



OPEN ACCESS

EDITED BY

Dawei Yin,
Baidu, China

REVIEWED BY

Yanjie Fu,
University of Central Florida, United States
Xianzhi Wang,
University of Technology Sydney, Australia

*CORRESPONDENCE

Nimesh Sinha
✉ nimesh280@gmail.com

SPECIALTY SECTION

This article was submitted to
Recommender Systems,
a section of the journal
Frontiers in Big Data

RECEIVED 20 June 2022

ACCEPTED 27 February 2023

PUBLISHED 22 March 2023

CITATION

Ma L, Sinha N, Cho JHD, Kumar S and Achan K
(2023) Personalized diversification of
complementary recommendations with user
preference in online grocery.
Front. Big Data 6:974072.
doi: 10.3389/fdata.2023.974072

COPYRIGHT

© 2023 Ma, Sinha, Cho, Kumar and Achan. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Personalized diversification of complementary recommendations with user preference in online grocery

Luyi Ma¹, Nimesh Sinha^{2*}, Jason H. D. Cho¹, Sushant Kumar¹ and Kannan Achan¹

¹Walmart Global Tech, Sunnyvale, CA, United States, ²DoorDash, San Francisco, CA, United States

Complementary recommendations play an important role in surfacing the relevant items to the customers. In the cross-selling scenario, some customers might present more exploratory shopping behaviors and prefer more diverse complements, while other customers show less exploratory (or more conventional) shopping behaviors and want to have a deep dive of less diverse types of complements. The existence of two distinct shopping behaviors reflects users' different shopping intents and requires complementary recommendations to be adaptable based on the user's shopping intent. Although many studies focus on improving the recommendations through post-processing techniques, such as user-item-level personalized ranking and diversification of recommendations, they fail to address such a requirement. First, many user-item-level personalization methods cannot explicitly model the preference of users in two types of shopping behaviors and their intent on the corresponding complementary recommendations. Second, most of the diversification methods increase the heterogeneity of the recommendations. However, users' intent on conventional complementary shopping requires more homogeneity of the recommendations, which is not explicitly modeled. The present study tries attempts to solve these problems by the personalized diversification strategies for complementary recommendations. To address the requirement of modeling heterogenized and homogenized complementary recommendations, we propose two diversification strategies, heterogenization and homogenization, to re-rank complementary recommendations based on the determinantal point process (DPP). We use transaction history to estimate users' intent on more exploratory or more conventional complementary shopping. With the estimated user intent scores and two diversification strategies, we propose an algorithm to personalize the diversification strategies dynamically. We demonstrate the effectiveness of our re-ranking algorithm on the publicly available Instacart dataset.

KEYWORDS

diversification, re-ranking, recommender system (RS), complementary recommendation, personalization

1. Introduction

Recommender system is an essential part of the e-commerce business. Recommending relevant items to customers makes the shopping experience more comfortable and time-saving. Online grocery platforms also have a wide variety of recommendation systems placed at various sections of their websites to improve customer journey. One of the

important sections is related to complementary item recommendations for cross-selling. Given a query item, complementary recommendations show the query item's complements to customers who frequently co-purchase to fulfill a particular demand. For example, when a customer purchases a bag of hot dog, she/he might also want to purchase a bag of hot dog buns together. Showing hot dog bun for hot dog as a complementary item recommendation will improve the shopping experience.

However, it is non-trivial to effectively recommend complementary items for a given query item when the users show different co-purchase behaviors, i.e., more exploratory co-purchase or more conventional co-purchase as shown in Figure 1. When a user prefers more exploratory co-purchases, she/he might also prefer to see more heterogeneous item recommendations complementary to the query item because of the intent on exploration. When a user prefers more conventional co-purchases, she/he might favor less diversified or even homogeneous item recommendations complementary to the query item because of the intent on classic combination with a deep comparison. In this case, the diversity of complementary item recommendations should be adaptable to the co-purchase pattern (exploratory vs. conventional) with personalization. Such an adaptation requires not only modeling the diversification of complementary item recommendations for more exploratory shopping intent, but also properly homogenizing the recommendation for the conventional shopping intent. Furthermore, we need to personalize the adaptation for users by their shopping intent.

All these problems become more challenging for online grocery because grocery items are deeply involved in our daily life under so many co-purchase scenarios, such that co-purchase patterns are more diverse and flexible compared with other online marketplaces. For example, in online grocery, tortilla chips have many food complements, such as salsa dip, guacamole dip, and soft drink and also non-food complements such as chip bowl, while in online electronics e-commerce, television (TV) might only have fixed complements related to television shopping.

Diversity of item recommendations could be quantified by item attributes. One of the commonly used attributes is the hierarchical classification of an item in the taxonomy. Figure 2 presents an example of grocery item taxonomy, with *item department*, *item category*, *item type*, and *individual items*. While items from the same category (one level of item classification in the taxonomy) generally share a similar item functionality (e.g., items from the milk category), the department level classification summarizes the diversity of the customer shopping intent better because each department could represent an aspect of daily shopping. In the aforementioned example of tortilla chips, the customer needs to purchase items from multiple departments such as *Deli* and *Beverage*. In our case, we define the diversity of complementary items at the department level (i.e., only items from two different departments contribute to the increment in the recommendation diversity).

Many complementary item recommendation models mainly focused on learning the complementarity between items rather

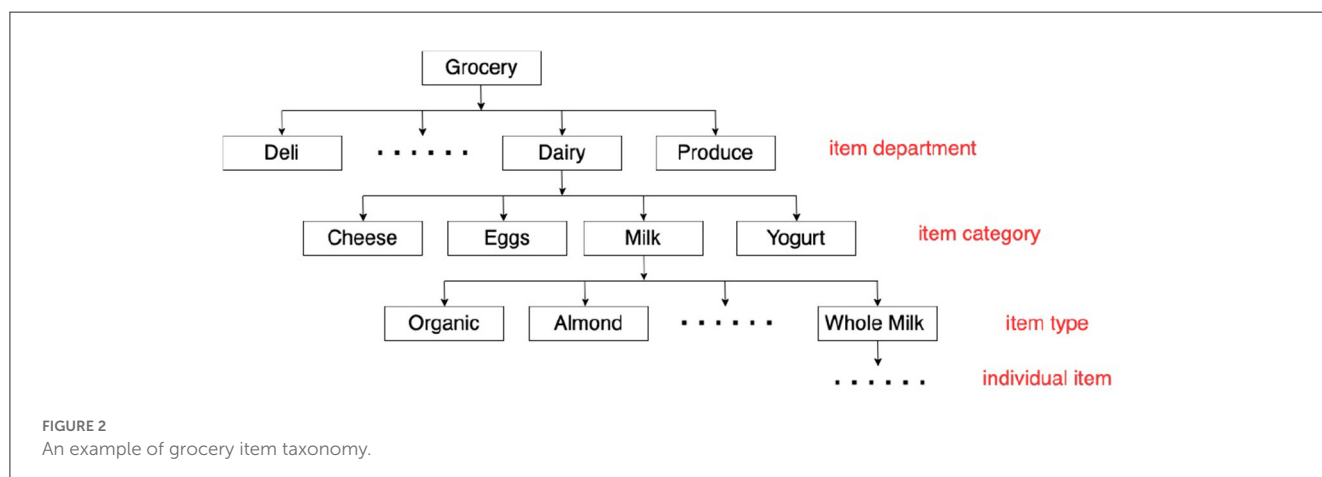
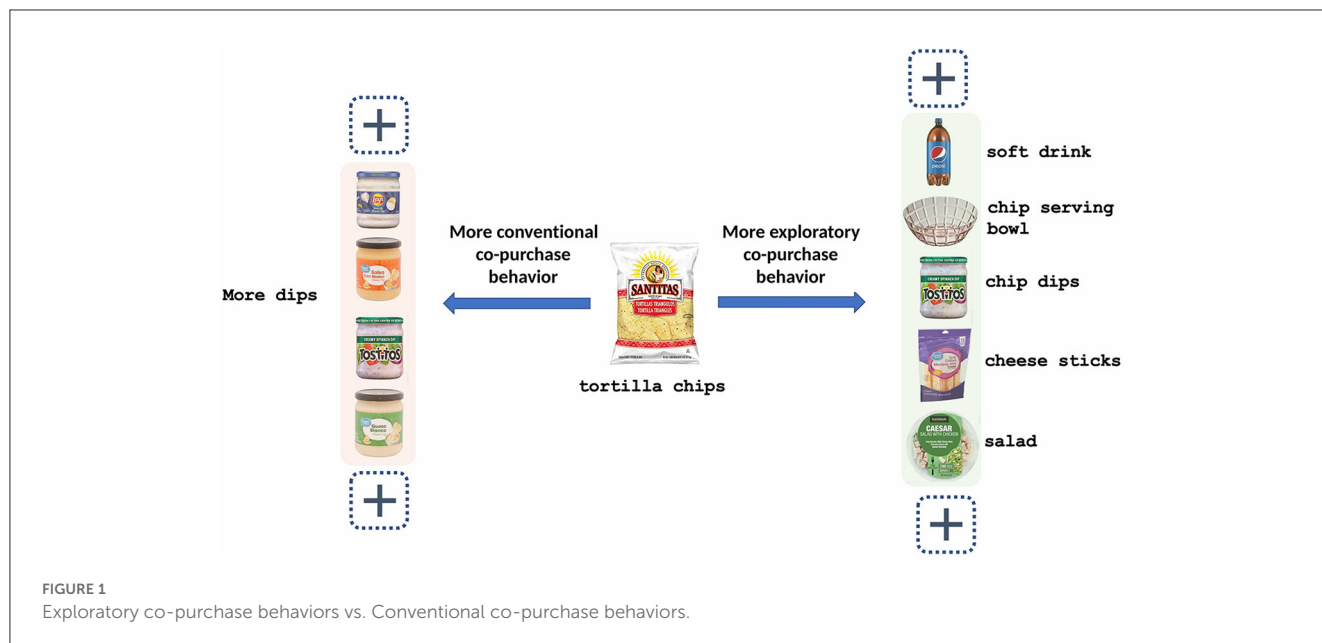
than the personalized adjustment of diversity of complementary item recommendations (McAuley et al., 2015; Barkan and Koenigstein, 2016; Wan et al., 2018; Wang et al., 2018; Zhang et al., 2018; Xu et al., 2019; Liu et al., 2020). Diversification of complementary item recommendations has been recently addressed in Hao et al. (2020) by considering the item type and categories. Unfortunately, it cannot distinguish the demand of users' shopping intent by surfacing more heterogeneous complementary items for exploratory shopping intent or more homogeneous complementary items for conventional shopping intent.

To address these challenges, we utilize the complementary item recommendations by the existing models and deploy the re-ranking strategy to balance out both the exploratory and the conventional complementary shopping intent. To illustrate the necessity and effectiveness of adjustable diversification of complementary item recommendations, we study two diversification strategies, **heterogenization** and **homogenization**. For heterogenization, we focus on diversified complementary item recommendations. We use the re-ranking strategy based on determinantal point process (DPP) to diversify our complementary item recommendations by the existing models. The more diversified complementary item recommendations can fit the exploratory shopping intent. For homogenization, we enforce the homogeneity of the complementary item recommendations by a re-ranking strategy based on a modified DPP. In this case, the modified DPP will encourage the homogeneity of the recommendations for conventional shopping intent.

To further address the personalized adjustment of diversification strategies, we estimate the user shopping intent (exploratory vs. conventional) by user shopping history. The estimated user shopping intent will guide the recommender system to select the proper diversification re-ranking (heterogenization vs. homogenization) for the complementary item recommendations and address the user shopping intent. We summarize our contributions as follows:

- We introduce the concept of exploratory and non-exploratory shopping demands from customer behavior to the modeling problem of complementary item recommendations, which has not been addressed before.
- We further address the requirement of personalizing the demand of exploratory and non-exploratory recommendations based on the diversity of recommendations and proposed a personalized ranking model for complementary item recommendations for the dynamic adjustment.
- We show the effectiveness of our proposed solution and conducted case studies on customer shopping intent on the publicly available dataset.

The rest of this article is structured as follows: we summarize the related articles in Section 2 and introduce the preliminaries of our model in Section 3. After that, we propose our model in Section 4. We provide the evaluation and result analysis in Section 5 and conclude our article in Section 6.



2. Related works

2.1. Complementary recommendations

Many studies have focused on the complementary item recommendations. Embedding-based methods, such as Barkan and Koenigstein (2016) and Wan et al. (2018), collaboratively learn the complementary item relationship from the co-purchase data. Another way of using the co-purchase data is to construct the co-purchase graph and apply graph neural networks on it (McAuley et al., 2015; Wang et al., 2018; Liu et al., 2020). They use the co-purchase records as labels for link predictions based on the distance between item embeddings. In addition to vector item embeddings, Gaussian embedding is also explored in Ma et al. (2021) to address the noise in the co-purchase data for better complementarity learning. In addition to the co-purchase data, many types of auxiliary data are incorporated into the modeling, such as the multimodal data of items (Zhang et al., 2018) and the shopping context (Xu et al., 2019). Diversified complementary recommendation is studied in Hao et al. (2020) by leveraging the

product-type information to improve the diversity. However, it focuses on the diversified recall process rather than the ranking process as our article targets.

In our study, we leverage the **triple2vec** in Wan et al. (2018) to learn the complementary item embedding due to its effectiveness in learning the item vector embeddings from the co-purchase data.

2.2. Recommendation diversification

For a long time, not much importance was given to diversity in the recommendations, as it is challenging to achieve both high accuracy and diversity at the same time. This is called *accuracy diversity dilemma* (Liu et al., 2012). Novelty and diversity of items have been improved by penalizing accuracy (Diez et al., 2019). Diversity has also been captured in an entropy regularizer (Qin and Zhu, 2013). Post-processing methods for diversity have been proposed to improve the personalized recommendations generated by collaborative filtering (Adomavicius and Kwon, 2012).

Determinantal point process (DPP) has been used for making personalized diversified recommendations and DPP models are probabilistic models with a lot of applications (Kulesza and Taskar, 2012). They have 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 (Wilhelm et al., 2018). 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 (Liu et al., 2019). However, they fail to give much attention to maintaining the delicate balance between the requirement of distinct diversity strategies for the exploratory and conventional shopping intents. Our proposed method focuses on the combined re-ranking strategy for exploratory and conventional user shopping intent on complementary recommendations.

3. Preliminaries

In this section, we first revisit the base model for complementary item recommendations, **triple2vec** (Wan et al., 2018), and for generating the item embedding used for diversification. We choose **triple2vec** as our baseline model for complementary item recommendation because we do not assume that transaction data (i.e., product IDs and user IDs) are the only available input due to its high accessibility for e-commerce systems and that there are no additional contexts such as click/view and user profiles (e.g., age and gender). Then, we introduce DPP for recommendation diversification and its basic setting.

3.1. Skip-gram-based item embedding and triple2vec

Skip-gram-based methods for item embedding leverage the item co-occurrence signal (e.g., co-purchase of items). Models for complementary item recommendations such as McAuley et al. (2015) and Barkan and Koenigstein (2016) exactly use the item co-occurrence signal to model item complementarity. **triple2vec** in Wan et al. (2018) introduced the cohesion of (*item*, *item*, and *user*) triplets that reflect the 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 downstream applications.

In **triple2vec**, a triplet of (q, r, u), $q \in V, r \in V, u \in U$, represents the user-item and the item-item relationship, 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 Equation (1), where f_q, g_r are two sets of representations for items (q, r) and h_u is the user embedding.

$$s_{q,r,u} = \overbrace{f_q^T g_r}^x + \underbrace{f_q^T h_u + g_r^T h_u}_y \quad (1)$$

x and y in Equation (1) indicate the item-to-item complementarity and user-to-item compatibility, respectively. The loss function \mathcal{L} in Equation (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 $R = \{r_1, \dots, r_m\}$ with the highest score of $f_q^T g_r$.

3.2. Recommendation diversification and determinantal point process

Improving the diversity of recommendations benefits the recommender systems because it introduces novelty and better topic coverage (Ziegler et al., 2005). Many studies on diversification follow the setting of bi-criterion optimization problem, which balances the relevance (between the query and recalled elements) and diversity (Wu et al., 2019). Particularly, diversity can be further divided into two types, (1) individual diversity¹ and (2) aggregate diversity² (Wu et al., 2019). We focus on the individual diversity in this study to adjust the diversity of complementary recommendations, given a user's intent.

The determinantal point process (DPP) is a probabilistic model that is good at modeling repulsion. The recent study (Chen et al., 2018) applies DPP to diversification of item recommendations and develops the fast greedy MAP inference to generate diversified recommendations. Our study is based on DPP with the fast greedy MAP inference in Chen et al. (2018). We introduce details of DPP and the fast greedy MAP inference following the notation in Chen et al. (2018). For the rest of our article, we denote the fast greedy MAP inference as **FG-MAP**.

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 the 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.

Formally, the DPP on a discrete set $Z = \{1, 2, \dots, M\}$ 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} does not give zero probability to the empty set, there exist a square, positive semidefinite (PSD) and real matrix $\mathbf{L} \in \mathbb{R}^{M \times M}$, which satisfies Equation (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 the sub-matrix extracted from \mathbf{L} based on elements in Y . Equation (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 Equation (4).

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

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

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

When Z becomes the item pools for complementary recommendations $R = \{r_1, \dots, r_m\}$ recalled by the item representations (i.e., item embedding learned by **triple2vec**), the 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 study, we adapt DPP and **FG-MAP**, with \mathbf{L} defined in Equation (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 items embedding from the **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) .

4. Diversification strategies

As aforementioned, the diversification strategy for complementary item recommendations in online grocery can fall into two types, heterogenization and homogenization, for exploratory and conventional complementary shopping intent, respectively. In this section, we first introduce the proposed diversification strategies based on DPP. Later, we present our user shopping intent modeling and the selection of diversification strategies with personalization.

³ For more details of the fast greedy MAP inference algorithm, please refer to [Chen et al. \(2018\)](#).

4.1. Strategy 1: Heterogenization

The heterogenization strategy for complementary item recommendation can be achieved by increasing the diversity in the complementary recommendations R recalled by a complementary item recommendation model, i.e., **triple2vec**. It could fulfill the users' intent on exploratory shopping by showing more diverse recommendations. 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 do not 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 embeddings. Equation (7) shows the modified objective function for diversification re-rank.

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

At t th iteration, $R_{t,d} := R_{t-1,d} + [r_j]$, 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.

4.2. Strategy 2: Homogenization

The homogenization strategy is different from the heterogenization strategy. We need to surface more items that 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 $R = \{\text{eggs, cheese, bread, margarine, banana, sausage, yogurt, cereal}\}$. If we want to address the homogenization strategy, the re-ranked recommendations could be $R_s = \{\text{eggs, cheese, margarine, yogurt, banana, bread, sausage, cereal}\}$ ⁴. We encourage more homogeneousness in this strategy while keeping the complementary relationship between recommendations and the query item. R_s surfaces more items under the Dairy & Eggs domain such as cheese and yogurt. The homogenization strategy can be promoted by the similarity of items among the recall set of complementary recommendations. Unlike diversification for the heterogenization strategy which is diverging the item relationship, boosting the homogeneity in the recommendations is more stable. We can mine candidate items

⁴ However, the diversification re-rank aforementioned could result in $R_d = \{\text{eggs, banana, cheese, bread, sausage, cereal, margarine, yogurt}\}$.

in a bigger recall set. Formally, we recall extra complementary items $R_x = \{r_{m+1}, \dots, r_n\}$ 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 Equation (8) and apply DPP to 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 Equation (7), and we can have a new re-ranking logic on the extended item pool $R + R_x$, shown in Equation (9).

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

Here, the parameter β is used to control the degree of similarity between recommendations. At t th iteration of Equation (9), $R_{t,s} := R_{t-1,s} + [r_h]$, where $R_{0,s} = []$. Both Equations (7), (9) can be optimized by the FG-MAP algorithm mentioned in Chen et al. (2018)⁵.

4.3. User intent modeling and diversification strategy selection

Only having two re-ranking strategies is not enough because we need to figure out the selection of two diversified re-ranking strategies for a certain user. We leverage the heuristic that users who usually add more diverse items during the next- k purchases would prefer more exploratory complementary shopping with the heterogenization strategy, while users who commonly add less diverse items during the next- k purchases would prefer more conventional complementary shopping with the homogenization strategy.

Formally, given a query item q at time t and a list of next- k items $B_q = \{b_{t+1}, \dots, b_{t+k}\}$ purchased by a user u , we leverage the taxonomy information $\text{tax}(\cdot)$ ⁶ to estimate how much diversity the user u prefers. Let $B_{T,q} = [\text{tax}(b_{t+1}), \dots, \text{tax}(b_{t+k})]$ 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 estimate the degree of diversity for the query item q and the user u in Equation (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, as shown in Equation (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 $\text{tax}(q) = \text{dept}_i$.

$$z_{u,\text{dept}_i} = \frac{1}{N} \sum_{\{q|\text{tax}(q)=\text{dept}_i\}} z_{u,q} \tag{11}$$

⁵ Algorithm 1 in Chen et al. (2018).

⁶ $\text{tax}(\cdot)$ returns the department of the input item.

We treat the score z_{u,dept_i} as the user intent score of exploratory complementary shopping and apply a threshold $T \in [0, 1]$ to binarize z_{u,dept_i} learnt from the training data. If $z_{u,\text{dept}_i} < T$, the user u prefers the more conventional complementary shopping under the department dept_i , otherwise, the user u might prefer more exploratory 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 heterogenization and homogenization strategies to develop a dynamic re-ranking algorithm for complementary item recommendations (summarized in Algorithm 1). It provides either diversified re-ranking strategy based on the user intent on a certain department dept_i of the query item q .

```

Require:  $u, q, R, R_x, \alpha, \beta, T, z_0, k$ 
Ensure:
1:  $R_{out} = []$ 
2:  $\text{dept}_i = \text{tax}(q)$ 
3: if  $z_{u,\text{dept}_i}$  available then
4:   use  $z_{u,\text{dept}_i}$ 
5: else
6:    $z_{u,\text{dept}_i} = z_0$ 
7: end if
8: if  $z_{u,\text{dept}_i} > T$  then
9:   use  $R, \alpha$  to compute  $R_d$  by Equation (7) and
     FG-MAP with  $k$  iterations (heterogenization
     strategy)
10:   $R_{out} := R_d$ 
11: else
12:  use  $R + R_x, \beta$  compute  $R_s$  by Equation (9)
     and FG-MAP with  $k$  iterations (homogenization
     strategy)
13:   $R_{out} := R_s$ 
14: end if
15: return  $R_{out}$  as the re-ranked complementary
     recommendations for  $u$  and  $q$ 

```

Algorithm 1. Dynamic re-ranking of complementary recommendation with user intent.

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} .

5. Evaluation

In this section, we evaluate our proposed solution on the publicly available Instacart dataset (Instacart, 2017). We also conduct a parameter analysis of re-ranking performance with different T .

5.1. Evaluation setting

The Instacart dataset (Instacart, 2017) has 49,677 distinct items, 134 distinct aisles, 21 distinct departments, and 206,209 distinct users. We train the **triple2vec** model on the Instacart training

dataset, with an embedding dimension of 100, a batch size of 128, an initial learning rate of 0.05, and a 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 Equation 10). When evaluating the re-ranking strategies, we compare the results before and after the re-rank. Given a query item q , a user u , and recommendations R, R_x generated by the **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) heterogenization only, (2) homogenization only, and (3) combining heterogenization and homogenization strategies with user intent scores dynamically on the task of next-item prediction. The reason why we focus on the 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 on the top recommendations. If we consider bigger k , it is likely to introduce diversity in recommendations. Here, we define $R = \{r_1, r_2, r_3, r_4, r_5\}$ and $R_x = \{r_6, r_7, r_8, r_9, r_{10}\}$ to cooperate the metrics of Hit-Rate@5 and NDCG@5.

Since it is a novel study on exploratory vs. non-exploratory user behaviors for complementary item recommendations, it is hard to find proper baselines. We choose three baseline models for comparison. (1) As aforementioned, we only consider the transaction data due to its high accessibility, we consider the raw recommendations from **triple2vec** for pure complementary item recommendations. (2) The second baseline model is the diversified recommendation by DPP for the pure heterogenization strategy. (3) Similarly, we use $T = 1$ to force homogenization and generate our third baseline model for comparisons.

To further understand the trade-off between heterogenization and homogenization strategies, 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.

5.2. Evaluation results

We evaluate our re-ranking strategies on the Instacart evaluation dataset and the detailed results are shown in Table 1. The heterogenization strategy improves the Hit-Rate@5 and NDCG@5 compared with the raw recommendations, while only using homogenization 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 the next-5 items under the same department. The evaluation results show a better performance for covering users who prefer conventional complementary shopping with the homogenization strategy.

Note that, only applying the homogenization strategy 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 (see Section 5.3 for more details). If a user is not interested in the first recommended item, this user will likely not be interested in the following recommendations because they are similar. The heterogenization strategy improves this feature

TABLE 1 Detailed results of next-item prediction.

	Hit-Rate@5	NDCG@5
Raw recommendation (triple2vec)	0.05581	0.03216
$T = 0$ (heterogenization only by DPP)	0.05581	0.03379
$T = 0.1$	0.05612	<u>0.03380</u>
$T = 0.2$	<u>0.05625</u>	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
$T = 1$ (homogenization only)	0.05261	0.03127

Hit-Rate@5 and NDCG@5 are shown. The highest score is highlighted by underline.

by surfacing different complementary items to the top. Now, the re-ranked recommendations are more likely to hit this user's interests. Combining these two strategies together actually covers the requirements of exploratory and conventional complementary shopping intents from users.

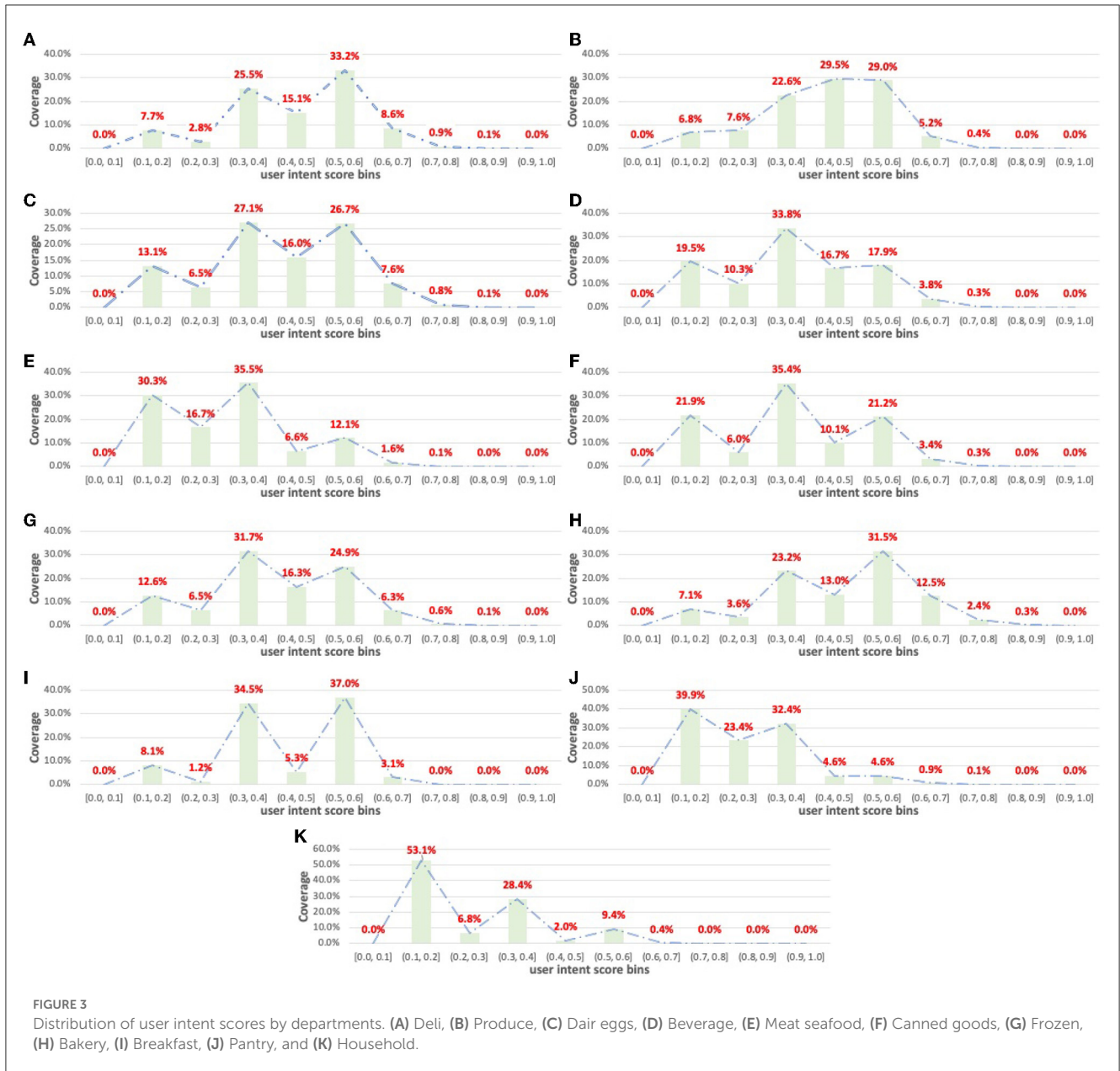
In summary, our results show that combining two strategies dynamically improves the overall performance, compared with only using a single diversification strategy.

5.3. User intent modeling

To further indicate the necessity of personalized diversification strategies for complementary item recommendations, we visualize the distribution of user intent score $z_{u,dept_i}$ by departments. Figure 3 summarizes the distribution. We can see that for a given department, the user intent scores distribute differently. For example, the majority of the user intent scores in Deli and Produce departments follow in the range [0.3, 0.6], which indicates that users tend to shop more diverse items when the query items are from these departments. Dairy and Beverage have similar results such as Deli and Produce. However, for departments such as Household and Pantry, the majority of the user intent scores are in the range [0.1, 0.4]. Compared with other departments, the users tend to purchase more homogeneous items when the query items are from Household and Pantry, which is reasonable because these departments usually cover most of the department-related shopping demands and correlate less with other departments. The aforementioned observation addresses the requirement of tuning the diversification of complementary item recommendations with personalization.

6. Conclusion and future work

We focus on the re-ranking of complementary recommendations in online grocery and point out the exploratory and conventional complementary shopping intents from users. To



fulfill these two user intents, we propose two re-ranking strategies, heterogenization and homogenization, based on DPP on the raw complementary recommendations and dynamically combine two re-rankings as a final solution to improve the performance. We demonstrate the effectiveness of our solution on the publicly available Instacart dataset.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

NS helped the experiments. JC, SK, and KA help the iteration of the research ideas and the design during this research. All authors contributed to the article and approved the submitted version.

Conflict of interest

LM, JC, SK, and KA were employed by Walmart Global Tech. NS was employed by DoorDash.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Adomavicius, G., and Kwon, Y. (2012). Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Trans. Knowl. Data Eng.* 24, 896–911. doi: 10.1109/TKDE.2011.15
- Barkan, O., and Koenigstein, N. (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, volume 1688 of CEUR Workshop Proceedings* (Boston, MA: CEUR-WS.org).
- Chen, L., Zhang, G., and Zhou, E. (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* (Montreal, QC), 5627–5638.
- Diez, J., Martínez-Rego, D., Alonso-Betanzos, A., Luaces, O., and Bahamonde, A. (2019). Optimizing novelty and diversity in recommendations. *Prog. Artif. Intell.* 8, 101–109. doi: 10.1007/s13748-018-0158-4
- Hao, J., Zhao, T., Li, J., Dong, X. L., Faloutsos, C., Sun, Y., et al. (2020). "P-companion: a principled framework for diversified complementary product recommendation," in *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, eds M. d'Aquin, S. Dietze, C. Hauff, E. Curry, and P. Cudré-Mauroux (Ireland: ACM), 2517–2524.
- Instacart (2017). *Instacart Market Basket Analysis*. Available online at: <https://www.kaggle.com/c/instacart-market-basket-analysis/>
- Kulesza, A., and Taskar, B. (2012). Determinantal point processes for machine learning. *CoRR*, abs/1207.6083. doi: 10.1561/9781601986290
- Liu, J., Shi, K., and Guo, Q. (2012). Solving the accuracy-diversity dilemma via directed random walks. *CoRR*, abs/1201.6278. doi: 10.1103/PhysRevE.85.016118
- Liu, Y., Gu, Y., Ding, Z., Gao, J., Guo, Z., Bao, Y., et al. (2020). "Decoupled graph convolution network for inferring substitutable and complementary items," in *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, eds M. d'Aquin, S. Dietze, C. Hauff, E. Curry and P. Cudré-Mauroux (Ireland: ACM), 2621–2628.
- Liu, Y., Zhang, Y., Wu, Q., Miao, C., Cui, L., Zhao, B., et al. (2019). Diversity-promoting deep reinforcement learning for interactive recommendation. *CoRR*, abs/1903.07826. doi: 10.48550/arXiv.1903.07826
- Ma, L., Xu, J., Cho, J. H. D., Körpeoglu, E., Kumar, S., and Achan, K. (2021). "NEAT: a label noise-resistant complementary item recommender system with trustworthy evaluation," in *2021 IEEE International Conference on Big Data (Big Data)*, Y. Chen, H. Ludwig, Y. Tu, U. M. Fayyad, X. Zhu, X. Hu, S. Byna, X. Liu, J. Zhang, S. Pan, V. Papalexakis, J. Wang, A. Cuzzocrea, and C. Ordonez (Orlando, FL: IEEE), 469–479.
- McAuley, J. J., Pandey, R., and Leskovec, J. (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: ACM.), 785–794.
- Qin, L., and Zhu, X. (2013). "Promoting diversity in recommendation by entropy regularizer," in *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence* (Beijing: IJCAI/AAAI), 2698–2704.
- Wan, M., Wang, D., Liu, J., Bennett, P., and McAuley, J. J. (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: ACM), 1133–1142.
- Wang, Z., Jiang, Z., Ren, Z., Tang, J., and Yin, D. (2018). "A path-constrained framework for discriminating substitutable and complementary products in e-commerce," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, WSDM 2018*, eds Y. Chang, C. Zhai, Y. Liu, and Y. Maarek (Marina Del Rey, CA: ACM), 619–627.
- Wilhelm, M., Ramanathan, A., Bonomo, A., Jain, S., Chi, E. H., and Gillenwater, J. (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: ACM), 2165–2173.
- Wu, Q., Liu, Y., Miao, C., Zhao, Y., Guan, L., and Tang, H. (2019). Recent advances in diversified recommendation. *CoRR*, abs/1905.06589. doi: 10.48550/arXiv.1905.06589
- Xu, D., Ruan, C., Körpeoglu, E., Kumar, S., and Achan, K. (2019). Product knowledge graph embedding for e-commerce. *CoRR*, abs/1911.12481. doi: 10.1145/3336191.3371778
- Zhang, Y., Lu, H., Niu, W., and Caverlee, J. (2018). "Quality-aware neural complementary item recommendation," in *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018*, eds S. Pera, M. D. Ekstrand, X. Amatriain, and J. O'Donovan (Vancouver, BC: ACM), 77–85.
- Ziegler, C., McNee, S. M., Konstan, J. A., and Lausen, G. (2005). "Improving recommendation lists through topic diversification," in *Proceedings of the 14th International Conference on World Wide Web, WWW 2005* (Chiba: ACM), 22–32.