International Journal of Information & Computation Technology.
ISSN 0974-2239 Volume 2, Number 3 (2012), pp. 141-151

http://www. ripublication.com

Aggregate Diversity Improvement in Recommender System through Ranking Approach

Doddegowda B.J., Sanjay Kumar, Rabindranath S. and Nagaraj K.

Department of Computer Science & Engineering
AMC Engineering College, Bangalore, India
E-mail: bjdgowda10@gmail.com sksanjay2@gmail.com
raben_s@yahoo.com rajahridaya@gmail.com

Abstract

Recommender system gives recommendations for its user and one of the important quality aspects of the recommendations is diversity. In this paper, effort is put for improving the aggregate diversity, diversity across all users, in recommendations of the recommender system using ranking approach. In presented ranking approach, to improve diversity in recommendations, because to make user to interact with system consistently since it is e-commerce oriented system, work is done under predefined assumptions: 1) products already popular or declared as excellent will not be considered and 2) products already declared as bad will not be considered in recommendation process. The effort utilizes collaborative filtering technique for filtering purpose, in collaborative filtering both users based and item based methods are used one after the other. The proposed approach broadly works in two phases, First phase, it will list out all the similar products which are collected by the result of collaborative filtering and remove the products from list according the assumptions. In second, it assigns ranks to the products in the list. Products with higher rate count will get low rank and vice versa and from the ranked products list system will recommend the information items or products to the user and while recommending it should consider about similarity of the products also. Here, popular products are not considered hence, long tail products will get opportunity to be recommended and in turn improves the aggregate diversity in recommendations.

Keywords— Recommender System, Collaborative Filtering, Diversity, User based, Item based, Ranking, Recommendations

Introduction

Recommender system, an information filtering technology, commonly used on e-commerce web sites that use a collaborative filtering to present information on items and products that are likely to be of interest to the reader or user. To presenting the recommendations, the recommender system will use details of the registered user's profile and opinions and habits of their whole community of users and compare the information to reference characteristics to present the recommendations. Systems that attempt to predict items, e.g., movies, music, books, that a user may be interested in and Systems that help people find information that will interest them, by facilitating social / conceptual connections or other means. Recommender system is absolute necessity because of growth of the original information in the world every year. The size of the original information is in terms of Exabyte, the Fig .1 shows the scenario why there is a need of recommender system and it is because of many options are present in the world. When a particular person wants to buy something but he do not able to make a decision among numerous options, there exactly the recommender system will recommend some products or information items to particular person and those recommended to the person are relevant to the his interest. By recommending products to person recommender system minimizes the effort for finding the relevant products.

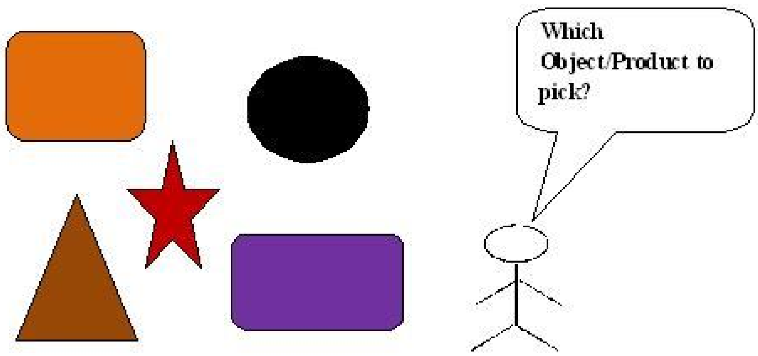

Fig.1 Necessity of Recommender System

Businesses increases amount of information that customer must process before they are able to select which products meet their needs. Effective solution for this information overload problem is recommender system. To recommend products based on different constraints like, top sellers on site, past behaviour of the user and all these techniques are part of personalization on a site. To achieve personalization on site means, system adapts itself to each customer. Recommender system is needed to reach individual personalization for each user or customer. Usually presence of recommender system will make following things and more to them get happen.

- Automates quotes like: I like this book; you might be interested in it, I saw this movie, you'll like it, don't go see that movie!

- Many of the top commerce sites use recommender systems to improve sales
- Users may find new books, music, or movies that was previously unknown to them
- Also can find the opposite for e.g.: movies or music that will definitely not be enjoyed

Opportunity for Recommendations

A good recommendation engine can make a difference for any online business. This is because there are two fundamental activities online - Search and Browse. When a consumer knows exactly what he is looking for, he searches for it. But when he is not looking for anything specific, he browses. It is the browsing that holds the golden opportunity for a recommendation system, because the user is not focused on finding a specific thing – he is open to suggestions. The overview of the recommender system is shown in the Fig.2.

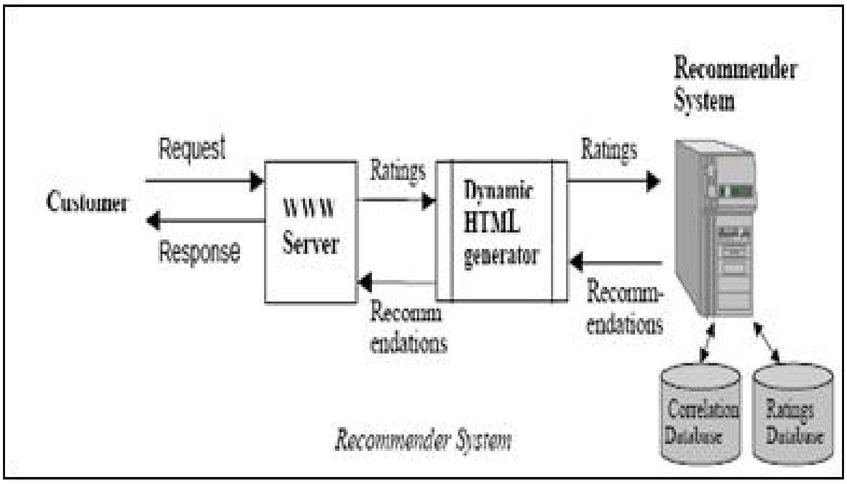

Fig.2 Overview of Recommender systems

Quality Aspects

Accuracy, it is a highly focused quality aspect of the recommendation in almost all the works on recommender system till date. Accurate recommendation list generally consist related or similar products and which are already popular or high rated. It states that how products are similar to each other.

Diversity, in all the previous studies and works it is not well considered quality aspect of recommendation. Diversity of recommendation is recently considered as important as accuracy. Some studies argue that diversity is much required compare to the accuracy. Diversity states how products are different to each other. There are two types of diversity like individual diversity and aggregate diversity. In diversity, there are two types: Individual Diversity, Number if distinct products recommended to user and Aggregate Diversity, Number of different products recommended to all users [2].

Assumptions of the Approach

- Product belong to ‘excellent’ group of average rating is not considered in recommendation, because in that group products are already popular and users can easily find them so recommendation effort will be worthless.
- Product belong to ‘bad’ group of average rating is not considered in recommendation, because in that group products are already declared as bad and users can not go through or invest their money on them so recommendation effort will be worthless.

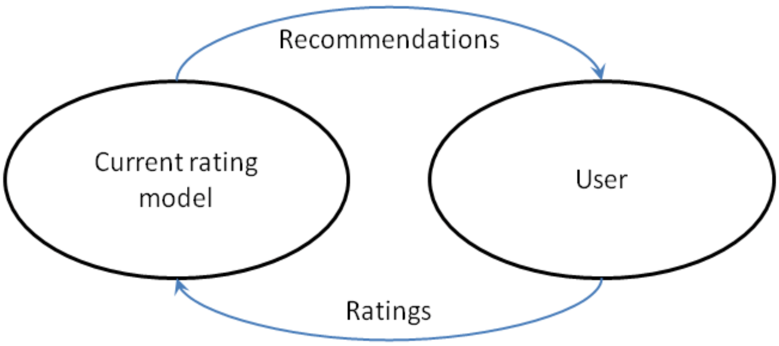

Fig.3 Recommendations on Ratings

The proposed approach to achieve aggregate diversity in recommendations uses the rating activities or purchase activities of the users. On that basis system will collect relavant details to recommend relevant products to the user. The process of recommendations based on ratings is shown in the Fig.3.

Related work

Recommendation Algorithms

Let *U* be the set of users and *I* be the set of items available in the recommender system. Then, the usefulness or utility of any item i to any user u can be denoted as *R (u,i)*, which usually is represented by a rating (on a numeric, ordinal, or binary scale) that indicates how much a particular user likes a particular item. Thus, the job of a recommender system in the rating estimation phase, is to use known ratings as well as other information that might be available) to estimate ratings for items that the users have not yet consumed. For clarity, use *R (u,i)* to denote the actual rating that user u gave to item i, and *R*(u,i)* for the system-estimated rating for item i that user u has not rated before [5].

Collaborative Filtering

It is a technique used by some recommender systems. Collaborative Filtering (CF) has two senses, a narrow one and a more general one. In general, collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets.

Collaborative filtering methods have been applied to many different kinds of data including sensing and monitoring data - such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data - such as financial service institutions that integrate many financial sources; or in electronic commerce and web 2.0 applications where the focus is on user data, etc. In the newer, narrower sense, collaborative filtering is a method of making automatic predictions about the interests of a user by collecting preferences or taste information from many users.

The underlying assumption of the collaborative filtering approach is that if a person A has the same opinion as a person B on an issue, A is more likely to have B's opinion on a different issue x than to have the opinion on x of a person chosen randomly. For example, a collaborative filtering recommendation system for television tastes could make predictions about which television show a user should like given a partial list of that user's tastes.

Note that these predictions are specific to the user, but use information gleaned from many users. This differs from the simpler approach of giving an average (non-specific) score for each item of interest, for example based on its number of votes. Collaborative filtering technique is classified into three categories: memory based, content based and hybrid. Here, use of memory based collaborative filtering is takes place.

E-Commerce Application

Recommender systems are being used by an ever-increasing number of E-commerce sites to help consumers find products to purchase. What started as a novelty has turned into a serious business tool. Recommender systems use product knowledge – either hand-coded knowledge provided by experts or “mined” knowledge learned from the behaviour of consumers – to guide consumers through the often-overwhelming task of locating products they will like. In this article authors present an explanation of how recommender systems are related to some traditional database analysis techniques. They examine how recommender systems help E-commerce sites increase sales. Some works identify five commonly used E-commerce recommender application models, describe several open research problems in the field of recommender systems, and examine privacy implications of recommender systems technology. The five models are listed below [1].

- Helping new and Infrequent Visitors: Broad Recommendation List
- Building Credibility through Community: Customer Comments and Ratings
- Inviting Customers Back: Notification Services
- Cross-Selling: Product-Associated Recommendations
- Building Long-Term Relationships: Deep Personalization

Recommender systems enhance E-commerce sales in three ways

Converting Browsers into Buyers

Visitors to a Web site often look over the site without purchasing anything. Recommender systems can help consumers find products they wish to purchase.

Increasing Cross-sell:
Recommender systems improve cross-sell by suggesting additional products for the customer to purchase. If the recommendations are good, the average order size should increase. For instance, a site might recommend additional products in the checkout process, based on those products already in the shopping cart.

Building Loyalty:
In a world where a site's competitors are only a click or two away, gaining consumer loyalties is an essential business strategy. Recommender systems improve loyalty by creating a value-added relationship between the site and the customer. Sites invest in learning about their customers, use recommender systems to workable that learning, and present custom interfaces that match consumer needs. Consumers repay these sites by returning to the ones that best match their needs [1].

Standard Ranking Approach
Typical recommender systems predict unknown ratings based on known ratings, using any traditional recommendation technique such as neighborhood-based or matrix factorization CF techniques. Then, the predicted ratings are used to support the user's decision-making. In particular, each user *u* gets recommended a list of top-N items, $L_N(u)$, selected according to some ranking criterion. More formally, item i_x is ranked ahead of item i_y (i.e., $ix < iy$) if $rank(i_x) < rank(i_y)$, where $rank: I \to R$ is a function representing the ranking criterion [4]. The vast majority of current recommender systems use the predicted rating value as the ranking criterion:

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

The power of -1 in the above expression indicates that the items with highest predicted ratings R*(u, i) are the ones being recommended to user. Note that, by definition, recommending the most highly predicted items selected by the standard ranking approach is designed to help improve recommendation accuracy, but not recommendation diversity. Therefore, new ranking criteria are needed in order to achieve diversity improvement. Since recommending best selling items to each user typically leads to diversity reduction, recommending less popular items intuitively should have an effect towards increasing recommendation diversity [2].

Recommendation Method Description
Each category discussed here represents a family of algorithms and approaches. The raw retrieval system provides customers with a search interface through which they can query a database of items. In this case, recommendation is a "binary," syntactic process whereby the system recommends whatever the customer has requested. While not technically a recommender application, such an application may appear as one to customers.

Applications that value personality over personalization may create sets of recommendations that have been manually selected by editors, artists, critics, and other experts. These "human recommenders" identify items based on their own tastes, interests, and objectives and create a list of recommended items available to community members. Their recommendations are often accompanied by text comments that help other customers evaluate and understand the recommendation. This process most closely mimics traditional critics and editors, including both potential insight and potential bias.

Some cases where personalization is impractical or unnecessary, recommender applications can very efficiently provide statistical summaries of community opinion. These summaries include within-community popularity measures and aggregate or summary ratings. It recommends the non-personalised products and it not so efficient on internet and this kind of products can be found in any physical stores easily.

Recommendations based on the syntactic properties of the items and customer interests in those properties use attribute-based recommendation technologies though the simplest attribute-based recommendation is raw retrieval, true recommenders that use attributes model customer interests beyond a simple query. Other attribute-based recommenders use customer profiles that indicate likes or dislikes making recommendations to the customer.

Other applications use item-to-item correlation [3] [4] to identify items frequently found in association with items in which a customer has expressed interest. Association may be based on co-purchase data, preference by common customers, or other measures. In its simplest implementation, item-to-item correlation can be used to identify matching items for a single item, such as other clothing items that are commonly purchased with a pair of pants.

More powerful systems match an entire set of items, such as those in a customer's shopping cart, to identify appropriate items to recommend. Item-to-item correlation recommender applications usually use current purchases or other current interests rather than long-term customer history, which make them particularly well-suited for recommending gifts.

Finally, recommender systems using user-to-user correlation [3] recommend products to a customer based on the correlation between that customer and other customers who have purchased products from the E-commerce site. This technology is often called “collaborative filtering” because it originated as an information filtering technique that used group opinions to recommend information items to individuals Though they use the word correlation in the name of this technique, thus hinting at nearest-neighbour techniques based on linear correlation, the technique can be implemented with many other technologies as well.

One important issue when considering the recommendation method is whether the computation can be performed entirely online, while the Web store is interacting with the customer, or whether parts of the computation must be performed offline for performance reasons.

Online recommendations are preferred because they can respond immediately to the consumer’s preferences. Most of the recommender processes mentioned above can be performed entirely online. Raw retrieval, manual selection, statistical summarization, and attribute-based are all simple computations that are usually performed during customer interaction. Item-to-item correlation and user-to-user correlation are computationally more intensive and often require an offline component to prepare a model that can be executed efficiently online. One challenge in designing the model-building is to ensure the resulting online system is as responsive as possible to interactive input from the user.

Proposed Approach

This section of the paper explains how exactly proposed approach works. Here, two methods of collaborative filtering used and they are user based and item based collaborative filtering [4] to filter the users and products according to the similarity.

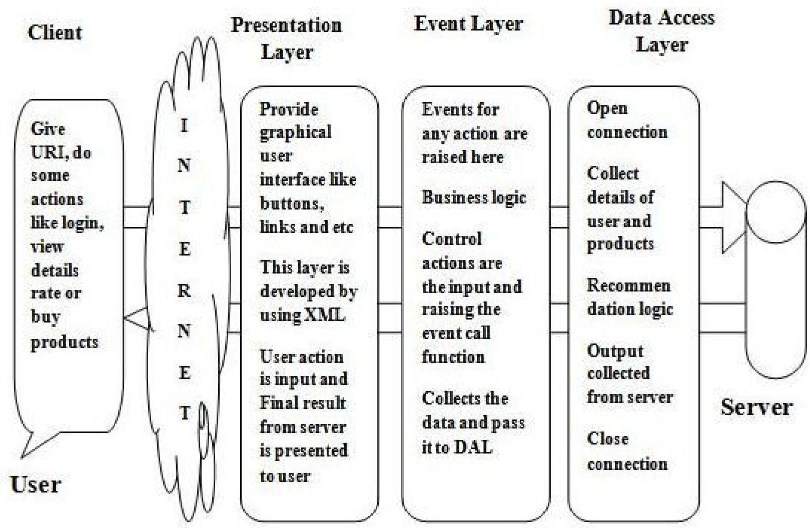

Fig.4 Architecture of Recommender System

First, once user gives rating or place an purchasing order for a product; system collects all required details about the particular user and product. User details are like, user identification number, current rate he has given, his history where all the products for which he gave rating or he purchased, etc. Product details are like, category of the product, sub-category of the product, attributes of the products, average rating of products, etc.

Second, system finds similar users, by comparing the users who are all gave same rating as the current user given to a current product. To find the similar user cosine similarity check will be followed as given in formula (1). If many users are rated the same products or purchased it, system will consider only top-n highly similar user, where similarity is the result of the cosine similarity.

Third, once after finding the similar users system will find the similar products. Similar products are from the similar users' history. All products collected from similar users' history and those are similar to current product. If similar products too many then system selects only top-n highly similar products and here also, similarity between products checked by cosine similarity as given in formula (1).

$$\mathrm{Sim}(x,y)=\cos(x',y')= x'.y'/\|x'\|_2 \times \|y'\|_2 = \sum_{i\in Ixy} r_{x,i}\, r_{y,i} / \sqrt{\sum_{i\in Ixy} r_{x,i}^2}\, \sqrt{\sum_{i\in Ixy} r_{y,i}^2} \qquad (1)$$

Fourth, after finding the similar products system will remove the products according to assumptions and make sure that any similar product should not be in current user's history and no duplicate product in selected similar products. After that system will assign ranks to the products in similar products list. Ranks are according to the total number rating given to that particular product so far. If product is having high total number of ratings then it is ranked as low and vice versa.

Sixth, system will select top-n products from ranked list started with first rank in list. Those top-n products recommended to the user.

The technical architecture of the proposed approach is shown in Fig.4. Here, it clearly shows that data access layer will directly interacts with data server. Data is collected from the users' activities and related data for user and products are collected from data store of the server.

When there is increase in diversity of the recommendation, there is decrease in accuracy of the recommendations because diversity includes many different products. Fig.5 shows how the increase of the diversity will decrease the accuracy of the recommendations.

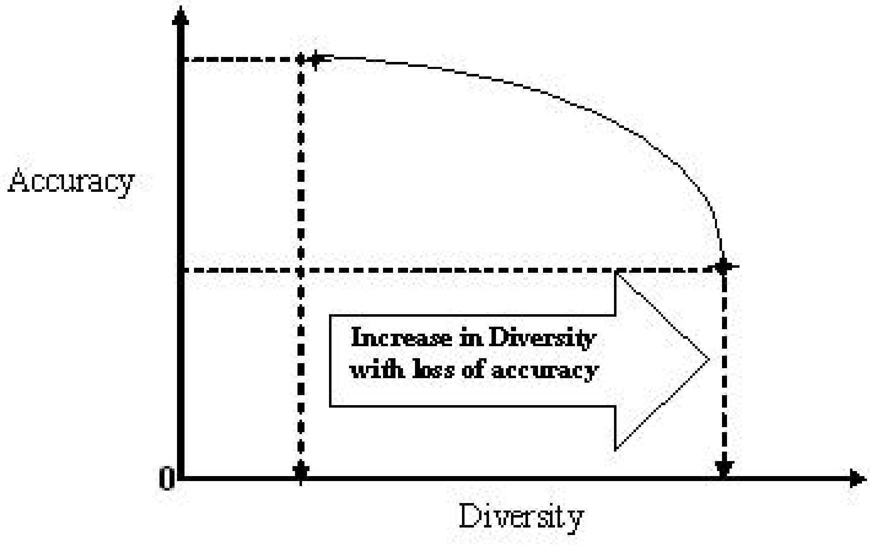

Fig.5 Increase of Diversity

If diversity increases means it actually states that there are many long tail products are added in the recommendations. Once many and many long tail products are taken into the consideration for recommendation in recommender system then there is loss of precision or accuracy. The variations of long tail products and accuracy of the recommendation is shown in the Fig.6.

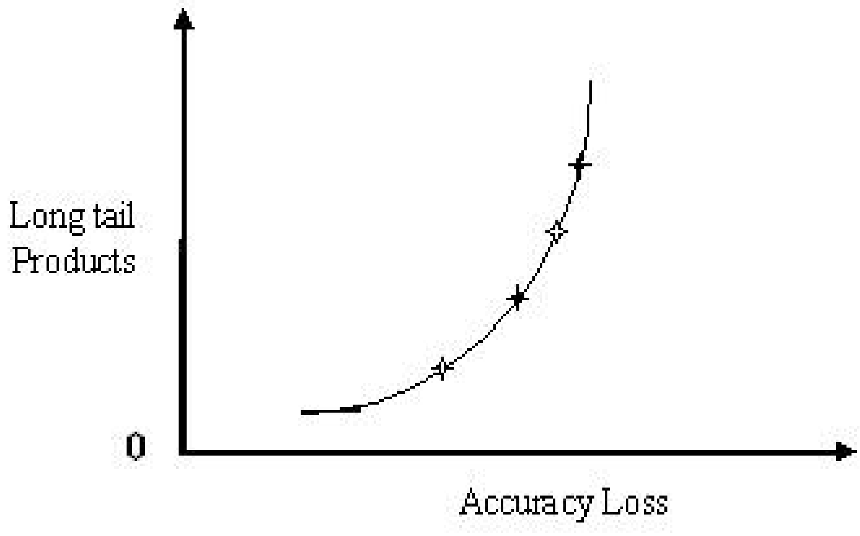

Fig.6 Accuracy Loss and more Long Tail Products

Conclusion

The proposed approach achieves aggregate diversity in recommendations. By applying this approach, long tail products get an opportunity to become recommended to the user. Long tail products get an opportunity because all the highly popular or products which are having high total count of ratings are ranked low and products which are less popular are ranked high, then long tail products are ranked high and products from that ranked list are recommended. This approach also give exposure the e-business since it is recommending products which different across all users and relevant to the each user.

References

[1] J. Ben Schafer, Joseph A. Konstan, John Riedl, "E-Commerce Recommendation Applications", Page 1 to 24, E-commerce Recommender Applications, Group Lens Research Project Department of Computer Science and Engineering University of Minnesota

[2] Gediminas Adomavicius, *Member, IEEE,* and YoungOk Kwon, "Improving Aggregate Recommendation Diversity Using Ranking Based Techniques", *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, 2009 IEEE*

[3] Greg Linden, Brent Smith, and Jeremy York. "Amazon.com

[4] Recommendations" *Amazon.com, Industry Report,* JANUARY• FEBRUARY 2003, Published by the IEEE Computer Society, *IEEE INTERNET COMPUTING*

[5] Jun Wang, Arjen P. de Vries, Marcel J.T. Reinders, "Unifying User based And Item based Collaborative Filtering Approaches by Similarity Fusion", SIGIR'06, August 6–11, 2006, Seattle, Washington, USA.

[6] Gediminas Adomavicius, YoungOk Kwon, "Maximizing Aggregate Recommendation Diversity: A Graph-Theoretic Approach", Workshop on Novelty and Diversity in Recommender Systems (DiveRS 2011), held in conjunction with ACM RecSys 2011. October 23, 2011, Chicago, Illinois, USA.

[7] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl ACM Transactions on Information Systems, Vol. 22, No. 1, January 2004. 'Evaluating Collaborative Filtering Recommender Systems'

[8] A. Kumar and Dr. P. Thambidurai, Global Journal of Computer Science and Technology, Vol. 9 Issue 5 (Ver. 2.0), January 2010 'Collaborative Web Recommendation Systems-A Survey Approach'

[9] Thomas Sandholm and Hang Ung, CaRR 2011, February 13, 2011, Stanford, CA, USA. 'Real-time, Location-aware Collaborative Filtering of Web Content'

[10] J. Ben Schafer, Joseph A. Konstan, John Riedl 'E-Commerce Recommendation Applications'