User Segmentation for Controlling Recommendation Diversity

Farzad Eskandanian
Center for Web Intelligence,
DePaul University
Chicago, IL 60604
feskanda@depaul.edu

Bamshad Mobasher
Center for Web Intelligence,
DePaul University
Chicago, IL 60604
mobasher@cs.depaul.edu

Robin Burke
Center for Web Intelligence,
DePaul University
Chicago, IL 60604
rburke@cs.depaul.edu

ABSTRACT

The quality of recommendations is known to be affected by diversity and novelty in addition to accuracy. Recent work has focused on methods that increase diversity of recommendation lists. However, these methods assume the user preference for diversity is constant across all users. In this paper, we show that users' propensity towards diversity varies greatly and argue that the diversity of recommendation lists should be consistent with the level of user interest in diverse recommendations. We introduce a user segmentation approach in order to personalize recommendation according to user preference for diversity. We show that recommendations generated using these segments match the diversity preferences of users in each segment. We also discuss the impact of this segmentation on the novelty of recommendations.

Keywords

Recommendation diversity, Performance evaluation metrics, Novelty, Collaborative Filtering

1. INTRODUCTION

Although there are many methods in the literature that can be used to increase diversity in recommendations [1], only a few have mentioned the varying degrees of interest users have for diverse recommendation results [2]. One can imagine two extreme cases of this interest: one user likes to receive as recommendations only science fiction movies made within the last 10 years; another user likes a more diverse set of movies from many genres in her recommendation list. Obviously, any attempt to increase the diversity of recommendation list is likely to generate poor results for the first user with limited interests.

We measure a user's preference for diversity as a function of the diversity of items that the user has rated, and segment the users into groups based on their scores. Recommendations for each group can be generated independently using any one of a variety of standard recommendation techniques. We show that such recommendations have a level of diversity that matches the interest of the segment's users.

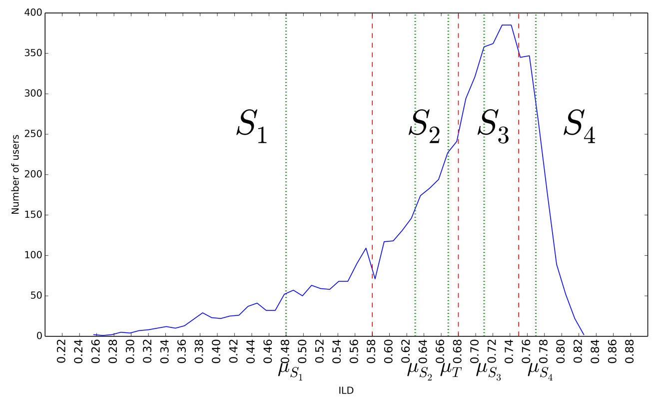

Figure 1: ILD Distribution of User Profiles.

2. DEFINITIONS

Let U and I be the sets of users and items, respectively. The lists of recommendations is denoted as R. R_u is the recommendation items for user $u \in U$ and user profile I_u is the list of items that u has rated. Diversity is the measure of dissimilarity between items in a set. For this purpose, we use average pairwise distance of items in a set as Intra-List Distance (ILD) [4].

$$ILD(L) = \frac{1}{|L|(|L|-1)} \sum_{i \in L} \sum_{j \in L} d(i,j) \tag{1}$$

In addition to diversity, we can measure the impact of user segmentation on the novelty or catalog coverage of recommendation lists. We define novelty as the average distance from the items in user profile to the items in recommendations.

$$Nov(I_u, R_u) = \frac{1}{|R_u||I_u| - min(|R_u|, |I_u|)} \sum_{i \in R_u} \sum_{j \in I_u} d(i,j) \tag{2}$$

We also consider the popularity of items in the recommendation lists. Popularity of an item i is defined by

$$Pop(i) = \frac{|U_i|}{\max_{j \in I}(|U_j|)}$$

where U_i is the set of users who have rated item i.

3. EXPERIMENTS AND DISCUSSIONS

We used MovieLens 1M[1] data set for analysis and evaluation of the proposed method. We create a term vector for each movie using the genre information in the dataset and measure the distance between movies, $d(i, i)$, as the cosine of two genre vectors. After the ILD value for each user has been computed, the next step is to define intervals for segmenting the user profiles. Figure 1 shows the distribution of ILD values across the MovieLens user profiles. The figure shows that there are a relatively small number of users with low ILD values, rising to a peak around 0.74 and falling off rapidly thereafter.

We divided the range of ILD values into four segments, shown graphically in Figure 1 and also in Table 1. The figure shows the boundaries of each segment and the mean ILD, μ_{s_k}, for $k = 1, 2, 3, 4$. Note that segment 3 is larger than the others, which reflects the large number of users with this range of diversity in their profiles.

We generated our re using three recommendation models (two neighborhood based models and one using matrix factorization, BPRMF (Bayesian Personalized Ranking with Matrix Factorization) [3]) using the whole dataset, as well as using each segment separately.

Table 2 shows the results of these experiments in terms of precision and recall, diversity, novelty and popularity. We expected to find that diversity would be increased when for segments with higher preference for diversity and that effect is clearly present in ILD values for all the recommendation algorithms. As we move from segments with low diversity to those with higher diversity, the ILD values of the resulting recommendations are monotonically increasing.

We expected to find that popularity is monotonically decreasing. That is, the segments containing users with diverse profiles would produce recommendations outside of the "short head" of highly popular items and the more diverse the users, the more obscure the recommendations. This effect is not seen. Instead, popularity increases between segments 1 and 2 and decreases afterwards. One explanation for this phenomenon is that segment 1 users are actually niche users with a strong interest in a single movie genre and as a result, their profiles do not contain many of the typical "blockbuster" films. Outside of segment 1, the expected effect is seen across all remaining segments. We will explore this phenomenon further in future work.

A trade-off between precision and recall is observed in Table 2. As ILD_{S_i} increases, $Precision_{S_i}$ increases and $Recall_{S_i}$ decreases. Increase in ILD_u suggests that a user u is interested in movies from a variety of genres. The number of hits (items in the recommendation list) for this user also increases because more movies are considered relevant recommendations. However, there are more movies in the catalog that can match the user's interests so achieving good recall of just those items that the user rated is more difficult.

Table 2 also shows that the novelty of recommended items increase along with the increase in diversity. So, segmentation based on diversity, not only preserves the user's propensity towards diverse recommendations, but also results in a corresponding change in the the level of recommendation novelty.

User Segment	users	movies	ratings	ILD intervals
Segment 1	1048	3192	62940	[0.0, 0.58)
Segment 2	1517	3320	119525	[0.58, 0.68)
Segment 3	2288	3515	301444	[0.68, 0.75)
Segment 4	1187	3579	313845	[0.75, 1.0]

Table 1: User Segments

Recommender	Segments	Precision	Recall	ILD	Nov	Pop
User-kNN	All users	0.244	0.199	0.527	0.685	0.0223
Item-kNN	All users	0.204	0.151	0.469	0.659	0.0047
BPRMF	All users	0.261	0.208	0.482	0.657	0.0216
User-kNN	Segment 1	0.158	0.226	0.332	0.474	0.0221
	Segment 2	0.183	0.209	0.489	0.651	0.0320
	Segment 3	0.245	0.185	0.569	0.734	0.0290
	Segment 4	0.352	0.133	0.611	0.783	0.0219
Item-kNN	Segment 1	0.137	0.191	0.324	0.471	0.0039
	Segment 2	0.164	0.180	0.457	0.636	0.0063
	Segment 3	0.208	0.142	0.545	0.722	0.0059
	Segment 4	0.294	0.099	0.608	0.779	0.0044
BPRMF	Segment 1	0.160	0.233	0.313	0.456	0.0165
	Segment 2	0.190	0.221	0.480	0.637	0.0279
	Segment 3	0.264	0.198	0.563	0.728	0.0265
	Segment 4	0.364	0.137	0.605	0.780	0.0174

Table 2: Recommendation Results

4. CONCLUSIONS

This work examines the consequences of segmenting user populations by diversity, as a means of personalizing user interest in and tolerance for diversity. We show that interest in diversity varies widely across users, with a distinct peak and users with preferences both low and high.

Our division of the user population into four segments is a simple but effective method for increasing diversity for those segments of the population interested in such diversity and decreasing it for those with less interest. The expected effects on diversity and novelty are seen across three different recommendation algorithms.

We plan to explore these effects in future work in additional datasets and algorithms, as well as alternate methods for personalizing diversity.

[1] http://grouplens.org/datasets/movielens/

5. REFERENCES

[1] G. Adomavicius and Y. Kwon. Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering*, 24(5):896–911, 2012.

[2] K. Kapoor, V. Kumar, L. Terveen, J. A. Konstan, and P. Schrater. I like to explore sometimes: Adapting to dynamic user novelty preferences. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 19–26. ACM, 2015.

[3] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*, pages 452–461. AUAI Press, 2009.

[4] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, pages 22–32. ACM, 2005.