

CLUSTERING APPROACH BASED ON FEATURE WEIGHTING FOR
RECOMMENDATION SYSTEM IN MOVIE DOMAIN

AMIR HOSSEIN NABIZADEH RAFSANJANI

A thesis submitted in fulfillment of the
requirements for the award of the degree of
Master of Science (Information Technology - Management)

Faculty of Computing
Universiti Teknologi Malaysia

JULY 2013

To my beloved Father and Mother

ACKNOWLEDGEMENT

I am heartily expressing my greater gratefulness to Allah s.w.t for His blessing and strength that He blessed to me during the completion of this research.

In preparing this dissertation, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main dissertation supervisor, Professor Dr. Naomie Salim, for encouragement, guidance, critics and friendship. Without her continued support and interest, this dissertation would not have been the same as presented here. Furthermore, I would like to thank my dear friend Mr. Amin Minouei who helped me during application development phase of my research.

I appreciate my family especially my father and mother because of their support to finish this study.

ABSTRACT

The advancement of the Internet has brought us into a world that represents a huge amount of information items such as movies, web pages, etc. with fluctuating quality. As a result of this massive world of items, people get confused and the question “Which one should I select?” arises in their minds. Recommendation Systems address the problem of getting confused about items to choose, and filter a specific type of information with a specific information filtering technique that attempts to present information items that are likely of interest to the user. A variety of information filtering techniques have been proposed for performing recommendations, including content-based and collaborative techniques which are the most commonly used approaches in recommendation systems. This dissertation introduces a new recommendation model, a feature weighting technique to cluster the user for recommendation top-n movies to avoid new user cold start and scalability problem. The distinctive point of this study lies in the methodology used to cluster the user and the methodology which is utilized to recommend movies to new users. The model makes it possible for the new users to define a weight for every feature of movie based on its importance to the new user in scale of one (with an increment of 0.1). By using these weights, it finds nearest cluster of users to the new user and suggests him the top-n movies (with the highest rate and most frequency) which are reviewed by users that are in the targeted cluster. Rating and Movie dataset were used during this study. Firstly, purity and entropy are applied to evaluate the clusters and then precision, recall and F1 metrics are used to assess the recommendation system. Eventually, the results of accuracy testing of proposed model are compared with two traditional models (OPENMORE and Movie Magician Hybrid) and based on the evaluation the level of preciseness of the proposed model is more better than Movie Magician Hybrid but worse than OPENMORE.

ABSTRAK

Kemajuan internet telah membawa kita ke dalam dunia yang mewakili sejumlah besar barangan maklumat seperti filem dengan kualiti yang berubah. Hasil daripada dunia barangan ini secara besar-besaran, orang keliru dan persoalan "Mana satu yang perlu saya pilih?" timbul dalam fikiran mereka. RS menangani masalah kekeliruan dalam memilih maklumat dan menapis jenis maklumat tertentu dengan teknik khusus penapisan maklumat yang cuba untuk membentangkan maklumat yang mungkin menarik minat pengguna. Pelbagai teknik penapisan maklumat telah dicadangkan untuk melaksanakan cadangan, termasuk teknik berasaskan kandungan dan kerjasama yang merupakan pendekatan yang paling biasa digunakan dalam RSs. Disertasi ini memperkenalkan, satu ciri teknik pemberat untuk kelompok pengguna bagi mencadangkan filem top-n untuk menghindar permulaan yang dingin bagi pengguna baru dan masalah berskala daripada sistem penentu. Perkara tersendiri kajian ini terletak pada kaedah yang digunakan untuk kelompok pengguna dan kaedah yang digunakan untuk mengesyorkan filem kepada pengguna baru. Model ini membolehkan pengguna baru untuk menentukan berat bagi setiap ciri filem berdasarkan kepentingan untuk pengguna baru dalam skala satu (dengan peningkatan sebanyak 0.1). Dengan menggunakan berat, ia mendapati kelompok terdekat dari pengguna kepada pengguna yang baru dan mencadangkan top-n filem (dengan kadar yang paling tinggi dan kekerapan yang paling). Dataset penilaian dan filem telah digunakan semasa kajian ini. Pertama, purity dan entropy digunakan untuk menilai kelompok dan kemudian precision, recall dan F1 metrics digunakan untuk menilai RS. Akhirnya, keputusan ujian ketepatan model yang dicadangkan berbanding dengan dua model tradisional (OPENMORE dan Movie Magician hibrid) dan berdasarkan penilaian tahap ketepatan model yang dicadangkan adalah lebih baik daripada Movie Magician Hibrid tetapi lebih teruk daripada OPENMORE.

TABLE OF CONTENTS

| CHAPTER | TITLE | PAGE |
|----------|--------------------------------------|----------|
| | ACKNOWLEDGEMENT | i |
| | ABSTRACT | ii |
| | ABSTRAK | vi |
| | TABLE OF CONTENTS | viiv |
| | LIST OF TABLES | v |
| | LIST OF FIGURES | ixii |
| | | |
| 1 | INTRODUCTION | 1 |
| | 1.1 Background of Study | 1 |
| | 1.2 Statement of the Problems | 4 |
| | 1.3 Purpose of the Study | 5 |
| | 1.4 Objectives of the Study | 5 |
| | 1.5 Research Questions | 6 |
| | 1.6 Significance of the Study | 6 |
| | 1.7 Scope | 6 |
| | | |
| 2 | LITERATURE REVIEW | 8 |
| | 2.1 Introduction | 8 |
| | 2.2 Collaborative Filtering (CF) | 10 |
| | 2.2.1 Memory-based CF Techniques | 11 |
| | 2.2.1.1 Similarity computation | 12 |
| | 2.2.2 Model-Based CF Techniques | 13 |
| | 2.2.2.1 Clustering Model | 13 |
| | 2.2.2.2 Feature weighting clustering | 14 |

| | | |
|----------|--|-----------|
| 2.2.2.3 | K-Means Clustering | 15 |
| 2.3 | Implicit and Explicit Ratings | 17 |
| 2.3.1 | Explicit Rating | 17 |
| 2.3.2 | Implicit Rating | 17 |
| 2.4 | Advantages and Disadvantages of CF | 18 |
| 2.5 | Content Base Filtering (CBF) | 20 |
| 2.5.1 | Advantages and Disadvantages of CBF | 21 |
| 2.6 | Hybrid Method | 22 |
| 2.6.1 | Weighted Hybrid Recommender | 23 |
| 2.6.2 | Switching Hybrid Recommender | 24 |
| 2.6.3 | Mixed Hybrid Recommender | 26 |
| 2.6.4 | Feature Combination Hybrid Recommender | 27 |
| 2.6.5 | Cascade Hybrid Recommender | 28 |
| 2.6.6 | Feature Augmentation Hybrid | 29 |
| 2.6.7 | Meta level Hybrid Recommender | 30 |
| 2.7 | Summary | 31 |
| 3 | RESEARCH METHODOLOGY | 32 |
| 3.1 | Introduction | 32 |
| 3.2 | Research design | 33 |
| 3.3 | Literature review | 35 |
| 3.4 | Dataset | 36 |
| 3.4.1 | Movies Dataset Structure | 36 |
| 3.4.2 | Rating Dataset Structure | 37 |
| 3.5 | Data cleaning | 37 |
| 3.6 | Dataset modifying | 38 |
| 3.6.1 | Rating Dataset | 38 |
| 3.6.2 | Movie Dataset | 38 |
| 3.7 | Movie features extraction | 39 |
| 3.8 | Movie features selection | 41 |
| 3.9 | Feature weighting | 41 |
| 3.10 | Data Clustering | 44 |
| 3.10.1 | K-means Clustering | 44 |
| 3.11 | System Evaluation | 45 |

| | | |
|----------|--|------------|
| 3.12 | Writing Report | 47 |
| 3.13 | Different between proposed and Debnath Model | 47 |
| 3.14 | Summary | 48 |
| 4 | RECOMMENDATION MODEL | 49 |
| 4.1 | Introduction | 49 |
| 4.2 | Proposed Model | 51 |
| 4.3 | Phase 1 and 2 - Calculation W_F, W_U and W_T | 53 |
| 4.3.1 | W_F Calculation | 53 |
| 4.3.2 | W_U and W_T Calculation | 55 |
| 4.3.3 | Users' array weights | 63 |
| 4.4 | Phase 3- User based clustering | 64 |
| 4.4.1 | Cluster members determination | 69 |
| 4.5 | Phase 4 – New user cluster | 70 |
| 4.6 | System Evaluation | 72 |
| 4.7 | Conclusion | 76 |
| 5 | DISCUSSION | 78 |
| 5.1 | Introduction | 78 |
| 5.2 | Discussion | 82 |
| 5.3 | Research contribution | 83 |
| 5.4 | Limitations and Restrictions | 84 |
| 5.5 | Future Works | 85 |
| 5.6 | Conclusion | 85 |
| | REFERENCES | 86 |
| | APPENDIX A | 94 |
| | APPENDIX B | 95 |
| | APPENDIX C | 96 |
| | APPENDIX D | 102 |

LIST OF TABLES

| TABLE NO. | TITLE | PAGE |
|------------------|--|-------------|
| 2.1 | Compare Implicit and Explicit approach | 18 |
| 2.2 | Summarization of Hybridization methods | 31 |
| 3.1 | Movie features | 40 |
| 3.2 | Feature distance measure | 41 |
| 3.3 | User and Movie ID | 42 |
| 3.4 | Differences between proposed and Debnath model | 48 |
| 4.1 | Distance measure | 53 |
| 4.2 | Feature value of movie A&B with corresponding distance | 54 |
| 4.3 | Movie feature weight | 55 |
| 4.4 | Genre's weight | 56 |
| 4.5 | Writer's weight | 57 |
| 4.6 | Year's weight | 58 |
| 4.7 | Director's weight | 59 |
| 4.8 | Company's weight | 60 |
| 4.9 | Actor1's weight | 61 |
| 4.10 | Actor2's weight | 62 |
| 4.11 | Actor3's weight | 63 |
| 4.12 | Users' array weights | 64 |
| 4.13 | Initial Cluster Centers for two clusters | 65 |
| 4.14 | Final Cluster Centers for two clusters | 65 |
| 4.15 | Initial Cluster Centers for seven clusters | 66 |
| 4.16 | Final Cluster Centers for seven clusters | 67 |
| 4.17 | Cluster member determination | 69 |
| 4.18 | Distance calculation example | 70 |

| | | |
|------|---|----|
| 4.19 | Novelties of model and deliverables of each phase | 71 |
| 4.20 | Precision, Recall analysis | 73 |
| 4.21 | F1 analyses | 74 |

LIST OF FIGURES

| FIGURE NO. | TITLE | PAGE |
|------------|---|------|
| 2.1 | Literature review taxonomy | 10 |
| 2.2 | K-means clustering algorithm (MacQueen, 1967) | 16 |
| 2.3 | Content-based filtering (Semeraro, 2010) | 20 |
| 2.4 | Hybrid recommendation system (Rodríguez <i>et al.</i> , 2010) | 22 |
| 2.5 | Fab architecture (Balabanović and Shoham, 1997) | 23 |
| 2.6 | The P-Tango architecture (Claypool <i>et al.</i> , 1999) | 24 |
| 2.7 | The Daily Learner architecture (Tran and Cohen, 2000) | 25 |
| 2.8 | PTV System Architecture (Smyth and Cotter, 2001) | 26 |
| 2.9 | The Entree restaurant recommender: initial screen | 28 |
| 3.1 | Research design | 33 |
| 3.2 | Phase 1 | 36 |
| 3.3 | Overview of Rating dataset | 37 |
| 3.4 | Overview of pre-processing phase | 39 |
| 3.5 | Social graph | 43 |
| 3.6 | K-means flowchart | 45 |
| 4.1 | Overview of recommendation system | 50 |
| 4.2 | Proposed Model | 52 |
| 4.3 | Visualization of two clusters | 66 |
| 4.4 | Visualization of seven clusters | 68 |
| 4.5 | Precision, Recall graph | 73 |
| 4.6 | F1 graph | 74 |
| 4.7 | Precision, Recall and F1 graph | 75 |
| 4.8 | Precision, Recall and F-measure for Different Methodologies | 76 |

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Recommendation is becoming one of the most important methods to provide documents, merchandises, and co-operators to response user requirements in providing information, trade, and services that are for society (community services), whether through mobile or on the web (Jameson *et al.*, 2002).

The internet started improving and maturing with unbelievable pace mid creation the epoch of Web 2.0. Therefore, a lot of opportunities emerge like partaking with other people in information, knowledge and ideas. This work causes coming up the social network such as face book. In current times, the musicians who are dilettante can be famous really faster than musicians who do not use this innovation by uploading their music that everybody can listen to their music. For another instance, writers can share their opinions and their creations with other people in all over the world. The merchants and also trading can find their customers more and more and earn their income from the Internet. A lot of shops and markets emerged and opened up in the internet. Currently, every person who work and use of World Wide Web (WWW) can buy nearly everything that require in every time and in everywhere. One of the best characteristics that the internet has is unlimitation space so everyone does not have limitation in getting space or on the other hand this space is endless. In spite of these benefits and characteristics, people face several new problems in World Wide Web.

The quantity of data and information has been increasing daily which causes overloading of information and data. At this time, finding the customers' requirements and tendencies became important as this problem changed into the big problem. One of the innovations which helped people a lot are the engines for search (search engines) and they were somewhat as a solution for this problem.

However, the information could not be personalized by these engines. Thus, the system developers introduced a solution for this problem that named recommendation system. This system is used to sort and filter information, data, and objects. Recommendation systems utilize users' idea of a society or community to assist for realization effectively users' tendency and also demand in a society from a possibly onerous set of selections (Resnick and Varian, 1997).

The main aim of recommendation system is creating significant suggestions and recommendations information, products or objects for users' society that users could interest in those. For instance, book recommendation on Amazon site, Netflix which recommend movies by using recommendation systems to identify users' tendencies and subsequently, attract users more and more (Melville and Sindhwani, 2010).

There are a lot of different methods and algorithms which can assist recommendation systems to create recommendations that are personalized. All of the recommendation approaches can be divided in these three famous categories:

- Content-based recommending: This method suggests and recommends objects and information which are comparable in content to objects that the users have interested previously, or compared and matched to the users' characteristics.
- Collaborative Filtering (CF): Collaborative Filtering systems suggested and recommended objects and information to a user according to the history valuation of all users communally.

- Hybrid methods: Hybrid methods are a combination of Content-based recommending and Collaborative Filtering (CF) methods (Melville and Sindhvani, 2010).

Content- based method recommends based on user's profile characteristics while Collaborative filter method only uses and assesses the traditional activities and finally, the effort of Hybrid method combines both methods (Melville and Sindhvani, 2010).

Content-based recommending and Collaborative Filtering (CF) are foundation of most new generation of recommendation systems. Recommendation systems entered in several fields like education, security, tourism and other fields. Semantic, context-aware and other methods are used by the new generation of recommender system to develop and make better their precision of recommendations. Currently, recommendation systems present suggestions and recommendations that are more personalized and specialized than previous ones (Asanov, 2011).

Collaborative filtering (CF) is designed to work on enormous database (Takács *et al.*, 2009). Goldberg *et al.* at 1992 used Collaborative filtering (CF) to introduce their filtering system that gives ability to customer for description their documents and e-mails (Goldberg *et al.*, 1992). Content-based recommendation algorithms can be seen as a extended work that is performed on filtering of information (Hanani *et al.*, 2001).

Previously, traditional recommendation systems depend only on models of the users that able to build from the "cold start" problem. The progress of latter decade, sharing data and internet connectivity make recommendation systems powerful to identify the users' tendency and bootstrap the model of the users from external resources like other recommendation systems.

On the other hand, in this study we want to propose a recommendation system method utilizing clustering technique which is based on movie feature weights.

Feature weighting or selection is a very important process to identify a significant subset of features from a data set. Removing redundant or irrelevant features can improve the generalization performance of ranking functions in information retrieval. A state of the art feature selection method, has been recently proposed, which exploits importance of each feature and similarity between every pair of features.

During this study has tried to propose a recommendation system method to remove the new user cold start and scalability problem by utilizing clustering technique which is based on movie feature weights.

1.2 Statement of the Problems

The quantity of data and information has been increasing daily which causes overloading of information and data. At this time, finding the customers' requirements and also customer precedence or their tendencies became important as this problem changed into the big problem. As it is obvious, when a customer faces with a huge number of items or information, he/she becomes confuse and thinks what should it do? Or what items is better than any others? Or even what is my interest? Then the procedure of search to find relevant or interest objects and items is usually takes long time and complicated while these days majority of people do not have enough time to do this work and also increasing competition on customers and markets among companies causes requirement for a new technology to identify users' demands and tendencies and also predict customers' expectation.

One of the ways to solve this problem is, utilizing of recommendation systems. Recommendation systems generate significant suggestions and recommendations of information, products or objects for users' society that users

interest in those. Then by this kind of systems, we can reduce time consuming and also the procedure of finding relevant or interest objects and items become easier and more efficient than before. This kind of system divided into three parts that include: Collaborative Filtering (CF), Content-based recommending and Hybrid methods.

During this study, we address two of the major problems suffered by recommendation system which is the new user's cold start and scalability. It means when recommendation systems cannot predict new user or customer precedence because of lack of adequate information and scalability means when the number of objects and customer increases, the traditional form of Collaborative Filtering (CF) suffers critical from scalability problem. Therefore, the results and recommendations that suggest to customer are in low quality and are not accurate. Thus, trying to find a solution for these problems can be a good motivation for this study.

1.3 **Purpose of the Study**

At this part of research we want to focus on the purposes of this investigation. The most important aim of this research is trying to propose a method to remove new user cold start and scalability from recommendation system by utilizing clustering technique which is based on movie feature weights.

1.4 **Objectives of the Study**

In this part of study, we want to mention two objectives of this research and items that we mostly focus on those:

1. To predict new users to clusters based on users' preferred features.
2. To design a recommendation system based on preferred movies of similar users in the same feature-weighted cluster.

1.5 **Research Questions**

In this part of study, we want to mention two questions of this research that we want to answer those until at the end of this study.

1. How can the weights of movie features be used to cluster and assigns users in recommendation system?
2. How can we recommend movies based on feature-weighted clusters?

1.6 **Significance of the Study**

Although, there is no proof that shows the current result of recommendation systems being useless or not useful for customer. But the findings of this study are very important to improve results of recommendation system that help customer with less consumption of time and energy, and find better and more efficient result that relevant to their precedence. With the information at hand, the recommendation system separated into three kinds: Collaborative Filtering (CF), Content-based recommending and Hybrid methods. During this study has tried to propose a method to remove new user cold start and scalability from recommendation system by utilizing clustering technique which is based on movie feature weights

1.7 **Scope**

During this research, we have several objectives: To identify the weight of movie features that can be used for RS to enhance the accuracy of clustering, to propose a clustering technique based on feature weighting with respect to new user cold start and the scalability problem of the RS, to develop and evaluate the proposed method by user. In general the most important goal of this study is proposing a method to remove new user cold start and scalability from recommendation system by utilizing clustering technique which is based on movie feature weights.

Hence, at first we try to analyze all methods that recommendation system used like CF, CBF, Hybrid and etc. To achieve this goal, we investigate research papers' that have performed and published on recommendation system and also semantic similarity. Afterwards, we try to classify and categorize those.

There are several sites which provide journals and papers about recommendation system and ontology that we most focus on those to acquire research papers and journals. All of the articles that are perusing collected from these kinds of Websites to comprehensive ontology and recommendation systems' techniques, deficiencies.

Then in this study we mostly focus on the previous papers and journals that have performed on recommendation system and also feature weighting techniques and methods. The data set of this study is movie, that's why we mostly focus on the research the performed on this kind of data set to comprehend more about our work and also to design a practical and good model to improve the result of recommendation system.

REFERENCES

- Adomavicius, G. and Tuzhilin, A. (2004). Recommendation Technologies: Survey of Current Methods and Possible Extensions.
- Ahmad Wasfi, A. M. (1998). Collecting User Access Patterns for Building User Profiles and Collaborative Filtering. Proceedings of the 1998 ACM Proceedings of the 4th international conference on Intelligent user interfaces. ACM, 57-64.
- Ansari, A., Essegai, S. and Kohli, R. (2000). Internet Recommendation Systems. Journal of Marketing research. 37(3), 363-375.
- Asanov, D. (2011). Algorithms and Methods in Recommender Systems.
- Balabanović, M. (1998). Exploring Versus Exploiting When Learning User Models for Text Recommendation. User Modeling and User-Adapted Interaction. 8(1), 71-102.
- Balabanović, M. and Shoham, Y. (1997). Fab: Content-Based, Collaborative Recommendation. Communications of the ACM. 40(3), 66-72.
- Basu, C., Hirsh, H. and Cohen, W. (1998). Recommendation as Classification: Using Social and Content-Based Information in Recommendation. Proceedings of the 1998a John Wiley & Sons LTD Proceedings of the national conference on artificial intelligence. John Wiley & Sons LTD, 714-720.
- Bhaidani, S. (2008). Recommender System Algorithms. University of Toronto.
- Birtürk, A. (2010). A Content Based Movie Recommendation System Empowered by Collaborative Missing Data Prediction. MIDDLE EAST TECHNICAL UNIVERSITY.
- Birukov, A., Blanzieri, E. and Giorgini, P. (2005). Implicit: A Recommender System That Uses Implicit Knowledge to Produce Suggestions.
- Blanco-Fernández, Y., Pazos-Arias, J. J., Gil-Solla, A., Ramos-Cabrera, M., López-Nores, M., García-Duque, J., Fernández-Vilas, A., Díaz-Redondo, R. P. and Bermejo-Muñoz, J. (2008). A Flexible Semantic Inference Methodology to

- Reason About User Preferences in Knowledge-Based Recommender Systems. *Knowledge-Based Systems*. 21(4), 305-320.
- Breese, J. S., Heckerman, D. and Kadie, C. (1998). Empirical Analysis of Predictive Algorithms for Collaborative Filtering. *Proceedings of the 1998 Morgan Kaufmann Publishers Inc. Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 43-52.
- Burke, R. (1999). Integrating Knowledge-Based and Collaborative-Filtering Recommender Systems. *Proceedings of the 1999 Proceedings of the Workshop on AI and Electronic Commerce*. 69-72.
- Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*. 12(4), 331-370.
- Burke, R. D., Hammond, K. J. and Yound, B. (1997). The Findme Approach to Assisted Browsing. *IEEE Expert*. 12(4), 32-40.
- Chiu, C.-C. and Tsai, C.-Y. (2007). A Weighted Feature C-Means Clustering Algorithm for Case Indexing and Retrieval in Cased-Based Reasoning. *New Trends in Applied Artificial Intelligence*. (pp. 541-551) Springer.
- Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M. (1999). Combining Content-Based and Collaborative Filters in an Online Newspaper. *Proceedings of the 1999 Citeseer Proceedings of ACM SIGIR Workshop on Recommender Systems*. Citeseer.
- Cordeiro de Amorim, R. and Mirkin, B. (2012). Minkowski Metric, Feature Weighting and Anomalous Cluster Initializing in K-Means Clustering. *Pattern Recognition*. 45(3), 1061-1075.
- Debnath, S. (2008). *Machine Learning Based Recommendation System*. Indian Institute of Technology.
- Debnath, S., Ganguly, N. and Mitra, P. (2008). Feature Weighting in Content Based Recommendation System Using Social Network Analysis. *Proceedings of the 2008 ACM Proceedings of the 17th international conference on World Wide Web*. ACM, 1041-1042.
- Deng, Y., Wu, Z., Tang, C., Si, H., Xiong, H. and Chen, Z. (2010). A Hybrid Movie Recommender Based on Ontology and Neural Networks. *Proceedings of the 2010 IEEE Computer Society Proceedings of the 2010 IEEE/ACM Int'l*

- Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing. IEEE Computer Society, 846-851.
- Dovidio, J. F., Kawakami, K. and Gaertner, S. L. (2002). Implicit and Explicit Prejudice and Interracial Interaction. *Journal of personality and social psychology*. 82(1), 62.
- Frias-Martinez, E., Chen, S. Y. and Liu, X. (2009). Evaluation of a Personalized Digital Library Based on Cognitive Styles: Adaptivity Vs. Adaptability. *International Journal of Information Management*. 29(1), 48-56.
- Frias-Martinez, E., Magoulas, G., Chen, S. and Macredie, R. (2006). Automated User Modeling for Personalized Digital Libraries. *International Journal of Information Management*. 26(3), 234-248.
- Garcia, R. and Amatriain, X. (2010). Weighted Content Based Methods for Recommending Connections in Online Social Networks. *Proceedings of the 2010 Workshop on Recommender Systems and the Social Web, Barcelona, Spain*. 68-71.
- George, T. and Merugu, S. (2005). A Scalable Collaborative Filtering Framework Based on Co-Clustering. *Proceedings of the 2005 IEEE Data Mining, Fifth IEEE International Conference on*. IEEE, 4 pp.
- Georgiou, O. and Tsapatsoulis, N. (2010). Improving the Scalability of Recommender Systems by Clustering Using Genetic Algorithms. *Artificial Neural Networks–Icann 2010*. (pp. 442-449) Springer.
- Goldberg, D., Nichols, D., Oki, B. M. and Terry, D. (1992). Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*. 35(12), 61-70.
- Goldberg, K., Roeder, T., Gupta, D. and Perkins, C. (2001). Eigentaste: A Constant Time Collaborative Filtering Algorithm. *Information Retrieval*. 4(2), 133-151.
- Hanani, U., Shapira, B. and Shoval, P. (2001). Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction*. 11(3), 203-259.
- Hangal, S., MacLean, D., Lam, M. S. and Heer, J. (2010). All Friends Are Not Equal: Using Weights in Social Graphs to Improve Search. *Proceedings of the 2010 4th ACM Workshop on SONAM*.

- Hofmann, T. (2004). Latent Semantic Models for Collaborative Filtering. *ACM Transactions on Information Systems (TOIS)*. 22(1), 89-115.
- Huang, C. and Yin, J. (2010). Effective Association Clusters Filtering to Cold-Start Recommendations. *Proceedings of the 2010 IEEE Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on. IEEE*, 2461-2464.
- Jameson, A., Konstan, J. and Riedl, J. (2002). *Ai Techniques for Personalized Recommendation*. Tutorial notes, *AAAI*. 2.
- Jawaheer, G., Szomszor, M. and Kostkova, P. (2010). Comparison of Implicit and Explicit Feedback from an Online Music Recommendation Service. *Proceedings of the 2010 ACM Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems*. ACM, 47-51.
- Johnson, S. C. (1967). Hierarchical Clustering Schemes. *Psychometrika*. 32(3), 241-254.
- Kim, H. N., Ji, A. T., Ha, I. and Jo, G. S. (2010). Collaborative Filtering Based on Collaborative Tagging for Enhancing the Quality of Recommendation. *Electronic Commerce Research and Applications*. 9(1), 73-83.
- Kim, T.-H. and Yang, S.-B. (2005). An Effective Recommendation Algorithm for Clustering-Based Recommender Systems. *Ai 2005: Advances in Artificial Intelligence*. (pp. 1150-1153) Springer.
- Kirmemis, O. and Birturk, A. (2008). A Content-Based User Model Generation and Optimization Approach for Movie Recommendation. *Proceedings of the 2008 Workshop on ITWP*.
- Lam, S. K. and Riedl, J. (2004). Shilling Recommender Systems for Fun and Profit. *Proceedings of the 2004 ACM Proceedings of the 13th international conference on World Wide Web*. ACM, 393-402.
- Lang, K. (1995). News Weeder: Learning to Filter Netnews. *Proceedings of the 1995 MORGAN KAUFMANN PUBLISHERS, INC. MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-. MORGAN KAUFMANN PUBLISHERS, INC.*, 331-339.

- Li, Q. and Kim, B. M. (2003). Clustering Approach for Hybrid Recommender System. Proceedings of the 2003 IEEE Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on. IEEE, 33-38.
- Linden, G., Smith, B. and York, J. (2003). Amazon. Com Recommendations: Item-to-Item Collaborative Filtering. Internet Computing, IEEE. 7(1), 76-80.
- Littlestone, N. and Warmuth, M. K. (1989). The Weighted Majority Algorithm. Proceedings of the 1989 IEEE Foundations of Computer Science, 1989., 30th Annual Symposium on. IEEE, 256-261.
- MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the 1967 California, USA Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. California, USA, 14.
- Medhi, D. G. and Dakua, J. (2005). Moviereco: A Recommendation System. Proceedings of the 2005 WEC (2). 70-73.
- Melville, P. and Sindhvani, V. (2010). Recommender Systems. Encyclopedia of machine learning. 829-838.
- Miller, B. N., Konstan, J. A. and Riedl, J. (2004). Pocketlens: Toward a Personal Recommender System. ACM Transactions on Information Systems (TOIS). 22(3), 437-476.
- Moghaddam, S. G. and Selamat, A. (2011). A Scalable Collaborative Recommender Algorithm Based on User Density-Based Clustering. Proceedings of the 2011 IEEE Data Mining and Intelligent Information Technology Applications (ICMiA), 2011 3rd International Conference on. IEEE, 246-249.
- Mooney, R. J. and Roy, L. (2000). Content-Based Book Recommending Using Learning for Text Categorization. Proceedings of the 2000 ACM Proceedings of the fifth ACM conference on Digital libraries. ACM, 195-204.
- Nichols, D. M. (1997). Implicit Rating and Filtering. In Proc. 5th DELOS Workshop on Filtering and Collaborative Filtering. 3.
- Oard, D. W. and Kim, J. (1998). Implicit Feedback for Recommender Systems. Proceedings of the 1998 Proceedings of the AAAI Workshop on Recommender Systems. 81-83.
- Oard, D. W. and Marchionini, G. (1998). A Conceptual Framework for Text Filtering Process.

- Pazzani, M. J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review*. 13(5), 393-408.
- Pine, I. (1995). J.; Peppers, D.; and Rogers, M. "Do You Want to Keep Your Customers Forever?". *Harvard business review*. 73(2).
- Rashid, A. M., Albert, I., Cosley, D., Lam, S. K., McNee, S. M., Konstan, J. A. and Riedl, J. (2002). Getting to Know You: Learning New User Preferences in Recommender Systems. *Proceedings of the 2002 ACM Proceedings of the 7th international conference on Intelligent user interfaces*. ACM, 127-134.
- Reichheld, F. F. (1993). Loyalty-Based Management. *Harvard business review*. 71(2), 64-73.
- Reichheld, F. F. and Sasser, W. E. (1990). Zero Defections: Quality Comes to Services. *Harvard business review*. 68(5), 105-111.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews. *Proceedings of the 1994 ACM Proceedings of the 1994 ACM conference on Computer supported cooperative work*. ACM, 175-186.
- Resnick, P. and Varian, H. R. (1997). Recommender Systems. *Communications of the ACM*. 40(3), 56-58.
- Rodríguez, R. M., Espinilla, M., Sánchez, P. J. and Martínez-López, L. (2010). Using Linguistic Incomplete Preference Relations to Cold Start Recommendations. *Internet Research*. 20(3), 296-315.
- Salamó, M., Escalera, S. and Radeva, P. (2009). Quality Enhancement Based on Reinforcement Learning and Feature Weighting for a Critiquing-Based Recommender. *Case-Based Reasoning Research and Development*. (pp. 298-312) Springer.
- Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2001). Item-Based Collaborative Filtering Recommendation Algorithms. *Proceedings of the 2001 ACM Proceedings of the 10th international conference on World Wide Web*. ACM, 285-295.
- Sarwar, B. M., Karypis, G., Konstan, J. and Riedl, J. (2002). Recommender Systems for Large-Scale E-Commerce: Scalable Neighborhood Formation Using Clustering. *Proceedings of the 2002 Proceedings of the fifth international conference on computer and information technology*.

- Schafer, J., Frankowski, D., Herlocker, J. and Sen, S. (2007). Collaborative Filtering Recommender Systems. *The adaptive web*. 291-324.
- Schein, A. I., Popescul, A., Ungar, L. H. and Pennock, D. M. (2002). Methods and Metrics for Cold-Start Recommendations. *Proceedings of the 2002 ACM Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 253-260.
- Semeraro, G. (2010). Content-Based Recommender Systems Problems, Challenges and Research Directions. *INTELLIGENT TECHNIQUES FOR WEB PERSONALIZATION & RECOMMENDER SYSTEMS (ITWP 2010)*. 89.
- Shardanand, U. and Maes, P. (1995). Social Information Filtering: Algorithms for Automating “Word of Mouth”. *Proceedings of the 1995 ACM Press/Addison-Wesley Publishing Co. Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 210-217.
- Smyth, B. and Cotter, P. (2001). Personalized Electronic Program Guides for Digital Tv. *Ai Magazine*. 22(2), 89.
- Su, X. and Khoshgoftaar, T. M. (2009). A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence*. 2009, 4.
- Sun, D., Li, C. and Luo, Z. (2011). A Content-Enhanced Approach for Cold-Start Problem in Collaborative Filtering. *Proceedings of the 2011 IEEE Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011 2nd International Conference on*. IEEE, 4501-4504.
- Symeonidis, P., Nanopoulos, A. and Manolopoulos, Y. (2007). Feature-Weighted User Model for Recommender Systems. *User Modeling 2007*. (pp. 97-106) Springer.
- Takács, G., Pilászy, I., Németh, B. and Tikk, D. (2009). Scalable Collaborative Filtering Approaches for Large Recommender Systems. *The Journal of Machine Learning Research*. 10, 623-656.
- Terveen, L. and Hill, W. (2001). Beyond Recommender Systems: Helping People Help Each Other. *HCI in the New Millennium*. 1, 487-509.
- Tran, T. and Cohen, R. (2000). Hybrid Recommender Systems for Electronic Commerce. *Proceedings of the 2000 Proc. Knowledge-Based Electronic*

- Markets, Papers from the AAAI Workshop, Technical Report WS-00-04, AAAI Press.
- Uluyagmur, M., Cataltepe, Z. and Tayfur, E. (2012). Content-Based Movie Recommendation Using Different Feature Sets. Proceedings of the 2012 Proceedings of the World Congress on Engineering and Computer Science.
- Van Setten, M. (2005). Supporting People in Finding Information: Hybrid Recommender Systems and Goal-Based Structuring.
- Wang, W., Zhang, D. and Zhou, J. (2012). Coba: A Credible and Co-Clustering Filterbot for Cold-Start Recommendations. Practical Applications of Intelligent Systems. (pp. 467-476) Springer.
- Wanner, L. (2004). Introduction to Clustering Technique. IULA. 32.
- Werthner, H., Hansen, H. R. and Ricci, F. (2007). Recommender Systems. Proceedings of the 2007 IEEE System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on. IEEE, 167-167.
- Yanxiang, L., Deke, G., Fei, C. and Honghui, C. (2013). User-Based Clustering with Top-N Recommendation on Cold-Start Problem. Proceedings of the 2013 IEEE Intelligent System Design and Engineering Applications (ISDEA), 2013 Third International Conference on. IEEE, 1585-1589.
- Yu, H., Oh, J. and Han, W.-S. (2009). Efficient Feature Weighting Methods for Ranking. Proceedings of the 2009 ACM Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 1157-1166.