

Recommendation Techniques to Improve Diversity and Novelty Based on User Behaviour

K Punyavathi, Jyothi P

Abstract— Recommender Systems are the powerful technologies for overcoming the information overload in the World Wide Web and to get the personalized recommendations. They can be benefit to both the consumers and to the business. Consumers can be benefited by finding the relevant items to them and the business individuals can be benefited by increasing their sales. The organizations do not study the behavior of a customer or his intension on a product. Since the customer can look into various products in the catalog it is necessary to study his behavior and identify his interest in a product or type of product. This paper aims at recommending items based on user behavior and similarity. Traditional recommender systems typically rank the relevant items in a descending order of their predicted ratings for each user and then recommend top N items, resulting in high accuracy. Many recommender systems only focuses on improving the accuracy of recommendations, but the other factors of recommendations like novelty to improve the quality are often overlooked. The proposed approaches consider additional factors, like item popularity, to increase recommendation diversity and novelty.

Index Terms—Recommender systems, recommendation diversity, collaborative filtering.

1 INTRODUCTION

In the contemporary world, the requirement to find relevant data has been identified as a problem and provided with solution by the recommender systems. Over the last 10-15 years, recommender systems technologies have been introduced to help people deal with these vast amounts of information [1]. The e-commerce applications such as Amazon and Netflix have been widely using these recommender systems. Recommendation system relies on rating which is estimated based on the consumed product and the products that are yet to be consumed. Recommender systems typically try to predict the ratings of unknown items, by using other users' ratings, and recommend top N items with the high ratings[1]. There have been many algorithms proposed that can improve accuracy of recommendations. The accuracy of recommendations alone may not be enough to find the most relevant items for each user [1]. The goal of recommender system is to provide a user with highly personalized items and more diverse recommendation that suggests more number of items to the users. There have been many studies on recommendation methods that can increase the

diversity of recommendation sets for given user. These studies measure recommendation diversity from an individual user's perspective (individual diversity)[3]. High individual diversity of recommendations doesn't necessarily imply high aggregate diversity [1].

Both individual and aggregate diversity (higher diversity) can come at the expense of accuracy. There is a tradeoff between accuracy and diversity because high accuracy may often be obtained by recommending the most popular items to users, which lead to the reduction in diversity. Higher diversity can be achieved by trying to uncover and recommend highly personalized items for each user [1].

2 REVIEW OF LITERATURE

2.1 Classification of Recommender Systems:

Recommender systems are usually classified into three categories based on their approach to recommendation: content-based, collaborative, and hybrid approaches [2]. Content-based recommender systems recommend items similar to the ones the user preferred in the past. Content-based filtering

uses the assumption that items which have similar features will get similar ratings [2]. The Collaborative Filtering (CF) approach is widely used in recommender systems. *Filtering* stands for filtering of information. *Collaborative* covers the fact that the information that is being used to filter the collection is being supplied by all the users of the system. Collaborative filtering assumes that people with similar tastes will rate things similarly [5]. Collaborative filtering recommender systems recommend items that users with similar preferences (i.e., "neighbors") have liked in the past. Hybrid approaches combine both content-based and collaborative methods [2]. Recommender systems also be classified based on their recommendation approach as heuristic and model based. Heuristic techniques typically calculate recommendations based directly on the previous user activities (e.g., transactional data or rating values). The commonly used technique is a neighborhood based approach that finds nearest neighbors that have tastes similar to those of the target user [3]. Recommender systems generally perform the two tasks in order to provide recommendations to each user. First, the ratings of items are estimated using the available information and some recommendation algorithm. Second, the system finds the items that gives maximum utility to users and recommends them to the user. The proposed ranking approach is designed to improve the recommendation diversity in the second task of finding the best items for each user [2].

Some Recommender systems need information about the use, products or both to provide recommendations. The data may be collected explicitly or implicitly. Data given by user called as

-
- Student, Punyavathi K is currently pursuing master's degree program in Computer Science and Engineering in Aurora's Scientific Technological and Research Academy, JNTUH, Hyderabad, India,
 - Jyothi P is currently working as a Senior Asst. Professor in the department of Information Technology in Aurora's Scientific Technological and Research Academy, JNTUH, Hyderabad, India,

explicit data. To provide explicit data, a rating is given on a scale form one to five, where one represents the least rating and five represents the highest rating. So there is no condition that each user should give rating to the products. Some users may not wish to provide rating for the items they have bought or viewed, they don't spend their time to give the rating to the items. And there is no need for a customer to register before he searches an item. For this kind of users another source is needed to get the ratings, one approach to this problem is to use the implicit ratings i.e., watching the behavior of the customer.

The following diagram shows the recommendation process.

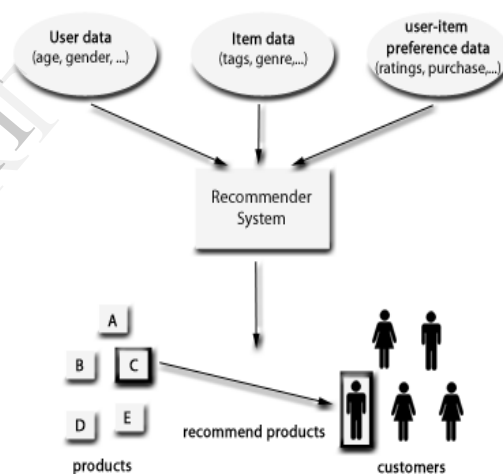


Figure: Recommendation process

2.2 Accuracy of Recommendations:

Many recommendation techniques have been developed over the last few years; the metrics used to measure the accuracy are statistical accuracy metrics and decision-support measures [6]. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are the examples of statistical accuracy metrics. How much accurately the system predicts the exact rating of the specific item is measured by these metrics. Precision, recall and F-Measure are the examples of decision support measures.

Precision (the percentage of truly “high” ratings among those that were predicted to be “high” by the recommender system), recall (the percentage of correctly predicted “high” ratings among all the ratings known to be “high”), and F-measure, which is a harmonic mean of precision and recall[1]. The ratings are given on 1-5 rating scale. This value denotes that the item is good and liked by some user.

$$\text{Precision-in-top-N} = \frac{\sum_{u \in U} \text{correct}(L_N(u))}{\sum_{u \in U} |L_N(u)|}$$

Where U is the list of users in the recommender system, I is the set of available items, and L is the List of Items

Only accuracy of recommendations is not enough to find the relevant items to users. The recommendations that are recommended to users should not only accurate but also they must be useful to the users. And the recommendation quality is not only measured by accuracy, so some other factors also considered measuring the quality. For that we consider the diversity of recommendations.

2.3 Diversity of Recommendations:

Diversity can be measured two ways in the recommender systems. They are individual and aggregate diversity. Individual diversity produces the unique item to the user, that item is accurately relevant to the search for people, but user won't satisfy with the unique item even the result is suitable for user[7]. Most of the studies concentrated on increasing the individual diversity, that can be calculated from each user's recommendation list (e.g., an average dissimilarity between all pairs of items recommended to a given user)[4]. These techniques focused on providing recommendations which are not similar for the same user. For this intra-list similarity metric is used and item novelty is other technique to measure the individual diversity. On the other hand

aggregate diversity has some potential importance for providing diverse recommendations in both users and business perspective.

The goal is the technique is to provide multiple results for the same user, but accuracy is a major failure. Our goal is to provide recommendations with diversity with accuracy [7]. To measure the diversity based on the top N items list recommended to users, the metric as follows,

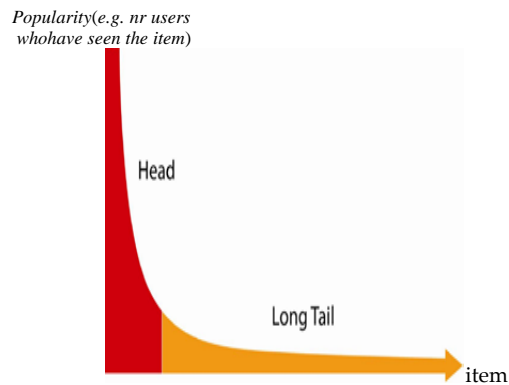
$$\text{Diversity-in-top-N} = \frac{1}{|U|} \sum_{u \in U} |L_N(u)|$$

Both individual and aggregate diversity (higher diversity) can come at the expense of accuracy. There is a tradeoff between accuracy and diversity because high accuracy may often be obtained by recommending the most popular items to users, which lead to the reduction in diversity. Higher diversity can be achieved by trying to uncover and recommend highly personalized items for each user [1].

2.4 Long Tail Items:

The items recommended to users, we considered only the items that were predicted above the rating threshold to assure the acceptable level of accuracy, as is typically done in recommender systems [2]. Among these candidate items for each user, we identified the item rated by most users (i.e., the item with the largest number of known ratings) as a popular item, and the item which have less ratings (i.e., the item with the smallest number of known ratings) as a long-tail item[2]. The Pareto Principle, sometimes called the 80/20 rules, states that a small proportion (e.g., 20%) of products in a market often generate a large proportion (e.g., 80%) of sales [3].

The long tail brings dual benefits for increasing companies' profit: (1) compared with popular items, long tail items embrace relatively large marginal profit, which expands the long tail market can bring much more profit [3].



It is generally accepted in Economics that the economical profit of a completely competitive market is nearly zero. The head market, full of bestselling products is an example of such highly competitive market with little profit [3].

The following figure shows the accuracy diversity tradeoff

Quality Metric:	Accuracy	Diversity
	Top-1 recommendation of:	
Popular Item (item with the largest number of known ratings)	82%	49 distinct items
"Long-Tail" Item (item with the smallest number of known ratings)	68%	695 distinct items

Figure: accuracy- diversity tradeoff

The above figure shows that the recommendation of popular items leads to 82% of accuracy but the gives only 49 distinct items i.e., less diversity and the recommendation of long tail items leads to 695 distinct items with measurable accuracy i.e., 68%.

So the recommendation of long tail items also produces some profit also based on Pareto Principle.

3. Recommendation Re-ranking:

3.1 Traditional Ranking Approach:

Typical recommender systems predict unknown ratings based on known ratings, using any traditional recommendation technique [1]. Given all of the predictions for each user, in the

recommendation phase, the system selects the items that maximize a user's utility. Formally, item ix is ranked ahead of item iy (i.e., $ix < iy$) if $rank(ix) < rank(iy)$, where $rank: I \rightarrow \mathbf{R}$ is a function representing the ranking criterion. Typical recommender systems rank the candidate items by their *predicted rating values* and recommend the most highly predicted N items to each user because users are typically only interested in several of the most relevant recommendations [1]. This is referred to as the *standard ranking approach* and we define the ranking function as

$$rank_{\text{Standard}}(i) = R^*(u, i)^{-1}.$$

The power of -1 in the above expression indicates that the items with the *highest*-predicted (as opposed to the lowest-predicted) ratings $R^*(u, i)$ are the ones being recommended to the user. The standard ranking approach and it shares the motivation with the widely used probability ranking principle in information retrieval literature that ranks the documents in order of decreasing probability of relevance [7].

The standard ranking approach increases the accuracy but reduces the diversity. To increase the diversity in proposed approach ranks the items using the popularity.

3.2 Popularity -Based-Item Ranking

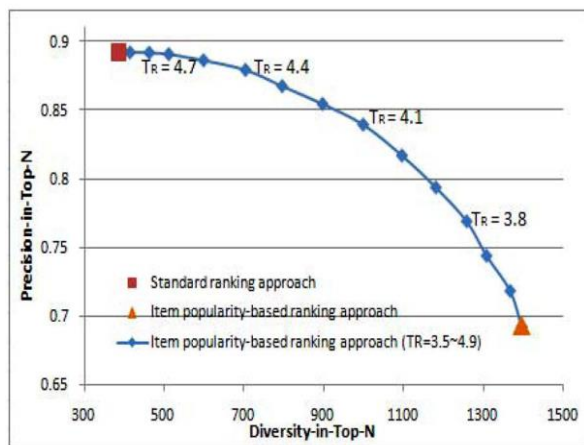
In this item ranking technique the items are ranked based on the popularity where popularity means the number of known ratings for each item. The technique ranks items from lowest to highest according to popularity of items. Item-popularity based ranking function can be written as follows:

$$\text{Rank}_{\text{ItemPop}(i)} = |U(i)|, \text{ where } U(i) = \{u \in U \mid \exists R(u, i)\}.$$

3.3 Tradeoff between Accuracy and Diversity

There exists a tradeoff between accuracy and diversity. Accuracy and diversity are inversely proportional to each other. If accuracy increases diversity decreased and vice versa. In this example, the item-based collaborative filtering technique is used to predict unknown ratings. So to balance the accuracy and diversity threshold is used, users can select the threshold for getting more diverse results. Among these candidate items for each user, we identified items that were rated by most users (i.e., items with the largest number of known ratings) as popular items, and items that were rated by the least number of users (i.e., items with the smallest number of known ratings) as long-tail items.

In the below figure there is a comparison between standard ranking approach with item popularity ranking approach where the accuracy and diversity will vary according to the threshold value. In the standard ranking approach the accuracy is 90% and gives only 350 items, and with item popularity approach the accuracy is 69% and 1400 items are recommended. So to balance between these accuracy and diversity the threshold should be varied. When the threshold is at 3.5-4.9, the results will vary to produce diverse results with comparable level of accuracy loss.



MovieLens data, item-based CF (50 neighbors), top-5 item recommendation

Figure: comparing standard ranking approach with item popularity ranking

Item popularity-based ranking approach is parameterized with “ranking threshold” $TR \in [TH, Tmax]$ to provide user the ability to choose a certain level of recommendation accuracy[1]. In the above graph the accuracy and diversity will be changed with different threshold values. If threshold is decreased accuracy is increased so less no. of items recommended and if threshold increases more no. of items recommended so that diversity is more. In particular, given any ranking function $rankX(i)$, the ranking threshold TR is used for creating the parameterized version of this ranking function, $rankX(i, TR)$, which is formally defined as:

$$Rank_{\alpha}(i, TR) = \begin{cases} rank_{\alpha}(i), & \text{if } (R^*(u,i) \in [TR, Tmax]), \\ \alpha_u + rank_{Standard}(i), & \text{if } R^*(u,i) \in [TH, TR] \end{cases}$$

Where $I_u^*(TR) = \{i \in I / R^*(u, i) \geq TR\}$, $\alpha_u = \max_{i \in I_u^*(TR)} rank_{\alpha}(i)$

It is known fact that in real life applications users won't accept the accuracy loss so, by using this parameterized approach the users will get the more diverse results with limited accuracy loss.

4. NOVELTY OF RECOMMENDATIONS:

So far we have discussed about the accuracy and diversity in recommender systems. But there are many factors that influence the quality of recommender systems like novelty, serendipity, utility, robustness, privacy, and trust etc. The typical recommender systems focused on accuracy and relevance as targets for satisfying the user information need. However, there is an increasing concern for the need of something more than accuracy to maximize the practical utility and the effective value of the retrieved information [8]. In particular the concepts of diversity and novelty are being increasingly recognized as important ingredients of information value in many

application domains. The issues related to diversity are presented in section 3.2 and section 3.3.

Now it is the time to discuss about novelty of recommender systems. Novelty is defined as the quality of the system that avoids the redundancy. For example in information retrieval if the system recommends two documents to users which are having similar content, it gives the little marginal utility from one another. The typical recommender systems focuses on the accuracy of recommendations that leads to some limitations while evaluating the recommender systems. Beyond accuracy there are many factors that effect the user satisfaction. To evaluate recommender systems we can consider other factors like novelty and user needs and expectations.

To provide automatic recommendations, diversity and novelty are the desirable features. The novelty of a piece of information generally refers to how different it is with respect to the information that is seen previously, by the particular users or group of user community [8]. Novelty is especially relevant to long tail effect. Diversity can be applied to set of items which show the items that are different from each other. This is related to novelty, when a set is diverse, each item is "novel" with respect to the rest of the set. A system that generates novel results tends to generate diverse results for each user over time and also enhances the global "diversity of sales" from the system perspective. It is worth to make a distinction between individual diversity and aggregate diversity.

5. ITEM NOVELTY MODELS:

In this paper we propose two item novelty models; Popularity based item novelty and distance based item novelty. In a generic sense, item novelty can be defined as the difference between an item and "what has been observed" in some context. In popularity based item novelty, the notion of item discovery enables a formulation of this principle as the probability that an item was not observed before.

$$nov(i|\theta) = 1 - p(seen|i, \theta)$$

The contextual variable θ here represents any element on which item discovery may depend, or relative to which we may want to particularize novelty. That includes a specific user, a group of users, vertical domains, sources of item discovery – such as searching, browsing, past or alternative recommendations, friends, advertisements, etc. In general terms, $p(seen|i, \theta)$ reflects a factor of item popularity, whereby high novelty values correspond to long-tail items few users have interacted with, and low novelty values correspond to popular head items.

As an alternative to the popularity-based view, we consider a similarity-based model i.e., distance based item novelty model where item novelty is defined by a distance function between the item and a context of experience.

6 CONCLUSION AND FUTURE WORK

The application can now be extended to all the ecommerce websites. Websites such as e-learning can also benefit where in with suitable modifications a learners interest in the content can be found and addressed. This paper presents the details of diversity and novelty of recommendations and the techniques to improve them. And in future research the other factors of recommendation quality also be reflected like serendipity etc., and also more techniques can be developed to improve the diversity and novelty beyond the popularity.

This work gives rise to several interesting directions for future research.

1. The application can be used in mobiles.
2. Updated products info can be notified to the customers.
3. Search suggestions can be included.

REFERENCES

- [1] Gediminas Adomavicius, Member, IEEE, and YoungOk Kwon "Improving Aggregate Recommendation Diversity using Ranking Based Techniques".
- [2] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 6, pp. 734- 749, June 2005.
- [3] C. Anderson, the Long Tail. Hyperion, 2006.
- [4] C-N. Ziegler, S.M. McNee, J.A. Konstan, and G. Lausen, "Improving Recommendation Lists through Topic Diversification," Proc. 14th Int'l World Wide Web Conf., pp. 22-32, 2005.
- [5] Chris Anderson, "Recommender systems for e-shops"
- [6] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J. Riedl, "Evaluating Collaborative Filtering Recommender Systems," ACM Trans. Information Systems, vol. 22, no. 1, pp. 5-53, 2004.
- [7] S.Abirami, A.Carolin Arockia Mary, V.R.Azhaguramyaa, "improving aggregate recommendation diversity using top k-queries" International Journal of Engineering Research & Technology (IJERT).
- [8] Saúl Vargas Sandoval, Novelty and Diversity Enhancement and Evaluation in Recommender Systems.
- [9] Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004, January). Evaluating Collaborative Filtering Recommender Systems. *ACM Transactions on Information Systems*, 22(1), 5-53.

PUNYAVATHI K received her B.Tech Degree in Computer Science and Engineering from Sai Spurthi Institute of Technology, JNTUH, Hyderabad, India in 2011. Now she is pursuing M.tech in Computer Science from Aurora's Scientific Technological and Research Academy, JNTUH, Hyderabad, India.

JYOTHI P received her M.tech Degree in Information Technology from Aurora's Scientific Technological and Research Academy, JNTUH, Hyderabad, India in 2010. She is currently working as a Senior Asst. Professor in the department of Information Technology in Aurora's Scientific Technological and Research Academy, JNTUH, Hyderabad, India.