Mohit Patel
SJSU ID: 014501461

Different approaches/models and problems in recommender systems: A Meta-Analytic review

A Project Report
Presented to
The faculty of the Department of Computer Engineering
San Jose State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Software Engineering

by
Mohit Patel
SJSU ID: 014501461
December 2020

Abstract

Different approaches/models and problems in recommender systems: A Meta-Analytic review
by Mohit Patel

Due to the adoption of internet and growth in the information available on the internet, a lot of the companies are trying to build models that can recommend products to their customers effectively. However, with the rapidly growing data and changing customer preferences, the existing recommender systems face problems such as 'Cold Start', 'Sparsity' and 'First Time Rater'. In this article, we identify the problems that existing recommender systems face and propose solutions such as demographic information and hybrid recommendation technique for overcoming these problems. The purpose of this study is to analyze the different approaches to recommender systems and devise a solution to overcome the identified problems. The anticipated result is a recommender system model that is free of all the problems identified in this study. Breach of user privacy is an anticipated issue of the provided solution. The major impact of this study is the identification of how different approaches to recommender systems tackle different problems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction:

Recommendation system is one of the hottest topics in machine learning. The focal research question of this article is: 'Which approach to recommender systems works best?'. The identification of problems such as 'Cold Start', 'Sparsity' and 'Freshness' is the sub-topic of this study. The scope of this study is to analyze the different approaches to recommender systems and design a solution that can overcome the identified problems. The paper starts with the analysis of different approaches to recommender systems: content-based filtering recommender system, item-item collaborative filtering recommender system, user-user collaborative filtering recommender system and demographic recommender system [1]. The analysis of different approaches to recommender systems is followed by the identification of different problems in the recommender system. These problems are briefly introduced in the paper and a justification of this study is drawn out. Then the description and goals of this study are laid out. The proposed solutions to the identified problems are listed in the technical aspect section of the paper. The paper ends with a conclusion which includes the major findings, limitations of the study and recommendation for future study.

To locate the literature for this study, the internet was searched for databases and articles using the keywords- recommender system, problems in recommender system, approaches to recommender system, trends of recommender system, content based filtering recommender system, collaborative filtering recommender system and solutions to recommender system.

2. Literature Review:

Internet has become a significant part of our life. Almost all commodities ranging from products to services are available on the internet today. Also, each day several new items are being made available to us on the internet. The convenience of the availability of this massive collection of items on the internet brings its own set of problems. For a user, it becomes extremely difficult to search for an item from this massive database. A recommender system is an obvious solution to this problem. The recommender system recommends products that the user would like based on predefined algorithms [2]. Along with providing the user with a pleasant experience, the recommender system also helps businesses by increasing the efficiency of targeting products to their customers.

2.1 Problems in the recommender system

2.1.1 Cold Start

The cold start of new user problem arises whenever a new user is added to the recommender system. Since the user is new to the system, the system does not have any

information regarding their likes or dislikes. This happens because the user of the recommender system has not rated any items [3]. Another category of the cold start problem is the cold start of new items. Cold start of new item problem arises whenever a new item or product is added to the database of the recommender system. Since this is a new item, the recommender system will not have any ratings for that product and will have a hard time recommending that product to any user.

2.1.2 Sparsity

Most businesses that use recommender systems will usually have a myriad of items or products to offer. Normally, any give user of the recommender system will have only rated a few of these items. This leads to the recommender system having a sparse user-items matrix which makes it difficult for the recommender system to find neighbors or patterns in the system [4]. This is known as the problem of Sparsity or lack of knowledge.

2.1.3 Scalability

With the growth of any business, the number of users in their system as well as the products or items that they have to offer also grows. Recommender system needs to process this growing information in order to make accurate and meaningful recommendations. However, as the size of the data increases, the computation time of the recommender system also increases [5]. This is the problem of scalability.

2.1.4 Freshness

'Freshness' is another problem that the recommender system faces. For example, in a movie recommender system, if a user likes the comedy genre then the system will recommend the user movies from the comedy genre. This restricts the system to recommend movies from other genres to the user [6]. The user might get bored watching the same genre of movie and would like some fresh recommendation of movies.

2.2 Approaches to recommender systems:

2.2.1 Content based filtering recommender system

Content based filtering recommender system makes recommendations on the basis of the features of the items in the database [Fig. 1]. This approach does not have to depend on the data of other users in the system. In this approach, the data of the items is gathered initially. Then, the gathered data is processed to extract features that will be useful in making recommendations to the users.

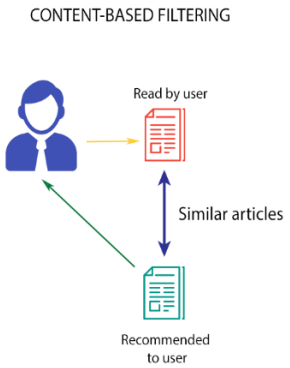


Fig. 1: Content-based filtering methodology overview

For example, in a book recommender system, the author, title, cost etc. of the book are some of the common content information that will be extracted from the data. Using these extracted features, book recommendations will then be made to the user [7]. An advantage of content-based filtering recommender system is that it does not require the data of other users. This helps to overcome the 'First Rater' problem, which is mentioned in the later part of this study.

2.2.2 Collaborative filtering recommender system

In user-user collaborative filtering recommender system, users that have similar taste as the active user are found. Then, the active user is recommended items that other similar users have liked. This way, the user-user collaborative filtering approach saves the system from representing items in terms of features [8].

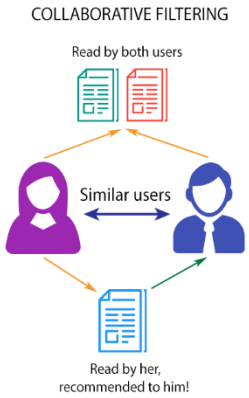


Fig. 2: Collaborative filtering methodology overview

In item-item collaborative filtering recommender system, for a given user, the correlation of the items that they have rated are computed against items that they have not rated. Then the ratings of highly correlated items are aggregated to make a recommendation to that user [Fig. 2]. The disadvantage of the collaborative filtering approach is that the recommendation of an item cannot be made unless that item is either rated by another highly correlated user or highly correlated with other items in the database.

2.2.3 Demographic filtering recommender system

Demographic filtering recommender system makes the assumption that users falling in the same demographic group will have similar taste. This approach collects the demographic information of the user and then makes recommendation of items that are popular in that demographic [9]. The major advantage of this approach is that it does not require the data of the ratings of the user. This approach makes recommendations based on few observations which makes it faster than other approaches. However, this advantage comes at invading the privacy of the personal information of the user.

2.3 Analysis of approaches to recommender system

The results obtained from analyzing the different approaches to recommender systems can be seen in [Table 1].

	Collaborative filtering	**Content based filtering**	**Demographic filtering**
Recommendation based on	A number of similar users	Single user	Users in that demographic
Problems	Cold Start, Sparsity, Scalability	Over-specialization, Freshness	Invades privacy, No-specialization
Advantages	Serendipitous recommendations	User independence	Simple, fast

Table 1: Comparative analysis of different approaches to recommender systems

The traditional recommender system faces several problems not limited to ‘Cold Start’ and ‘Sparsity’. The synthesis drawn from the literature would be that the different approaches to recommender systems, namely content based filtering, collaborative filtering and demographic filtering when implemented individually suffer from several problems. A possible approach to overcome the problems identified in the literature would be to build a hybrid model- a combination of multiple approaches to recommender systems. Problems in the traditional recommender system and its different approaches would be the convergent findings from the literature. The divergent finding would be the lack of a model that overcomes these problems.

This article closes the gap between the convergent and divergent findings by proposing a hybrid recommender system that can overcome the problems faced by traditional recommender systems.

3. Project Justification:

Recommender systems need data to make recommendations. For a recommender system to make accurate recommendations, this data needs to be consistent and meaningful. However, that is not usually the case. Most of the time the data is changing, growing or lacking. Recommender systems need effective measures to combat this nature of data. Also, the preferences of user are changing. This change in preference also needs to be handled by the recommender system. Traditional recommender systems are surrounded by many such problems. These problems hinder the accuracy of the system. Recommender systems are imperative to businesses and users all around the world and hence their recommendations need to be effective and meaningful. This justifies the purpose of the conducted study.

4. Goals of the project:

Recommender systems have been around for some time now. The goal of this project is to provide solutions for problems with the traditional recommender system. Over the years people have come up with several approaches on how to build a recommender system, namely, content-based filtering, user-user collaborative filtering, item-item collaborative filtering and demographic filtering recommender systems. The goal of this project will be achieved by identifying the problems with the recommender system and analyzing the different approaches to recommender system. The study will also cover the ethical aspect of user privacy breach.

5. Technical Aspect:

After identifying the problems in the traditional recommender system and analyzing the different approaches to recommender systems, the following are the solutions to each of the identified problems.

5.1 Demographic Information

To combat the problem of Cold Start, demographic filtering approach can be used. The user's demographic information can be collected by either tracking their location or having them explicitly list out their location [5]. Once the demographic information of the user is obtained,

then recommendations of the highly rated items in that demographic can be recommended to the user.

5.2 Matrix Factorization

To solve the problem of sparsity in a recommender system, latent factor models algorithms can be used. In latent factor models algorithms, firstly, dimensionality reduction is done for items as well as users. Secondly, patterns are identified in this reduced dimension space which avoids the problem of sparsity [10]. Matrix factorization methods do an excellent job at combating the sparsity problem of recommender system.

5.3 Parallelization

The problem of scalability is rather difficult to combat. This is because all the approaches to recommender systems face this problem. One way to solve this problem would be by using parallelization techniques [9]. The algorithm 'Gellyfish' is an excellent choice for this problem. Another way to solve the problem of scalability would be to improve the performance of the system by either upgrading the hardware of the system or making certain computations offline.

6. Ethical Aspect:

Recommender systems need to collect data from users in order to make recommendations tailored to that user's preferences. Being the immediate stakeholders for my project, any person or organization looking to increase the efficiency of targeting their products to customers would find that privacy is a major ethical issue with recommender systems. "Personal data collected by Recommender Systems should be kept safe and piracy must be neglected and should be uninfluenced and unmodified" [10]. Recommender systems need to take care of the value and risk of personal information ethical aspect while being developed. The recommender system should only collect information that is going to be useful in making accurate recommendations. Also, if possible, the recommender system should discard the personal information of users if that information is no longer required.

The zero-knowledge proof in cryptography is an interactive protocol that verifies information without disclosing the underlying data. This revolutionizes the way data is interacted with. In transactions involving zero-knowledge proof, one system tries to prove something to the other system without ever revealing the data. Recommender systems can use this technique to keep the information about its users private. The recommender system also needs to handle various security and privacy attacks. Techniques such as zero knowledge proof and cryptosystem can be used to handle such attacks.

7. Conclusion:

This study investigated the approaches to recommender systems discussing the problems and advantages of each. The major findings of this study would be the design recommendations of using demographic information, matrix factorization, parallelization and zero knowledge proof in a recommender system. These design recommendations helped solve the problems that exist in traditional recommender systems and achieve the goal of this research. The lack of information regarding the integration of these solutions would be a limitation of this study. For future research, I would like to study the integration of these solutions into a recommender system.

References:

1. XiaoYan Shi, HongWu Ye and SongJie Gong, "A Personalized Recommender Integrating Item-Based and User-Based Collaborative Filtering," 2008 International Seminar on Business and Information Management.

2. A. Al-Doulat, “Surprise and Curiosity in A Recommender System,” in 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), 2018.

3. K. Shah, A. Salunke, S. Dongare, and K. Antala, “Recommender systems: An overview of different approaches to recommendations,” in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

4. A. Pal, P. Parhi, and M. Aggarwal, “An improved content based collaborative filtering algorithm for movie recommendations,” 2017 Tenth International Conference on Contemporary Computing (IC3), 2017.

5. S. Jain, A. Grover, P. S. Thakur, and S. K. Choudhary, “Trends, problems and solutions of recommender system,” in International Conference on Computing, Communication & Automation, 2015.

6. F. Mansur, V. Patel, and M. Patel, “A review on recommender systems,” in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

7. A. S. Tewari, A. Kumar, and A. G. Barman, “Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining,” in 2014 IEEE International Advance Computing Conference (IACC), 2014.

8. G. Gupta and R. Katarya, “Recommendation Analysis on Item-based and User-Based Collaborative Filtering,” 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2019.

9. M. Lerato, O. A. Esan, A.-D. Ebunoluwa, S. M. Ngwira, and T. Zuva, “A survey of recommender system feedback techniques, comparison and evaluation metrics,” in 2015 International Conference on Computing, Communication and Security (ICCCS), 2015.

10. N. Sodera and A. Kumar, “Open problems in recommender systems diversity,” in 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017.

Appendix A:

1) XiaoYan Shi, HongWu Ye and SongJie Gong, "A Personalized Recommender Integrating Item-Based and User-Based Collaborative Filtering," 2008 International Seminar on Business and Information Management.

   This source is peer reviewed. Also, the fact that this study was conducted by people in the Zhejiang Business Technology Institute gives credibility to this source. The source provides insight on the performance of a recommender system that uses both item and user based collaborative filtering.

   The main insights from this source material would be the benefits of using both item-item and user-user collaborative filtering techniques of recommender system. The design considerations needed for integrating these two techniques would be an emerging finding that influenced the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

   The information provided in this study has changed my thinking about using just one approach to recommender system for building my model. Now, I will use multiple approaches together.

2) A. Al-Doulat, “Surprise and Curiosity in A Recommender System,” in 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), 2018.

   This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Software and Information Systems of a reputable university gives credibility to this source. The source provides insight on the factors to keep in mind while designing recommender systems.

   The main insights from this source material would be the understanding of the considerations one should keep in mind while designing recommender systems. The design considerations needed for recommender systems are an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

   The information provided in this study has changed my thinking about the central question of my research by shedding light to some important considerations while designing recommender systems.

3) K. Shah, A. Salunke, S. Dongare, and K. Antala, “Recommender systems: An overview of different approaches to recommendations,” in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Science of a reputable university gives credibility to this source. The source provides insight on the abstract area of my research.

The main insights from this source material would be the understanding of different approaches to implementing recommender systems. The designing of different techniques to recommender systems is an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

Instead of changing my thinking about the central question of my research and related sub-topics, the information in this study made sure that I am in the right direction.

4) A. Pal, P. Parhi, and M. Aggarwal, "An improved content based collaborative filtering algorithm for movie recommendations," 2017 Tenth International Conference on Contemporary Computing (IC3), 2017.

This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Science of a reputable university gives credibility to this source. The source provides insight on the architecture area of my research.

The main insights from this source material would be the understanding of content-based filtering recommender systems. The working of content based collaborative filtering algorithm is an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

The information in this study has improved my understanding of content-based filtering algorithm which will help me in providing a better solution to the central question of my research.

5) S. Jain, A. Grover, P. S. Thakur, and S. K. Choudhary, "Trends, problems and solutions of recommender system," in International Conference on Computing, Communication & Automation, 2015.

This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Application of a reputable university gives credibility to this source. The source provides insight on the conclusion area of my research.

The main insights from this source material would be the understanding of solutions to the problems of recommender systems. The solutions provided to some of the problems of recommender systems are an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

The information found in this study has educated me on where to start the designing of my solution to the central question of my research.

6) F. Mansur, V. Patel, and M. Patel, "A review on recommender systems," in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

   This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Engineering of a reputable university gives credibility to this source. The source provides insight on the introduction area of my research.

   The main insights from this source material would be the understanding of recommender systems. The basic architecture of recommender systems that is laid out in the study is an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

   Instead of changing my thinking about the central question of my research and related sub-topics, the information in this study made sure that I am in the right direction.

7) A. S. Tewari, A. Kumar, and A. G. Barman, "Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining," in 2014 IEEE International Advance Computing Conference (IACC), 2014.

   This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Information Technology of a reputable university gives credibility to this source. The source provides insight on the technical aspect area of my research.

   The main insights from this source material would be the benefits of using content-based filtering, collaborative filtering and association rule mining in combination. The design considerations needed for integrating these techniques would be an emerging finding that influenced the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

   The information provided in this study has changed my thinking about using just one approach to recommender system for building my model. Now, I will use multiple approaches together.

8) G. Gupta and R. Katarya, "Recommendation Analysis on Item-based and User-Based Collaborative Filtering," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2019.

   This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Science and Engineering of a reputable university gives

credibility to this source. The source provides insight on the technical design area of my research.

The main insights from this source material would be the understanding of item-item and user-user based collaborative filtering recommender systems. The working of item-item and user-user based collaborative filtering recommender systems is an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

The information in this study has improved my understanding of item-item and user-user based collaborative filtering recommender systems which will help me in providing a better solution to the central question of my research.

9) M. Lerato, O. A. Esan, A.-D. Ebunoluwa, S. M. Ngwira, and T. Zuva, "A survey of recommender system feedback techniques, comparison and evaluation metrics," in 2015 International Conference on Computing, Communication and Security (ICCCS), 2015.

This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Systems Engineering of a reputable university gives credibility to this source. The source provides insight on the evaluation area of my research.

The main insights from this source material would be the understanding of evaluation metrics to compare recommender systems. The evaluation metrics portrayed in the study are an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

This study has shed light to some evaluation metrics that I intend on using to answer the central question of my study.

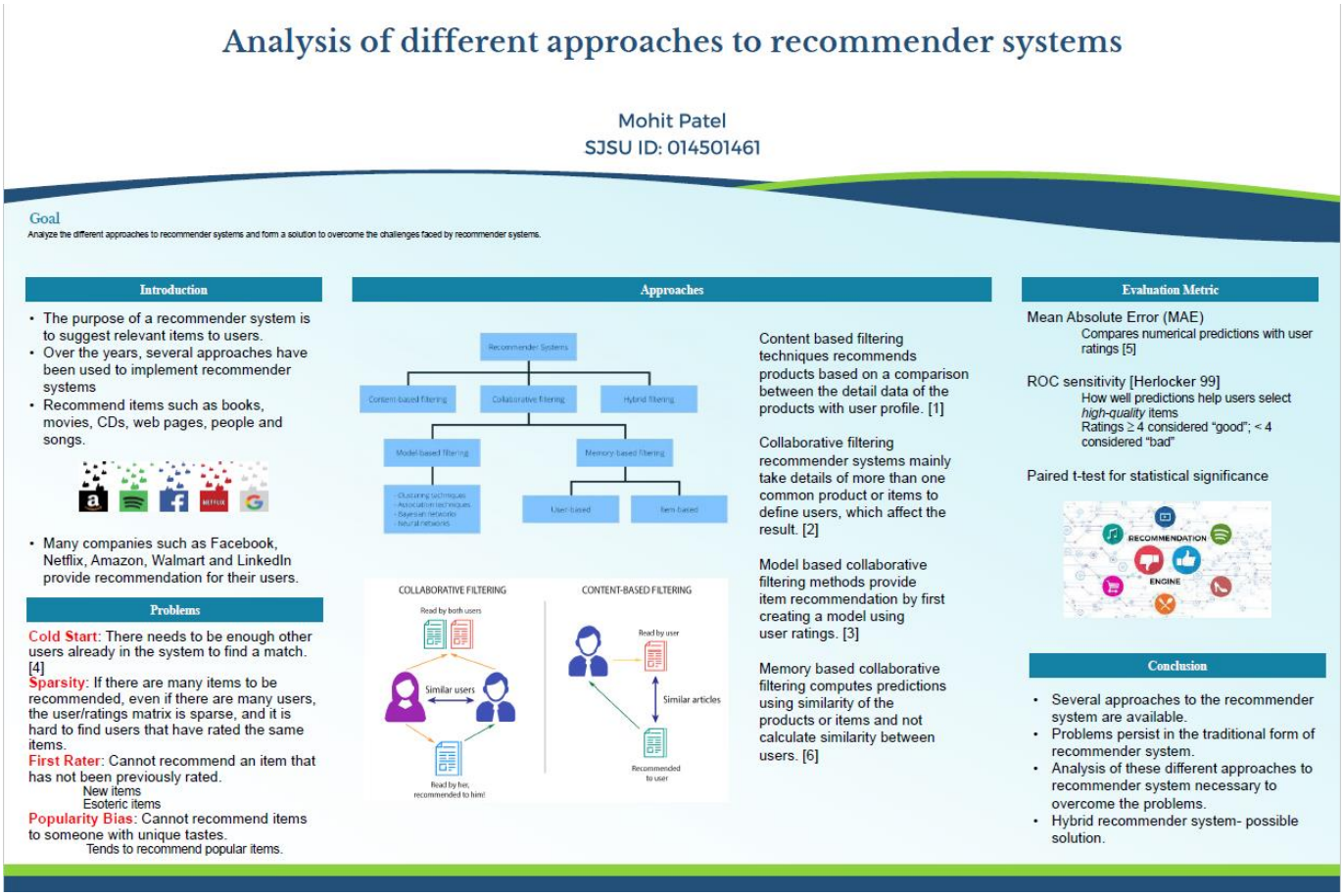
10) N. Sodera and A. Kumar, "Open problems in recommender systems diversity," in 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017.

This source is peer reviewed. Also, the fact that this study was conducted by people in the department of Computer Science and Engineering of a reputable university gives credibility to this source. The source provides insight on the problems area of my research.

The main insights from this source material would be the problems that the current recommender systems face. The problem of 'Cold Start' and 'Sparsity' are an emerging finding that influence the direction of my future study. I did not find any gaps for the scope of the study that was conducted in the source.

The information found in this study shed light to a lot of new problems that recommender systems face, and this has broadened the scope of the related sub-topics of my research.

Appendix B:



4a_poster_final.pdf

(Double click the icon to access the infographic poster)