Product Recommendations System Survey

Sahil Pathan
Sinhgad Academy of

Engineering, Pune, India.

Karan Panjwani
Sinhgad Academy of

Engineering, Pune, India.

Nitin Yadav
Sinhgad Academy of

Engineering, Pune, India

Shreyas Lokhande
Sinhgad Academy of

Engineering, Pune, India.

Bhushan Thakare
Asst. Prof. Sinhgad Academy of

Engineering, Pune, India.

ABSTRACT

Recommendation Systems are used to increase the growth of various online businesses. E-commerce players are utilizing such systems to get high sales. Such systems make use of statistics and data from user behaviour e.g. Purchase history, product ratings. So, decision to display a specific product from a specific category is taken after considering such parameters. In Hyper-Local based services (Locality Based) recommendation systems operate in a challenging environment. Such as, new customers have too much limited information associated, less purchase history, no product ratings etc. Secondly a large retailer have too much categories to choose from. Last, users tends have scattered data-less patterns. In order to handle such information mainly three methods are available: search-based methods, collaborative filtering and cluster models. These methods are more suitable in a vast user base environment. Whereas, in small scale environments a set of customers whose purchased and rated products overlaps with a current user's purchased and rated products are subject to a simple measurements.

Keywords
Cluster Model, Collaborative Filtering, Recommendation System, Search-based Methods.

1. INTRODUCTION

Recommendation systems are derived from human behavior which tends to trust products or services which are recommended to them. So if recommendations are provided from trusted source sales are definitely going to increase. Such systems are good with respect to marketing and sales. But the system needs to be cautious in terms of what products are listed as recommended because users tends to buy a product only when they found it related, to be more specific customers should think that they needed this product in the first place. In this paper a survey of some of the most reliable recommendation is provided. Reliability and scalability is also discussed.

2. ASSOCIATION RULE

This method relates association of two set of products, most likely bought at same transaction. Support and confidence are the two parameters to calculate association. Support measures the occurrence frequency while confidence measure the occurrence frequency.

$$s = \frac{number\ of\ transaction\ containing\ X \cup Y}{number\ of\ transactions}$$

Support indicates percent of transactions containing X ∪ Y. Whereas, Confidence: percent of transaction contains X also contains Y.

$$c = \frac{number\ of\ transations\ containing\ X \cup Y}{number\ of\ transaction\ containing\ X}$$

If a rule has high confidence is more important because it predicts accurate outcome. Association rules are of the form if X then Y. For example if 70% of those who buy books online also buy music online. These rules are used to target customers for marketing or product placing. Association rules are derived directly from historical data, as opposed to via customer surveys or any other ways.

These rules should be utilized with caution. Because the interpretation given by association rule suggests a strong occurrence relation between products not necessarily implies causality.

Once the frequent item sets from transactions in a database have been found, it is straightforward to generate strong association rules from them, where strong association rules satisfies both minimum support and minimum confidence. Support is an important measure because a rule that has very low support may occur by chance. A low support rule is not that much important for business. Confidence on the other hand measures the reliability of the inference made by rule. It provides the estimation of the conditional probability of Y given X.

3. COLLABORATIVE FILTERING

This method mainly depends upon other users with similar history buys similar set of products. For this nearest users are selected based on distance between them in k-dimensional space. Representation, nearest user formation and recommendation generation are the three main tasks of such systems.

If two customers a and b are thought of as two vectors in space, then proximity between two users is cosine of the angle between them,

$$cos(\vec{a}, \vec{b}) = (\vec{a} \cdot \vec{b}) / |\vec{a}| * |\vec{b}|$$

But if sample user set is small then chances are current selected user is less similar to existing users. Collaborative filtering method needs a large dataset for a close recommendation.

Number of products recommended is restricted to very less category and most viewed products from other categories are not recommended.
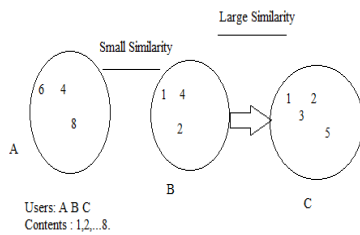

Fig. 1: Collaborative Filtering

4. SEARCH BASED METHOD

This kind of recommendation systems are too general ,because they are based on search keywords like in Hyper-local based service example grocery shopping if a user searched for kissan jam then search based recommendation system recommends product range from keywords: kissan+jam ,too general. How about considering that user will purchase jam and display product whole wheat bread for a perfect snack. Search queries are separated on the word basis and products from those searched words are considered.

By considering such search queries recommendation system is not utilized fully. Product listings are only shown on the basis of search queries. So the recommendation becomes too much category specific, user miss out the other popular products from different categories in which he might be interested in.

So popular products from same categories are recommended in such systems. So this recommendation system is too category specific and completely neglects intra category recommendation.

Search based method build keyword and category index to provide such recommendation.

5. CLUSTER MODELS

In this model users with similar interest are grouped together into class, the problem is solved as a classification problem. After assigning users to a particular class of similar interest then recommendation is generated accordingly. Using a similarity metric, the most similar users are grouped. But for initial stage a random training data set is taken and users are grouped. After this a continuously matching is performed with existing classes and a new merged class is formed.

When a final classes is generated customers similarity with strongly similar classes is taken into consideration for recommendation.
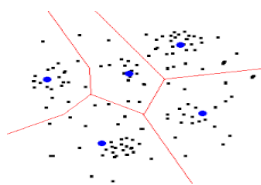

Fig. 2: Example of K-means Clustering

Cluster models have better online scalability and performance than collaborative filtering because they compare the user to a controlled number of classes than the entire user base.

6. ITEM TO ITEM COLLABORATIVE FILTERING

It is a method for collaborative filtering in which users purchased and rated items are matched instead of matching similar user base. A set of items is created that user tends to purchase together. Item to item matrix can be created for recommendation of similar set of items.In terms of scale if large number of items are present for computation, calculations should be performed before transactions. Because it's not compatible for such algorithms to handle such vast computation at that time span.

Algorithm works as, for every item in catalog X as well as a customer who bought from that catalog X AND every item in other catalog Y. A user record is recorded which matched above condition. Next for every item in catalog Y, similarity is calculated between X and Y. Such similarities are generally calculated using cosine measures.

7. CONCLUSION

Analyzing above all recommendation systems approaches the most suitable approach to hyper-local based service is combine selective methods according to our requirement. A hybrid method is most suitable for such kind of environment. Parameters like purchase history and ratings are core of collaborative filtering, a new parameter number of product views is added for calculation. So as to increase the probability to a specific recommended product for which a user is looking for. Clustering is done on the basis of three parameters purchase history, previous ratings, product views etc. After acquiring set of classes of similar interest association rule is applied to only those classes rather than whole bunch of purchased products. This dramatically increases overall integrity of a recommendation system. Because the point of view is narrowed down to the exact products which are of similar interest and provides support and confidence to that products. Recommendation systems provides an efficient way of targeted marketing which directly helps in growth of businesses. A best user experience is achieved by putting a more personalized view to each user.

With very little computation a vast sales growth is achieved so such system is a great tool.

8. REFERENCES

[1] J.B. Schafer, J.A. Konstan, and J. Reidl, "E-Commerce Recommendation Applications," Data Mining and Knowledge Discovery, Kluwer Academic, 2001, pp. 115-153.

[2] P. Resnick et al., "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," Proc. ACM 1994 Conf.Computer Supported Cooperative Work, ACM Press, 1994, pp. 175-186.

[3] B.M. Sarwarm et al., "Analysis of Recommendation Algorithms for E-Commerce,"ACM Conf. Electronic Commerce, ACM Press, 2000, pp.158-167.

[4] L. Ungar and D. Foster, "Clustering Methods for Collaborative Filtering," Proc. Workshop on Recommendation Systems, AAAI Press, 1998.

[5] B.M. Sarwar et al., "Item-Based Collaborative FilteringRecommendation Algorithms," 10th Int'l World

Wide Web   Conference, ACM Press, 2001, pp. 285-295.

[6] P.S. Bradley, U.M. Fayyad, and C. Reina, "Scaling Clustering Algorithms to Large Databases," Knowledge Discovery and Data Mining, Kluwer Academic, 1998, pp.   9-15.

[7] L. Ungar and D. Foster, "Clustering Methods for Collaborative Filtering," Proc. Workshop on Recommendation Systems, AAAI Press, 1998.

[8] M. Balabanovic and Y. Shoham, "Content-Based Collaborative Recommendation," Comm. ACM, Mar. 1997, pp. 66-72.

[9] G.D. Linden, J.A. Jacobi, and E.A. Benson, Collaborative Recommendations Using Item-to- Item Similarity Mappings, US Patent 6,266,649 (to Amazon.com), Patent and Trademark Office, Washington, D.C., 2001.

[10] Canny, J. F. (2002). Collaborative filtering with privacy via factor analysis. SIGIR (pp. 238–245). ACM.

[11] Carreira-Perpinan, M., & Hinton, G. (2005). On contrastive divergence learning. 10th Int. Workshop on Artificial Intelligence and Statistics (AISTATS'2005).