Fairness Aware Recommendations on Behance

Natwar Modani[1], Deepali Jain[1], Ujjawal Soni[2], Gaurav Kumar Gupta[3], and Palak Agarwal[4]

[1] BigData Experience Lab, Adobe Research {nmodani, deepjain}@adobe.com
[2] IIT Madras ujjawals@cse.iitm.ac.in
[3] IIT Roorkee gup22uec@iitr.ac.in
[4] IIT Kanpur palakag@iitk.ac.in

Abstract. Traditionally, recommender systems strive to maximize the user acceptance of the recommendations, while more recently, diversity and serendipity have also been addressed. In two-sided platforms, the users can have two personas, consumers who would like relevant and diverse recommendations, and creators who would like to receive exposure for their creations. If the new creators do not get adequate exposure, they tend to leave the platform, and consequently, less content is generated, resulting in lower consumer satisfaction. We propose a re-ranking strategy that can be applied to the scored recommendation lists to improve exposure distribution across the creators (thereby improving the *fairness*), without unduly affecting the relevance of recommendations provided to the consumers. We also propose a different notion of diversity, which we call representative diversity, as opposed to dissimilarity based diversity, that captures level of interest of the consumer in different categories. We show that our method results in recommendations that have much higher level of fairness and representative diversity compared to the state-of-art recommendation strategies, without compromising the relevance score too much. Interestingly, higher diversity and fairness leads to increased user acceptance rate of the recommendations.

1 Introduction

The typical objective of the recommender systems is to maximize the user acceptance of the recommendations, treating the acceptance of recommendation as a proxy to maximizing the utility from the consumers' point-of-view. Hence, the focus in all recommender systems has been to improve the prediction accuracy.

In a two-sided creative content discovery platform, e.g., Behance [1], the users can have two personas; consumers that consume the items, and creators who produce/supply the items. Such a platform needs to satisfy both the personas in order to be successful. Consumers satisfaction with the recommendations is based on, and can be measured by the traditional metrics (relevance of the recommendations, the level of diversity and chance of serendipitous discovery). On the other hand, the creators look for opportunities to reach out to a wide set of audience in order to be noticed and appreciated for their creations. If the creators (especially the new ones) do not get adequate exposure, they tend to leave the platform (or become inactive), and consequently, less content is generated on the platform, resulting in lower consumer satisfaction. Hence, for the two-sided platforms, while the relevance of the recommendations to the consumers remain a high priority, providing adequate exposures to the creators also plays an important role in creating a thriving community. The current state-of-art collaborative filtering techniques have been shown to favour popular items [5], thereby increasing the chances of new creators not receiving adequate exposure.

Diversity in the recommendations is recognized as an important consideration. The current notion of diversity is based on (dis)similarity of items, and hence, a uniform strategy is adopted for all consumers to introduce diversity in the recommendation results. Such strategies do not recognize the fact that different consumers have different level of interest in different categories of items. We propose a new notion of diversity, which we call 'representative diversity' that captures level of interest of the consumer in different categories.

In this paper, we focus on Behance [1], a creative content discovery platform. We propose a re-ranking strategy that can be applied to the scored recommendation lists to improve exposure distribution across the creators, without unduly affecting the relevance of recommendations provided to the consumers, and provides representatively diverse results. We define 'Creative Capital' as a notion of value of the creators, based on their contributions to the platform, measured in terms of number of projects created, number of views and appreciations received on their projects along with the recency of such events. 'Desired Exposure' is the ideal amount of exposure to be given to the creator based on the creative capital, and is defined as a sublinear function of contribution of the creators to address the fairness requirement. Fairness is defined as inverse of Jensen-Shannon Divergence (JS-Divergence) between the desired distribution and the actual obtained distribution of the exposures for the creators. Similarly, representative diversity is defined as inverse of JS-Divergence between the desired distribution and the actual obtained distribution of the exposures for the categories. We show that our method results in recommendation lists that have much higher level of fairness and representative diversity compliance compared to the state-of-art recommendation strategies, while the relevance score is not compromised too much. In fact, our experimental results on real data show that improvement in fairness and diversity tends to increase the user acceptance rate of the recommendations (which is the most relevant metric), even though the cumulative relevance score as assigned by the recommender systems is marginally lower.

2 Related Work

Over the years, many different recommendation techniques have been developed, mainly categorized into three types:

1. **Content-Based Filtering:** In these type of recommender systems, items (projects in case of Behance) with similar features to the ones already liked by the consumer are recommended [9]. For creative images of Behance, Fang et al., 2015 [6] have proposed a feature learning paradigm to learn image

similarities. Content-based filtering techniques are fair for all creators, i.e., projects of established popular creators as well as less popular (or upcoming creators) have equal chance of being recommended. But these techniques are limited to recommend items similar to those already liked by the consumer, hence less diverse and serendipitous recommendations.

2. **Collaborative Filtering:** These recommender systems predict relevant items to be recommended to a consumer using the history of items liked by other consumers. There is vast literature in Collaborative filtering (CF), including Item-based CF [10], user-based (nearest neighbors) CF [13], Matrix Factorization [8,7], and other techniques. CF techniques solves the problem of diverse and serendipitous recommendations to some extent. Though CF techniques perform better than content-based filtering, they tend to favor popular projects [5]. Since recommendations provide exposure to projects, this in-turn increases the likelihood of those projects being appreciated. This creates a clear rich-getting-richer scenario.
3. **Hybrid Recommender systems:** To improve performance of recommender systems, content-based and collaborative filtering techniques are sometimes combined in the form of Hybrid Recommender Systems [3]. These methods deal with the cold start problem better than collaborative filtering by recommending new items through content filtering. A major limitation of these systems is the requirement of rich content and meta-data of the items. Moreover, these systems tend to be computationally more complex than either of the two approaches and hence, less scalable.

Diversity [2,12,4] has also been considered in some research, but they focus on diversifying the recommendations and do not consider consumer's diversity preferences. To the best of our knowledge, fairness for the creators on a two-sided platform is not studied as yet. Our method of ensuring fairness resembles the idea of the lottery scheduling method in CPU time allocation [11].

3 Definitions

As we will work in the context of 'Behance' as the application domain, we will start by discussing it briefly. Behance [1] is a creative content discovery platform. Users of Behance can have two personas; creators, who create 'projects' and publish them on Behance, and consumers, who view projects created by the creators. The projects can have one or more of the 137 creative fields associated with them, which can be thought of as categories on Behance. Every click of a consumer on a project to open it is counted towards number of views on the project. The consumers can also appreciate projects, which is another metric associated with the project. Consumers can also follow their favorite creators. We will now define the various notions we will use in the rest of paper.

Positional Value: Since the recommendations are ranked lists, and items at the lower ranking are less likely to receive attention of the consumers, we associate a *positional value* with each rank in the recommendation list. We take the positional value for rank 1 as 1 unit, and determine the positional value in relative terms by observing the relative click-through rates. Due to lack of space, we will not present detailed results, but we observed a near exponential decrease in the click-through rate for the items in various positions. Accordingly, we take the positional value of rank k based on the best fit to the data as:

$$pv(k) = e^{-\frac{k-1}{45}} \tag{1}$$

Creative Capital: Now we define 'Creative Capital' for the Behance creators, which is a measure of their contribution to the platform factoring in the recency of contributions. One can imagine that the creators who create more projects contribute more to the platform. However, higher quality projects should carry more weight. The quality of the projects can be estimated by popularity of the projects, which can be captured in term of the number of views and appreciations. Further, since our focus here is on recommendations, projects that are recent, or have received views and/or appreciation recently should carry more weight than projects that are old, and have not received much user attention recently. Accordingly, we define 'Creative Capital' $C_u(t)$ as follows:

$$C_u(t) \; = \; \gamma \times C_u(t-1) + \omega_p \times \Delta n_p(t) + \omega_a \times \Delta n_a(t) + \omega_v \times \Delta n_v(t) \tag{2}$$

A creator u *earns* creative capital by creating projects or receiving views and appreciations for projects created by her. The creative capital at the previous time step $C_u(t-1)$ is decayed with by a factor γ and carried over. Here, ω_p, ω_a and ω_v are the weights of each project creation, appreciation and view respectively.Also, $\Delta n_p(t)$ is the number of projects created by this creator between $(t-1)$ and t. Similarly, $\Delta n_a(t)$ and $\Delta n_v(t)$ are the number of appreciations and views received on his projects from $(t-1)$ to t, respectively.

Desired Exposure Distribution: We had noted that due to favoring popular items, collaborative filtering techniques tend to create rich-gets-richer scenario. To avoid this situation (which is key to ensure fairness), we allocate the exposures to the creators based on a sub-linear function of their creative capital. Please note that we want the creators who contribute more to receive more exposures to maintain incentive compatibility (i.e., there should always be incentive to produce more of high quality work, assuming that having more exposure is the incentive), and hence, the exposures should be a monotonic function of the creative capital. Hence, we define the deserved exposure for a creator u as:

$$E_u = \theta \times {C_u}^{\alpha}, 0 < \alpha < 1 \tag{3}$$

where θ is a normalization factor such that $\sum E_u = 1$ for all users u.

Fairness: Let the amount of exposures provided to the creations of a creator u be denoted by A_u, and the desired exposure distribution for the creator be E_u. Then, we can think of fractional exposure provided to creators (by normalizing across all creators) and exposure distributions as probability distributions over the creators. We define a fairness of a recommender system as inverse of JS-Divergence between these two distributions. Low value of JS-Divergence means that the actual exposure distribution is close to the desired exposure distribution, and hence the system is fair (so the fairness score is high), and a high JS-Divergence implies that the actual exposure distribution is significantly different than the desired exposure distribution, and hence the system is not fair.

$$F = \frac{1}{JSD(E||A)} \tag{4}$$

where, $JSD(E||A)$ is Jensen-Shannon divergence between two probability distributions E and A.

Representative Diversity: Different consumers on a platform have different appetite for different categories of items. We allocate the exposure to be given to the items from a category g for a consumer based on their (normalized) interest in that category. One challenge in such a strategy is that the user may not have explored the items of the platform enough for us to learn her preferences completely. Hence, we keep the exposure allocation for the category as a weighted average of the consumer's preference for the category and global preference of the category. The weight is based on the number of observations available for the consumer. As we gather more and more data about the consumer's preference, the global preference's weight keeps decreasing.

$$E_g(u) = \beta \times (\lambda_u p_g^u + (1 - \lambda_u) G_g) \tag{5}$$

where $E_g(u)$ is the exposure fraction allocated to category g for consumer u, $0 \leq \lambda_u \leq 1$ is the degree of certainty about estimate of consumer u's preferences, p_g^u is the estimated preference of consumer u for category g, and G_g is the global preference for category g. Also, β is a normalizing factor to ensure that $\sum_g E_g(u) = 1$. Clearly, λ_u is a function of amount of data available about consumer u's preferences.

We define the diversity compliance of the recommender system for a consumer as inverse of JS-Divergence of the desired exposure distribution for the categories E^c and the actual exposure distribution A^c for that consumer.

$$DC(u) = \frac{1}{JSD(E^c(u)||A^c(u))} \tag{6}$$

The global diversity compliance is defined as

$$GDC = \sum_u \{W(u) \times DC(u)\} / \sum_u W(u) \tag{7}$$

where $W(u)$ is the importance of consumer u, which we take as the sum of positional value of all exposures provided to the user u.

A note about simplification: In Behance, a project can be created by collaboration amongst multiple creators. Also, the project can have multiple categories associated with it. In the above description, we have given all formula considering only the case where each project is created by one creator and is associated with one category. This is done for ease of reading. While implementing our system, we have assigned partial credit to the creators and categories for such projects. Our experimental results are given for partial credit assignments.

3.1 The Final Objective Function

Recall that our aim is to provide "*relevant* and *representatively diverse* recommendations to the consumers, that provide *fair exposure* to the creators". Hence, we define our overall objective function as a combination of the user relevance, fairness to creators and representative diversity across categories. Suppose the relevance of an item i for a consumer u is given by r_{ui}, which may be based on the underlying recommendation algorithm (e.g., Collaborative Filtering). We define the overall relevance R_u for the user u as $R_u = \sum_k pv(k) \times r_{ui}$, where $pv(k)$ is the positional value of rank k, and r_{ui} is the relevance of the item i, which is recommended in position k in the recommendation list. The final relevance score for the recommender system across all users R_{all} is given as

$$R_{all} = \sum_u W(u) \times R_u \tag{8}$$

where $W(u)$ is the importance of consumer u as in Equation (7), which we again take as the sum of positional value of all exposures provided to the user u.

Finally, we are ready to define our overall objective function:

$$O = (w_1 + R_{all})^{w_r} \times (w_2 + F)^{w_f} \times (w_3 + GDC)^{w_d} \tag{9}$$

This form of objective function ensures that none of the factors can be ignored completely. The various weights (w_1, w_r, w_2, w_f, w_3, w_d) control the importance of the different factors. We would like to give higher importance to relevance and fairness compared to the diversity, and hence we select $w_1 = 0, w_r = 1, w_2 = 0, w_f = 1$, and $w_3 = 1, w_d = 1$. This results in simplified objective function

$$O = R_{all} \times F \times (1 + GDC) \tag{10}$$

Given that, we would not know a-priori which consumers are likely to visit the platform on a given day, we would like to make the recommendations in such a way, that the solution has a high value of objective function on an ongoing basis, and not only at the end of one round of execution. In the next section, we give a heuristic approach for ongoing optimization of this objective function, as due to JS-Divergence in the objective function for our problem formulation, it is not possible to devise an efficient exact or approximation algorithm.

4 Algorithm for Generating Recommendations

We will first outline an optimization approach for a general resource allocation problem and then illustrate how to translate it to the present context of re-ranking recommendations.

Consider a set of resource requesters, along with a prespecified share of resource eligibility for each requester. The resource become available in chunks in an online fashion. When a resource chunk becomes available, it needs to be allocated to one requester (without dividing it). The goal is to allocate resource chunks in such a fashion, that at every time, the resource distribution over all requesters is as close to the prespecified resource eligibility share as possible.

We propose the following greedy algorithm to solve the given problem. For every resource chunk $r(t)$, calculate the value of allocating the resource to each requester u as

$$V_u = E_u \times \frac{(\sum_v A_v(t-1) + r(t))}{(A_u(t-1) + r(t))} \tag{11}$$

where E_u is the pre-specified share of resource eligibility for requester u, $A_u(t-1)$ is the already allocated resource units to requester u until time $(t-1)$. Now, there are two strategies possible. First is a deterministic strategy, where we allocate $r(t)$ to the requester such that the value is the highest. Second is a probabilistic strategy, where we allocate the resource to the requesters with probability equal to the normalized value.

One can see both the fairness and representative diversity as resource allocation problem described above. Our overall objective function is a combination of three components. Hence, we use this method to generate two of the factors which we use for the re-ranking strategy, while the third component is based on the relevance as assigned by the underlying recommendation algorithm.

First, we generate a rating or relevance scores $r_{u,i}$ using state-of-art collaborative filtering techniques for all project-consumer pairs. We also compute the global popularity ratings g_i for all projects as the average of all observed rating for the project. We then follow the following steps for recommending projects to each consumer u, for whom, we need to generate k_u recommendations:

1. Create a candidate pool of projects to recommend by taking all the projects for which the rating is positive (i.e., $r_{u,i} > 0$).
2. If the pool is smaller than the number of projects to be recommended, add all the other projects (ones with $r_{u,i} = 0$) to the pool.
3. Then calculate goodness of all the projects in pool as follows:

   $$G_{u,i} = r_{u,i} \times V_F(c(i)) \times V_D(g(i)) \tag{12}$$

   where, $V_F(c(i))$ is value of allocating the exposure to the creator of project i (refer Equation 11), $V_D(g(i))$ is the value of allocating the exposure to the category that project i belongs to, and $r_{u,i}$ is the relevance rating of project i to the user u as mentioned earlier.
4. Now select project with maximum goodness (we will call this as deterministic approach) or select a project probabilistically from the list with probability of selection equal to its normalized goodness (we will call this as probabilistic approach).
5. Remove the selected project from the list and continue recommending from remaining projects until all recommendations are done.

While computing the relevance, the rating of projects with $r_{ui} = 0$ is taken as $1/5$ of the lowest value of r_{ui} from the top-k projects.

5 Experimental Results

First, we will validate our motivation by analyzing the churn rate of the creators to show that creators who do not get adequate views in the beginning tend to churn with higher probability. We will show that the Creative Capital is a good metric to capture the contribution of the creators for the platform. Both these analysis are on the full Behance dataset. We will then describe our data set for recommendation re-ranking. We study the performance of various state-of-art collaborative filtering techniques to choose the baseline relevance assignment approach. We will compare the performance of the proposed approaches and baseline approaches on the three axis, fairness, relevance, and diversity. Finally, we will also evaluate the various approaches for the precision and recall based performance. Given the space limitation, we will not present detailed results and plot in all cases, and only quote the results in the running text.

5.1 Churn Rate Analysis

To illustrate the need to address the fairness, we have done an analysis of the churn rate of creators on Behance. A creator is said to have churned if he stops publishing any new projects. We calculated the number of creators who got only a small number of views/appreciations in their initial 12 months, and computed the churn rate as the fraction of creators who stopped publishing projects after this initial period. We found that the churn rate for creators who get up to 5 views during the initial 12 months is approximately 2.5 times than the creators who got at least 100 views in the first 12 months. However, the churn rate does not change significantly for the creators that received at least 100 views. If we assume that the relation between views received and churn rate remains the same, then the re-ranking strategy proposed in this paper that marginally reduces number of exposures for highly popular creators and distributes those among less popular creators for fairer exposure, results in 12% reduction in churn rate. This experiment clearly highlights the importance of giving fair opportunities to creators for their projects to be viewed to reduce churn-rate.

5.2 Creative Capital Analysis

As explained in Section 3 Equation (2), we computed the 'Creative Capital' as a function of number of projects created and number of views and appreciations received, along with recency of such signal. We used the following parameter values: $\gamma = 0.98$, $\omega_p = 50$, $\omega_a = 5$ and $\omega_v = 1$. These weights are inversely proportional to the relative frequency of occurrence of respective events in order to give equal importance to each of these. The intent of this metric was to capture the perceived contribution of the creators to the platform. Typically, people tend to follow the creators based on their contribution. As we did not use the follower information for defining this metric, we can use it for cross validating the metric. If the metric is indeed a good indicator of creator's contribution, the increase in number of followers should coincide with the increase in creative capital.

Accordingly, we calculated the Pearson Correlation Coefficient between increase in creative capital, $C_u(t) - C_u(t-1)$ and corresponding $\Delta n_f(t)$ (increase in the number of followers of u from $t-1$ to t). The average correlation was observed to be 0.7457 which establishes the validity of Creative Capital as a measure of worthiness of a creator.

5.3 Data Set

Behance has an active user base of multiple million users, with about one quarter of the users being creators. The number of projects created by these creators is also in millions. To evaluate the recommendation performance, we work with a sample of data that has 638 creators, having $2,000$ projects, and $1,400$ consumers. The total number of project views and appreciations were $28,000$ and $9,800$, respectively. We split the data such that approximately 80% views and appreciations go into train and 20% in test sets.

5.4 Baseline

As collaborative filtering techniques have been shown to outperform other recommendation approaches, we take collaborative filtering techniques as the baseline for comparison. Since there are many collaborative filtering techniques proposed in literature, we first conducted experiments to determine which of these techniques perform the best for our dataset. We implemented nearest neighbour, item-item and matrix factorization based collaborative techniques, and checked for accuracy of the recommendations provided. We found that item-item jaccard nearest neighbor based CF algorithm performed the best with approximately 5% better accuracy in top-k recommendations for a broad range of k. Hence, we take item-item CF as our main baseline and call it 'Traditional' baseline. We also take randomized strategy as baseline called 'Baseline Random', as randomness would likely result in high degree of fairness and diversity. To ensure that the recommendations are not completely irrelevant, we also created hybrid baselines called 'Baseline Hybrid', where first 50% of the recommendations are the ones with the highest predicted ratings and the rest are chosen randomly.

5.5 Fairness, Diversity and Relevance

We now evaluate the performance of our two approaches (probabilistic and deterministic), and the results are compared against traditional CF approach and other baselines.

First, we look into fairness. There are two aspects of fairness; first, the strategy should allocate the exposures to the creators in a manner consistent with the objective of giving fair exposure to all creators. Second, the recommendation algorithm should follow the exposure allocation while performing the recommendations. The left hand side of the plot in Figure 1 shows the number of people (on y-axis) who will be given a certain amount of exposure (on x-axis). Here,

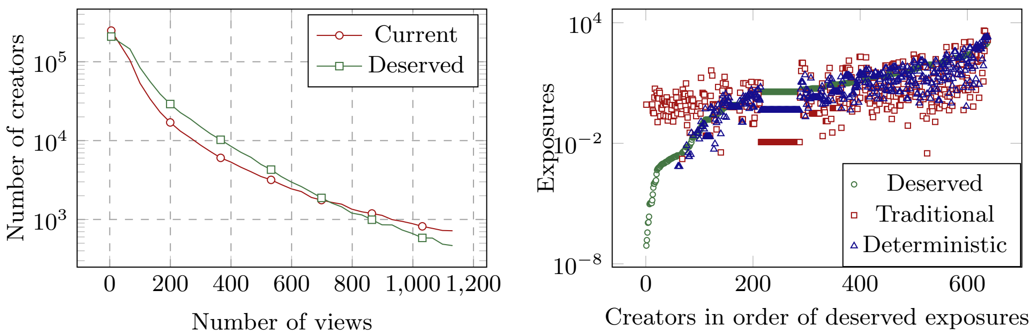

Fig. 1. Fairness results: Left - View Allocation; Right - Actual Exposure Distribution

we have taken $\alpha = 0.75$ for determining the deserved/allocated exposures (ref. Equation (3)). As one can see, the number of people who receive large number of exposures is reduced, and the number of people who get moderate exposures is increased. This shows that our method allocates the exposure in a fairer manner compared to the current state-of-art collaborative filtering techniques. The right hand side plot in the same Figure 1 shows the deserved exposure allocation, the exposure provided by our method, and the exposure provided by the collaborative filtering technique. One can clearly see that our method adheres much more closely to the allocated exposures as compared to the collaborative filtering. The correlation between the deserved and actual exposure provided by our deterministic method is 0.8682, whereas the correlation for the item-item CF with deserved exposure is 0.6573. This clearly shows that our method has good intent (left plot) and good execution (right plot) for fairness to creators. Table 1 reports the fairness numbers achieved by various methods, which clearly shows that our proposed approaches achieve nearly twice as good fairness compared to traditional and randomzied baselines.

Figure 2 compares the diversity in the categories of the projects recommended and the relevance for the consumers. The figure on the left shows that our models (especially probabilistic without beyond k) perform better than the traditional approach. The randomized baseline approaches are expected to perform well because picking random projects would lead to increase in diversity. The figure on the right shows that while our models perform very well on fairness and diversity fronts, as expected it lags behind in terms of relevance, as the improvement in fairness has been achieved at the cost of drop in relevance. However, we find that the average loss in relevance was about 9% only, whereas the average

Table 1. Fairness value achieved by various methods

Method	Baseline Random	Baseline Hybrid	Baseline IICF	Deterministic	Probabilistic
Fairness	4.11	3.85	2.97	6.56	6.03

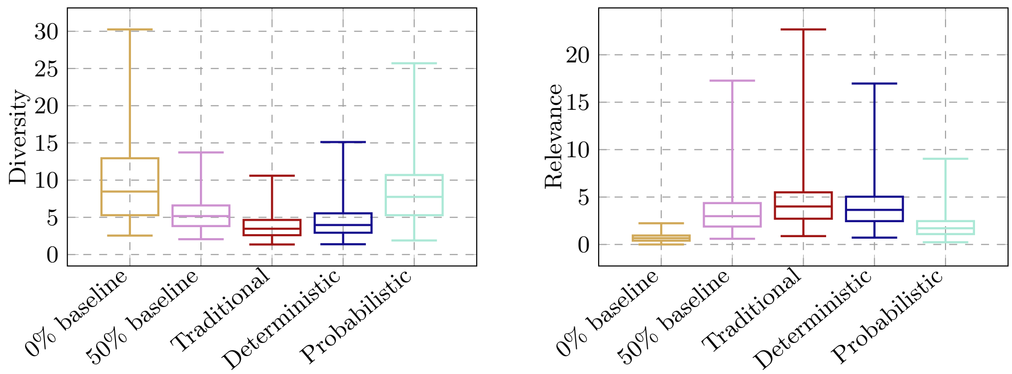

Fig. 2. Diversity and Relevance distribution

improvement in fairness was 97.1%, over and above the considerable increase in diversity. We also see that the random (and hybrid) baseline performs poorly on the relevance front even though it performed fairly well in terms of fairness. This means that randomized approach are not viable alternatives.

Finally, Figure 3 compares the precision and recall of the results for all the approaches. where precision and recall at cutoff k are defined as:

$$P(k) = |a \cap p_k|/k \qquad R(k) = |a \cap p_k|/|a|$$

where a is the set of projects that the consumer has appreciated and p_k is the set of top k projects recommended to the consumer. As we can see our models have higher precision and recall than the baseline models, including even the best performing CF technique. The deterministic approaches perform the best in general. The high precision and recall for the traditional method is expected.

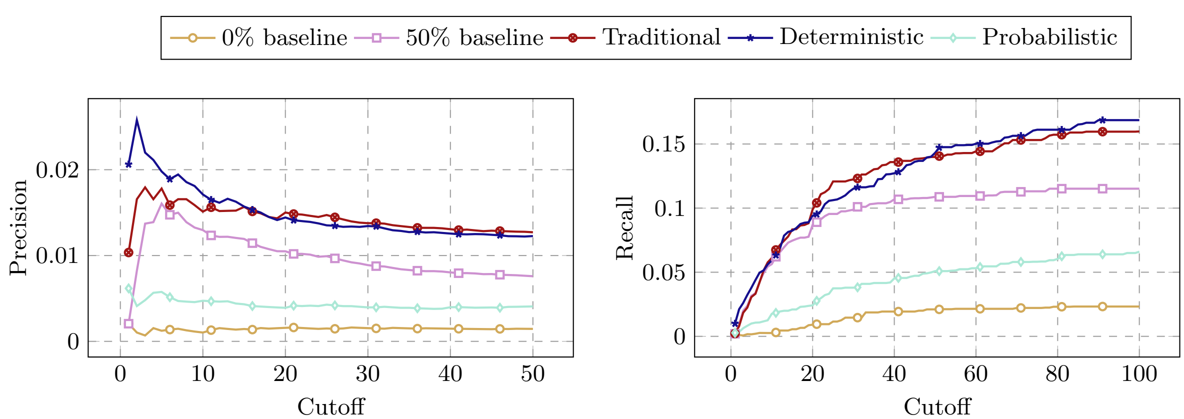

Fig. 3. Precision and recall results for different methods

6 Conclusions

In this paper, we addressed an important issue of fairness to the creators while providing relevant and diverse recommendations to the consumers on a two-sided platform. We showed that by sacrificing a small amount of relevance, one can achieve a much higher degree of fairness and diversity in the recommendations. Further, we also showed that in terms of the precision and recall, which are the most relevant metrics, our proposed approach outperforms the state-of-art collaborative filtering techniques. There are some interesting research directions as a follow up of this work, including more robust definition of Creative Capital and approximation gurantee algorithms.

References

1. http://www.behance.net.
2. G. Adomavicius and Y. Kwon. Toward more diverse recommendations: Item reranking methods for recommender systems. In *Workshop on Information Technologies and Systems*, 2009.
3. R. Burke. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370, 2002.
4. J. Carbonell and J. Goldstein. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 335–336. ACM, 1998.
5. Ò. Celma and P. Cano. From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proc. of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*, page 5. ACM, 2008.
6. C. Fang, H. Jin, J. Yang, and Z. Lin. Collaborative feature learning from social media. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 577–585, 2015.
7. Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 426–434. ACM, 2008.
8. Y. Koren, R. Bell, C. Volinsky, et al. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.
9. M. J. Pazzani and D. Billsus. Content-based recommendation systems. In *The adaptive web*, pages 325–341. Springer, 2007.
10. B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pages 285–295. ACM, 2001.
11. C. A. Waldspurger and W. E. Weihl. Lottery scheduling: Flexible proportional-share resource management. In *Proceedings of the 1st USENIX conference on Operating Systems Design and Implementation*, page 1. USENIX Association, 1994.
12. M. Zhang and N. Hurley. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2008 ACM conference on Recommender systems*, pages 123–130. ACM, 2008.
13. Z.-D. Zhao and M.-S. Shang. User-based collaborative-filtering recommendation algorithms on hadoop. In *Knowledge Discovery and Data Mining, 2010. WKDD'10. Third International Conference on*, pages 478–481. IEEE, 2010.