ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

DIVERSITY IMPROVEMENT IN RECOMMENDER SYSTEM

Antony Taurshia.A[1],S.Deepa Kanmani [2]

[1]Post-Graduate Student, Department of Computer Science and Engineering, Karunya University, India
[2] Assistant professor, Department of Computer Science and Engineering, Karunya University, India

***Abstract*— Recommender Systems are software tools and techniques for providing suggestions to a user. The suggestions relate to decision-making processes, like what items to buy, what music to listen to, or what online news to read. There are different ways in which recommendations can be made. The success of recommender system depends on the usefulness of the system. The usefulness can be measured in terms of accuracy, diversity, flexibility, serendipity and reliability. Most of the previous works have been focused only on improving recommendation accuracy. One of the important aspects like diversity has been never considered. In this paper several recommendation techniques have been explored. A graph based approach for maximizing diversity and item bundles for increasing capability have been proposed.**

***Index Terms*—Accuracy, aggregate diversity, individual diversity, recommendtation.**

I. INTRODUCTION

There are vast amount of information available and the recommender system is proven to be useful in extracting information and making useful recommendation to users. Recommender system plays an important role in electronic commerce. It increases sales by recommending items and it allows users to make decisions such as which item to buy. Item is termed as what the system recommends to users such as movies, books. For example Amazon.com uses recommender system for recommending books to users. Mostly recommender systems work based on the ratings given by user. Rating is user's preference for an item. Rating directly given by the user is called known rating.

The ratings of unrated items are predicted (i.e predicted rating) by some recommender system from the information available and then recommendation is made. For increasing diversity the rated items can be ranked and then recommended to users to maximize the user's utility.

There are several recommendation techniques like collaborative filtering, content based, hybrid. Content based recommender system uses history of user's preferences for recommending items. Collaborative filtering recommender system recommends items based on preferences of similar users. Hybrid recommender system uses a combination of recommender system for recommending items. So far used recommender system have focused only on improving recommendation accuracy. This causes overspecialization which leads to frustration of the user. Some of the researches have focused on improving individual diversity. A graph based approach is proposed which gives better diversity improvement compared to the ranking method. The concept of item bundles is also proposed which can increase the capacity of recommender system.

II. KEYWORDS

This consist of several keywords about the recommender system.

ACCURACY

Accuracy is how well a recommender system make predictions. Accuracy can be calculated as truly highly ranked items divided by highly ranked items.

AGGREGATE DIVERSITY

Aggregate diversity is the diversity in recommendation list across all users.

INDIVIDUAL DIVERSITY

Individual diversity is the diversity in the individual user's recommendation list.

RECOMMENDATION

Recommendation is the suggestion given by system to user like suggestion for books in Amazon.com and movies in Netflix.

III. RECOMMENDATION TECHNIQUES

This section briefly discusses several recommendation techniques.

A Content Based Recommendation

Content based recommender system [6] recommend items based on user profile information and item description. User's profile contains information like description about the type of items that interest the user and the history of user's interaction with the recommender system. Items information are stored in a database with its attributes. The key component of content based recommendation is classification learning algorithm that creates a user model from the user history. This recommender system is capable of introducing new items to user. It can also provide explanation for recommending items.

.
***A.Antony Taurshia**, CSE/KARUNYA UNIVERSITY*
***S.Deepakanmani** ASSISTANT PROFESSOR/CSE/KARUNYA UNIVERSITY*

The user has to fill profile details mandatorily in order to get recommendations.

B Collaborative Filtering Recommendation

Collaborative filtering recommender system [4] recommends items based on the past preferences of similar users. First the user gives the preferences by rating the items. Based on the users ratings the system finds the similar users. With the similar users the ratings of unrated items are predicted and recommended to users. There are several approaches of collaborative filtering technique. Active filtering uses peer-to-peer approach, people who have similar interest rate products. Passive filtering approach uses implicit information like user's action for recommendation. Item-based filtering system uses item-item relationship for recommendation.

The collaborative filtering technique can be memory based or model based.

Memory based CF: Heuristic based techniques recommend items based on the past activites of users.

Model based CF: This technique learns a predictive model based on the past user activities using statistical or machine learning model.

The system should have enough ratings for recommending items as it is fully dependent on rating. This is called ramp up problem.

C Knowledge Based Recommendation

This technique uses the knowledge about products and users needs for making recommendations. Recommendation is made by matching the similarity between user's preference and product description. It does not suffer from ramp up problem since it does not depend on the ratings given by user. This system needs a database and needs to be updated for making useful recommendations.

D Outside The Box Recommendation

The problem of overspecialization is overcome using OTB recommendations [3], [5] and helps to make fresh discoveries. This technique uses a concept called item region. Region (i.e the "box") is defined as the group of similar items. Regions are created based on similarity distances between items. Stickiness is user's familiarity to a region. Based on the stickiness the system finds items that are not familiar to the user and recommends those items. This technique increases the novelty.

E Hybrid Recommendation

Hybrid recommendation uses a combination of recommendation technique. Content based recommendation and collaborative filtering recommendation is the commonly used combination. Both the system suffers from ramp up problem. The disadvantages in both the techniques can be overcome by combining them in parallel or cascade. Both the rating and the profile data can be used for finding recommendations. This technique improves the recommendation accuracy but diversity is not considered.

IV. ITEM BUNDLING

The usefulness of recommender system can also be improved by recommending item as bundles (eg fairever and bangles as a bundle). Significant savings can be made by recommending item as bundles.

The bundles can be created using any association rules. There are different types of bundle like deterministic bundle and non deterministic bundle.

Once the user gives the preference, the item bundles are selected based on preference. Then the bundles are ranked based on savings.

$$sav_b = ub_c - b_c \qquad 5.1$$

Where ub_c is the unbundled cost and b_c is the bundled cost. Ranking items according to the savings drags the attention of users. The user can get benefited through savings and the company will also get more customers.

V. MAXIMUM BIPARTITE MATCHING

The Ranking Techniques improve the aggregate diversity to some extent. A Graph based approach can maximize diversity compared to the ranking techniques. Max-flow problem in graph theory can be used to improve diversity.

Let $G = (V, E)$ be a directed graph where that V is the set of vertices, and E is the set of directed edges, each of which connects two vertices with a single source node $s \in V$ and a single sink node $t \in V$. Each directed edge $e \in E$ has capacity $c(e) \in \mathbf{R}$ associated with it. Also, the amount of actual flow between two vertices is denoted by $f(e) \in \mathbf{R}$. The flow of an edge cannot exceed its capacity, and the sum of the flows entering a vertex must equal the sum of the flows exiting a vertex, except for the source and the sink vertices.

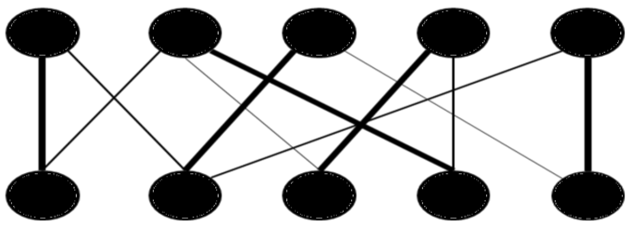

Fig 5.1 maximum bipartite matching

For recommender system let user and item be vertices. An edge from user to item exits if the item is relevant to user i.e the rating given by user should be greater than the threshold value. The max-flow approach increases the aggregate diversity where no user can be recommended more than N items and no item can be counted more than once.

Maximum flow problem can be extended to maximum bipartite matching problem. Let $G = (U, I; E)$ be a bipartite graph, where vertices represent users U and items I, and edges E represent the possible recommendations of items for users. A subset of edges M is a matching, if all edges in M are pair wise nonadjacent, i.e., any two edges in M share neither a user vertex nor an item vertex. A vertex is matched if it is adjacent to an edge that is in the matching. In Fig 5.1 let the top vertices be users and bottom vertices be items. The thick line represents a matching. The maximum matching of a bipartite graph is a matching with the largest possible number of edges.

VI. RESULT DISCUSSION

The graph-based approach is able to obtain substantial

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

diversity improvements at the given level of accuracy, compared to the ranking techniques, across all experiments including different datasets, different recommendation techniques, and different number of recommendations (N = 1, 5, 10). Item bundling and savings approach also increases the performance of recommender system.

REFERENCES

[1] Gediminas. A., and Youngok.K., "Overcoming Accuracy-Diversity Tradeoff in Recommender Systems: A Variance-Based Approach." Proceedings of the 18th Workshop on Information Technology and Systems (WITS'08), Paris, France, 2008.
[2] Gediminas. A., and Youngok. K., "Toward More Diverse Recommendations: Item Re-Ranking Methods for Recommender Systems." 19th Workshop on Information Technology and Systems (WITS'09), Phoenix, Arizona,2009.
[3] Zeinab. A., Sihem. A.Y., Lakshmanan. V.S., Sergei. V., Cong. Y., "Getting recommender systems to think outside the box." 285-288 RecSys, 2009.
[4] Herlocker. J.L., Konstan. J.A., Terveen. L.G., and Riedl. J., "Evaluating Collaborative Filtering Recommender Systems," ACM Transactions on Information Systems, 22(1), pp. 5-53, 2004.
[5] Amer-Yahia. S., Lakshmanan. L.V.S., Vassilvitskii. SS., Yu. C., "Battling Predictability and Overconcentration in Recommender Systems." IEEE Data Eng. Bulletin,33-40, 2009.
[6] Michael J. Pazzani and Daniel Billsus, "Content-Based Recommendation Systems. ", 325-341, 2007.
[7] Sangkeun. L, "A generic graph-based multidimensional recommendation framework and its implementations." WWW (Companion Volume) 161-166, 2011.
[8] Gediminas. A., and Youngok. K., "Maximizing Aggregate Recommendation Diversity: A Graph-Theoretic Approach" Workshop on Novelty and Diversity in Recommender Systems, held in conjunction with ACM RecSys, 2011.
[9] Fatih. A., Aysenur. B., "Enhancing Accuracy of Hybrid Recommender Systems through Adapting the Domain Trends."ACM RecSys'10 PRSAT Workshop, 2010.