Relevance Meets Diversity: A User-Centric Framework for Knowledge Exploration Through Recommendations

Erica Coppolillo[1,2,*], Giuseppe Manco[2] and Aristides Gionis[3]

[1]University of Calabria, Via Pietro Bucci, 11, Rende, Italy
[2]ICAR-CNR, Via Pietro Bucci, 8/c, Rende, Italy
[3]KTH Royal Institute of Technology, Brinellvägen 8, 114 28, Stockholm, Sweden

Abstract

Recommender systems aim to provide suggestions that are both *relevant* and *diverse*. Balancing these measures is challenging, as increased diversity often reduces relevance, lowering user engagement. Current algorithms combine relevance and diversity into a single objective, but they typically overlook user interaction with the recommended items. In this paper, we prioritize the *user*, integrating *relevance*, *diversity*, and *user behavior*. Our probabilistic user-behavior model assumes users continue engaging with the system if they find the recommendations relevant but may stop if relevance decreases. Therefore, to achieve high diversity, recommendations must be *both relevant and diverse*. We introduce a novel recommendation strategy using a copula function. Extensive evaluations on multiple datasets demonstrate that our strategy overcomes several state-of-the-art methods. Our implementation is publicly available[1].

Keywords

User Modeling, Diversity, Recommender Systems

1. Introduction

Recommender systems are crucial for helping users discover new information and expand their knowledge base [1, 2, 3]. Even though they generally focus on maximizing relevance, previous research highlighted how incorporating diversity into recommendations adds significant value for knowledge exploration [4, 5]; further preventing users from getting trapped in content "rabbit holes" on platforms like YouTube [6, 7, 8] or Reddit [9]. Current methods balance relevance and diversity by merging them into a single optimization objective but often ignore user behavior and interactions with recommended items.

We propose a novel framework prioritizing the user, where the interaction with the algorithm is seen as a *knowledge-exploration task*, guided by a *user-behavior model* that accounts for user preferences and patience in accepting or rejecting recommendations. Our goal is to maximize the knowledge a user gains, modeled through a *diversity* measure, and couple it with the user-

[1]https://github.com/EricaCoppolillo/EXPLORE

HI-AI@KDD, Human-Interpretable AI Workshop at the KDD 2024, 26th of August 2024, Barcelona, Spain

*Corresponding author.

†These authors contributed equally.

erica.coppolillo@unical.it (E. Coppolillo); giuseppe.manco@icar.cnr.it (G. Manco); argioni@kth.se (A. Gionis)

0000-0002-4670-8157 (E. Coppolillo); 0000-0001-9672-3833 (G. Manco); 0000-0002-5211-112X (A. Gionis)

behavior model, thus producing recommendations that are *both relevant and diverse* . We refer to this concept as *"knowledge exploration via recommendations"*.

In our framework, relevance governs the termination of exploration, while the overall quality is measured by diversity. We instantiate our model using two standard notions of diversity, coverage and pair-wise distances [10, 11, 12]. Finally, we propose a novel recommendation strategy that combines relevance and diversity by a copula function. We perform an extensive evaluation of the proposed framework and strategy using five benchmark datasets publicly available, and show that our strategy outperforms several state-of-the-art competitors.

Notice that the present work is an abbreviated version of the paper accepted to main track of the KDD2024 conference [13]. Please refer to the latter for more details.

The rest of the paper is structured as follows. Section 2 presents our problem definition and methodology. In Section 3, we illustrate our recommendation strategy. Results are reported in Section 4, and finally Section 5 concludes the paper and provides pointers for future extensions.

2. User Model and Problem Formulation

We consider a typical recommendation setting in which we have a set of m users $\mathcal{U}$ and a set of n items $\mathcal{I}$. We also consider a (black-box) function $\mathcal{R} : \mathcal{U} \times \mathcal{I} \to \mathbb{R}$ that provides us with a relevance score $\mathcal{R}(u, i)$, for each user $u \in \mathcal{U}$ and item $i \in \mathcal{I}$.

Item-to-item distance function. Given an item $i \in \mathcal{I}$, we denote by $\mathbf{x}_i$ the vector of *users* with $x_{iu} = 1$ if user u interacted with item i, and 0, otherwise. In addition, we consider a set of categories $\mathcal{C}$, and we define $\mathbf{y}_i$ to be a *category* vector, for item $i \in \mathcal{I}$, where $y_{ic} = 1$ if category c relates to item i, and 0 otherwise. Given two items $i, j \in \mathcal{I}$, we hence define their *distance* as the *weighted Jaccard distance*

$$d(i, j) = 1 - \frac{\sum_{w \in \mathcal{W}} \min\{z_{iw}, z_{jw}\}}{\sum_{w \in \mathcal{W}} \max\{z_{iw}, z_{jw}\}}, \tag{1}$$

where $\mathcal{W}$ is either the set of users $\mathcal{U}$ or the set of categories $\mathcal{C}$, and accordingly, $\mathbf{z}_i$ is the *user vector* or the *category vector* of item i.

Diversity. Given a set of items $\mathcal{X} \subseteq \mathcal{I}$, we define the *diversity* of the set $\mathcal{X}$. First, we define its *coverage-based diversity* as

$$\mathrm{div}_C(\mathcal{X}) = \frac{1}{|\mathcal{C}|} \left\| \bigvee\nolimits_{i \in \mathcal{X}} \mathbf{y}_i \right\|_0, \tag{2}$$

where $\| \cdot \|_0$ returns the number of non-zero entries of the binary vector $\bigvee_{i \in \mathcal{X}} \mathbf{y}_i$. It is worth highlighting that div_C favours larger $\mathcal{X}$ sizes, and naturally prefers items that individually provide extensive coverage.

Further, for a set of items $\mathcal{X} \subseteq \mathcal{I}$ with $|\mathcal{X}| \geq 2$, we define its *distance-based diversity* as

$$\mathrm{div}_D(\mathcal{X}) = \frac{1}{|\mathcal{X}| - 1} \sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{X}} d(i, j), \tag{3}$$

and we define $\mathrm{div}_D(\mathcal{X}) = 0$, if $|\mathcal{X}| < 2$. As with div_C, the div_D metric favors larger sets, in addition to favoring items whose distance is large to each other.

User model. We aim to evaluate the quality of a recommendation algorithm $\mathcal{S}$ in the context of the user response to items recommended by $\mathcal{S}$. We view the user-algorithm interaction as a dynamic knowledge-exploration process, that continues as long as the recommended items are of interest to the user. If the recommended items have low relevance for the user, they may (stochastically) decide to quit.

To formalize the exploration process, we propose the following *user model*:

1. The set of items the user interacts with during exploration is denoted by $\mathcal{X}$. Initially, $\mathcal{X}$ is empty.
2. In step t, the recommendation algorithm $\mathcal{S}$ generates a list of items $\mathbf{L}_t$ for the user to examine in a specified order.
3. At any point, the user may quit based on two factors: the relevance of the recommended items and their patience. If the user finds no interesting items in $\mathbf{L}_t$ or runs out of patience, they may end the exploration process.
4. If the user does not quit, they select an item i from $\mathbf{L}_t$ with a probability depending on the item's relevance. The item i is added to $\mathcal{X}$, and the exploration continues.
5. Upon quitting, the total score achieved by $\mathcal{S}$ is $\mathrm{div}(\mathcal{X})$, where div is a diversity function, either div_C or div_D. This score reflects the diversity of items the user interacted with. The final number of steps performed by the user is denoted as κ.

Item selection. We now discuss step (4) of our iterative knowledge-exploration user model. that is, the probability that a user selects an item i from $\mathbf{L}_t$. We first assume that a user does not quit the exploration, i.e., that they have enough patience to explore the whole $\mathbf{L}_t$ and that they find at least a relevant item within it (see next paragraph). In that case, the user selects an item i from $\mathbf{L}_t$ with probability proportional to the relevance of i for that user u, that is, $p_i = \frac{\mathcal{R}(u,i)}{\sum_{j \in \mathbf{L}_t} \mathcal{R}(u,j)}$. As noted before, the selected item i is added to the set of interacted items $\mathcal{X}$.

Quitting exploration. Last, we discuss step (3) in our user model, that is, how we model the probability that a user quits the exploration process. We assume a user examines the items in the list $\mathbf{L}_t$ sequentially. Upon examining an item $i \in \mathbf{L}_t$, the user decides with probability η_t to quit the exploration due to worn out at step t. We refer to this as the *weariness* probability (discussed below). If the user does not quit, they decide whether item i is interesting to explore. The latter is decided again stochastically with Bernoulli probability q_i, which is a function of the relevance score $\mathcal{R}(u,i)$[1]. The examination of the list $\mathbf{L}_t$ continues until the user decides to quit or decides that there is at least one item that is interesting to explore. Therefore, the probability that the user quits examining the list $\mathbf{L}_t$ without identifying any item to explore is

$$Q_t = \sum_{j=1}^{|\mathbf{L}_t|} \eta_t (1-\eta_t)^{j-1} \prod_{i=1}^{j-1} (1-q_i). \tag{4}$$

Finally, the weariness probability η_t is modeled by resorting to the discrete version of the Weibull distribution [14, 15], which has been previously used to model web page dwell times and session lengths in web page navigation [16]:

$$\eta_t = 1 - q^{(t+1)^\gamma - t^\gamma}, \tag{5}$$

[1]We obtain q_i by normalizing $\mathcal{R}(u,i)$ into the $[0,1]$ interval by considering the maximum relevance range.

where $q = e^{-1/\lambda^\gamma}$, $0 \leq q \leq 1$. For $\gamma = 1$, the weariness probability remains constant, while for $\gamma > 1$, the weariness probability increases over time — modeling the tiredness of the user[2].

We can use the properties of the Weibull distribution to obtain the expected number of steps in the exploration process, for the case that all recommended items are maximally relevant, i.e., $q_i = 1$ for all $i \in \mathbf{L}_t$. In this case, $Q_t = \eta_t$ for all t. The overall quitting probability Q_T is then

$$Q_T = \sum_{t=1}^{\infty} q^{t^\gamma} - q^{(t+1)^\gamma}. \tag{6}$$

The expected number of steps $\mathbb{E}[\text{steps}]$ examined by a user before quitting (or equivalently, the number of items in $\mathcal{X}$) is hence given by

$$\mathbb{E}[\text{steps}] = \sum_{t=1}^{\infty} t\Big(q^{t^\gamma} - q^{(t+1)^\gamma}\Big). \tag{7}$$

Khan et al. [17] show that it is bounded by the expectation $\mu = \lambda\,\Gamma(1 + 1/\gamma)$ of the Weibull distribution in the continuous setting [15] as $\mu < \mathbb{E}[\text{steps}] < \mu + 1$, which provides an algebraic relationship between the λ parameter and the admissible range for the expected number of steps. Note that, if the relevance of the recommended items is less than 1, the right-hand side of Equation (7) provides an upper bound on the expected number of steps during exploration.

3. Recommendation Strategy

In this section, we present our recommendation strategy for the proposed knowledge-exploration framework. The core of the problem is to construct a list of recommendations $\mathbf{L}_t$ of size $\|\mathbf{L}_t\| = k$ at the t-th step of exploration, for a given user $u \in \mathcal{U}$. We assume that $\mathcal{X}_t$ is the set of items that the user has interacted with at step t, where $\mathcal{X}_1 = \emptyset$. We define $\mathcal{J}_t = \mathcal{I} \setminus \mathcal{X}_t$ to be set of items that are available for recommendation, that is, all items except the ones that the user has already interacted with.

For a user u and each item in the candidate set $i \in \mathcal{J}_t$, we consider its relevance score $\mathcal{R}_i = \mathcal{R}(u, i)$ and its *marginal diversity*

$$\mathcal{T}_i = \text{div}(\mathcal{X}_t \cup \{i\}) - \text{div}(\mathcal{X}_t), \tag{8}$$

with respect to the interaction set $\mathcal{X}_t$, where $\text{div} \in \{\text{div}_D, \text{div}_C\}$. We denote $\mathcal{T}_i = \mathcal{D}_i$ when the distance diversity function div_D is used, and $\mathcal{T}_i = \mathcal{C}_i$ when the coverage diversity function div_C is used. Intuitively, $\mathcal{D}_i$ represents the distance of i from all the items in the interaction set $\mathcal{X}_t$, while $\mathcal{C}_i$ represents the additional coverage that i provides[3]. Given $\mathcal{P}_i \in \{\mathcal{R}_i, \mathcal{T}_i\}$, we also denote the normalization of the score $\mathcal{P}$ as $\widehat{\mathcal{P}_i} = (\mathcal{P}_i - \mathcal{P}_{min})/(\mathcal{P}_{max} - \mathcal{P}_{min})$, where $\mathcal{P}_{max}$ and $\mathcal{P}_{min}$ are the maximum and minimum values of $\mathcal{P}$, respectively, over all items in $\mathcal{X}_t$.

[2] For $\gamma < 1$, the weariness probability decreases over time.

[3] At the beginning of the exploration process (when $\mathcal{X}_t = \emptyset$), if $\mathcal{T}_i = \mathcal{D}_i$, the strategy samples a highly relevant item i_r so that $\mathcal{D}_i = d(i, i_r)$; if $\mathcal{T}_i = \mathcal{C}_i$, then $\mathcal{C}_i = \mathbf{y}_i$, thus picking the item that individually provides the highest coverage.

Our strategy for constructing the recommendation list $\mathbf{L}_t$ is to combine relevance and diversity into one score. For each item i with relevance $\mathcal{R}_i$ and diversity $\mathcal{T}_i$, we compute the combined score $\mathcal{Z}_i$ by adopting the Clayton copula function [18]

$$\mathcal{Z}_i = \left[\widehat{\mathcal{R}}_i^{-\alpha} + \widehat{\mathcal{T}}_i^{-\alpha} - 1\right]^{-1/\alpha}, \tag{9}$$

where $\alpha > 0$ is a regularization parameter. The list $\mathbf{L}_t$ is then formed by selecting the top-k items from $\mathcal{J}_t$ according to $\mathcal{Z}_i$.

We refer to this strategy as EXPLORE. When the distance diversity function is used we refer to it as EXPLORE-D, and when coverage diversity is used we refer to it as EXPLORE-C.

4. Experiments

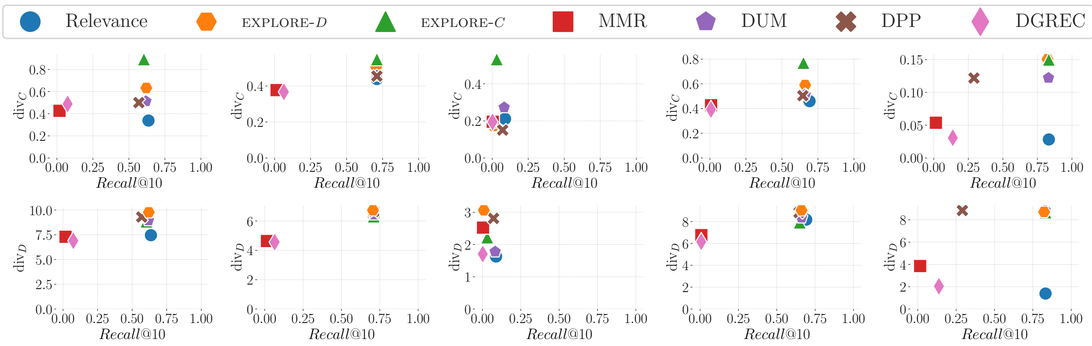
In this section, we assess the effectiveness of our strategy, either EXPLORE-D or EXPLORE-C, in balancing accuracy and diversity. We also evaluate it against several state-of-the-art competitors within the proposed knowledge-exploration framework.

Datasets and Competitors. We use the following five benchmark datasets, freely available online: **Movielens-1M**[4] [19], **Coat**[5] [20], **KuaiRec-2.0**[6] [21], **Netflix-Prize**[7] [22], and **Yahoo-R2**[8]. We evaluate our recommendation algorithm, EXPLORE, against several state-of-art competing strategies, namely: **Relevance**, a straightforward baseline which recommends the k most relevant items, **MMR** [11], **DUM** [10], **DPP** [12], and **DGREC** [23].

Setting. To evaluate the performance of the examined recommendation strategies, we divide user interactions into a training and a test set, following an 80-20% split ratio. When evaluating the accuracy, we only focus on the recommendation list generated in the initial exploration step, since it represents a lower bound of the system's overall accuracy. Regarding diversity instead, we consider the complete set of recommendation lists produced across all exploration steps. To calculate the relevance score $\mathcal{R}(u, i)$, we employ a black-box model in the form of a neural network based on matrix factorization [24]. For EXPLORE, we use a value of $\alpha = 0.5$ in the Clayton copula. We keep the length of the recommendation list fixed at 10, and vary the expected number of steps, $\mathbb{E}[\text{steps}]$, in the range of [5, 10, 20]. To assess recommendation quality, we use standard metrics: *Hit-Ratio* (*HR*), *Precision*, and *Recall*. Our experimental results are the average of 20 independent trials.

Results. We assess the performance of all our strategies in terms of recommendation quality and diversity. Figure 1 displays the scores for *Recall*@10 (on the x-axis) and diversity (on the y-axis) across all five datasets, either in terms of coverage (top-row) or distance (bottom-row). Our method, EXPLORE-C, clearly outperforms the other strategies, achieving a substantially higher diversity score while still delivering relevant recommendations. Similar considerations can be made for the distance-based variant, EXPLORE-D. Other results are in the Appendix.

[4] https://grouplens.org/datasets/movielens/1m/
[5] https://www.cs.cornell.edu/~schnabts/mnar/
[6] https://kuairec.com/
[7] https://www.kaggle.com/datasets/rishitjavia/netflix-movie-rating-dataset
[8] https://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67



(a) Movielens-1M (b) Coat (c) KuaiRec-2.0 (d) Netflix-Prize (e) Yahoo-R2

Figure 1: Trade-off between div_C (top) and div_D (bottom), and *Recall*@10, respectively, across all the datasets considered. The x-axis represents recommendation quality, while the y-axis indicates the diversity score.

5. Conclusion and Future Work

In this study, we addressed recommendation diversity by introducing a user-behavior model where relevance drives engagement. We developed a recommendation strategy that optimizes the delivery of diverse knowledge based on user behavior. Our experimental analysis confirms the effectiveness of this approach, though it remains open to further enhancements. First, the behavioral model can be refined to include more sophisticated scenarios, such as refreshing the list, guiding its composition, and incorporating dynamic adjustments to the weariness probability beyond temporal decay. Additionally, our model assumes the relevance score accurately reflects a user's interest in an item. However, since the relevance score is algorithmically computed and may not be entirely accurate, we can adapt the user behavior model by incorporating a random discount factor for the relevance of each item. Finally, the proposed strategy can be improved in several ways, such as integrating different distance measures or extending it to include additional metrics beyond diversity, like serendipity or fairness.

Acknolwedgements

This work has been partially funded by MUR on D.M. 351/2022, PNRR Ricerca, CUP H23C22000440007, and supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union – NextGenerationEU. Aristides Gionis is supported by the ERC Advanced Grant REBOUND (834862), the EC H2020 RIA project SoBigData++ (871042), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

[1] C. Wu, F. Wu, Y. Huang, X. Xie, Personalized news recommendation: Methods and challenges, ACM Trans. Inf. Syst. 41 (2023) 24:1–24:50.

[2] C. Hansen, C. Hansen, L. Maystre, R. Mehrotra, B. Brost, F. Tomasi, M. Lalmas, Contextual and sequential user embeddings for large-scale music recommendation, in: ACM RecSys, 2020, pp. 53–62.

[3] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, S. Wang, Learning graph-based POI embedding for location-based recommendation, in: Procs. ACM CIKM, 2016, pp. 15–24.

[4] J. L. Herlocker, J. A. Konstan, L. G. Terveen, J. T. Riedl, Evaluating collaborative filtering recommender systems, ACM Trans. Inf. Syst. 22 (2004) 5–53.

[5] B. Smyth, P. McClave, Similarity vs. diversity, in: Procs. Springer ICCBR, volume 2080, 2001, pp. 347–361.

[6] Z. Tufekci, Youtube, the great radicalizer, The New York Times 10 (2018) 2018.

[7] M. H. Ribeiro, R. Ottoni, R. West, V. A. Almeida, W. Meira Jr, Auditing radicalization pathways on youtube, in: Procs. ACM FAT*, 2020, pp. 131–141.

[8] M. Haroon, A. Chhabra, X. Liu, P. Mohapatra, Z. Shafiq, M. E. Wojcieszak, Youtube, the great radicalizer? auditing and mitigating ideological biases in youtube recommendations, CoRR abs/2203.10666 (2022).

[9] S. Phadke, M. Samory, T. Mitra, Pathways through conspiracy: The evolution of conspiracy radicalization through engagement in online conspiracy discussions, in: Procs. AAAI ICWSM, 2022, pp. 770–781.

[10] A. Ashkan, B. Kveton, S. Berkovsky, Z. Wen, Optimal greedy diversity for recommendation, in: Procs. AAAI IJCAI, 2015, pp. 1742–1748.

[11] J. Carbonell, J. Goldstein, The use of mmr, diversity-based reranking for reordering documents and producing summaries, in: Procs. ACM SIGIR, 1998, pp. 335–336.

[12] L. Chen, G. Zhang, H. Zhou, Fast greedy MAP inference for determinantal point process to improve recommendation diversity, in: Procs. NeurIPS, 2018, pp. 5627–5638.

[13] E. Coppolillo, G. Manco, A. Gionis, Relevance meets diversity: A user-centric framework for knowledge exploration through recommendations, ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024. doi:10.1145/3637528.3671949.

[14] H. Rinne, The Weibull Distribution: A Handbook (1st ed.), Chapman and Hall/CRC, 2008.

[15] J. G. Proakis, Probability, random variables and stochastic processes, IEEE Trans. Acoust. Speech Signal Process. 33 (1985) 1637.

[16] C. Liu, R. W. White, S. T. Dumais, Understanding web browsing behaviors through weibull analysis of dwell time, in: Procs. ACM SIGIR, 2010, pp. 379–386.

[17] M. Khan, A. Khalique, A. M. Abouammoh, On estimating parameters in a discrete weibull distribution, IEEE Transactions on Reliability 38 (1989) 348–350.

[18] D. G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.

[19] F. M. Harper, J. A. Konstan, The movielens datasets: History and context, ACM Trans. Interact. Intell. Syst. 5 (2016) 19:1–19:19.

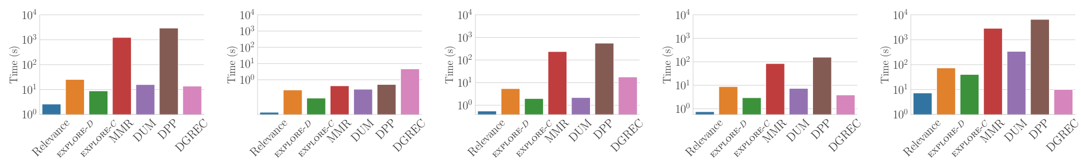
[20] T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, T. Joachims, Recommendations as

treatments: Debiasing learning and evaluation, in: Procs. JMLR ICML, volume 48, 2016, pp. 1670–1679.
[21] C. Gao, S. Li, W. Lei, J. Chen, B. Li, P. Jiang, X. He, J. Mao, T. Chua, Kuairec: A fully-observed dataset and insights for evaluating recommender systems, in: Procs. ACM CIKM, 2022, pp. 540–550.
[22] J. Bennett, S. Lanning, The netflix prize, 2007.
[23] L. Yang, S. Wang, Y. Tao, J. Sun, X. Liu, P. S. Yu, T. Wang, Dgrec: Graph neural network for recommendation with diversified embedding generation, in: Procs. ACM WSDM, 2023, pp. 661–669.
[24] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (2009) 30–37.

A. Additional Results

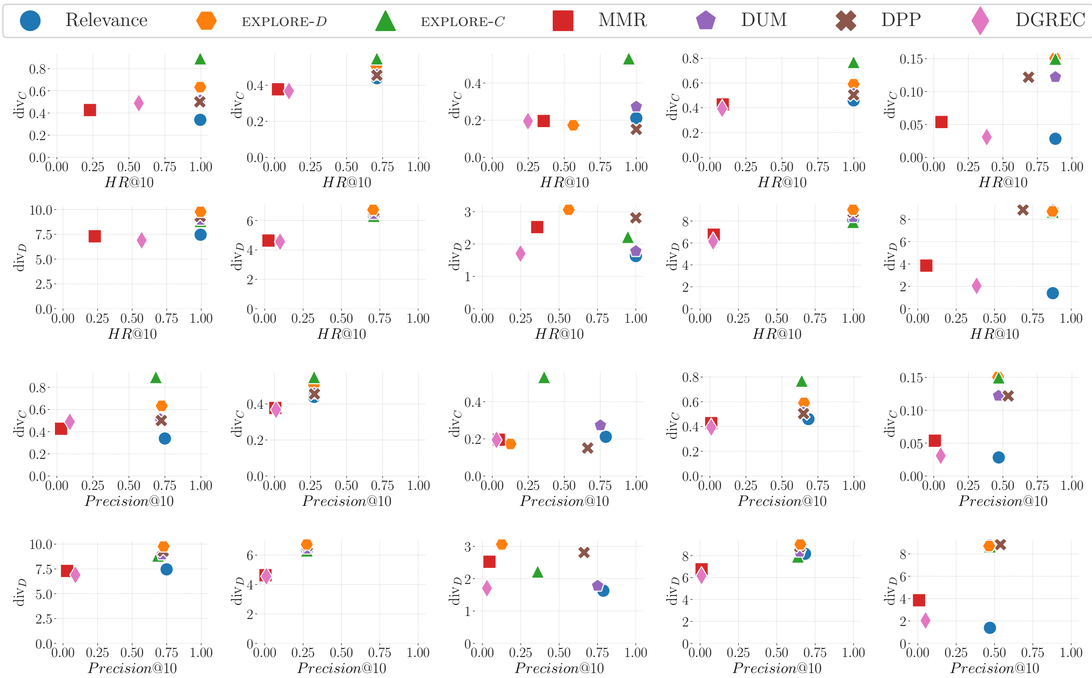
Table 1 presents a comprehensive analysis of div_D, div_C and κ when $\mathbb{E}[\text{steps}] = 5$. Additionally, we report the deviations from the maximum diversity scores in terms of distance and coverage, denoted as $\Delta_{\bar{D}}$ and $\Delta_{\bar{C}}$, along with $\Delta_{\mathbb{E}[\text{steps}]}$. We observe that our strategy, either EXPLORE-D or EXPLORE-C, consistently outperforms the competitors in terms of both div_D and div_C across all datasets. We also show how these values deviate from the expected maximum values. Notably, on the Movielens-1M dataset, their scores are very close to their maxima. Our strategy achieves significantly higher scores than the competitors on all datasets, especially in terms of coverage. Table 2 reports additional results for $\mathbb{E}[\text{steps}] \in [10, 20]$.

Figure 3 shows the trade-off between diversity and accuracy, computed in terms of *HR@10* and *Precision@10*. Figure 2 reports the timing (in second) needed to produce a recommendation list $\mathbf{L}_t$. Notably, while competitors such as MMR and DPP struggle with larger datasets, the timing of our strategy is basically constant across all the datasets.



(a) Movielens-1M (b) Coat (c) KuaiRec-2.0 (d) Netflix-Prize (e) Yahoo-R2

Figure 2: Timing for producing $\mathbf{L}_t$. The x-axis reports the strategies, while the y-axis the recommendation time (in seconds).



(a) Movielens-1M (b) Coat (c) KuaiRec-2.0 (d) Netflix-Prize (e) Yahoo-R2

Figure 3: Trade-off between either div_C or div_D and either *HR*@10 or *Precision*@10 across all the datasets. The x-axis shows the recommendation quality while the y-axis represents the diversity score.

Table 1

Results for $\mathbb{E}[\text{steps}] = 5$. Any best scores with a statistical significance $p < 0.05$ are highlighted in bold.

Dataset	Strategy	div_D	div_C	κ	$\Delta_{\bar{D}}$	$\Delta_{\bar{C}}$	$\Delta_{\mathbb{E}[\text{steps}]}$
Movielens-1M	Relevance	3.67	0.22	5.0	0.27	0.73	0.0
	EXPLORE-D	**4.91**	0.36	4.98	**0.03**	0.55	0.0
	EXPLORE-C	4.43	**0.71**	4.71	0.12	**0.11**	0.06
	MMR	3.96	0.29	4.57	0.22	0.64	0.09
	DUM	4.4	0.33	4.98	0.13	0.59	0.0
	DPP	4.59	0.31	4.99	0.09	0.61	0.0
	DGREC	3.36	0.37	4.49	0.33	0.54	0.1
Coat	Relevance	3.15	0.3	4.36	0.38	0.3	0.13
	EXPLORE-D	**3.48**	0.34	4.16	**0.31**	0.2	0.17
	EXPLORE-C	3.36	**0.35**	4.13	0.33	**0.18**	0.17
	MMR	2.43	0.26	3.54	0.52	0.39	0.29
	DUM	3.11	0.3	4.31	0.38	0.3	0.14
	DPP	3.28	0.31	4.33	0.35	0.27	0.13
	DGREC	2.2	0.24	3.2	0.56	0.44	0.36
KuaiRec-2.0	Relevance	0.76	0.13	4.81	0.81	0.74	0.04
	EXPLORE-D	**1.56**	0.11	3.54	**0.61**	0.78	0.29
	EXPLORE-C	1.08	**0.34**	4.09	0.73	**0.32**	0.18
	MMR	1.25	0.12	3.89	0.68	0.76	0.22
	DUM	0.83	0.17	4.8	0.79	0.66	0.04
	DPP	1.38	0.09	4.75	0.65	0.82	0.05
	DGREC	0.77	0.11	2.64	0.81	0.78	0.47
Netflix	Relevance	4.04	0.32	4.86	0.2	0.59	0.03
	EXPLORE-D	**4.62**	0.38	4.75	**0.09**	0.51	0.05
	EXPLORE-C	3.97	**0.6**	4.43	0.21	**0.22**	0.11
	MMR	3.56	0.3	4.19	0.3	0.61	0.16
	DUM	4.16	0.36	4.89	0.18	0.53	0.02
	DPP	4.38	0.34	4.88	0.13	0.56	0.02
	DGREC	3.0	0.26	3.74	0.41	0.66	0.25
Yahoo-R2	Relevance	0.66	0.02	**4.77**	0.87	0.77	**0.05**
	EXPLORE-D	**4.4**	**0.08**	4.49	**0.13**	**0.1**	0.1
	EXPLORE-C	4.38	**0.08**	4.47	**0.13**	**0.1**	0.11
	MMR	2.45	0.04	3.96	0.52	0.55	0.21
	DUM	4.38	0.07	4.72	**0.13**	0.21	0.06
	DPP	4.38	0.07	4.72	**0.13**	0.21	0.06
	DGREC	1.02	0.02	3.68	0.8	0.77	0.26

Table 2
Results with $\mathbb{E}[\text{steps}] \in [10, 20]$ across all the datasets. Best scores with statistical significance $p < 0.05$ are in bold.

Dataset	Strategy	$\text{div}_{\mathcal{D}}$	$\text{div}_{\mathcal{C}}$	κ	$\Delta_{\bar{\mathcal{D}}}$	$\Delta_{\bar{\mathcal{C}}}$	$\Delta_{\mathbb{E}[\text{steps}]}$
Movielens-1M	Relevance	7.45	0.34	10.02	0.26	0.64	0.0
	EXPLORE-$\mathcal{D}$	**9.77**	0.63	9.85	**0.03**	0.32	0.02
	EXPLORE-$\mathcal{C}$	8.84	**0.89**	9.7	0.12	**0.05**	0.03
	MMR	7.29	0.43	8.62	0.27	0.54	0.14
	DUM	8.95	0.51	10.03	0.11	0.45	0.0
	DPP	9.29	0.5	9.99	0.08	0.46	0.0
	DGREC	6.89	0.49	8.92	0.31	0.47	0.11
Coat	Relevance	6.38	0.44	8.64	0.36	0.35	0.14
	EXPLORE-$\mathcal{D}$	6.73	0.51	8.02	0.33	0.25	0.2
	EXPLORE-$\mathcal{C}$	6.31	**0.55**	7.81	0.37	**0.19**	0.22
	MMR	4.63	0.38	6.75	0.54	0.44	0.32
	DUM	6.44	0.46	8.65	0.36	0.32	0.13
	DPP	6.64	0.45	8.61	0.34	0.34	0.14
	DGREC	4.56	0.37	6.33	0.55	0.46	0.37
KuaiRec-2.0	Relevance	1.63	0.21	9.6	0.79	0.71	0.04
	EXPLORE-$\mathcal{D}$	**3.06**	0.17	6.86	**0.61**	0.77	0.31
	EXPLORE-$\mathcal{C}$	2.22	**0.53**	7.95	0.72	**0.28**	0.2
	MMR	2.52	0.2	7.34	0.68	0.73	0.27
	DUM	1.78	0.27	9.59	0.77	0.63	0.04
	DPP	2.81	0.15	9.54	0.64	0.8	0.05
	DGREC	1.71	0.19	5.45	0.78	0.74	0.45
Netflix	Relevance	8.17	0.46	9.73	0.19	0.47	0.03
	EXPLORE-$\mathcal{D}$	**9.03**	0.59	9.24	**0.1**	0.32	0.08
	EXPLORE-$\mathcal{C}$	7.92	**0.77**	8.69	0.21	**0.12**	0.13
	MMR	6.76	0.43	7.94	0.33	0.51	0.21
	DUM	8.36	0.51	9.72	0.17	0.41	0.03
	DPP	8.82	0.5	9.73	0.12	0.43	0.03
	DGREC	6.15	0.39	7.44	0.39	0.55	0.26
Yahoo-R2	Relevance	1.39	0.03	**9.49**	0.86	0.83	**0.05**
	EXPLORE-$\mathcal{D}$	8.71	**0.15**	8.73	0.13	**0.14**	0.13
	EXPLORE-$\mathcal{C}$	8.67	**0.15**	8.7	0.14	**0.14**	0.13
	MMR	3.86	0.05	7.36	0.62	0.71	0.26
	DUM	8.86	0.12	9.43	0.12	0.31	0.06
	DPP	8.84	0.12	9.41	0.12	0.31	0.06
	DGREC	2.04	0.03	7.35	0.8	0.83	0.26

(a) $\mathbb{E}[\text{steps}] = 10$.

Dataset	Strategy	$\text{div}_{\mathcal{D}}$	$\text{div}_{\mathcal{C}}$	κ	$\Delta_{\bar{\mathcal{D}}}$	$\Delta_{\bar{\mathcal{C}}}$	$\Delta_{\mathbb{E}[\text{steps}]}$
Movielens-1M	Relevance	14.96	0.48	20.04	0.25	0.51	0.0
	EXPLORE-$\mathcal{D}$	**18.96**	0.86	19.4	**0.05**	0.12	0.03
	EXPLORE-$\mathcal{C}$	16.81	**0.97**	19.76	0.16	**0.01**	0.01
	MMR	13.45	0.57	16.49	0.33	0.42	0.18
	DUM	17.98	0.7	**20.16**	0.1	0.29	**-0.01**
	DPP	18.6	0.71	20.02	0.07	0.28	0.0
	DGREC	13.91	0.63	17.64	0.31	0.36	0.12
Coat	Relevance	12.23	0.59	16.36	0.39	0.33	0.18
	EXPLORE-$\mathcal{D}$	12.85	0.7	15.48	0.36	0.21	0.23
	EXPLORE-$\mathcal{C}$	12.09	**0.77**	15.36	0.4	**0.13**	0.23
	MMR	8.79	0.53	13.12	0.56	0.4	0.34
	DUM	12.63	0.62	16.78	0.37	0.3	0.16
	DPP	13.06	0.62	16.84	0.35	0.3	0.16
	DGREC	9.31	0.53	12.84	0.54	0.4	0.36
KuaiRec-2.0	Relevance	3.62	0.32	19.02	0.77	0.65	0.05
	EXPLORE-$\mathcal{D}$	**6.16**	0.26	13.83	**0.61**	0.71	0.31
	EXPLORE-$\mathcal{C}$	4.55	**0.76**	15.09	0.71	**0.16**	0.25
	MMR	4.79	0.3	13.91	0.69	0.67	0.3
	DUM	3.86	0.39	19.09	0.75	0.57	0.05
	DPP	5.56	0.24	18.99	0.64	0.73	0.05
	DGREC	3.55	0.3	11.33	0.77	0.67	0.43
Netflix	Relevance	16.19	0.6	19.29	0.19	0.34	0.04
	EXPLORE-$\mathcal{D}$	17.38	0.77	17.98	**0.13**	0.15	0.1
	EXPLORE-$\mathcal{C}$	15.6	**0.87**	17.53	0.22	**0.04**	0.12
	MMR	12.79	0.56	15.34	0.36	0.38	0.23
	DUM	16.59	0.65	19.33	0.17	0.29	0.03
	DPP	**17.49**	0.66	19.36	**0.13**	0.27	0.03
	DGREC	12.11	0.53	14.59	0.4	0.42	0.27
Yahoo-R2	Relevance	3.0	0.04	**18.81**	0.85	0.88	**0.06**
	EXPLORE-$\mathcal{D}$	16.48	**0.28**	16.5	0.18	**0.17**	0.18
	EXPLORE-$\mathcal{C}$	16.43	**0.28**	16.46	0.18	**0.17**	0.18
	MMR	5.9	0.07	13.89	0.71	0.79	0.31
	DUM	17.59	0.19	18.73	0.12	0.44	**0.06**
	DPP	17.6	0.19	18.74	0.12	0.44	**0.06**
	DGREC	3.92	0.04	14.56	0.8	0.88	0.27

(b) $\mathbb{E}[\text{steps}] = 20$.