Diversified Utility Maximization for Recommendations

Azin Ashkan *, Branislav Kveton *, Shlomo Berkovsky **, Zheng Wen *
*Technicolor, United States
{azin.ashkan,branislav.kveton,zheng.wen}@technicolor.com
**CSIRO, Australia
shlomo.berkovsky@csiro.au

ABSTRACT

Consider the problem of recommending items to a group of users subject to the diversity of their tastes. The goal is to recommend a list of items, such that the interests of each user are covered. We cast this problem as maximizing a diversified utility function of the group, the optimal solution of which can be found greedily. We conduct a user study in order to evaluate the performance of the proposed method. Evaluation results show that our method represents an effective strategy compared to various settings in which a convex combination of utility and diversity is maximized.

1. INTRODUCTION

It is often the case that multiple users share a single account with an online recommender, e.g., family members may share the same account with a movie recommendation service on their TV. The use of a single account by multiple users poses a challenge in providing recommendations satisfying the spectrum of their tastes. One solution to this problem is to recommend a list of diverse movies, in order to cover a range of tastes and increase user satisfaction. However, there exists a tradeoff between increasing the list diversity and maintaining the utility of the results [4].

Let E be the ground set of L recommendable items and $w(e)$ be the utility of item $e \in E$. The goal is to recommend a subset of items $S \subseteq E$ with the highest utility and diversity for users. A common approach to diversified ranking is based on the notion of *maximal marginal relevance (*`MMR`*)* [1]. In this approach, utility and diversity are represented by two independent metrics, and marginal relevance is a convex combination of these metrics:

$$S_{\texttt{MMR}} = \underset{S \subseteq E : |S| = K}{\arg\max} \ (1 - \lambda) w(S) + \lambda f(S), \qquad (1)$$

where K is the cardinality of S, $\lambda \in [0, 1]$ is a parameter that balances the importance of utility and diversity, and $w(S)$ is the sum of utilities of all items in S. The utility w is a modular function of S, whereas the diversity f is typically a submodular function of S. Under these assumptions, the objective in Equation 1 is submodular in S. Therefore, a $(1 - 1/e)$-approximation to the optimal solution can be computed greedily [3].

In this paper, we consider a different objective function, the optimal solution of which can be found greedily. Our objective is to maximize the utility of recommending a list of items to a group of users subject to the diversity of their tastes. Items with high utility are expected to be included in the list as long as they have a contribution to the diversity of the list. The utility remains the primary concern, but it is subjected to maintaining the diversity.

2. DIVERSIFIED UTILITY MAXIMIZATION

The main idea of our approach is to maximize utility weighted by diversity and cover each increase in diversity by the item with the highest possible utility. The increase in diversity can be viewed as the probability that a user chooses the item, in the sense that items that are similar to the previously recommended items are less likely to be chosen. When the item is chosen, we would like to maximize the satisfaction of the user, i.e., the utility of the choice. Formally, our optimization problem is given by:

$$A^* = \arg\max_A \sum_{k=1}^{L} g_A(a_k) w(a_k), \qquad (2)$$

where $A = (a_1, \dots, a_L)$ is an ordered set of L items that we also call a *list*, and $g_A \in (\mathbb{R}^+)^L$ is a vector of gains in diversity, where:

$$g_A(a_k) = f(A_k) - f(A_{k-1}) \qquad (3)$$

is the gain associated with choosing item a_k after choosing items in A_{k-1}. The sets A_k and A_{k-1} are the first k and $k - 1$ items in list A, respectively. We refer to our approach as *diversified utility maximization* (`DUM`), since our objective is to maximize the utility weighted by the increases in diversity.

For a general function f, the problem in Equation 2 is NP-hard. However, when f is submodular, the problem can be solved optimally by a greedy algorithm [2]. The items are ordered in decreasing order of utility, $A^* = (a_1^*, \dots, a_L^*)$, where $w(a_1^*) \geq \dots \geq w(a_L^*)$, and they are added to the recommended list in this order. When $g_{A^*}(a_k^*) > 0$, item a_k^* is added to the list. Otherwise, the item is not added because it does not contribute to the diversity of the list.

The above solution is meaningful when the length of the recommended list, i.e., the number of non-zero entries in g_{A^*}, can be controlled. This is possible for a range of submodular functions. One such function is:

$$f(S) = \sum_{t \in \mathcal{T}} \min \left\{ \sum_{e \in S} \mathbb{1}\{\text{item } e \text{ covers topic } t\}, N_t \right\}. \qquad (4)$$

Here, $\mathcal{T}$ is a set of topics and N_t is an integer threshold for a topic t. For this f, the recommendation list is guaranteed to contain at most $\sum_{t \in \mathcal{T}} N_t$ items, such that each topic t is covered by at least N_t most relevant items in this topic.

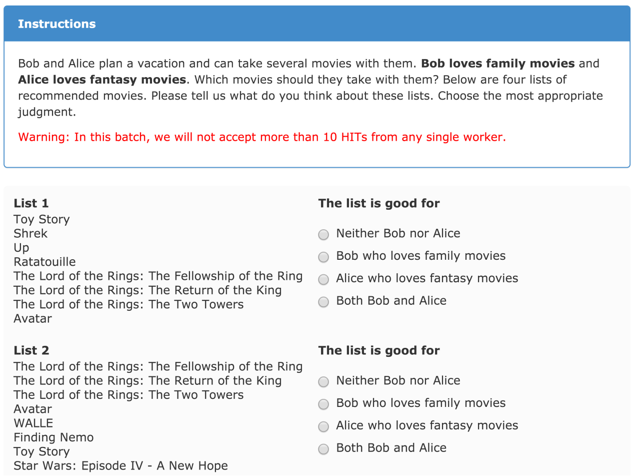
3. USER STUDY

We conduct a user study in the movie recommendation domain. The ground set E are $1,000$ most popular movies on IMDb[1]. The utility $w(e)$ of a movie e is approximated by its overall popularity, i.e., the number of people who rated e.

[1]http://www.imdb.com – The Internet Movie Database.

Table 1: An example scenario where DUM outperforms MMR.

DUM		MMR ($\lambda = 2/3$)		MMR ($\lambda = 1/3$)		MMR ($\lambda = 0.01$)	
The Dark Knight	action	Zombieland	horror action	The Dark Knight	action	The Dark Knight	action
LOTR 1	action	From Dusk Till Dawn	horror action	LOTR 1	action	LOTR 1	action
The Matrix	action	Dawn of the Dead	horror action	The Matrix	action	The Matrix	action
Inception	action	Resident Evil	horror action	Inception	action	Inception	action
The Shining	horror	The Dark Knight	action	LOTR 2	action	LOTR 2	action
Alien	horror	LOTR 1	action	The Dark Knight Rises	action	The Dark Knight Rises	action
Psycho	horror	The Matrix	action	LOTR 3	action	LOTR 3	action
Shaun of the Dead	horror	Inception	action	The Shining	horror	Avatar	action


Instructions

Bob and Alice plan a vacation and can take several movies with them. **Bob loves family movies** and **Alice loves fantasy movies**. Which movies should they take with them? Below are four lists of recommended movies. Please tell us what do you think about these lists. Choose the most appropriate judgment.

Warning: In this batch, we will not accept more than 10 HITs from any single worker.

List 1
Toy Story
Shrek
Up
Ratatouille
The Lord of the Rings: The Fellowship of the Ring
The Lord of the Rings: The Return of the King
The Lord of the Rings: The Two Towers
Avatar

The list is good for
Neither Bob nor Alice
Bob who loves family movies
Alice who loves fantasy movies
Both Bob and Alice

List 2
The Lord of the Rings: The Fellowship of the Ring
The Lord of the Rings: The Return of the King
The Lord of the Rings: The Two Towers
Avatar
WALLE
Finding Nemo
Toy Story
Star Wars: Episode IV - A New Hope

The list is good for
Neither Bob nor Alice
Bob who loves family movies
Alice who loves fantasy movies
Both Bob and Alice

Figure 1: An example questionnaire from the user study. We show only two lists out of four.

Our study consists of a set of tasks. In each task, we ask a Mechanical Turk[2] worker to consider a situation, where Bob and Alice go for a vacation and can take several movies with them. Bob and Alice prefer two different movie genres. The workers are asked to rate four lists of movies, based on how these movies are appropriate for Bob and Alice. An example questionnaire is shown in Figure 1.

One list is generated by DUM. The other three lists are generated by MMR, with $\lambda \in \left\{\frac{2}{3}, \frac{1}{3}, 0.01\right\}$. The length of all four lists is identical, and they are shown in random order to avoid position bias. Each task has different movie genres t_1 and t_2 preferred by Bob and Alice. We generate one task for each pair of 18 most popular genres on IMDb and assign each task to three workers. Thus, we obtain $3 \times 153 = 459$ ratings for each of the four lists.

In each task, the diversity function f is defined as in Equation 4. The topics are $\mathcal{T} = \{t_1, t_2\}$ and $N_{t_1} = N_{t_2} = 4$. For this setting, DUM recommends between four to eight movies such that each genre is covered by at least four movies. The functions w and f are normalized such that the maximum gain in each function is 1.

On average, workers spent 57.39 seconds on each task, i.e., 14.35 seconds for a list of at most 8 movies, which is reasonable to judge whether the list covers two genres. The results of the study are reported in Figure 2. The workers considered the lists generated by DUM to be suitable for both Bob and Alice in 74.5% of cases. This ratio is significantly higher than those of MMR with $\lambda = \frac{2}{3}$ and $\lambda = \frac{1}{3}$. The absolute improvement with respect to the best baseline, MMR with $\lambda = \frac{2}{3}$, is 9.6%, while the relative improvement is 14.8%. It is worth to note that in 70 combinations of t_1 and t_2 (45.8% of cases), all three workers unanimously rated DUM as appropriate for both Bob and Alice.

Since DUM did not dominate the MMR baselines across the board, it is important to identify cases, where it performs well. One such

[2]http://www.mturk.com – Amazon Mechanical Turk.

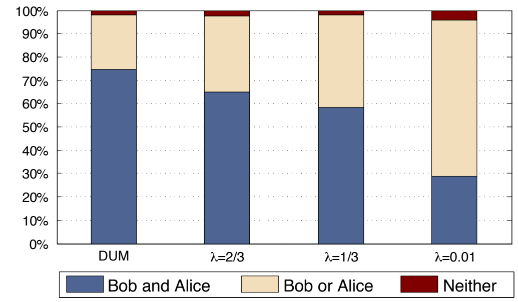


Figure 2: The percentage of times for each method when the recommendation results were suitable for: both Bob and Alice, only Bob or only Alice, and neither Bob nor Alice.

example, for *action* and *horror* movies, is shown in Table 1. MMR cannot solve this problem well for the following reasons. When λ is large, MMR first chooses most diverse movies that are *both* action and horror movies. These movies are less popular than horror movies that are *not* action. As a result, the list is only considered as a good representation of action movies, but not of horror movies. On the other hand, when λ is small, MMR tends to choose mostly action movies, because these movies happen to be more popular than horror movies. So the list is again a good representation of action movies only.

4. CONCLUSION

We propose a method to maximize the utility of recommended items subject to the diversity of users' tastes. This method guarantees that movies in the recommendation list cover various aspects of user tastes with high utility items. We conduct a user study showing the effectiveness of our method compared to models that maximize a convex combination of utility and diversity. In the future, we plan to apply DUM to other domains, such as document summarization and Web search result diversification. We also plan to extend our study to groups of more than two users and evaluate the performance of DUM for various combinations of genres.

5. REFERENCES

[1] J. Carbonell and J. Goldstein. The use of MMR, diversity based reranking for reordering documents and producing summaries. In *SIGIR*, pp. 335–336, 1998.

[2] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Optimization*, pp. 11–26, 2001.

[3] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming*, 14(1):265–294, 1978.

[4] T. Zhou, Z. Kuscsik, J. Liu, M. Medo, J. R. Wakeling, and Y. Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *National Academy of Sciences*, 107(10):4511–4515, 2010.