RECOMMENDER SYSTEMS

1Swathi Ghattamaneni, 2Sarvana Kumar K
(VIT university, vellore, India)
1ghattamaneni.swathi2013@vit.ac.in, 2ksaravanakumar@vit.ac.in

Abstract: Due to the information overload recommender systems are becoming more popular now a day. Generally recommender systems are classified into types’ content based and collaborative filtering. Content based approach provides recommendations based on users history of access/purchase (in case of e-commerce sites), whereas collaborative filtering provides recommendations based on the similar user interest. Both the approaches have their advantages and disadvantages. In our paper we provide a detailed study of various approaches for recommender systems, and problems with specified approaches and possible solutions.

Keywords: Recommender systems, content based approach, collaborative filtering, and hybrid recommendation systems.

1. Introduction: Recommendation systems make users decision making simple in this huge information centric world. For decisions like what to access? What to buy? People now a day mostly relay on these recommendation systems. Recommendation systems save lot of effort and time for the user [5]. For example if a user wants to purchase a book, then it will be difficult for him or her to go through all ratings and reviews in order to select the better option. So these recommender systems will make it easy for the user in critical decision making process. Generally speaking there are two types of these recommender systems they are content based and collaborative filtering [6]. Collaborative filtering can be further classified into model based approach and neighborhood based approach. Model based approach creates a constant model/profile based on the user’s preferences and ratings and provides recommendations based on that, conversely neighborhood approach uses user item ratings stored for predicting ratings for new items. On the other hand in content based approaches recommendations are made based on user’s demographic information. Content based recommendations are made by comparing the contents of user interests. In addition to these content based and collaborative filtering approaches several hybrid approaches are used to enhance the performance of the recommender systems, these approaches are based on the collaborative filtering approach but in addition content based profile is maintained for each user, rather than co-rated items these content based profiles are used to find out similar users. Despite of the advances recommender systems used now-a-days require improvements further for recommendations to be more accurate and effective so that they can be applicable in many more real time applications, those improvements are possible by predicting better methods for user behavior capturing. So recently most important requirement in recommender system is need to combine various recommendation techniques for achievement of peak performance i.e. in terms of accuracy. Hybrid recommendations can be done in several ways by predicting recommendations using collaborative and content based separately and then combining them for accuracy or by creating a unified model for recommendations to be done. Compared to the performance of content based approach [6,9,14] and collaborative filtering [18, 19, 20, 21] approach implemented independently the hybrid approaches proved to be far better for providing recommendations in various areas, one of the good example of this hybrid recommendation approach is Netflix [15, 17] where the technique of comparison of searching and watching habits of similar users is implemented. However there are also several approaches of hybridization in spite of that for accurate recommendations. In our paper we have studied various approaches to collaborative filtering, content based approach and hybrid approach and analyzed the pros and cons of each approach used for recommender systems. We have listed few of our research with detailed description of the technique adopted in order to provide recommendations and also the pros and cons of the approach adopted. Thus our paper provides a broad study of recommender systems. We have narrowed down the research for further implementation of an effective technique for file recommender systems in social networks, this will be our preceding work next to this survey, which implements a better technique for sharing of files in social networks to only specific interested users/friends among hundreds and thousands of friends in the list.

2. DEFINITIONS:

Content based filtering: content based filtering is also known as cognitive filtering. In this approach item recommendation is done based on comparison of content of items in user’s profile. In content based systems items are described using keywords. A profile is created for the user to indicate the types of the item liked by the user.

Collaborative filtering: One of the widely used techniques in recommender systems is Collaborative filtering. Collection and analysis of
Hybrid Recommendation systems: Hybrid recommendation system is the combination of content based and collaborative approach. Even though content based and collaborative approaches provide accurate results for recommendation systems, but still there are few drawbacks with these techniques. To overcome the drawbacks of content based and collaborative approaches, hybrid recommendation systems are implemented, which is a combination of content based and collaborative approach.

Gray Sheep problem: Gray sheep problem is the problem related to users whose interests don’t match with other users for providing recommendations. Such problem can be dealt with content based approach.

Performance Measures: Evaluation is the important step for analyzing the effectiveness of recommender systems. Performance measures are the measures of the effectiveness of recommender systems. Several performance measuring techniques such as mean squared error, root mean squared error and Netflix are used for performance analysis.

3. LITERATURE SURVEY:

Virtual user approach for group recommendation systems using precedence relations.

Problem: Recommendation systems are very much useful for several users as they make users choice making simple. Several techniques are used for finding the recommended items for the users (like movies, web articles, purchase items), but all the techniques lack in one or the other aspect for providing best recommender system to the user. Recommender systems are classified into two types personal recommender systems and group recommender systems [2, 3, 4, 5]. Major problem is constructing group recommender systems from personal recommender systems. For the construction of group recommender systems from personal recommender systems, one major approach used now days is merging recommendations [2, 3] of individual group members and formation of common recommendation list based on the ratings of individual users. But this strategy is not best as the common recommendation list will be having items/articles that are recommended by all the members of the group but will not consider the items recommended by max members that is an item will be just ignored if it is not rated by at least one user in that group but rated as recommended item by remaining nine members in the group so merging approaching cannot handle the problem of diversities or disagreements of group members [1].

Proposed approach: The proposed approach focuses on creation of group recommender system from personal recommender system by using precedence mining technique. Our proposed technique constructs a virtual user which represents the entire group members interests here precedence mining technique is used to find individual group members interests. Our proposed approach deals with all kinds of items like items used by all the members of the group, some members of the group and no members of the group. When an item is used by all the members of the group then it is given weight 1, if at all an item is not used by any members of the group then it is given 0 weight, for items accessed by some members of the group, score is calculated, if score is greater than threshold then that item will be included in virtual profile.

Observation: A group based recommender system is developed. Virtual user is created by considering individual member interests of the group, based on that virtual user recommendation prediction are made. Cold stack problem is efficiently handled by using precedence mining technique. Though precedence mining technique is applied cold stack problem is not completely resolved.

Using community structure to control information sharing in online social networks.

Problem: Online social networks are widely spread now a day and are acting as wide range of information sources. Huge amounts of information are shared through online social network in the form of news articles blogs and posts. All this information results in information overload problems. For example a file related to a particular subject shared by a teacher is useful to only students of that particular subject and will be a scrap to others in the community. So in such a case access to the information must be provided based on relativity.

Proposed approach: In this paper by Amin Ranjbar, Muthucumarumaheswaran [7] an approach for information sharing is proposed where a community named as x-my community is created and members of the community is decided based on their interaction with the user i.e. the users
that are frequently connected to the user. Based on the information relevancy a sub graph\[7,8,9\] is created each time when information is shared. The sub graph contains the users that are relevant to the data that is shared. The sub graph varies each time based on the shared information relevance.

Observation: The research provides a group based information recommender systems by creating a community of users that likely to have access of information shared by the user and then sub grouping based on the user’s information interests. But the proposed approach fails to handle the changing conditions. The user’s interests and interaction priorities vary over time the proposed approach is unable to handle this aspect.

- Information over load for the users is reduced.
- Ease of access.
- User interests may change from time to time proposed approach cannot handle this problem.
- Each time when an article is shared sub graph need to be created.

Personalized recommendation of popular blog articles for mobile applications.

Problem: content based or collaborative filtering techniques have their own advantages and disadvantages. Because of the drawbacks such as scalability, shilling attacks, cold start synonyms, grey sheep the performance of the recommender systems implemented using content based or collaborative filtering techniques are not the best. Combination of both these techniques will lead to a hybrid approach which provides better results compared to these individual approaches implemented individually.

Proposed approach: The proposed approach constructs a novel hybrid approach\[10\] that takes in to consideration prediction of user’s preference by collaborative filtering, degree of attention of internet articles and degree of popularity of that topic, degree of attention of internet articles means the web articles that grab more attention of users are likely to be recommended and more popular the topic the more it attracts the user attention.

Observation: The proposed hybrid approach provides better recommendation system compared to other techniques implemented individually. Even though the proposed approach undergoes several stages of filtering but it cannot handle one of the leading problem of the recommendation systems cold start problem. It can be said as the better approach but not the best.

- Since the hybrid approaches under goes several stages of filtering the proposed approach provides better results compared to each technique applied individually.
- Cold start problem cannot be solved by such technique.
- Limited content analysis.

Recommendation using neighborhood methods with preference relation based similarity.

Problem: Collaborative filtering is the most famous and widely used approach for recommender systems, key property of such system is that users are required to offer only an overall rating for each item consumed. In other words users cannot express their preference for individual criteria on a given item. However practical decision problems are often characterized by multiple criteria. Furthermore, it is helpful to improve recommendation accuracy by incorporating multiple criteria that can affect user preference in to the recommender.

Proposed approach: This paper contributes to propose the similarity measures using the strict preference relation for neighborhood methods in single-criterion and multi-criterion \[12, 13\] collaborative filtering. In contrast to traditional similarity measure, the new and simple measure not only estimates the overall strength of one user’s preference over the other, but is asymmetric. Because many users of recommendation applications are typically interested in only the few high ranked items recommendations, the goal of the proposed approach is to recommend correctly a set of a few relevant items to each user.

Observation: The proposed multi priority based approach provides better recommendation system compared to other techniques implemented individually. Even though the proposed approach several criteria’s for rating users preferences but it cannot handle one of the leading problem of the recommendation systems cold start problem.

- By taking multiple rating criteria’s in to consideration all the aspects of the items and users opinions on all those aspects can be well covered which in turn provides better recommendation systems.
- Cold start problem cannot be solved by such technique.
- Only most liked ones will be given preference.

Capturing scholar’s knowledge from heterogeneous resources for profiling in Recommender systems.

Problem: Scholar’s article recommendation systems are the systems that recommend articles to the users based on various factors such as home page information, publication history, research interests etc. But there may be a possibility that each scholar will have different tendency of selecting articles based on his background knowledge levels acquired. One more problem is research interests may change over time or one scholar may be interested in one or more topics at a time i.e. multi-disciplinary \[14\] scholars in order to suggest articles for these scholars’ domain or topic aware
recommendation is necessary. Current approaches fail to provide solutions for these approaches.

**Proposed approach:** Proposed approach is the sequence of steps starting with analyzing the domain of the content and exploring resources in that domain followed by keyword extraction from the knowledge of that domain then keywords are merged up to form a unified keywords. After identification of unified keywords clustering of the keywords is done to solve the problem of multi-dimensional researchers k-means clustering algorithm has been used for this purpose.

**Observation:** The proposed approach provides solution for the fact of changing interests of scholars and also suggests recommendations to multi-dimensional researcher’s i.e. scholars researching on different researching area at same time.

- Supports recommendation of articles for multi-dimensional research scholars.
- More accuracy compared to other techniques.
- Involves more operational costs.
- Hierarchical sequence of documents is not recommended as specified.

4. **Generic Architecture:**

```
Collect data  Refine the data collected  Cluster users browsing history
```

```
Suggest recommended products  Filter based on the ratings or user preferences  Predict related items to the browsed
```

**Fig.1 Content based filtering**

**Fig.2 Collaborative filtering**

5. **Findings:**

- Virtual user approach for group recommendation systems using precedence relations
  - Virtual user by threshold method is used where if an item is consumed by all the members of the group is given weight 1 score is the weight of that item. if overall weight Is greater than threshold item will be in the recommended list
  - In virtual user by weight for individual user is created and all the group individual members weights are merged here if an item is consumed by user it is given weight 1 otherwise 0.
  - Comparatively both the methods yield better results for movie recommendation of a group but virtual user by threshold seemed to be simple and accurate.

- Using community structure to control information sharing in online social networks.
  - A community named x-my community is computed from a list of friends of users in OSN based on the probability of willingness of the user to share with particular friend. Here the social network is considered as a graph \( SG = (U, E) \) where \( U \) is users and \( E \) is edges, the probability i.e. the willingness of user is denoted by \( P \).
  - This approach provides better solution for users changing relations with other users in sharing information but every time when a user is willing to share information community need to be computed , and if the user need to share with the sub group of the community then sub group also need to be computed every time. Computational burden is more.
• **Personalized recommendation of popular blog articles for mobile applications**
  - A hybrid approach is proposed to suggest web articles for mobile users which involve topic based collaborative filtering where topics related to users past topics is recommended but only collaborative filtering has drawbacks of never browsed articles will never get a chance like wise to overcome these degree of attention of blog articles is calculated which indicates mobile users are interested to those articles to which majority internet readers pay attention, this internet reader attention is known through click rate. To this method predictive popularity degree of articles is computed to further improve recommendations.
  - The proposed hybrid method provides better solution for web article recommendation for mobile users. Especially proposed approach is designed considering the interests of internet users but the drawbacks of web articles recommendation system for internet users will be applicable for web article recommendation of mobile users such as newly introduced articles are not covered.

• **Recommendation using neighborhood methods with preference relation based similarity**
  - Similarity measure using strict preference is done in order to recommend few but highly acceptable items to users here single and multi-criteria ratings are considered for recommendation of products.
  - This multi-criteria ratings method provides better results compared to simple collaborative filtering approach however it cannot solve the problem of newly introduced items without any ratings.

• **Capturing scholar’s knowledge from heterogeneous resources for profiling in Recommender systems**
  - A hierarchical approach is proposed for scholarly articles recommendation systems which involves context and resource analysis here scholarly interactions are recorded as use cases for finding out the sources.
  - This method is followed by key term extraction to find out theme or topics of articles followed by key term unification and clustering finally the key term are evaluated based on these term profile is created.
  - The proposed approach seemed to be the best approach for scholarly article recommendation. It can also be applied to web page or file recommender systems with few changes for better results.

**Conclusion:** Recommender systems plays major role in various disciplines such as entertainment, research, online shopping etc. The use of these recommender systems is increasing day by day with the increase of large amounts of data. In our paper we have studied and listed various approaches used for recommender systems and the advantages and drawbacks of each approach is listed. From the above study we predicted the drawbacks to be overcome in order to implement best recommender system. Our next research will be on this aspect, we are trying to implement these findings on file recommender systems in social networks by constructing a hybrid approach that involves creation of profile that must satisfy the threshold in order to access the shared file, here the profile is combination of various factors user browsing behaviors, ratings ,views, shares etc.

**References:**


[15] Hybrid Recommender systems from Wikipedia, the free encyclopedia.


