Diversification of Complementary Item Recommendations with User Preference in Online Grocery

Luyi Ma
WalmartLabs
Sunnyvale, California
Luyi.Ma@walmartlabs.com

Nimesh Sinha
WalmartLabs
Sunnyvale, California
Nimesh.Sinha@walmartlabs.com

Jason H.D. Cho
WalmartLabs
Sunnyvale, California
hcho@walmartlabs.com

Sushant Kumar
WalmartLabs
Sunnyvale, California
skumar4@walmartlabs.com

Kannan Achan
WalmartLabs
Sunnyvale, California
kachan@walmartlabs.com

ABSTRACT

Complementary item recommendations play an important role in surfacing the relevant items to the customers in online grocery platform. Many research works focus on improving the recommendations through post-processing techniques such as user-item-level personalization and diversification of recommendations. In online grocery, complementary recommendations can fall into two types, cross-domain and within-domain. Cross-domain complementary recommendations refer to products that belong to two distinct categories, for instance, for query item hot dog bun, steak and barbecue sauce are cross-domain complementary items. On the other hand, for the same query item, chicken sausage and beef sausage form within-domain complementary recommendations as they are from delicatessen. Although diversification of recommendations can promote cross-domain complementary recommendations, it can't capture the within-domain complementary recommendations. Within-domain complementary recommendations can not be replaced by similarity. Users also show their preferences on cross-domain or within-domain complementary items during online shopping, which can be indicated by their shopping behavior. However, many user-item-level personalization methods can't explicitly model the preference of users on cross-domain and within-domain complementary recommendations. To stress the requirement of modeling cross-domain and within-domain complementary recommendations with personalization, we propose our re-ranking solution to provide fine-grained control of diversification on the complementary recommendations. We use transaction history to estimate users' preference for the level of diversity in the complementary recommendations and combine it to our re-ranking solution. We demonstrate the effectiveness of our re-ranking algorithm on the publicly available Instacart dataset.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/1122445.1122456

CCS CONCEPTS

• **Information systems → Information retrieval diversity**; **Similarity measures**.

KEYWORDS

Recommender systems, Diversity, Similarity, Re-ranking

ACM Reference Format:
Luyi Ma, Nimesh Sinha, Jason H.D. Cho, Sushant Kumar, and Kannan Achan. 2018. Diversification of Complementary Item Recommendations with User Preference in Online Grocery. In ,. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Recommender system is an essential part of e-commerce business. Recommending relevant items to customers make their experience more comfortable, hassle-free, and time-saving. Online grocery platforms also have wide variety of recommendation systems placed at various sections of their websites to improve customer journey. One such important section is complementary recommendations, where the customers can select different items from a recommendation pool. Complementary recommendations, in particular, show products that customers frequently co-purchase. However, there is no real definition of complementary items in online grocery. A large number of items like bread, cereal, sugar, fruits, coffee, etc can be complementary recommendations for the query item milk.

In online grocery, items from both within-domain and cross-domain are bought as complementary recommendations. Domain can be defined in terms of any taxonomy hierarchy level in grocery. In our case, we have defined domain as department level for the grocery items. Cross-domain recommendations refer to the items which belong to different and diverse departments, while within-domain recommendations refer to the items which belong to the similar departments. Note that the all the recommendations either they are coming from cross-domain or within-domain are complementary to the query item. For example, suppose the query item is tortilla chips, which has salsa dip, guacamole dip, and soft drink as its complementary recommendations. A customer can buy salsa dip and guacamole dip together from the recommendations which are from within-domain or similar departments, while another customer can buy salsa dip and soft drink together which are from cross-domain or different departments. All the three items - salsa dip, guacamole dip, and soft drink are complementary to the query item tortilla chips, but the recommendation pool in itself can have a mixture of such within-domain and cross-domain complementary items. It depends on the user intent to buy within-domain or cross domain recommendations together.

Most of the recommendation systems optimize on accuracy and diversity. It is not surprising, then, that a lot of researchers focus on both personalization to improve recommendation relevance for each individual user, and diversification to boost novelty. Diversification can promote the cross-domain complementary recommendations, but we need personalization to control the degree of diversification. Balancing out diversification and personalization is a delicate task. Without careful calibration, one can overshadow within-domain complementary recommendations (too much diversification), and vice versa.

To address these challenges, we utilize existing complementary recommendations and deploy re-ranking strategy to balance out within-domain and cross-domain recommendations. We further combine user behavior to guide the diversification process. Some users may prefer more within-domain recommendations, and vice versa.

We use re-ranking strategy based on Determinantal Point Process (DPP), combining the cross-domain and wihin-domain re-ranking with user intent to maintain the balance between the cross-domain and within-domain complementary recommendations. We perform the re-ranking on top of complementary recommendation pool because we want to ensure complementary relationship between the query item and recommended items. We also combine the user behavior by utilizing their historical shopping pattern to model the preferential balance between within-domain and cross-domain complementary recommendations.

2 RELATED WORKS

Many past research works have focused on applying personalization in the recommendation systems. These works focus on modeling user preference to increase accuracy of the recommendations and give better customer experience. Techniques like matrix factorization and collaborative filtering have been widely used [13] [1]. Many research works focus on the complementary items by inferring item-item relations [3]. These complementary items have been personalized using both item relations and user preferences [17]. In online grocery platform, models like **triple2vec** have been used to personalize the complementary recommendations using (*item, item, user*) triplets [6] [14].

For a long time, there was not much importance given to diversity in the recommendations as it is challenging to achieve both high accuracy and diversity at the same time. This is called as *accuracy diversity dilemma* [9]. Novelty and diversity of items have been improved by penalizing accuracy [5]. Diversity has also been captured in entropy regularizer [12]. Post-processing methods for diversity have been proposed to improve the personalized recommendations generated by collaborative filtering [2]. Determinantal Point Process (DPP) has been used for making personalized diversified recommendations. DPP model are elegant probabilistic models which have a lot of applications. [8]. It has been incorporated with a tunable parameter allowing the users to smoothly control the level of diversity in recommendations and also, applied to large scale scenarios with faster inference [15]. Deep reinforcement learning has utilized DPP to promote diversity to generate diverse, while relevant item recommendations. DPP kernel matrix is maintained for each user, which is constructed from two parts: a fixed similarity matrix capturing item-item similarity, and the relevance of items dynamically learnt through an actor-critic reinforcement learning framework [10].

Most of the existing works revolve around tackling the challenges of personalization, and diversity in complementary recommendations. They don't give much stress on maintaining the delicate balance between within-domain and cross-domain complementary items. Our proposed method focuses on combined re-ranking strategy for within-domain and cross-domain complementary recommendations including user intent.

3 MODEL

In this section, we first revisit the skip-gram-based embedding learning framework [14] and generating the item embedding used for diversification. Then, we propose our algorithm under the setting of DPP to tackle both cross-domain and within-domain complementary recommendations. We estimate the user intent of cross-domain and within-domain complementary item recommendation based on transaction data and combine it with our DPP-based algorithm to provide complementary item recommendation with user preference (personalization).

3.1 Background: Skip-gram-based Item Embedding and *triple2vec*

As mentioned before, skip-gram based methods for item embedding leverage the item co-occurrence signal. Models for complementary item recommendations like [11], [3] exactly use the item co-occurrence signal to model item complementarity. **triple2vec** in [14] introduced the cohesion of $(item, item, user)$ triplets which reflect co-purchase of two items by the same user in the same basket. This technique improves the performance of complementary item recommendations and **triple2vec** achieves the state-of-the-art performance. As we focus on the post-processing of the recommendations, we decide to leverage the item representations learned by **triple2vec** to generate item pools for down-stream applications.

In **triple2vec**, triplets of $\{(q, r, u) \mid q \in V, r \in V, u \in U\}$ represents user-item and item-item relationships where V is the set of items and U is the set of users. Here, q and r are two items purchased by the user u in the same basket. Particularly, we refer q to the query item and r to the recommended item. The relationship between q and r can be viewed in the way that r is the recommended complementary item for the query item q. The cohesion of (q, r, u) in **triple2vec** is computed by Eq. 1, where f_q, g_r are two sets of representations for items (q, r) and h_u is the user embedding.

$$s_{q,r,u} = \overbrace{f_q^T g_r}^{x} + \underbrace{f_q^T h_u + g_r^T h_u}_{y} \tag{1}$$

x and y in Eq. 1 indicate the item-to-item complementarity and user-to-item compatibility respectively. The loss function $\mathcal{L}$ in Eq.

2 computes the likelihood of all possible triplets $\mathcal{T}$ and is optimized to learn representations of items and users.

$$\mathcal{L} = \sum_{q,r,u \in \mathcal{T}} (\log p(r|q,u) + \log p(q|r,u) + \log p(u|q,r)) \quad (2)$$

Here, $p(r|q,u) = \frac{\exp(q,r,u)}{\sum_{r'} \exp(q,r',u)}$, $p(q|r,u) = \frac{\exp(q,r,u)}{\sum_{q'} \exp(q',r,u)}$ and $p(u|q,r) = \frac{\exp(q,r,u)}{\sum_{u'} \exp(q,r,u')}$.

We leverage **triple2vec** to learn item representations and generate item pools of complementary recommendations for downstream processes. To recall the item pool of complementary recommendations, we consider the inner product score $f_q^T g_r$ for two items q, r. For each query item q, we select a pool of items $R = \{r_1, ..., r_m\}$ with the highest score of $f_q^T g_r$.

3.2 Recommendation Diversification and Determinantal Point Process

Improving diversity of recommendations benefits the recommender systems because it introduces novelty and better topic coverage [18]. Many works in diversification follow the setting of bi-criterion optimization problem which balances the relevance (between the query and recalled elements) and diversity [16]. Particularly, diversity can be further divided into two types, (1) individual diversity [1] and (2) aggregate diversity [2] [16]. We focus on the individual diversity in this work to adjust the diversity of complementary recommendations given a user's intent.

The determinantal point process (DPP) is a probabilistic model which is good at modeling repulsion. The recent work [4] applies DPP on diversification of item recommendations and develop fast greedy MAP inference to generate diversified recommendations. Our work is based on DPP with the fast greedy MAP inference in [4]. Thus, we introduce details of DPP and the fast greedy MAP inference following the notation in [4]. For the rest of our paper, we denote the fast greedy MAP inference as **FG-MAP**.

Formally, a DPP on a discrete set $Z = \{1, 2, ..., M\}$ is a probability measure $\mathcal{P}$ on $2^{|Z|}$ number of subsets of Z, where $|Z|$ is the number of elements in Z. Because the empty set is also a subset of Z, when $\mathcal{P}$ doesn't give zero probability to the empty set, there exists a square, positive semidefinite (PSD) and real matrix $\mathbf{L} \in \mathbb{R}^{M \times M}$ which satisfies Eq. 3 for each subset $Y \subseteq Z$.

$$\mathcal{P}(Y) \propto \det(\mathbf{L}_Y), \ \mathbf{L}_Y \in \mathbb{R}^{|Y| \times |Y|} \quad (3)$$

$\mathbf{L}$ serves as a kernel matrix indexed by the elements in Z and $\det(\mathbf{L}_Y)$ is the determinant of sub-matrix extracted from $\mathbf{L}$ based on elements in Y. Eq. 3 indicates that the probability of a subset Y is proportional to the determinant of the corresponding sub-matrix of the PSD kernel. The MAP inference of the aforementioned DPP $\mathcal{P}$ on Z is defined in Eq. 4.

$$Y_{map} = \arg\max_{Y \subseteq Z} \mathbf{L}_Y \quad (4)$$

Unlike other inference on DPP, the MAP inference of DPP is NP-hard. The algorithm **FG-MAP** approximates the MAP inference in a greedy approach. Eq. 5 shows how to greedily select the next candidate item j which is added to the existing growing subset $Y_g \subseteq Z$ built from the previous iterations. After the current iteration, Y_g grows and $Y_g \coloneqq Y_g \bigcup \{j\}$ [3].

$$j = \arg\max_{i \in Z \setminus Y_g} \log\det(\mathbf{L}_{Y_g \bigcup \{i\}}) - \log\det(\mathbf{L}_{Y_g}) \quad (5)$$

When Z becomes the item pools for complementary recommendations $R = \{r_1, ..., r_m\}$ recalled by the item representations (i.e., item embedding learned by **triple2vec**), a DPP on R maximizes the $\mathcal{P}(Y)$ and diversifies the recommendations by selecting r_i from R iteratively. Now the kernel matrix $\mathbf{L}$ could be initialized by the item-to-item similarity matrix based on the item embedding. In our work, we adapt DPP and **FG-MAP**, with $\mathbf{L}$ defined in Eq. 6.

$$\mathbf{L} = \frac{1 + H^T H}{2}, \ H \equiv \{g_{r_i} | r_i \in R\} \quad (6)$$

H is a sub-matrix of the item embedding for the item pools R recalled by the **triple2vec** model. g_{r_i} is normalized embedding of item r_i and the value of $H^T H$ is shifted to ensure $\mathbf{L}$ is PSD. We only use one set of item embedding from **triple2vec** model to compute item similarity as the distance between f_q and g_r from two sets of embedding represents the complementarity of (q, r).

3.3 Cross-domain and Within-domain Complementary Item Recommendation by DPP with User Intent

As we mentioned in Session 1, complementary recommendations in online grocery can fall into two types, (1) cross-domain and (2) within-domain. Cross-domain complementary recommendations refers to products that belong to two distinct categories, for instance, steak and barbecue sauce, given the query item of hot dog bun. Chicken sausage and beef sausage, on the other hand form within-domain complementary recommendations as they are from delicatessen The within-domain complementary recommendations can not be covered by similar item recommendations and also not explicitly covered by existing complementary models like [14] [3].

Cross-domain Complementary Recommendations. The cross-domain complementary recommendations can be achieved by increasing the diversity in the complementary recommendations R recalled by a complementary item recommendation model, i.e., **triple2vec**. We first generate R to ensure complementary recommendations and then re-rank items in R to surface more diverse but relevant items to the top. If we don't conduct diversification within the pool of pre-selected complementary items, the diversification logic could easily bias irrelevant items. We can re-rank the items in R by modifying **FG-MAP** into bi-criterion optimization. Specifically, we consider the score $S_{q,r_i} = \frac{1+f_q^T g_{r_i}}{2}$ for complementarity of (q, r_i), where f_q and g_{r_i} are normalized item embedding. Eq. 7 shows the

[1] Individual diversity refers to the diversity of recommendations for a given user, or individual diversity focuses on the problem of how to maximize item novelty in face of already recommended ones when generating the recommendation list.

[2] Aggregate diversity refers to the diversity of recommendations across all users, or aggregate diversity can be viewed as a problem of how to improve the ability of a recommender system to recommend long-tail items.

[3] For more details of the Fast Greedy MAP inference algorithm, please refer to [4].

modified objective function for diversification re-rank.

$$r_j = \arg\max_{r_i \in R \setminus R_d} \underbrace{\alpha S_{q,r_i}}_{\text{complementarity}} + \underbrace{(1-\alpha)\left(\log\det(\mathbf{L}_{R_d+[r_i]}) - \log\det(\mathbf{L}_{R_d})\right)}_{\text{increment of diversification}} \tag{7}$$

At t-th iteration, $R_{t,d} \coloneqq R_{t-1,d} + [r_j]$ where $R_{0,d} = []$ and $R_{t-1,d} + [r_j]$ means the newly selected recommendation r_j by the diversification re-rank is inserted at the end of the current item list $R_{t-1,d}$. The weight α controls the amount of diversity introduced to the re-ranked item list. Each selected item r_j can maximize the combined score of diversity and complementarity. The re-ranked item list R_d will surface more diversified recommendation to the top compared with the original item list R in which items are simply sorted by the score S_{q,r_i} in descending order.

Within-domain Complementary Recommendations. The within-domain complementary recommendations is different from the cross-domain complementary recommendations. We need to surface more items which are related to the query items but under the same topic instead of diverse results. For example, assume a query item Milk has a list of recommendations R = {Eggs, Cheese, Bread, Margarine, Banana, Sausage, Yogurt, Cereal}. If we want to stress the within-domain complementary recommendations, the re-ranked recommendations could be R_s = {Eggs, Cheese, Margarine, Yogurt, Banana, Bread, Sausage, Cereal} [4]. We encourage more homogeneousness in the within-domain complementary recommendations. R_s surfaces more items under the Dairy & Eggs domain such as Cheese, Yogurt. The within-domain complementary recommendations can be promoted by similarity of items in the recall set of complementary recommendations. Unlike diversification for cross-domain complementary recommendations which is diverging the item relationship, boosting within-domain complementary recommendations by similarity is more stable. We can mind candidate items in a bigger recall set. Formally, we recall extra complementary items $R_x = \{r_{m+1}, ..., r_n\}$ and insert them at the end of R. The new item list becomes $R + R_x$. To force the similarity between recommendations, we modify the kernel $\mathbf{L}$ in DPP by Eq. 8 and apply DPP on the new dissimilarity matrix $\mathbf{L}'$

$$\mathbf{L}' = \mathbf{1} + \text{diag}(\mathbf{L}) - \mathbf{L} \tag{8}$$

, where $\text{diag}(\mathbf{L})$ is a diagonal matrix with all entries in the main diagonal equal to the diagonal of $\mathbf{L}$ and $\mathbf{1}$ is a square matrix with all entries equal to 1. Plug $\mathbf{L}'$ into Eq. 7, and we can have a new re-ranking logic on the extended item pool $R + R_x$, shown in Eq. 9 .

$$r_h = \arg\max_{r_i \in (R+R_x) \setminus R_s} \underbrace{\beta S_{q,r_i}}_{\text{complementarity}} + \underbrace{(1-\beta)\left(\log\det(\mathbf{L}'_{R_s+[r_i]}) - \log\det(\mathbf{L}'_{R_s})\right)}_{\text{increment of similarity between recommendations}} \tag{9}$$

Here, the parameter β is used to control the degree of similarity between recommendations. At t-th iteration of Eq. 9, $R_{t,s} \coloneqq R_{t-1,s} + [r_h]$ where $R_{0,s} = []$.

Both Eq. 7 and 9 can be optimized by the **FG-MAP** algorithm mentioned in [4] [5]

User Intent Modeling. Only having cross-domain and within-domain re-ranking strategies is not enough because we need to figure out when to use cross-domain re-ranking and when to use within-domain re-ranking. We leverage the heuristic that users who prefer cross-domain complementary recommendations for a query item might add more diverse items during the next-k purchases while users who prefer within-domain complementary recommendations for a query item might add less diverse items during the next-k purchases.

Formally, given a query item q at time t and a list of next-k items $B_q = \{b_{t+1}, ..., b_{t+k}\}$ purchased by the user u, we leverage the taxonomy information $tax(\cdot)$ [6] to estimate how much diversity the user u prefers. Let $B_{T,q} = \{tax(b_{t+1}), ..., tax(b_{t+k})\}$ be the list of departments of the next-k items purchased by the user and $|B_{T,q}|$ be the number of unique elements in $B_{T,q}$. We can compute an estimation of degree of diversity for the query item q and the user u in Eq. 10.

$$z_{u,q} = \frac{|B_{T,q}|}{k} \tag{10}$$

However, the score $z_{u,q}$ is at user-item level and not stable due to the sparsity issue. We then extend it to a score at user-department level to reduce the sparsity, shown in Eq. 11, where $dept_i$ is the department i and the score $z_{u,dept_i}$ is an average of score $z_{u,q}$ for any query items satisfying $tax(q) = dept_i$.

$$z_{u,dept_i} = \frac{1}{N} \sum_{\{q|tax(q)=dept_i\}}^{N} z_{u,q} \tag{11}$$

We can use a threshold $T \in [0, 1]$ to binarize $z_{u,dept_i}$ learnt from the training data. If $z_{u,dept_i} < T$, the user u prefers the more converged scope of complementary items for the query items under the department $dept_i$, otherwise, the user u might prefer more diversified complementary items because u tends to add items from different departments during the next-k purchases. We can combine the score $z_{u,dept_i}$ with the re-ranking strategies of cross-domain and within-domain complementary recommendation to develop a dynamic re-ranking algorithms (shown in Algorithm 1) It provides either cross-domain re-ranking strategy or within-domain re-ranking strategy based on the user intent on a certain department $dept_i$ of the query item q.

We add z_0 as a default value for cold departments of query items which are not seen in the history. z_0 could be initialized by the average of all $z_{u,dept_i}$.

4 EVALUATION

In this section, we evaluate our proposed solution on the publicly available Instacart Dataset [7]. We also conduct a parameter analysis of re-ranking performance with different T.

[4]The diversification re-rank aforementioned could result in R_d = {Eggs, Banana, Cheese, Bread, Sausage, Cereal, Margarine, Yogurt}

[5]The algorithm 1 in [4] .

[6]$tax(\cdot)$ returns the department of the input item.

Algorithm 1 Dynamic Re-ranking of Complementary Recommendation with User Intent

Require: $u, q, R, R_x, \alpha, \beta, T, z_0$, k;
Ensure:
$R_{out} = []$
$dept_i = tax(q)$
if $z_{u,dept_i}$ available **then**
  use $z_{u,dept_i}$
else
  $z_{u,dept_i} = z_0$
end if
if $z_{u,dept_i} > T$ **then**
  use R, α to compute R_d by Eq. 7 and **FG-MAP** with k iterations (cross-domain)
  $R_{out} := R_d$
else
  use $R + R_x$, β compute R_s by Eq. 9 and **FG-MAP** with k iterations (within-domain)
  $R_{out} := R_s$
end if
return R_{out} as the re-ranked complementary recommendations for u and q

4.1 Evaluation Setting

The Instacart dataset [7] has 49,677 distinct items, 134 distinct aisles, 21 distinct departments and 206,209 distinct users. We train **triple2vec** model on the Instacart training dataset, with embedding dimension of 100, batch size of 128, initial learning rate of 0.05 and Stochastic Gradient Descent optimizer. We also compute $z_{u,dept_i}$ for each pair of $(u, dept_i)$ for the next-5 purchase ($k = 5$ in Eq. 10). When evaluating the re-ranking strategies, we compare results before and after the re-rank. Given a query item q, a user u and recommendations R, R_x generated by **triple2vec** model, we compute the Hit-Rate@5 and Normalized Discounted Cumulative Gain (NDCG@5) for raw complementary recommendation R and the re-ranked complementary recommendations by (1) cross-domain, (2) within-domain and (3) combined strategy with user intent, on the task of next-item prediction. The reason why we focus on next-5 purchase is because the user intent might last for a short period and we want to study the impact of two different complementary recommendations in the top recommendations. If we consider bigger k, it is likely to introduce diversity for recommendations. Here, we define $R = \{r_1, r_2, r_3, r_4, r_5\}$ and $R_x = \{r_6, r_7, r_8, r_9, t_{10}\}$ to cooperate the metrics of Hit-Rate@5 and NDCG@5. To use cross-domain re-ranking only, we set T to be 0. Similarly, we use $T = 1$ to force within-domain re-ranking. To further understand the trade-off between cross-domain re-ranking and within-domain re-ranking, we evaluate the combined strategy with $T \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. We use $\alpha = \beta = 0.01$ for evaluations.

4.2 Evaluation Results

We evaluate our re-ranking strategies on the Instacart evaluation dataset and the detailed results are shown in Table 1. The cross-domain re-ranking strategy improves the Hit-Rate@5 and NDCG@5 compared with the raw recommendations, while only using within-domain re-ranking strategy reduces the performance. Combining both re-ranking strategies together with a proper T improves the overall performance. Particularly, $T = 0.2$ achieves the best Hit-Rate@5 with and $T = 0.1$ achieves the best NDCG@5. This result is reasonable because $T = 0.2$ means on average users purchase next 5 items under the same department. The evaluation result show better performance for covering users who prefer within-domain recommendations.

Table 1: Detailed Results of Next-Item Prediction

	Hit-Rate@5	NDCG@5
raw recommendation	0.05581	0.03216
T = 0 (cross-domain)	0.05581	0.03379
T = 0.1	0.05612	0.03380
T = 0.2	0.05625	0.03377
T = 0.3	0.05612	0.03371
T = 0.4	0.05558	0.03318
T = 0.5	0.05388	0.03214
T = 0.6	0.05259	0.03133
T = 0.7	0.05261	0.03128
T = 0.8	0.05261	0.03127
T = 0.9	0.05261	0.03127
T = 1 (within-domain)	0.05261	0.03127

Note that, pure within-domain re-ranking reduces both Hit-Rate@5 and NDCG@5. It might be because only showing complementary recommendations in a narrow scope is likely to miss users' interests. If a user is not interested in the first recommended item, this user will be likely not interested in the following recommendations because they are similar. Cross-domain re-ranking improves this by surfacing different complementary items to the top. Now the re-ranked recommendations are more likely to match this user's interests. Combining these two strategies together actually covers the requirements of cross-domain and within-domain complementary recommendations. Those users who prefer within-domain recommendations are now covered by the within-domain re-ranked strategy.

Our result shows that combining cross-domain and within-domain re-ranking strategies improves the overall performance, which indicates simple diversification of recommendations could be further improved by fine-grained re-ranking.

5 CONCLUSION AND FUTURE WORK

We focus on the re-ranking of complementary recommendations in Online Grocery and point out the (1) cross-domain and (2) within-domain requirements in complementary recommendations. To fulfill these two requirements, we propose a re-ranking solution based on DPP on the raw complementary recommendations by combining cross-domain and within-domain re-rankings dynamically. We demonstrate the effectiveness of our solution on the publicly available Instacart dataset.

There is scope of future work in the following directions: (1) a better estimation or representation of the users' intent on cross-domain and within-domain strategy, (2) applying techniques such as reinforcement learning or multi-armed bandit.

REFERENCES

[1] Mohamed Hussein Abdi, George Onyango Okeyo, and Ronald Waweru Mwangi. 2018. Matrix Factorization Techniques for Context-Aware Collaborative Filtering Recommender Systems: A Survey. *Computer and Information Science* 11, 2 (2018), 1–10. https://doi.org/10.5539/cis.v11n2p1

[2] Gediminas Adomavicius and YoungOk Kwon. 2012. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques. *IEEE Trans. Knowl. Data Eng.* 24, 5 (2012), 896–911. https://doi.org/10.1109/TKDE.2011.15

[3] Oren Barkan and Noam Koenigstein. 2016. Item2vec: Neural Item Embedding for Collaborative Filtering. In *Proceedings of the Poster Track of the 10th ACM Conference on Recommender Systems (RecSys 2016), Boston, USA, September 17, 2016 (CEUR Workshop Proceedings)*, Vol. 1688. CEUR-WS.org. http://ceur-ws.org/Vol-1688/paper-13.pdf

[4] Laming Chen, Guoxin Zhang, and Eric Zhou. 2018. Fast Greedy MAP Inference for Determinantal Point Process to Improve Recommendation Diversity. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada*. 5627–5638. http://papers.nips.cc/paper/7805-fast-greedy-map-inference-for-determinantal-point-process-to-improve-recommendation-diversity

[5] Jorge Díez, David Martínez-Rego, Amparo Alonso-Betanzos, Oscar Luaces, and Antonio Bahamonde. 2019. Optimizing novelty and diversity in recommendations. *Prog. Artif. Intell.* 8, 1 (2019), 101–109. https://doi.org/10.1007/s13748-018-0158-4

[6] Valeria Fionda and Giuseppe Pirrò. 2019. Triple2Vec: Learning Triple Embeddings from Knowledge Graphs. *CoRR* abs/1905.11691 (2019). arXiv:1905.11691 http://arxiv.org/abs/1905.11691

[7] The instacart online grocery shopping dataset 2017. Accessed on Dec. 2017. https://www.instacart.com/datasets/grocery-shopping-2017.

[8] Alex Kulesza and Ben Taskar. 2012. Determinantal point processes for machine learning. *CoRR* abs/1207.6083 (2012). arXiv:1207.6083 http://arxiv.org/abs/1207.6083

[9] Jian-Guo Liu, Kerui Shi, and Qiang Guo. 2012. Solving the accuracy-diversity dilemma via directed random walks. *CoRR* abs/1201.6278 (2012). arXiv:1201.6278 http://arxiv.org/abs/1201.6278

[10] Yong Liu, Yinan Zhang, Qiong Wu, Chunyan Miao, Lizhen Cui, Binqiang Zhao, Yin Zhao, and Lu Guan. 2019. Diversity-Promoting Deep Reinforcement Learning for Interactive Recommendation. *CoRR* abs/1903.07826 (2019). arXiv:1903.07826 http://arxiv.org/abs/1903.07826

[11] Julian J. McAuley, Rahul Pandey, and Jure Leskovec. 2015. Inferring Networks of Substitutable and Complementary Products. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015*. ACM, 785–794. https://doi.org/10.1145/2783258.2783381

[12] Lijing Qin and Xiaoyan Zhu. 2013. Promoting Diversity in Recommendation by Entropy Regularizer. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. IJCAI/AAAI, 2698–2704. http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6511

[13] Panagiotis Symeonidis and Andreas Zioupos. 2016. *Matrix and Tensor Factorization Techniques for Recommender Systems*. Springer. https://doi.org/10.1007/978-3-319-41357-0

[14] Mengting Wan, Di Wang, Jie Liu, Paul Bennett, and Julian J. McAuley. 2018. Representing and Recommending Shopping Baskets with Complementarity, Compatibility and Loyalty. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*. ACM, 1133–1142. https://doi.org/10.1145/3269206.3271786

[15] Mark Wilhelm, Ajith Ramanathan, Alexander Bonomo, Sagar Jain, Ed H. Chi, and Jennifer Gillenwater. 2018. Practical Diversified Recommendations on YouTube with Determinantal Point Processes. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*. ACM, 2165–2173. https://doi.org/10.1145/3269206.3272018

[16] Qiong Wu, Yong Liu, Chunyan Miao, Yin Zhao, Lu Guan, and Haihong Tang. 2019. Recent Advances in Diversified Recommendation. *CoRR* abs/1905.06589 (2019). arXiv:1905.06589 http://arxiv.org/abs/1905.06589

[17] Yin Zhang, Haokai Lu, Wei Niu, and James Caverlee. 2018. Quality-aware neural complementary item recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018, Vancouver, BC, Canada, October 2-7, 2018*, Sole Pera, Michael D. Ekstrand, Xavier Amatriain, and John O'Donovan (Eds.). ACM, 77–85. https://doi.org/10.1145/3240323.3240368

[18] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. 2005. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web, WWW 2005, Chiba, Japan, May 10-14, 2005*. ACM, 22–32. https://doi.org/10.1145/1060745.1060754