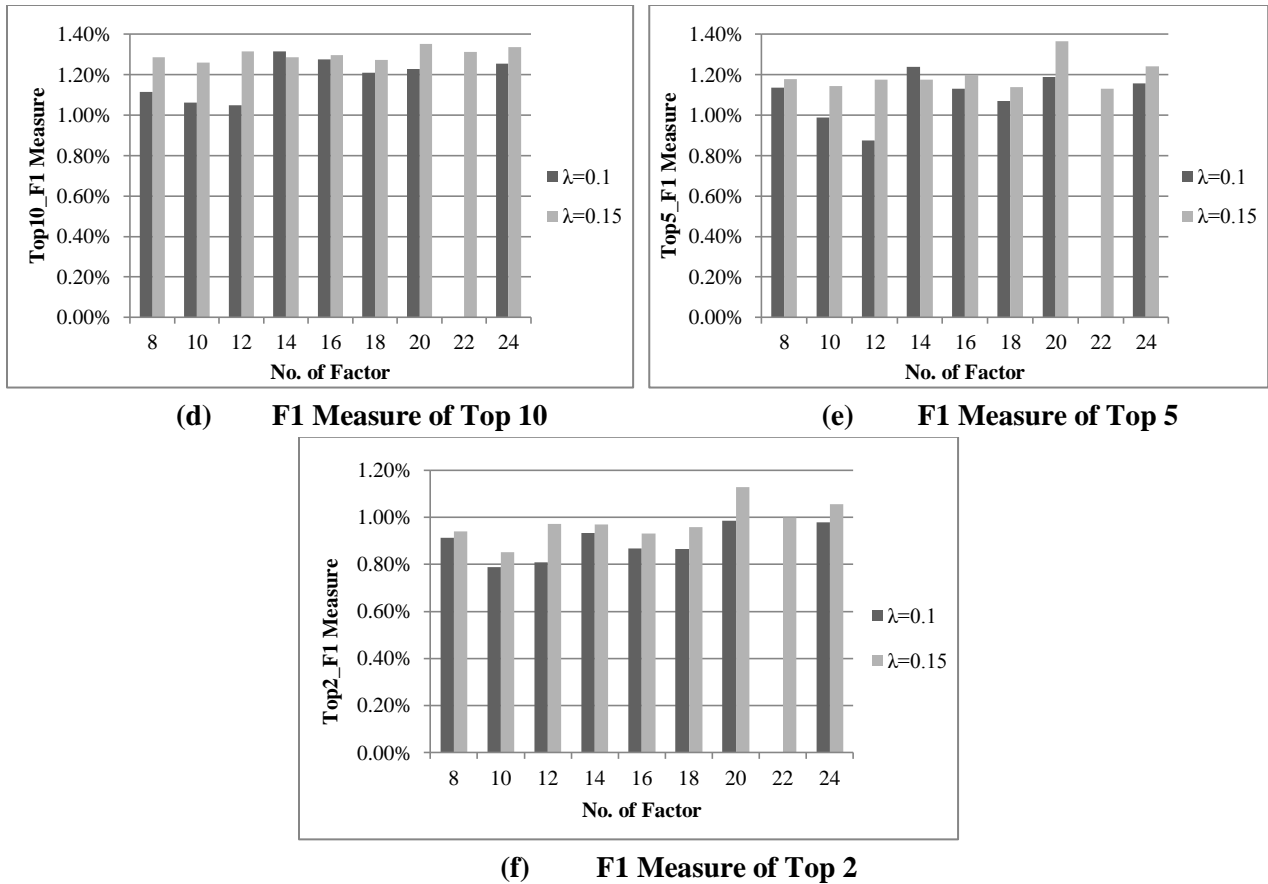


Figure 83. Comparison of Matrix Factorization Recommendations based on Watching Networks and Group Memberships (GMemWatchSVD)

A.3 FINDING GOOD SETTINGS FOR COLLABORATION NETWORK-BASED RECOMMENDATIONS

Settings for SVD Recommendations based on Anonymous Peers (CFSVD)

Figure 84 shows the results of matrix factorization recommendations based on anonymous peers (i.e. CFSVD). 3 varying numbers of latent factors (5, 7, and 10 factors) with $\lambda = 1.5$ were tested. According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .03, p = .96$ for top 10; $F = .01, p = .99$ for top 5; $F = .01, p = .99$ for top 2). In spite of the statistical insignificance, the graph shows that the F1 values of

SVD having with 15 factors produced the slightly higher F1 measure than the others, in particular in higher ranks.

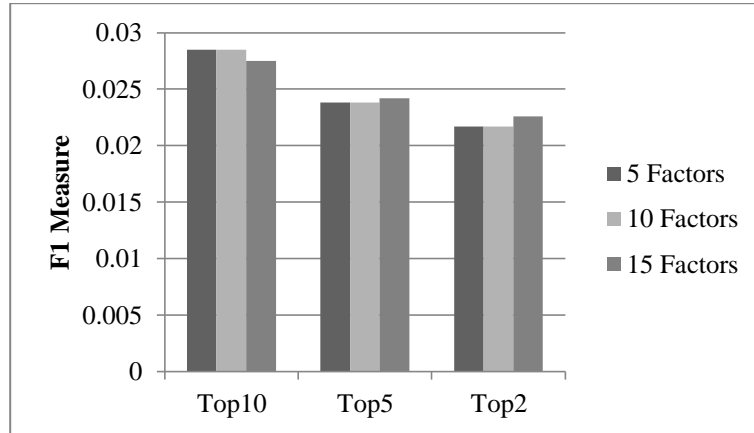


Figure 84. Comparison of Matrix Factorization Recommendations based on Users’ Anonymous Peers

Settings for SVD Recommendations based on Collaborators (SNSVD_Coll)

Figure 85 shows the results of matrix factorization recommendations based on users’ collaborators (i.e. SNSVD_Coll). 4 varying numbers of latent factors (3, 5, 7, and 10 factors) with $\lambda = 1.5$ were tested. According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .06, p = .98$ for top 10; $F = .11, p = .95$ for top 5; $F = .05, p = .98$ for top 2). In spite of the statistical insignificance, the graph shows that the F1 values of SVD having with 3 factors produced the higher F1 measure than the others in higher ranks.

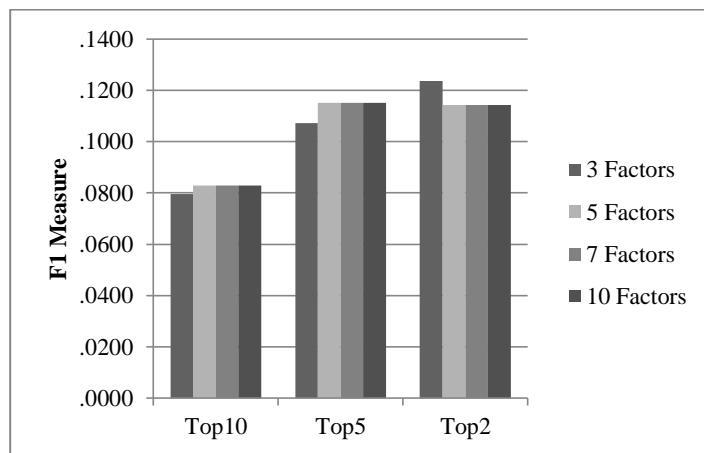


Figure 85. Comparison of Matrix Factorization Recommendations based on Users’ Collaborators

Settings for SVD Recommendations based on CN3 Social Connections (SNSVD_CN3)

Figure 86 shows the results of matrix factorization recommendations based on users' CN3 social connections (i.e. SNSVD_CN3). 4 varying numbers of latent factors (3, 5, 7, and 10 factors) with $\lambda = 1.5$ were tested. According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .01, p = .98$ for top 10; $F = .00, p = .99$ for top 5; $F = .03, p = .99$ for top 2). In spite of the statistical insignificance, the graph shows that the F1 values of SVD having with 10 factors produced the higher F1 measure than the others in higher ranks.

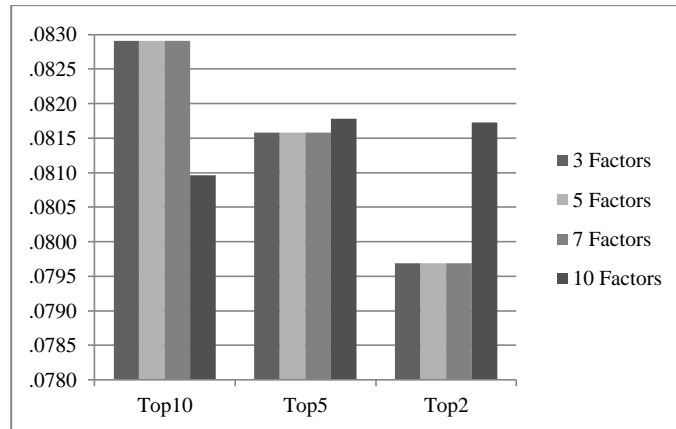


Figure 86. Comparison of Matrix Factorization Recommendation based on Users' CN3 Social Connections

Settings for SVD Recommendations based on Both Collaborators and CN3 Social Connections (SNSVD_All)

Figure 87 shows the results of matrix factorization recommendations based on users' collaborators and CN3 social connections (i.e. SNSVD_All). 4 varying numbers of latent factors (3, 5, 7, and 10 factors) with $\lambda = 1.5$ were tested. According to the One-way ANOVA and Scheffé pairwise test, regardless of the ranks, there were no significant differences in the F1 measure across all settings ($F = .00, p = .99$ for top 10; $F = .12, p = .95$ for top 5; $F = .08, p = .97$ for top 2). In spite of the statistical insignificance, the graph shows that the F1 values of SVD having with 3 factors produced the higher F1 measure than the others in higher ranks.

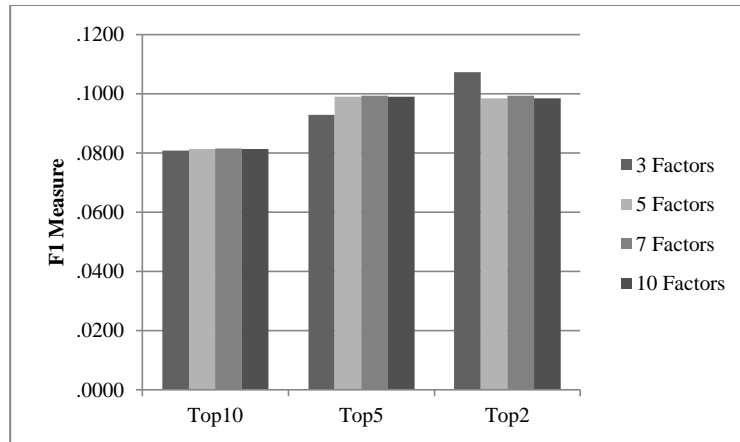


Figure 87. Comparison of Matrix Factorization Recommendation based on Both Collaborators and CN3 Social Connections

Appendix B

IMPLEMENTATION OF INFORMATION SIMILARITY AND RECOMMENDATION ALGORITHMS

In this dissertation, most of the information similarities and recommendation algorithms were implemented using third-party applications, except the cases where collecting and cleaning data sources. In particular, the following applications were used.

- *Lucene* (v3.2): This application was to process text-related information (i.e. case normalization and stemming), to build word vectors and to compute the Singular Vector Decomposition.
- *Mahout* (v0.6): This was core application to implement information similarities and most of the recommendation algorithms.
- *R* with *igraph* package: This application was to compute all kinds of social properties, in both edge level and vertex level.

BIBLIOGRAPHY

- [1] Abbassi, Z., S. Amer-Yahia, L.V.S. Lakshmanan, S. Vassilvitskii, and C. Yu, *Getting recommender systems to think outside the box*, in *Proceedings of the third ACM conference on Recommender systems*. 2009, ACM: New York, New York, USA. p. 285-288.
- [2] Abbattista, F., F. Calefato, D. Gendarmi, and F. Lanubile, *Shaping Personal Information Spaces from Collaborative Tagging Systems*. 2010. p. 728-735.
- [3] Adamic, L.A. and E. Adar, *Friends and neighbors on the Web*. *Social Networks*, 2003. **25**(3): p. 211-230.
- [4] Adomavicius, G. and A. Tuzhilin, *Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions*. *IEEE Trans. on Knowl. and Data Eng.*, 2005. **17**(6): p. 734-749.
- [5] Akehurst, J., I. Koprinska, K. Yacef, L. Pizzato, J. Kay, and T. Rej. CCR: a content-collaborative reciprocal recommender for online dating. in *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Three*. 2011: AAAI Press.
- [6] Akther, A., H.-N. Kim, M. Rawashdeh, and A. El Saddik, *Applying Latent Semantic Analysis to Tag-Based Community Recommendations*, in *Advances in Artificial Intelligence*, L. Kosseim and D. Inkpen, Editors. 2012, Springer Berlin Heidelberg. p. 1-12.
- [7] Al-Oufi, S., H.-N. Kim, and A.E. Saddik, *A group trust metric for identifying people of trust in online social networks*. *Expert Syst. Appl.*, 2012. **39**(18): p. 13173-13181.
- [8] Al-sharawneh, J. and M. Williams. Credibility-aware Web-based Social Network Recommender: Follow the Leader. in *Proceedings of the 2nd ACM RecSys Workshop on Recommender Systems and the Social Web*. 2010. Barcelona, Spain.
- [9] Alan Page, F., *The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations*. *Psychological Review*, 1992. **99**(4): p. 689-723.
- [10] Alani, H., S. Dasmahapatra, K. O'Hara, and N. Shadbolt, *Identifying Communities of Practice through Ontology Network Analysis*. *IEEE Intelligent Systems*, 2003. **18**(2): p. 18-25.
- [11] Anagnostopoulos, A., R. Kumar, and M. Mahdian, *Influence and correlation in social networks*, in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2008, ACM: Las Vegas, Nevada, USA. p. 7-15.

- [12] Ardissono, L., A. Goy, G. Petrone, M. Segnan, and P. Torasso, *INTRIGUE: Personalized recommendation of tourist attractions for desktop and handset devices*. Applied Artificial Intelligence, 2003. **17**(8): p. 687-714.
- [13] Au-Yeung, C.M., N. Gibbins, and N. Shadbolt. A Study of User Profile Generation from Folksonomies. in *Proceedings of the WWW 2008 Social Web and Knowledge Management, Social Web Workshop*. 2008. Beijing, China.
- [14] Avesani, P., P. Massa, and R. Tiella, *A trust-enhanced recommender system application: Moleskiing*, in *Proceedings of the 2005 ACM symposium on Applied computing*. 2005, ACM: Santa Fe, New Mexico. p. 1589-1593.
- [15] Baatarjav, E.-A., S. Phithakkitnukoon, and R. Dantu, *Group Recommendation System for Facebook*, in *Proceedings of the OTM Confederated International Workshops and Posters on On the Move to Meaningful Internet Systems: 2008 Workshops: ADI, AWeSoMe, COMBEK, EI2N, IWSSA, MONET, OnToContent + QSI, ORM, PerSys, RDDS, SEMELS, and SWWS*. 2008, Springer-Verlag: Monterrey, Mexico. p. 211-219.
- [16] Backstrom, L., E. Bakshy, J. Kleinberg, T.M. Lento, and I. Rosenn. Center of Attention: How Facebook Users Allocate Attention Across Friends. in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. 2011. Barcelona, Spain.
- [17] Backstrom, L., D. Huttenlocher, J. Kleinberg, and X. Lan, *Group formation in large social networks: membership, growth, and evolution*, in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006, ACM: Philadelphia, PA, USA. p. 44-54.
- [18] Bakshy, E., B. Karrer, and L.A. Adamic, *Social influence and the diffusion of user-created content*, in *Proceedings of the tenth ACM conference on Electronic commerce*. 2009, ACM: Stanford, California, USA. p. 325-334.
- [19] Blei, D.M., A.Y. Ng, and M.I. Jordan, *Latent dirichlet allocation*. the Journal of machine Learning research, 2003. **3**: p. 993-1022.
- [20] Bonhard, P. and M. Sasse, *'Knowing me, knowing you' — Using profiles and social networking to improve recommender systems*. BT Technology Journal, 2006. **24**(3): p. 84-98.
- [21] Bonhard, P., M.A. Sasse, and C. Harries, *"The devil you know knows best": how online recommendations can benefit from social networking*, in *Proceedings of the 21st British CHI Group Annual Conference on HCI 2007: People and Computers XXI: HCI...but not as we know it - Volume 1*. 2007, British Computer Society: University of Lancaster, United Kingdom. p. 77-86.

- [22] Bourke, S., K. McCarthy, and B. Smyth, *Power to the People: Exploring Neighborhood Formations in Social Recommender Systems*, in *The Fifth ACM Conference on Recommender Systems*. 2011: Chicago, IL, USA.
- [23] Brand, M., *Fast low-rank modifications of the thin singular value decomposition*. *Linear Algebra and its Applications*, 2006. **415**(1): p. 20-30.
- [24] Breese, J.S., D. Heckerman, and C. Kadie. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. in *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98)*. 1998.
- [25] Breslin, J., *The Future of Social Networks on the Internet: The Need for Semantics*, D. Stefan, Editor. 2007. p. 86-90.
- [26] Brown, J., A. Broderick, and N. Lee, *Word of mouth communication within online communities: Conceptualizing the online social network*. *Journal of Interactive Marketing*, 2007. **21**(3): p. 2-20.
- [27] Brusilovsky, P., D. Parra, S. Sahebi, and C. Wongchokprasitti, *Collaborative information finding in smaller communities: The case of research talks*. *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, 2010 6th International Conference on, 2010: p. 1-10.
- [28] Brzozowski, M.J., T. Hogg, and G. Szabo, *Friends and foes: ideological social networking*, in *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. 2008, ACM: Florence, Italy. p. 817-820.
- [29] Burke, R., *Hybrid web recommender systems*, in *The adaptive web: methods and strategies of web personalization*. 2007, Springer-Verlag. p. 377-408.
- [30] Cai, X., M. Bain, A. Krzywicki, W. Wobcke, Y. Kim, P. Compton, and A. Mahidadia, *Collaborative Filtering for People to People Recommendation in Social Networks AI 2010: Advances in Artificial Intelligence*, J. Li, Editor. 2011, Springer Berlin / Heidelberg. p. 476-485.
- [31] Carley, K., *A Theory of Group Stability*. *American Sociological Review*, 1991. **56**(3): p. 331-354.
- [32] Caverlee, J., L. Liu, and S. Webb. Socialtrust: tamper-resilient trust establishment in online communities. in *JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*. 2008. Pittsburgh PA, PA, USA: ACM.
- [33] Cetina, K.K., *Sociality with Objects*. *Theory, Culture & Society*, 1997. **14**(4): p. 1-30.
- [34] Chen, Y.-L., L.-C. Cheng, and C.-N. Chuang, *A group recommendation system with consideration of interactions among group members*. *Expert Syst. Appl.*, 2008. **34**(3): p. 2082-2090.

- [35] Chia, P.H. and G. Pitsilis, *Exploring the Use of Explicit Trust Links for Filtering Recommenders: A Study on Epinions. com*. Information and Media Technologies, 2011. **6**(3): p. 871-883.
- [36] Cialdini, R.B. and N.J. Goldstein, *Social influence: compliance and conformity*. Annu Rev Psychol, 2004. **55**: p. 591-621.
- [37] Crandall, D., D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri, *Feedback effects between similarity and social influence in online communities*, in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2008, ACM: Las Vegas, Nevada, USA. p. 160-168.
- [38] Cremonesi, P., Y. Koren, and R. Turrin, *Performance of recommender algorithms on top-n recommendation tasks*, in *Proceedings of the fourth ACM conference on Recommender systems*. 2010, ACM: Barcelona, Spain. p. 39-46.
- [39] Crossen, A., J. Budzik, and K.J. Hammond, *Flytrap: intelligent group music recommendation*, in *Proceedings of the 7th international conference on Intelligent user interfaces*. 2002, ACM: San Francisco, California, USA. p. 184-185.
- [40] Debnath, S., N. Ganguly, and P. Mitra, *Feature weighting in content based recommendation system using social network analysis*, in *Proceeding of the 17th international conference on World Wide Web*. 2008, ACM: Beijing, China. p. 1041-1042.
- [41] Deerwester, S., S. Dumais, G. Furnas, Thomas, and R. Harshman, *Indexing by latent semantic analysis*. Journal of the American Society for Information Science, 1990. **41**: p. 391-407.
- [42] Desrosiers, C. and G. Karypis, *A Comprehensive Survey of Neighborhood-based Recommendation Methods*, in *Recommender Systems Handbook*, R. Francesco, et al., Editors. 2011, Springer. p. 107-144.
- [43] Deutsch, M. and H. Gerard, *A study of normative and informational social influences upon individual judgement*. Journal of Abnormal & Social Psychology, 1955. **51**(3): p. 629-36.
- [44] DuBois, J., Golbeck, J. Kleint, J. & Srinivasan Improving Recommendation Accuracy by Clustering Social Networks with Trust. in *Proc. Of ACM RecSys 2009 Workshop on Recommender Systems and the Social Web*. 2009. New York, New York.
- [45] Dugan, C., M. Muller, D.R. Millen, W. Geyer, B. Brownholtz, and M. Moore, *The dogear game: a social bookmark recommender system*, in *Proceedings of the 2007 international ACM conference on Supporting group work*. 2007, ACM: Sanibel Island, Florida, USA. p. 387-390.
- [46] Dunning, T., *Accurate methods for the statistics of surprise and coincidence*. Comput. Linguist., 1993. **19**(1): p. 61-74.

- [47] Elliott, S., *Report Details Rise of Social Media*, in *The New York Times*. 2011: New York, NY.
- [48] Firan, C.S., W. Nejdl, and R. Paiu, *The Benefit of Using Tag-Based Profiles*, in *Proceedings of the 2007 Latin American Web Conference*. 2007, IEEE Computer Society. p. 32-41.
- [49] Fortunato, S., *Community detection in graphs*. Physics Reports, 2010. **486**(3-5): p. 75-174.
- [50] Friedkin, N.E., *Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study*. American Sociological Review, 1993. **58**(6): p. 861-872.
- [51] Garcia, R. and X. Amatriain. Weighted Content Based Methods for Recommendation Connections in Online Social Networks. in *Proceedings of the 2nd ACM RecSys Workshop on Recommender Systems and the Social Web*. 2010. Barcelona, Spain.
- [52] Godoy, D. and A. Amandi. Hybrid Content and Tag-based Profiles for Recommendation in Collaborative Tagging Systems. in *Latin American Web Conference, 2008. LA-WEB '08*. 2008.
- [53] Golbeck, J., *The dynamics of Web-based social networks: Membership, relationships, and change*. First Monday, 2007. **12**(11).
- [54] Golbeck, J., *Introduction to Computing with Social Trust*, in *Computing with Social Trust*, J. Golbeck, Editor. 2009, Springer: London. p. pp. 1 - 5.
- [55] Golbeck, J. and J. Hendler, *Accuracy of Metrics for Inferring Trust and Reputation in Semantic Web-Based Social Networks*, in *Engineering Knowledge in the Age of the Semantic Web*, E. Motta, et al., Editors. 2004, Springer Berlin / Heidelberg. p. 116-131.
- [56] Golbeck, J. and J. Hendler. FilmTrust: movie recommendations using trust in web-based social networks. in *Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE*. 2006.
- [57] Golder, S.A. and B.A. Huberman, *Usage patterns of collaborative tagging systems*. J. Inf. Sci., 2006. **32**(2): p. 198-208.
- [58] Granovetter, M., *The Strength of Weak Ties: A Network Theory Revised*. Sociological Theory, 1983. **1**: p. 201 ~ 233.
- [59] Groh, G. and C. Ehmig, *Recommendations in taste related domains: collaborative filtering vs. social filtering*, in *Proceedings of the 2007 international ACM conference on Supporting group work*. 2007, ACM: Sanibel Island, Florida, USA. p. 127-136.
- [60] Gruzd, A., K. Staves, and A. Wilk, *Tenure and Promotion in the Age of Online Social Media*, in *ASIST 2011 Annual Meeting* 2011: New Orleans, LA, USA.

- [61] Guan, Z., C. Wang, J. Bu, C. Chen, K. Yang, D. Cai, and X. He, *Document recommendation in social tagging services*, in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA. p. 391-400.
- [62] Guo, G., J. Zhang, and D. Thalmann, *A simple but effective method to incorporate trusted neighbors in recommender systems*. *User Modeling, Adaptation, and Personalization*, 2012: p. 114-125.
- [63] Guy, I., I. Ronen, and E. Wilcox, *Do you know?: recommending people to invite into your social network*, in *Proceedings of the 14th international conference on Intelligent user interfaces*. 2009, ACM: Sanibel Island, Florida, USA. p. 77-86.
- [64] Guy, I., N. Zwerdling, D. Carmel, I. Ronen, E. Uziel, S. Yogev, and S. Ofek-Koifman. *Personalized Recommendation of Social Software Items Based on Social Relations*. in *the 2009 ACM conference on Recommender systems, RecSys '09*. 2009. New York, NY: ACM.
- [65] Halpin, H., V. Robu, and H. Shepherd, *The complex dynamics of collaborative tagging*, in *Proceedings of the 16th international conference on World Wide Web*. 2007, ACM: Banff, Alberta, Canada. p. 211-220.
- [66] Hammond, T., T. Hannay, B. Lund, and J. Scott, *Social Bookmarking Tools (I)*. *D-Lib Magazine*, 2005. **11**(4).
- [67] Herlocker, J.L., J.A. Konstan, A. Borchers, and J. Riedl, *An algorithmic framework for performing collaborative filtering*, in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. 1999, ACM: Berkeley, California, United States. p. 230-237.
- [68] Hotho, A., R. Jäschke, C. Schmitz, and G. Stumme, *Information Retrieval in Folksonomies: Search and Ranking*. 2006. p. 411-426.
- [69] Hotho, A., R. Jäschke, C. Schmitz, and G. Stumme, *Trend Detection in Folksonomies*. 2006. p. 56-70.
- [70] Huang, Z., D. Zeng, and H. Chen, *A Comparison of Collaborative-Filtering Recommendation Algorithms for E-commerce*. *IEEE Intelligent Systems*, 2007. **22**(5): p. 68-78.
- [71] Hung, C.-C., Y.-C. Huang, J. Hsu, and D. Wu. *Tag-based User Profiling for Social Media Recommendation*. in *Workshop on Intelligent Techniques for Web Personalization and Recommender Systems, AAAI 2008*. 2008.
- [72] Huston, T.L. and G. Levinger, *Interpersonal attraction and relationships*. *Annual Review of Psychology*, 1978. **29**: p. 115-156.

- [73] Jamali, M. and M. Ester. *TrustWalker*: a random walk model for combining trust-based and item-based recommendation. in *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2009. Paris, France: ACM.
- [74] Jamali, M. and M. Ester, *A matrix factorization technique with trust propagation for recommendation in social networks*, in *Proceedings of the fourth ACM conference on Recommender systems*. 2010, ACM: Barcelona, Spain. p. 135-142.
- [75] James, K., S. Lycett, A. Wipat, and J. Hallinan, *Multiple Gold Standards Address Bias in Functional Network Integration*, in *Newcastle University, Computing Science Technical Reports Series*. 2011, Newcastle University.
- [76] Jandel, M. and M. Elahi, *Tribal taste: mobile multiagent recommender system*, in *Proceedings of the 13th international conference on Intelligent user interfaces*. 2009, ACM: Sanibel Island, Florida, USA. p. 489-490.
- [77] Jannach, D., M. Zanker, A. Felfernig, and G. Friedrich, *Recommender Systems: An Introduction*. 2011: Cambridge University Press, New York, New York.
- [78] Jansen, B.J., M. Zhang, K. Sobel, and A. Chowdury, *Twitter power: Tweets as electronic word of mouth*. *Journal of the American Society for Information Science and Technology*, 2009. **60**(11): p. 2169-2188.
- [79] Jäschke, R., L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, *Tag recommendations in social bookmarking systems*. *AI Communications*, 2008. **21**(4): p. 231-247.
- [80] Java, A., X. Song, T. Finin, and B. Tseng, *Why we twitter: understanding microblogging usage and communities*, in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. 2007, ACM: San Jose, California. p. 56-65.
- [81] Kalman, D., *A Singularly Valuable Decomposition: The SVD of a Matrix*. *The College Mathematics Journal*, 1996. **27**(1): p. 2-23.
- [82] Kalnikaite, V. and S. Whittaker, *Social summarization: does social feedback improve access to speech data?*, in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 2008, ACM: San Diego, CA, USA. p. 9-12.
- [83] Kautz, H., B. Selman, and M. Shah, *Referral Web: combining social networks and collaborative filtering*. *Commun. ACM*, 1997. **40**(3): p. 63-65.
- [84] Kazama, J.i. and K. Torisawa. *Exploiting Wikipedia as External Knowledge for Named Entity Recognition*. in *Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 2007.

- [85] Konstas, I., V. Stathopoulos, and J.M. Jose, *On social networks and collaborative recommendation*, in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. 2009, ACM: Boston, MA, USA. p. 195-202.
- [86] Koren, Y., R. Bell, and C. Volinsky, *Matrix Factorization Techniques for Recommender Systems*. Computer, 2009. **42**(8): p. 30-37.
- [87] Krestel, R., P. Fankhauser, and W. Nejdl, *Latent dirichlet allocation for tag recommendation*, in *Proceedings of the third ACM conference on Recommender systems*. 2009, ACM: New York, New York, USA. p. 61-68.
- [88] Kutty, S., L. Chen, and R. Nayak, *A people-to-people recommendation system using tensor space models*, in *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. 2012, ACM: Trento, Italy. p. 187-192.
- [89] Kwak, H., C. Lee, H. Park, and S. Moon, *What is Twitter, a social network or a news media?*, in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA. p. 591-600.
- [90] Lam, C., *SNACK: incorporating social network information in automated collaborative filtering*, in *Proceedings of the 5th ACM conference on Electronic commerce*. 2004, ACM: New York, NY, USA. p. 254-255.
- [91] Lam, S.K. and J. Riedl, *Shilling recommender systems for fun and profit*, in *Proceedings of the 13th international conference on World Wide Web*. 2004, ACM: New York, NY, USA. p. 393-402.
- [92] Lee, D., *How to measure the information similarity in unilateral relations: the case study of <i>Delicious</i>*, in *Proceedings of the International Workshop on Modeling Social Media*. 2010, ACM: Toronto, Ontario, Canada. p. 1-4.
- [93] Lee, D. and P. Brusilovsky. *Improving Recommendations Using Watching Networks in a Social Tagging System*. in *The Proceedings of IConference 2011*. 2011. Seattle, WA, USA.
- [94] Lee, D.H. and P. Brusilovsky. *Does Trust Influence Information similarity?* in *Proceedings of Workshop on Recommender Systems & the Social Web, the 3rd ACM International Conference on Recommender Systems*. 2009. New York, NY, USA.
- [95] Lee, D.H. and P. Brusilovsky, *Exploring Social Approach to Recommend Talks at Research Conferences in 8th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*. 2012: Pittsburgh, PA., USA.
- [96] Lee, D.H. and P. Brusilovsky, *Interest Similarity of Group Members: the Case study of Citeulike*, in *Proceedings of the WebSci10*. 2010: Raleigh, NC: US.

- [97] Lee, D.H. and P. Brusilovsky, *Social networks and interest similarity: the case of <i>CiteULike</i>*, in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*. 2010, ACM: Toronto, Ontario, Canada. p. 151-156.
- [98] Lee, D.H. and P. Brusilovsky. Social Networks and Interest Similarity: The Case of CiteULike. in *Proceedings of the 21th ACM conference on Hypertext and hypermedia*. 2010. Toronto, Canada: ACM.
- [99] Lee, D.H. and P. Brusilovsky, *Using self-defined group activities for improving recommendations in collaborative tagging systems*, in *Proceedings of the fourth ACM conference on Recommender systems*. 2010, ACM: Barcelona, Spain. p. 221-224.
- [100] Lee, D.H., P. Brusilovsky, and T. Schleyer, *Recommending Collaborators using Social Features and MeSH terms*, in *ASIST 2011 Annual Meeting* 2011: New Orleans, LA, USA.
- [101] Lee, D.H. and T. Schleyer, *MeSH term explosion and author rank improve expert recommendations*. AMIA Annu Symp Proc, 2010. **2010**: p. 412-6.
- [102] Li, J., J. Tang, J. Zhang, Q. Luo, Y. Liu, and M. Hong, *EOS: expertise oriented search using social networks*, in *Proceedings of the 16th international conference on World Wide Web*. 2007, ACM: Banff, Alberta, Canada. p. 1271-1272.
- [103] Li, J. and O. Zaiane, *Combining Usage, Content, and Structure Data to Improve Web Site Recommendation E-Commerce and Web Technologies*, K. Bauknecht, M. Bichler, and B. Pröll, Editors. 2004, Springer Berlin / Heidelberg. p. 313-315.
- [104] Li, X., L. Guo, and Y.E. Zhao, *Tag-based social interest discovery*, in *Proceeding of the 17th international conference on World Wide Web*. 2008, ACM: Beijing, China. p. 675-684.
- [105] Lin, C.-Y., N. Cao, S.X. Liu, S. Papadimitriou, J. Sun, and X. Yang. SmallBlue: Social Network Analysis for Expertise Search and Collective Intelligence. in *IEEE International Conference on Data Engineering*. 2009.
- [106] Lin, N., *Building a Network Theory of Social Capital*. Connections, 1999. **22**(1): p. 28-51.
- [107] Liu, B., *Informational Retrieval and Web Search*, in *Web Data Mining: Exploring Hyperlinks, Contents and Usage Data*, B. Liu, Editor. 2007, Springer: New York. p. pp. 183 - 236.
- [108] Liu, F. and H.J. Lee, *Use of social network information to enhance collaborative filtering performance*. Expert Systems with Applications, 2010. **37**(7): p. 4772-4778.
- [109] Liu, H. and P. Maes, *InterestMap: Harvesting Social Network Profiles for Recommendations*, in *Beyond Personalisation*. 2004.

- [110] Lumbreras, A. and R. Gavalda, *Applying trust metrics based on user interactions to recommendation in social networks*, in *Social Knowledge Discovery and Utilization Workshop*. 2012.
- [111] Lund, B., T. Hammond, M. Flack, and T. Hannay, *Social Bookmarking Tools (II)*. D-Lib Magazine, 2005. **11**(4): p. 1-1.
- [112] Ma, H., H. Yang, M.R. Lyu, and I. King, *SoRec: social recommendation using probabilistic matrix factorization*, in *Proceeding of the 17th ACM conference on Information and knowledge management*. 2008, ACM: Napa Valley, California, USA. p. 931-940.
- [113] Ma, H., T.C. Zhou, M.R. Lyu, and I. King, *Improving Recommender Systems by Incorporating Social Contextual Information*. ACM Trans. Inf. Syst., 2011. **29**(2): p. 1-23.
- [114] Magureanu, S., N. Dokoohaki, S. Mocarizadeh, and M. Matskin, *Design and Analysis of a Gossip-based Decentralized Trust Recommender System*, in *4th ACM Recsys Workshop on Recommender Systems and the Social Web*. 2012.
- [115] Maltz, D. and K. Ehrlich, *Pointing the way: active collaborative filtering*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1995, ACM Press/Addison-Wesley Publishing Co.: Denver, Colorado, United States. p. 202-209.
- [116] Manning, C.D., P. Raghavan, and H. Schütze, *Introduction to information retrieval*. 2008: Cambridge University Press.
- [117] Marks, S.R., *Intimacy in the Public Realm: The Case of Co-workers*. Social Forces, 1994. **72**(3): p. 843-858.
- [118] Massa, P. and P. Avesani, *Trust-Aware Collaborative Filtering for Recommender Systems*. Vol. 3290. 2004. 492-508.
- [119] Massa, P. and P. Avesani, *Trust-aware recommender systems*, in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007, ACM: Minneapolis, MN, USA. p. 17-24.
- [120] Masthof, J., *Computationally Modeling Trust: An Exploration*, in *SociUM: Adaptation and Personalisation in Social Systems: Groups, Teams and Communities: Workshop Proceedings*. 2007: Corfu, Greece.
- [121] Masthoff, J., *Group Recommender Systems: Combining Individual Models*, in *Recommender Systems Handbook*, F. Ricci, et al., Editors. 2011, Springer Science+Business Media, LLC. p. pp. 677 - 702.

- [122] Matsuo, Y. and H. Yamamoto, *Diffusion of recommendation through a trust network*, in *International AAAI Conference on Weblogs and Social Media*. 2007: Boulder, Colorado, USA.
- [123] McCarthy, K., L. McGinty, B. Smyth, and M. Salamó, *The Needs of the Many: A Case-Based Group Recommender System*. 2006. p. 196-210.
- [124] McDonald, D.W., *Recommending collaboration with social networks: a comparative evaluation*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 593-600.
- [125] McDonald, D.W. and M.S. Ackerman, *Expertise recommender: a flexible recommendation system and architecture*, in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 2000, ACM: Philadelphia, Pennsylvania, United States. p. 231-240.
- [126] McPherson, J.M., P.A. Popielarz, and S. Drobnic, *Social Networks and Organizational Dynamics*. *American Sociological Review*, 1992. **57**(2): p. 153-170.
- [127] McPherson, M., S. Lovin, and J. Cook, *BIRDS OF A FEATHER: Homophily in Social Networks*. *Annual Review of Sociology*, 2001. **27**: p. 415-445.
- [128] Mehta, B., T. Hofmann, and W. Nejdl, *Robust collaborative filtering*, in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007, ACM: Minneapolis, MN, USA. p. 49-56.
- [129] Mei, Q., D. Zhou, and K. Church. Query Suggestion Using Hitting Time. in *The 17th ACM conference on Conference on information and knowledge management: CIKM '08*. 2008. Napa Valley, CA, USA: ACM Press.
- [130] Melville, P., R.J. Mooney, and R. Nagarajan, *Content-boosted collaborative filtering for improved recommendations*, in *Eighteenth national conference on Artificial intelligence*. 2002, American Association for Artificial Intelligence: Edmonton, Alberta, Canada. p. 187-192.
- [131] Millen, D., J. Feinberg, and B. Kerr, *Social Bookmarking in the Enterprise*. *Queue*, 2005. **3**(9): p. 28-35.
- [132] Monge, P.E. and N.S. Contractor, *Homophily, Proximity, and Social Support Theories*, in *Theories of Communication Networks*. 2003, Oxford: New York, NY. p. 223 - 239.
- [133] Monge, P.E. and N.S. Contractor, *Network Concepts, Measures, and the Multitheoretical, Multilevel Analytic Framework*, in *Theories of Communication Networks*. 2003, Oxford: New York, NY. p. 29 - 77.

- [134] Mooney, R.J. and L. Roy, *Content-based book recommending using learning for text categorization*, in *Proceedings of the fifth ACM conference on Digital libraries*. 2000, ACM: San Antonio, Texas, United States. p. 195-204.
- [135] Newman, M., *Networks: An Introduction*. 2010, USA: Oxford University Press.
- [136] Newman, M.E., *Social Networks*, in *Networks: An Introduction*, O.N. York, Editor. 2010: New York, NY. p. 36 - 62.
- [137] Nielsen. *State of the Media: The Social Media Report*. 2011 [cited 2011 August 20]; Available from: <http://blog.nielsen.com/nielsenwire/social/>.
- [138] O'Connor, M., D. Cosley, J.A. Konstan, and J. Riedl, *PolyLens: a recommender system for groups of users*, in *Proceedings of the seventh conference on European Conference on Computer Supported Cooperative Work*. 2001, Kluwer Academic Publishers: Bonn, Germany. p. 199-218.
- [139] O'Donovan, J. and B. Smyth, *Is trust robust?: an analysis of trust-based recommendation*, in *Proceedings of the 11th international conference on Intelligent user interfaces*. 2006, ACM: Sydney, Australia. p. 101-108.
- [140] O'Donovan, J. and B. Smyth, *Trust in recommender systems*, in *Proceedings of the 10th international conference on Intelligent user interfaces*. 2005, ACM: San Diego, California, USA. p. 167-174.
- [141] O'Doherty, D., S. Jouili, and P. Van Roy. Trust-based recommendation: an empirical analysis. in *Submitted to: Proceedings of the Sixth ACM SIGKDD Workshop on Social Network Mining and Analysis SNA-KDD, Beijing, China, ACM*. 2012.
- [142] Olsson, T. Decentralized social filtering based on trust. in *In Working Notes of the AAAI-98 Recommender Systems Workshop*. 1998.
- [143] Palau, J., M. Montaner, B. López, and J.L. de la Rosa, *Collaboration Analysis in Recommender Systems Using Social Networks*. 2004. p. 137-151.
- [144] Palla, G., I. Derenyi, I. Farkas, and T. Vicsek, *Uncovering the overlapping community structure of complex networks in nature and society*. *Nature*, 2005. **435**(7043): p. 814-818.
- [145] Park, M.-H., H.-S. Park, and S.-B. Cho, *Restaurant Recommendation for Group of People in Mobile Environments Using Probabilistic Multi-criteria Decision Making Computer-Human Interaction*, S. Lee, et al., Editors. 2008, Springer Berlin / Heidelberg. p. 114-122.

- [146] Pazos Arias, J.J., A. Fernández Vilas, R.P. Díaz Redondo, G. Groh, S. Birnkammerer, and V. Köllhofer, *Social Recommender Systems Recommender Systems for the Social Web*. 2012, Springer Berlin Heidelberg. p. 3-42.
- [147] Pazzani, M. and D. Billsus, *Content-Based Recommendation Systems*, in *The Adaptive Web: Methods and Strategies of Web Personalization*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Editors. 2007, Springer: Berlin, Germany. p. 325-341.
- [148] Pera, M.S. and Y.-K. Ng, *With a Little Help from My Friends: Generating Personalized Book Recommendations Using Data Extracted from a Social Website*, in *Proceedings of the 2011 IEEE/WIC/ACM Joint Conference on Web Intelligent (WI'11)*. 2011: Lyon, France. p. pp. 96-99.
- [149] Pilaszy, I., D. Zibriczky, and D. Tikk, *Fast als-based matrix factorization for explicit and implicit feedback datasets*, in *Proceedings of the fourth ACM conference on Recommender systems*. 2010, ACM: Barcelona, Spain. p. 71-78.
- [150] Pitsilis, G. and S.J. Knapkog. Social Trust as a solution to address sparsity-inherent problems of recommender systems. in *Proc. Of ACM RecSys 2009 Workshop on Recommender Systems and the Social Web*. 2009. New York, New York.
- [151] Pizzato, L., T. Rej, J. Akehurst, I. Koprinska, K. Yacef, and J. Kay, *Recommending people to people: the nature of reciprocal recommenders with a case study in online dating*. User Modeling and User-Adapted Interaction: p. 1-42.
- [152] Pizzato, L., T. Rej, T. Chung, I. Koprinska, and J. Kay, *RECON: a reciprocal recommender for online dating*, in *Proceedings of the fourth ACM conference on Recommender systems*. 2010, ACM: Barcelona, Spain. p. 207-214.
- [153] Postmes, T., R. Spears, K. Sakhel, and D. de Groot, *Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior*. Personality and Social Psychology Bulletin, 2001. **27**(10): p. 1243-1254.
- [154] Preece, J. and D. Maloney-Krichmar, *Online communities: focusing on sociability and usability*, in *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*. 2003, L. Erlbaum Associates Inc. p. 596-620.
- [155] Roth, M., A. Ben-David, D. Deutscher, G. Flysher, I. Horn, A. Leichtberg, N. Leiser, Y. Matias, and R. Merom, *Suggesting friends using the implicit social graph*, in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2010, ACM: Washington, DC, USA. p. 233-242.
- [156] Salton, G., A. Wong, and C.S. Yang, *A vector space model for automatic indexing*. Commun. ACM, 1975. **18**(11): p. 613-620.

- [157] Sanchez, F., M. Barrilero, S. Uribe, F. Alvarez, A. Tena, and J. Menendez, *Social and Content Hybrid Image Recommender System for Mobile Social Networks*. Mobile Networks and Applications: p. 1-14.
- [158] Sarwar, B., G. Karypis, J. Konstan, and J. Riedl, *Incremental Singular Value Decomposition Algorithms for Highly Scalable Recommender Systems*, in *Fifth International Conference on Computer and Information Science*. 2002.
- [159] Schaefer, J.B., D. Frankowski, J. Herlocker, and S. Sen, *Collaborative Filtering Recommender Systems*, in *The Adaptive Web: Methods and Strategies of Web Personalization*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Editors. 2007, Springer: Berlin, Germany. p. 291-324.
- [160] Segaran, T., *Programming collective intelligence: building smart web 2.0 applications*. 2007: O'Reilly.
- [161] Shani, G. and A. Gunawardana, *Evaluating Recommender Systems*. Microsoft Research Report MSR-TR-2009-159, Nov, 2009.
- [162] Shi, Y., M. Larson, and A. Hanjalic. Towards Understanding the Challenges Facing Effective Trust-Aware Recommendation. in *Proceedings of the 2nd ACM RecSys Workshop on Recommender Systems and the Social Web*. 2010. Barcelona, Spain.
- [163] Singla, P. and M. Richardson, *Yes, there is a correlation: - from social networks to personal behavior on the web*, in *Proceeding of the 17th international conference on World Wide Web*. 2008, ACM: Beijing, China. p. 655-664.
- [164] Sinha, R. and K. Swearingen. Comparing Recommendations Made by Online Systems and Friends. in *In Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*. 2001.
- [165] Stecher, R., G. Demartini, C. Niederl, and G. Demartini, *Social recommendations of content and metadata*, in *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services*. 2008, ACM: Linz, Austria. p. 92-97.
- [166] Steck, H., *Item popularity and recommendation accuracy*, in *Proceedings of the fifth ACM conference on Recommender systems*. 2011, ACM: Chicago, Illinois, USA. p. 125-132.
- [167] Suchanek, F.M., M. Vojnovic, and D. Gunawardana, *Social tags: meaning and suggestions*, in *Proceeding of the 17th ACM conference on Information and knowledge management*. 2008, ACM: Napa Valley, California, USA. p. 223-232.
- [168] Sugiyama, K. and M.-Y. Kan, *Scholarly paper recommendation via user's recent research interests*, in *Proceedings of the 10th annual joint conference on Digital libraries*. 2010, ACM: Gold Coast, Queensland, Australia. p. 29-38.

- [169] Swamynathan, G., C. Wilson, B. Boe, K. Almeroth, and B.Y. Zhao, *Do social networks improve e-commerce?: a study on social marketplaces*, in *Proceedings of the first workshop on Online social networks*. 2008, ACM: Seattle, WA, USA. p. 1-6.
- [170] Symeonidis, P., A. Nanopoulos, and Y. Manolopoulos, *Tag recommendations based on tensor dimensionality reduction*, in *Proceedings of the 2008 ACM conference on Recommender systems*. 2008, ACM: Lausanne, Switzerland. p. 43-50.
- [171] Symeonidis, P., E. Tiakas, and Y. Manolopoulos. *Transitive Node Similarity for Link Prediction in Social Networks with Positive and Negative Links*. in *2010 ACM conference on Recommender systems, RecSys '10*. 2010. Barcelona, Spain: ACM.
- [172] Szomszor, M.N., I. Cantador, and H. Alani, *Correlating user profiles from multiple folksonomies*, in *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*. 2008, ACM: Pittsburgh, PA, USA. p. 33-42.
- [173] Terveen, L. and D.W. McDonald, *Social matching: A framework and research agenda*. *ACM Trans. Comput.-Hum. Interact.*, 2005. **12**(3): p. 401-434.
- [174] Tintarev, N. and J. Masthoff, *Effective explanations of recommendations: user-centered design*, in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007, ACM: Minneapolis, MN, USA. p. 153-156.
- [175] Trier, U. *The DBLP Computer Science Bibliography*. 2012 [cited 2012 January 31]; Available from: <http://www.informatik.uni-trier.de/~ley/db/>.
- [176] Turner, J.C., *Social influence*. 1991: Brooks/Cole.
- [177] Turner, J.C. and K.J. Reynolds, *Self-Categorization Theory*, in *Handbook of social psychology*, S.T. Fiske, D.T. Gilbert, and G. Lindzey, Editors. 2010, John Wiley & Sons.
- [178] Victor, P., C. Cornelis, M. de Cock, and M. Cock, *Social Recommender Systems Trust Networks for Recommender Systems*. 2011, Atlantis Press. p. 91-107.
- [179] Victor, P., C. Cornelis, M. De Cock, and P. Pinheiro da Silva, *Gradual trust and distrust in recommender systems*. *Fuzzy Sets and Systems*, 2009. **160**(10): p. 1367-1382.
- [180] von Ahn, L., *Games with a Purpose*. *IEEE Computer Magazine*, 2006. **June 2006**: p. 96 - 98.
- [181] Walter, F., S. Battiston, and F. Schweitzer, *A model of a trust-based recommendation system on a social network*. *Autonomous Agents and Multi-Agent Systems*, 2008. **16**(1): p. 57-74.

- [182] Walter, F.E., S. Battiston, and F. Schweitzer, *Personalised and dynamic trust in social networks*, in *RecSys*, D.B. Lawrence, et al., Editors. 2009, ACM. p. 197-204.
- [183] Walther, J.B., B. Van Der Heide, S.-Y. Kim, D. Westerman, and S.T. Tong, *The Role of Friends' Appearance and Behavior on Evaluations of Individuals on Facebook: Are We Known by the Company We Keep?* *Human Communication Research*, 2008. **34**(1): p. 28-49.
- [184] Wang, J., Z. Zhao, J. Zhou, H. Wang, B. Cui, and G. Qi, *Recommending Flickr groups with social topic model*. *Inf. Retr.*, 2012. **15**(3-4): p. 278-295.
- [185] Wang, Y. and V. Varadharajan. Role-based Recommendation and Trust Evaluation. in *The 9th IEEE International Conference on E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007. CEC/EEE 2007*. 2007.
- [186] Wellman, B., *Computer networks as social networks*. *Science*, 2001. **293**(5537): p. 2031-4.
- [187] Wikipedia. *Digital Bibliography & Library Project*. 2012 [cited 2012 January 31]; Available from: http://en.wikipedia.org/wiki/Digital_Bibliography_%26_Library_Project.
- [188] Wongchokprasitti, C., P. Brusilovsky, and D. Parra, *Conference Navigator 2.0: Community-Based Recommendation for Academic Conferences*, in *Proceedings of Workshop on Social Recommender Systems at the International Conference on Intelligent User Interfaces (IUI 2010)*. 2010: Hong Kong, China.
- [189] Xu, Y., X. Guo, J. Hao, J. Ma, R.Y.K. Lau, and W. Xu, *Combining social network and semantic concept analysis for personalized academic researcher recommendation*. *Decision Support Systems*, (0).
- [190] Xue, G.-R., C. Lin, Q. Yang, W. Xi, H.-J. Zeng, Y. Yu, and Z. Chen, *Scalable collaborative filtering using cluster-based smoothing*, in *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. 2005, ACM: Salvador, Brazil. p. 114-121.
- [191] Yang, S.J.H. and I.Y.L. Chen, *A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network*. *Int. J. Hum.-Comput. Stud.*, 2008. **66**(1): p. 36-50.
- [192] Yuan, Q., S. Zhao, L. chen, Y. Liu, S. Ding, X. Zhang, and W. Zheng. *Augmenting Collaborative Recommender by Fusing Explicit Social Relationships*. in *ACM RecSys'09 Workshop on Recommender Systems and the Social Web*. 2009. NYC, NY, USA.
- [193] Zanardi, V. and L. Capra, *Social ranking: uncovering relevant content using tag-based recommender systems*, in *Proceedings of the 2008 ACM conference on Recommender systems*. 2008, ACM: Lausanne, Switzerland. p. 51-58.

- [194] Zhang, F.-u. Analysis of Trust-Based E-Commerce Recommender Systems Under Recommendation Attacks. 2007.
- [195] Zhang, S., W. Wang, J. Ford, F. Makedon, and J. Pearlman, *Using Singular Value Decomposition Approximation for Collaborative Filtering*, in *Proceedings of the Seventh IEEE International Conference on E-Commerce Technology*. 2005, IEEE Computer Society. p. 257-264.
- [196] Zhao, S., N. Du, A. Nauerz, X. Zhang, Q. Yuan, and R. Fu, *Improved recommendation based on collaborative tagging behaviors*, in *Proceedings of the 13th international conference on Intelligent user interfaces*. 2008, ACM: Gran Canaria, Spain. p. 413-416.
- [197] Zhou, D., E. Manavoglu, J. Li, C.L. Giles, and H. Zha, *Probabilistic models for discovering e-communities*, in *Proceedings of the 15th international conference on World Wide Web*. 2006, ACM: Edinburgh, Scotland. p. 173-182.
- [198] Zhou, X., Y. Xu, Y. Li, A. Josang, and C. Cox, *The state-of-the-art in personalized recommender systems for social networking*. *Artif. Intell. Rev.*, 2012. **37**(2): p. 119-132.
- [199] Zhou, Y., D. Wilkinson, R. Schreiber, and R. Pan, *Large-Scale Parallel Collaborative Filtering for the Netflix Prize*, in *Proceedings of the 4th international conference on Algorithmic Aspects in Information and Management*. 2008, Springer-Verlag: Shanghai, China. p. 337-348.
- [200] Ziegler, C.-N. and J. Golbeck, *Investigating interactions of trust and interest similarity*. *Decis. Support Syst.*, 2007. **43**(2): p. 460-475.
- [201] Ziegler, C.-N. and G. Lausen. Analyzing Correlation between Trust and User Similarity in Online Communities. in *Trust Management: Second International Conference, iTrust 2004*. 2004. Oxford, UK: Springer.
- [202] Ziegler, C.-N., G. Lausen, and L. Schmidt-Thieme, *Taxonomy-driven computation of product recommendations*, in *Proceedings of the thirteenth ACM international conference on Information and knowledge management*. 2004, ACM: Washington, D.C., USA. p. 406-415.