Article

ReRec: A Divide-and-Conquer Approach to Recommendation Based on Repeat Purchase Behaviors of Users in Community E-Commerce

Jun Wu [1,2], Yuanyuan Li [1], Li Shi [2], Liping Yang [1], Xiaxia Niu [1] and Wen Zhang [3,*]

1   School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China; wujun@mail.buct.edu.cn (J.W.); lyyletter@126.com (Y.L.); m17856381770@163.com (L.Y.); 15136212624@163.com (X.N.)
2   College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China; simon_shl@126.com
3   College of Economics and Management, Beijing University of Technology, Beijing 100124, China
*   Correspondence: zhangwen@bjut.edu.cn

Abstract: Existing studies have made a great endeavor in predicting users' potential interests in items by modeling user preferences and item characteristics. As an important indicator of users' satisfaction and loyalty, repeat purchase behavior is a promising perspective to extract insightful information for community e-commerce. However, the repeated purchase behaviors of users have not yet been thoroughly studied. To fill in this research gap from the perspective of repeated purchase behavior and improve the process of generation of candidate recommended items this research proposed a novel approach called ReRec (Repeat purchase Recommender) for real-life applications. Specifically, the proposed ReRec approach comprises two components: the first is to model the repeat purchase behaviors of different types of users and the second is to recommend items to users based on their repeat purchase behaviors of different types. The extensive experiments are conducted on a real dataset collected from a community e-commerce platform, and the performance of our model has improved at least about 13.6% compared with the state-of-the-art techniques in recommending online items (measured by F-measure). Specifically, for active users, with $w = 1$ and $N_{(U_A)} \in [5, 25]$, the results of ReRec show a significant improvement (at least 50%) in recommendation. With α and σ as 0.75 and 0.2284, respectively, the proposed ReRec for unactive users is also superior to (at least 13.6%) the evaluation indicators of traditional Item CF when $N_{(U_B)} \in [6, 25]$. To the best of our knowledge, this paper is the first to study recommendations in community e-commerce.

Keywords: ReRec; community e-commerce; repeat purchase; user behavior modeling; recommendation system

1. Introduction

Community e-commerce, which combines the features of traditional e-commerce and mobile commerce, is a representative of community economy [1] and marks the rise of a new commercial ideology. Generally speaking, community e-commerce refers to a novel business model that takes communities as service units and provides a more convenient manner in online shopping than traditional e-commerce for community residents [2,3]. On the one hand, unlike traditional e-commerce that provides products and services all over the world or a country, community e-commerce focuses on a relatively stable group of consumers in a local area as a compatible complement for B2B, B2C and C2C models. On the other hand, like traditional e-commerce, the huge amount of online information and items brings about a heavy burden for online consumers, the users of community e-commerce also suffer from the endless choices and decisions in online shopping and the merchants in community e-commerce are still struggling to predict the interests of users in

online items beforehand, in order to manage their inventories. For this reason, it is urgent to develop a recommendation system for community e-commerce platforms to predict the items that a user may possibly purchase in the near future based on the user's purchase history [4].

In community e-commerce, it is a usual case that a user would purchase the same item repeatedly and periodically. In the scenario of traditional e-commerce, these items will not be recommended to the user repeatedly in the future. However, with the focus on limited number of users in a local area, the recommendation for repeat purchase is crucial for the success of community e-commerce. For instance, by observing user behaviors on the community e-commerce platform T-app (see Section 5.1), we find that from 1 January 2018 to 1 April 2019, among 955 users who have made purchases on T-app, 58.74% have repeat purchases. For these users with repeat purchase, their average repurchase is 3.61 times, and 10.33% of them repurchase the same item six times. In an extreme case, we find that one user has repurchased the same item up to 43 times during the investigated time duration. Among all the 105 types of items, 82 (78.10%) have been repurchased by users. Therefore, it can be seen that repeated purchase behavior is an essential user characteristic that should be paid enough attention to when community e-commerce platforms make recommendation plans.

Existing studies have proposed many recommendation algorithms to predict users' potential interests in items by characterizing user preferences and item characteristics, e.g., the nearest neighborhood based recommendation algorithm [5–7], the matrix factorization based recommendation algorithm [8,9] and the context aware recommendation algorithm [10,11]. Clearly, the basic idea of these algorithms is straightforward—that if a user purchased an item in the past, he or she will also purchase similar items, or items purchased similar users at that time, in the future. However, if an item has already been purchased by a user, then the item will not be recommended by theses algorithms to the user. That is to say, the repeated purchase behavior of users has not yet been thoroughly studied. To fill in this research gap, this paper proposes a novel approach called ReRec (Repeat purchase Recommender) for recommending items to users in community e-commerce. To the best of our knowledge, this paper is the first to conduct item recommendation in community e-commerce. For industrial applications, the proposed method can help manage and identify loyal users and segment users and to improve customer relationship management (CRM) processes. In addition, for managers, this method can also help them formulate precision marketing strategies, recognize the market, and advance the sustainable development of products.

Specifically, ReRec comprises two components. The first component is to model the repeat purchase behaviors of different types of user. This research models the repeat purchase behaviors of the users in community e-commerce. based on their activity in the community and the stability of their interests in items, in a divide-and-conquer manner, using these categories: active users with stable interest (ASI), active users with unstable interest (AUSI), inactive users with stable interest (IASI) and inactive users with unstable interest (IAUSI). The second component is to recommend items to users based on repeat purchase behaviors. This research proposes the ReRec approach in four variants to deal with different types of users and interests, i.e., recommendation for active users with stable interest (ReRec-ASI), recommendation for active users with unstable interest (ReRec-AUSI), recommendation for inactive users with stable interest (ReRec-IASI) and recommendation for inactive users with unstable interest (ReRec-IAUSI). Finally, extensive experiments based on a real community e-commerce platform are conducted and the experimental results demonstrate that the proposed ReRec approach outperforms state-of-the-art techniques significantly.

The rest of this paper is organized as follows. Section 2 states the problem. Section 3 presents related works. Section 4 proposes the ReRec approach. Section 5 conducts the experiments. Section 6 concludes the paper and indicates future work.

2. Problem Statement

The problem studied in this paper is one of recommendation for repeat purchase in community e-commerce, which is different from that of traditional recommendation, such as collaborative filtering [5,6,12]. Essentially, this research can formulate the problem as follows. Assume that there are a set of users as $U = \{u_k | 1 \leq k \leq m\}$, and a set of items as $I = \{i_s | 1 \leq s \leq n\}$ in community e-commerce. The historical sales data until time t is recorded as a matrix $R_{UI}^t = \left\{ r_{u_k i_s}^t \big| 1 \leq k \leq m, 1 \leq s \leq n \right\}$, where $r_{u_k i_s}^t$ is the number of cumulative purchases of the user u_k of the item i_s at t. Note that the user u_k has purchased the item i_s repeatedly and periodically. Let $\widetilde{R}_{u_k i_s}^{t+1}$ be the possibility that the user u_k purchases the item i_s on $t + 1$. We need to speculate the possibilities of user u_k purchasing all the possible items i_s ($1 \leq s \leq n$) on $t + 1$, i.e., $\widetilde{R}_{u_k i_s}^{t+1}$ for all the items i_s on $t + 1$. After deriving the $\widetilde{R}_{u_k i_s}^{t+1}$, it sorts all the possibilities in descending order for user u_k, and uses the top N items as the recommendation list to him or her.

3. Related Works

3.1. Nearest Neighborhood Based Recommendation

On the aspect of nearest neighborhood recommendation, the user-based nearest neighbor method and item-based nearest neighbor method are usually adopted. Resnick et al. [13] propose user-based collaborative filtering to recommend internet news to readers according to readers' rating scores of the internet news. This algorithm firstly calculates the similarity between users, and then for a given user it recommends items that are of interest to similar users to him or her. Considering the large number of items in a recommender system, Sarwar et al. [14] propose item-based recommendation to compute and store items' similarities beforehand in the system and use these similarities in real time when needed to produce a recommendation list for a user. The basic idea of the item-based algorithm is to assume that people will like items that are similar to those items they have purchased before. Since a user has purchased an item in history, he or she would also purchase similar items in the future. The item-based algorithm is very similar to the user-based algorithm. More details about user-based collaborative filtering and item-based collaborative filtering approaches can be found in the available literature [5–7,12,15]. The advantage of the nearest neighborhood algorithms is that they are easy to implement in real practice because of their simple mathematical form and consolidated intuitiveness. However, due to the sparse nature of the historical purchasing data, it is difficult to measure similarities between users and items [8]. Moreover, because the users' interests in items can change very frequently, it makes the computation complexity of real-time recommendation intractable [4].

3.2. Matrix Factorization Based Recommendation

On the aspect of matrix factorization based recommendation, SVD (Single Value Decomposition), SVD++ and NMF (non-negative Matrix Factorization) are the most representative techniques. SVD is a basic matrix decomposition method used in the recommender systems proposed by Chen et al. [9] and Brand [16]. It decomposes the original matrix R with higher dimensions into three matrix multiplication forms with lower dimensions, which brings convenience to matrix calculation and storage. Specifically, SVD decomposes the rating matrix $R_{m \times n}$ into three matrices: left singular vector $P_{m \times n}$, right singular vector $Q_{m \times n}$ and singular value diagonal matrix $S_{m \times n}$ as in Equation (1). Both P and Q matrices are orthogonal and matrix S is a diagonal matrix composed of singular values where all the singular values are aligned in descending order from the largest to the smallest. For all the singular values $S_{ii} \geq 0$, the rank of the rating matrix R is a, and the number of ranks that can be taken is $\{a_h | 1 \leq h \leq \min(m, n)\}$.

$$R = PSQ^T \tag{1}$$

SVD++ is an extension of traditional SVD that takes into account both explicit and implicit information for recommendation [17]. Here, explicit information refers to the users'

rating of an item, and implicit information refers to the users' implicit feedback, such as browsing, buying, and clicking history [8]. The prediction rating $\hat{\gamma}_{ui}$ of SVD++ is defined in Equation (2).

$$\hat{\gamma}_{ui} = b_{ui} + q_i^T(p_u) = \mu + b_u + b_i + q_i^T(p_u) \tag{2}$$

The prediction rating $\hat{\gamma}_{ui}$ is composed of two parts: one is the deviation of different users to different products b_{ui}, the other is the product of the user preference vector p_u and the product feature vector q_i, where μ denotes the benchmark value in the score, b_u is the deviation value of user rating, and b_i is the score deviation of the product. These parameters need to be trained to obtain specific values.

As for NMF, the rating matrix R is approximated by the product of two low-dimensional matrices P and Q, as shown in Equation (3). The NMF problem is non-convex and is usually solved by the gradient descent method [18].

$$R = P^T Q \tag{3}$$

The advantage of the matrix decomposition method is that the users' preference in the item is regarded as the product of two components, i.e., as the users' latent vector representing the user preference and the item's latent vector representing the item's characteristics. Both the user's latent vector and the item's latent vector can be stored in the memory of the recommender system in advance, so it is convenient to compute and predict the user's preference in the item in real time. However, the matrix factorization method also has some defects. Because most view the user item rating matrix from a global perspective and perform matrix decomposition, their performance will be affected due to the large scale of the original user project scoring matrix and the sparse data.

3.3. Context-Aware Recommendation

The collaborative filtering algorithm for recommendation only considers the interactive information between users and items, such as the users' rating matrix for items. Meanwhile, other information, such as contextual situation information during interactive behavior, is generally not considered. A context-aware recommender system (CARS) is used to recommend items to users based on relevant contextual information such as time, weather and location. Contextual information can improve the performance of recommendation and user satisfaction when it is combined with the recommendation algorithm. Gorgoglione et al. [19] report that the context-aware recommendation system can achieve more accurate recommendation by adding contextual information in the experiments, and this recommendation system can significantly increase the platform profit and users' stickability. Time information can consist of the time when users purchase, comment, search or perform other behaviors, or the time of the season or holiday. For instance, around the time of the Dragon Boat Festival in China, users may have a higher preference for rice dumplings than usual.

There are also some studies showing that reasonable use of time information can improve the performance of the recommendation algorithm. Zimdars et al. [20] make use of time series forecasting in collaborative filtering for recommendation. Campos et al. [21] find that there is a time-dependent characteristic of user behaviors in online shopping. For instance, the same user may have different preference patterns on different dates, months and seasons. Liang et al. [22] propose the Time SVD algorithm to integrate four kinds of time-affected factors into time functions and they find that the performance of the Time SVD algorithm is significantly better than that of the traditional SVD algorithm. Qin et al. [23] claim that users of different professions have obvious differences in understanding items, and there is an important relationship between user hierarchy classification and user interest. Traditional collaborative filtering algorithms do not consider change in users' interests. However, in real practice, users' interests are constantly changing with time and the influence of the environment. Therefore, some studies introduce the concept of user interest drift [24,25]. Chen et al. [26] provide a matrix decomposition optimization

model that is constructed to think about the score matrix and combines time information and the original score matrix to improve the recommendation efficiency. Wu et al. [27] include the time factor in order optimize the weights of users' ratings based on time and user similarities.

4. The Proposed Approach

4.1. The Overview of the ReRec Approach

The overall structure of the proposed ReRec approach is shown in Figure 1. As can be seen, the proposed ReRec approach is composed of two components, i.e., repeat purchase behavior analysis and item recommendation. Before the analysis, this research collects the user-item purchase records as a basic data matrix, i.e., the original user-item interaction matrix. In the original user-item interaction matrix, the row label is user ID, the column label is item ID and the element is the cumulative purchase quantity of an item by the corresponding user at time t. Then, users are classified according to activeness and user-items are classified according to stableness. As shown in the yellow area of Figure 1, users are partitioned by mathematical modeling as the active and inactive users, and the items are partitioned as stable and unstable interest. The nodes with black circles denote the users. The nodes with blue circles denote the items. The nodes with red circles denote user–item interaction. The nodes with dotted circles denote the immediate process. In addition, the partition process can be visualized as the user partition matrix and the item partition matrix derived from the original user–item interaction. As for the user partition matrix, a user ID with yellow indicates an active user and a user ID with green indicates an inactive user. As for the item partition matrix, an item ID with red indicates the stable interest of its user and an item ID with blue indicates the unstable interest of its user. Results of the combined user and item classification can be seen in the joint user-item partition matrix.
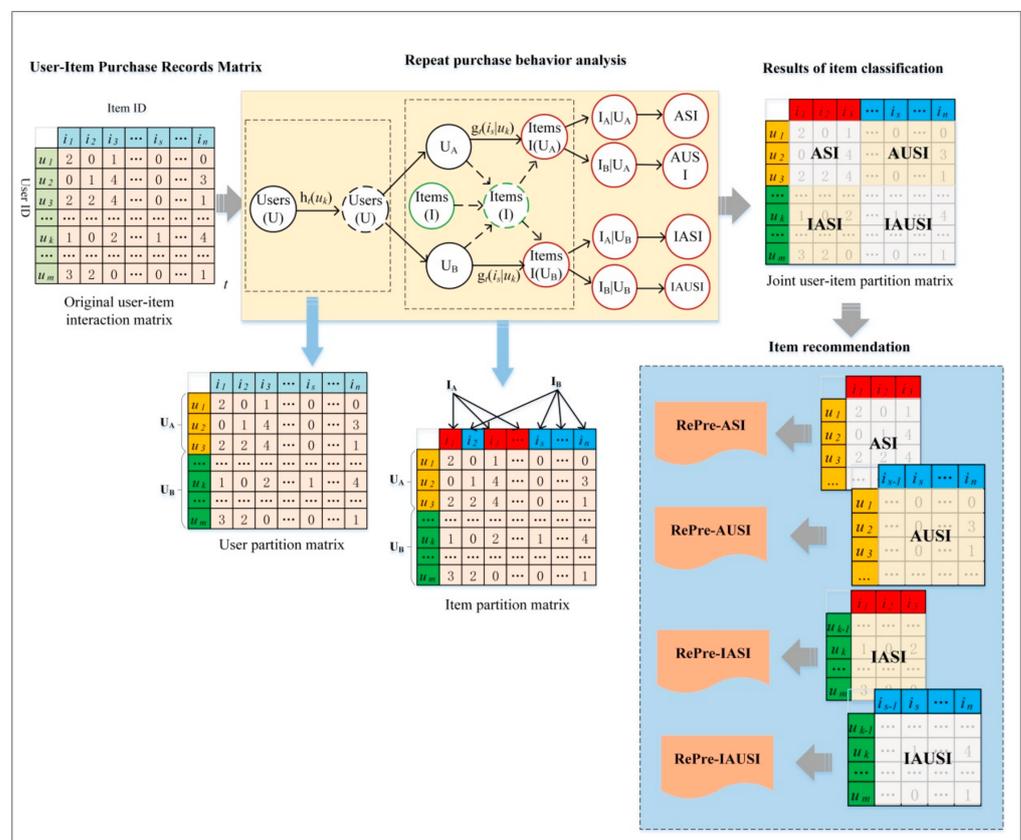

Figure 1. The overall structure of the proposed ReRec approach.

With the above joint user-item partition matrix at hand, this research conducts the item recommendation by using a divide-and-conquer approach. That is, it partitions the repeat purchase behaviors of users into four types: active users with stable interest (ASI), inactive users with stable interest (IASI), active users with unstable interest (AUSI), and inactive users with unstable interest (IAUSI). Furthermore, this research proposes the ReRec recommendation algorithm with its four variants to deal with the repeat purchase behaviors of the four types one by one: the ReRec-ASI approach, the ReRec-IASI approach, the ReRec-AUSI approach and the ReRec-IAUSI approach, which are shown in the blue area of Figure 1.

4.2. Repeat Purchase Behavior Modeling

As community e-commerce focuses on the residents in the local community, the characteristics of user purchase data are different from that of the large-scale e-commerce platform, such as Alibaba, JD and Amazon. Firstly, the consumer group for community e-commerce is relatively stable. That is, the users of community e-commerce are local residents in a limited area such as a residential area, an office area or a campus. Secondly, the number of item types in a community e-commerce is relatively small. Therefore, it can study the characteristics of user-item interactions in a finer granularity than that of the traditional recommendation algorithms and this research holds that the study of fine-grained interactions between users and items is beneficial for improvement of the recommendation algorithm. For this purpose, this research classifies and studies the repeat purchase behaviors of users based on their historical purchase data.

4.2.1. The Classification Models

As the activeness of users is related to the transaction volume of the users' base over time [28], this research adopts a mathematical modeling method to model the behaviors of users along with user-item by purchase volume and the length of time in using community e-commerce.

The mathematical models of user classification are shown in Equations (4)–(6), where $h_t(u_k)$ is the user activeness. $h_t(u_k)$ is positively related to the number of item types purchased by users u_k at time t and the number of days of user u_k when using community e-commerce. This research standardizes these factors to eliminate inconsistent dimensions. Equations (4)–(6) can divide all the users of the e-commerce platform into two types, as active users and inactive users.

$$h_t(u_k) = \frac{h_t'(u_k) - min\{h_t'(u_k)\}}{max\{h_t'(u_k)\} - min\{h_t'(u_k)\}}, u_k \in U \tag{4}$$

$$h_t'(u_k) = \frac{N_{type}^t(u_k) - min\{N_{type}^t(u_k)\}}{max\{N_{type}^t(u_k)\} - min\{N_{type}^t(u_k)\}} * \frac{\Delta t(u_k) - min\{\Delta t(u_k)\}}{max\{\Delta t(u_k)\} - min\{\Delta t(u_k)\}} \tag{5}$$

$$\Delta t(u_k) = t_{last}^{u_k} - t_{start}^{u_k} \tag{6}$$

Meanwhile, the mathematical models of item classification are shown in Equations (7)–(9), where $g_t(i_s|u_k)$ is the interest stableness. $g_t(i_s|u_k)$ is positively related to the total number of item i_s purchased by user u_k before time t and the time interval between the last purchase of user u_k as well as the earliest purchase of item i_s. This research also standardized these factors to eliminate inconsistent dimensions. Equations (7)–(9) can divide users' interests in items into stable interest and unstable interest.

The symbolic definitions of the classification models are shown in Table 1. This research defines the four types of user repeat purchase behaviors based on user activeness and item stableness.

$$g_t(i_s|u_k) = \frac{g_t'(i_s|u_k) - min\{g_t'(i_s|u_k)\}}{max\{g_t'(i_s|u_k)\} - min\{g_t'(i_s|u_k)\}}, u_k \in U, i_s \in I \tag{7}$$

$$g'_t(i_s|u_k) = \frac{N^t_{num}(i_s|u_k) - min\{N^t_{num}(i_s|u_k)\}}{max\{N^t_{num}(i_s|u_k)\} - min\{N^t_{num}(i_s|u_k)\}} * \frac{\Delta t(i_s|u_k) - min\{\Delta t(i_s|u_k)\}}{max\{\Delta t(i_s|u_k)\} - min\{\Delta t(i_s|u_k)\}} \tag{8}$$

$$\Delta t(i_s|u_k) = t^{u_k i_s}_{last} - t^{u_k i_s}_{start} \tag{9}$$

Table 1. Symbolic definition.

Index	Symbols	Definition Description
1	u_k	User u_k , $u_k \in U$, $U = \{u_1, u_2, \cdots, u_k, \cdots, u_m\}$
2	i_s	Item i_s , $i_s \in I$, $I = \{i_1, i_2, \cdots, i_s, \cdots, i_n\}$
3	$h'_t(u_k)$	The activeness of user u_k (in using community e-commerce) at time t
4	$h_t(u_k)$	The activeness of user u_k at time t after standardization
5	$N^t_{type}(u_k)$	The number of item types purchased by users u_k at time t
6	$\Delta t(u_k)$	The number of days of user u_k in using community e-commerce
7	$t^{u_k}_{last}$	The last time that user u_k purchased an item in using community e-commerce
8	$t^{u_k}_{strat}$	The first time that user u_k purchased an item in using community e-commerce
9	$g'_t(i_s|u_k)$	The stability of user u_k purchasing item i_s after standardization
10	$g_t(i_s|u_k)$	The stability of user u_k purchasing item i_s
11	$N^t_{num}(i_s|u_k)$	The total number of item i_s purchased by user u_k before time t
12	$\Delta t(i_s|u_k)$	The time interval between the last purchase of user u_k and the earliest purchase of item i_s
13	$t^{u_k i_s}_{last}$	The last time user u_k purchasing item i_s
14	$t^{u_k i_s}_{start}$	The first time user u_k purchasing item i_s
15	n_{U_A}	The number of users in U_A
16	n_{U_B}	The number of users in U_B

4.2.2. User–Item Interaction

The user interacts with the item when a purchase record occurs. This section defines the user–item interactions theoretically by using mathematical modelling. It calculates the activeness of a user by mathematical model $h_t(u_k)$, and the stableness of user-item by $g_t(i_s|u_k)$. The following shows the definitions of active user and inactive user, and the definitions of stableness interest and unstableness interest.

Definition 1. *Assume that U_A denotes a set of active users. If a user u_k from U_A uses the community e-commence software for a relatively long time and purchases a variety of items, the user is an active user, where if $h_t(u_k) \geq \delta$, user $u_k \in U_A$; δ is the threshold of user activeness, $\delta \in (0, 1)$, and δ is decided by the cumulative distribution of $h_t(u_k)$ of all users.*

Definition 2. *Assume that U_B denotes a set of inactive users. If a user u_k uses the threshold of user activeness for a relatively short time or purchases fewer types of item, the user is an inactive user, where if $h_t(u_k) < \delta$, user $u_k \in U_B$. So the set of all users U consists of U_A and U_B, i.e., $U = U_A \cup U_B$.*

Definition 3. *Let $I_A(u_k)$ be a set of stable interests of user u_k. If the number of item i_s purchased by user u_k is relatively large and the time span of the purchase behavior is long, item i_s is users with stable interest, where $g_t(i_s|u_k) \geq \theta$, $i_s(u_k) \in I_A(u_k)$. θ is the threshold of user-item stableness, $\theta \in (0,1)$, and the value of θ is decided by the cumulative distribution of $g_t(i_s|u_k)$ of all items.*

Definition 4. *Let $I_B(u_k)$ be a set of unstable interests of user u_k. If the number of items i_s purchased by user u_k is relatively small or the time span of the purchase behavior is short, item i_s is users with unstable interest, where $g_t(i_s|u_k) < \theta$ and $i_s(u_k) \in I_B(u_k)$. Thus, the set of all items for user u_k i.e., $I(u_k)$ consists of $I_A(u_k)$ and $I_B(u_k)$, $I(u_k) = I_A(u_k) \cup I_B(u_k)$.*

With the above definitions, the user–item interactions can be divided into four categories. The notation $(I_A|U_A)$ denotes the active users with stable interest (ASI). The notation $(I_B|U_A)$ denotes the active users with unstable interest (AUSI). The notation $(I_A|U_B)$ denotes the inactive users with stable interest (IASI). The notation $(I_B|U_B)$ denotes the inactive users with unstable interest (IAUSI). Figure 2 shows the classification process as a whole. Mathematical functions of the classification are shown as Equation (10).

$$
i_s(u_k) \in \begin{cases}
I_A|U_A & , if\, h_t(u_k) \geq \delta \ and \ g_t(i_s|u_k) \geq \theta \\
I_B|U_A & , if\, h_t(u_k) \geq \delta \ and \ g_t(i_s|u_k) < \theta \\
I_A|U_B & , if\, h_t(u_k) < \delta \ and \ g_t(i_s|u_k) \geq \theta \\
I_B|U_B & , if\, h_t(u_k) < \delta \ and \ g_t(i_s|u_k) < \theta
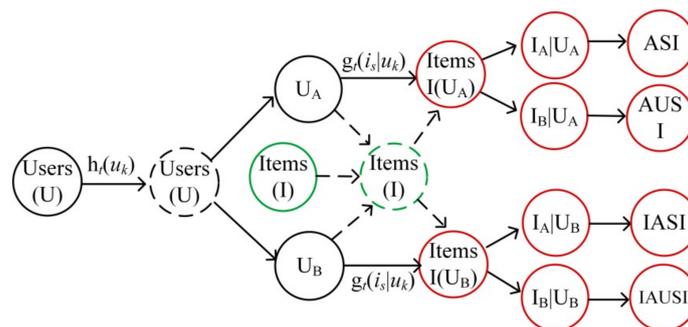\end{cases}
\tag{10}
$$

Figure 2. The classification structure of user–item interactions.

4.3. Item Recommendation

4.3.1. Model of ReRec-ASI

The overall interests of ASI users remain active and they have stable interests in items in $I_A|U_A$. This improves the algorithm upon the repurchase cycle of items. Generally, when a user has just purchased an item, the possibility of repeating the purchase immediately is very low. However, as time goes on, with the user running out of the item, he/she is more likely to make repeated purchase. For this reason, this research could prioritize the recommendation of the item to the user. This research develops a time incentive factor $w_\alpha\left(t^{u_k i_s}\right)$ based on relationship of the last purchase time and the repurchase cycle to improve the user–KNN recommendation algorithm. Due to users in ASI having stable purchase interests, it assumes that their stable interests do not change over time, and the time incentive factor $w_\alpha\left(t^{u_k i_s}\right)$ is a periodic piecewise constant function. The model of the time incentive function is as shown in Equation (11).

$$
w_\alpha\left(t^{u_k i_s}\right) = \begin{cases}
-w, & t^{u_k i_s}_{last} \leq t^{u_k i_s} < t^{u_k i_s}_{last} + \alpha T^{i_s} \\
w, & t^{u_k i_s}_{last} + \alpha T^{i_s} \leq t^{u_k i_s} < t^{u_k i_s}_{last+1}
\end{cases}
\tag{11}
$$

Here, $t^{u_k i_s}_{last+1} = t^{u_k i_s}_{last} + T^{i_s}$, T^{i_s} is the repurchase cycle of item i_s, αT^{i_s} is the best time to recommend from time $t^{u_k i_s}_{last}$ to the next purchase time $t^{u_k i_s}_{last+1}$, and α is a lead-time factor and $\alpha \in (0,1)$.

To be specific, as users in ASI have stable purchasing interest and obvious repeat purchase behavior, it considers a periodic time incentive factor for item recommendation in ASI. That is, the time incentive factor changes with the repurchase cycle. In particular, if the last time user u_k purchases item i_s is time $t_{last}^{u_k i_s}$, he/she will purchase item i_s repeatedly at time $t_{last}^{u_k i_s} + T^{i_s}$, i.e., $t_{last+1}^{u_k i_s}$. When the recommendation time $t^{u_k i_s} \in \left[t_{last}^{u_k i_s}, t_{last}^{u_k i_s} + \alpha T^{i_s} \right)$, it is very unlikely for user u_k to make a repeat purchase. Thus, a negative time incentive factor $-w$ should be combined with the recommendation algorithm. However, when the recommendation time $t^{u_k i_s}$ is close to the next time of repeat purchase $t_{last+1}^{u_k i_s}$, and $t^{u_k i_s} \in \left[t_{last}^{u_k i_s} + \alpha T^{i_s}, t_{last+1}^{u_k i_s} \right)$, it is very likely for user u_k to repeat purchase item i_s. Thus, a positive time incentive factor w should be combined with the recommendation algorithm. This time incentive process is carried out periodically with repeated purchase.

Next, this research employs cosine similarity to calculate the similarity between users. The similarity between user u_k and user $u_{k'}$ at time t is shown in Equation (12):

$$\text{sim}(u_k, u_{k'})_t = \frac{\boldsymbol{u_k} * \boldsymbol{u_{k'}}}{\|\boldsymbol{u_k}\| * \|\boldsymbol{u_{k'}}\|} \tag{12}$$

where $\boldsymbol{u_k}$, $\boldsymbol{u_{k'}}$ are the vectors of historical purchase records of user u_k and user $u_{k'}$ before time t, respectively. The $\widetilde{R}_{u_k i_s}^{t+1}$ function of this kind of items is established as Equation (13).

$$\widetilde{R}_{u_k i_s}^{t+1} = \sum_{\substack{u_{k'} \in U_A \\ i_s \in I_A | U_A}} q_{u_{k'} i_s}^t * \text{sim}(u_k, u_{k'})_t + x_{u_k i_s} w_\alpha \left(t^{u_k i_s} \right) + \left(1 - x_{u_k i_s} \right) \overline{w_\alpha \left(t^{u_{k'} i_s} \right)} \tag{13}$$

Here, $q_{u_{k'} i_s}^t$ is the cumulative purchase of item i_s by user $u_{k'}$ at time t. $\text{sim}(u_k, u_{k'})_t$ is the similarity between user u_k and user $u_{k'}$ at time t. $w_\alpha \left(t^{u_k i_s} \right)$ is the time incentive factor if user u_k purchased item i_s at time t. If user u_k did not purchase item i_s before time t, it uses $\overline{w_\alpha \left(t^{u_{k'} i_s} \right)}$ to incentive the recommendation process, where $w_\alpha \left(t^{u_{k'} i_s} \right)$ is the time incentive factor by users $u_{k'}$ in U_A except user u_k and $\overline{w_\alpha \left(t^{u_{k'} i_s} \right)}$ is the average time incentive factor by all other users $u_{k'}$. $\overline{w_\alpha \left(t^{u_{k'} i_s} \right)}$ is established as Equation (14). $x_{u_k i_s}$ is a 0–1 variable.

$$\overline{w_\alpha \left(t^{u_{k'} i_s} \right)} = \frac{\sum_{\substack{k' \neq k \\ u_{k'} \in U_A}} x_{u_{k'} i_s} w_\alpha \left(t^{u_{k'} i_s} \right)}{\sum_{\substack{k' \neq k \\ u_{k'} \in U_A}} x_{u_{k'} i_s}} \tag{14}$$

Here, it regulates $x_{u_k i_s} = \begin{cases} 1 & \textit{if user } u_k \textit{ ever purchased item } i_s, \\ 0 & \textit{else} \end{cases}$.

4.3.2. Model of ReRec-AUSI

The overall interests of the AUSI users remain active, but they purchase items in $I_B | U_A$ of their random interest, where the activeness of users is more than threshold δ but the stableness of user-item interest is less than threshold θ. The repeat purchase behavior of users is not significant. Hence, the proposed ReRec-ASI based on the repeat purchase cycle of items will be invalid for item recommendation in AUSI. For this reason, this research considers the recommendation algorithm for the AUSI users by combining the user-KNN algorithm and the one-time hot-sale index, assuming that items with higher one-time hot-sale index in AUSI may be preferred by users. In particular, one-time hot-sale index of item i_s, denoted by $\tau_{i_s}^t$, refers to an index that is the largest single sales quantity before time t of item i_s, after the range standardized calculation. The calculation of $\tau_{i_s}^t$ is as Equation (15), where $c_{max}^{i_s t}$ is the largest one-time sales at time t of item i_s, $max \sum_{i_{s'} \in I_B | U_A} c_{max}^{i_{s'} t}$ is the largest

$c_{max}^{i_{s'}t}$ among all the $c_{max}^{i_{s'}t}$ of items in $I_B|U_A$, and $min\sum_{i_{s'}\in I_B|U_A} c_{max}^{i_{s'}t}$ is the smallest $c_{max}^{i_{s'}t}$ among all the $c_{max}^{i_{s'}t}$ of items in $I_B|U_A$. The bigger the largest single sales quantity, the greater the one-time hot-sale index. $\tau_{i_s}^t$ is a decimal between 0 and 1.

$$\tau_{i_s}^t = \frac{c_{max}^{i_st} - min\sum_{i_{s'}\in I_B|U_A} c_{max}^{i_{s'}t}}{max\sum_{i_{s'}\in I_B|U_A} c_{max}^{i_{s'}t} - min\sum_{i_{s'}\in I_B|U_A} c_{max}^{i_{s'}t}} \tag{15}$$

It is similar to the ReRec-ASI approach that this research considers the ReRec-AUSI method by adding one-time hot-sale index to the user-KNN recommendation algorithm. However, as users in AUSI have unstable interest, this research recognizes the similarity by reversing it from 1, and then multiplying by the cumulative purchase amount of other users for item i_s. The improved similarity can pledge that not only the recommended items were purchased by similar users, but also are not always recommended. This is in line with the characteristics of unstable purchase interest of users in AUSI. Moreover, combined with the one-time hot-sale index, the improved similarity will further better the hit rate of recommended items. The $\widetilde{R}_{u_k i_s}^{t+1}$ function of ReRec-AUSI is established as Equation (16), where $q_{u_{k'} i_s}^t$ is the cumulative purchase quantity of item i_s by user $u_{k'}$ at time t. $sim(u_k, u_{k'})_t$ is the similarity between user u_k and user $u_{k'}$ at time t based on KNN algorithm. $\tau_{i_s}^t$ is the one-time hot-sale index at time t of item i_s.

$$\widetilde{R}_{u_k i_s}^t = \sum_{\substack{u_{k'}\in U_A \\ i_s\in I_B|U_A}} q_{u_{k'} i_s}^t * (1 - sim(u_k, u_{k'})_t) + \tau_{i_s}^t \tag{16}$$

4.3.3. Model of ReRec-IASI

The overall interests of users in IASI remain inactive, but they purchase items in $I_A|U_B$ of their stable interest. This research improves the algorithm upon repurchase cycle of items. Especially, it is similar to the behavior of users in ASI in that when a user has just purchased an item the possibility of repeating the purchase immediately is very low, but, as time goes on, with the user running out of the item, he/she is more likely to make repeated purchase. However, as the users in IASI remain inactive, the proposed ReRec-ASI for active users will be invalid for item recommendation in IASI, and the similarity based on users is unreliable. For this reason, this research prioritizes the item-KNN recommendation algorithm by adding a time incentive factor. Considering the characteristic of users in IASI, it assumes that the trajectory of their purchasing interest conforms to the Eibinghaus forgetting curve [29] and the interest declines over time. So, similar but different from the time incentive function in ReRec-ASI is that the principal of function segmentation of time incentive factor $w_b(t^{u_k i_s})$ of ReRec-IASI is the same, but is improved by the Eibinghaus forgetting curve, and is a periodic piecewise exponential function. The model of the time incentive function is as shown in Equation (17).

$$w_b\left(t^{u_k i_s}\right) = \begin{cases} -e^{-\frac{t^{u_k i_s} - t_{last}^{u_k i_s}}{\sigma}} & , t_{last}^{u_k i_s} \le t^{u_k i_s} < t_{last}^{u_k i_s} + \alpha T^{i_s} \\ 1 - e^{-\frac{t^{u_k i_s} - t_{last}^{u_k i_s}}{\sigma}} & , t_{last}^{u_k i_s} + \alpha T^{i_s} \le t^{u_k i_s} < t_{last+1}^{u_k i_s} \end{cases} \tag{17}$$

Here, $t_{last+1}^{u_k i_s} = t_{last}^{u_k i_s} + T^{i_s}$, T^{i_s} is the repurchase cycle of item i_s. αT^{i_s} is the best time to recommend from time $t_{last}^{u_k i_s}$ to the next purchase time $t_{last+1}^{u_k i_s}$, where α is a lead-time factor and $\alpha \in (0,1)$. σ is the forgetting rate, and $\sigma \in (0,1)$.

To be specific, as users in IASI have stable purchasing interest in items and obvious repeat purchase behavior, this research considers a periodic time incentive factor to item recommendation in IASI, i.e., the time incentive factor according to improved Eibinghaus forgetting curve changes with the repurchase cycle. In particular, if the last time user u_k purchases item i_s is time $t_{last}^{u_k i_s}$, generally he/she will purchase item i_s repeatedly at

time $t_{last}^{u_k i_s} + T^{i_s}$, i.e., $t_{last+1}^{u_k i_s}$. When the recommendation time is $t^{u_k i_s} \in \left[t_{last}^{u_k i_s}, \ t_{last}^{u_k i_s} + \alpha T^{i_s} \right)$, it is very unlikely for user u_k to make a repeat purchase. For this reason, a negative time incentive factor $-e^{-\frac{t^{u_k i_s} - t_{last}^{u_k i_s}}{\sigma}}$ should be considered in the recommendation algorithm. However, when the recommendation time $t^{u_k i_s}$ is close to the next time of repeat purchase $t_{last+1}^{u_k i_s}$, where $t^{u_k i_s} \in \left[t_{last}^{u_k i_s} + \alpha T^{i_s}, \ t_{last+1}^{u_k i_s} \right)$, it is very likely for user u_k to repeat purchase item i_s, so a positive time incentive factor $1 - e^{-\frac{t^{u_k i_s} - t_{last}^{u_k i_s}}{\sigma}}$ should be considered in the recommendation algorithm. This time incentive process is also carried out periodically with repeated purchase.

Next, it uses cosine similarity to calculate the similarity between items. The similarity between item i_s and item $i_{s'}$ at time t is shown in Equation (18).

$$\text{sim}(i_s, i_{s'})_t = \frac{i_s * i_{s'}}{\parallel i_s \parallel * \parallel i_{s'} \parallel} \tag{18}$$

where i_s, $i_{s'}$ are the vectors of historical purchase records of item i_s and item $i_{s'}$ before time t, respectively. So, the $\widetilde{R}_{u_k i_s}^{t+1}$ function of this kind of items is established as Equation (19).

$$\widetilde{R}_{u_k i_s}^{t+1} = \sum_{i_{s'} \in I_A | U_B} q_{u_k i_{s'}}^t * sim(i_s, i_{s'})_t + x_{u_k i_s} w_b\left(t^{u_k i_s}\right) + \left(1 - x_{u_k i_s}\right) \overline{w_b\left(t^{u_{k'} i_s}\right)} \tag{19}$$

Here, $q_{u_k i_{s'}}^t$ is the cumulative purchase of item $i_{s'}$ by user u_k at time t. $sim(i_s, i_{s'})_t$ is the similarity between item i_s and item $i_{s'}$ at time t. $w_b\left(t^{u_k i_s}\right)$ is the time incentive factor when user u_k purchases item i_s at time t. If user u_k did not purchase item i_s before time t, this research uses $\overline{w_b\left(t^{u_{k'} i_s}\right)}$ to incentivize the recommendation process, where $w_b\left(t^{u_{k'} i_s}\right)$ is the time incentive factor by users $u_{k'}$ in U_B except user u_k. $\overline{w_b\left(t^{u_{k'} i_s}\right)}$ is the average value of the time incentive factor when user $u_{k'}$ who is not user u_k, purchases item i_s at time t. $x_{u_k i_s}$ is a 0–1 variable and it is modeled as Equation (20).

$$x_{u_k i_s} = \left\{ \begin{array}{ll} 1 & \textit{if user } u_k \textit{ ever purchased item } i_s, \\ 0 & \textit{else} \end{array} \right. \tag{20}$$

4.3.4. Model of ReRec-IAUSI

The overall interests of the IAUSI users remain inactive and they usually purchase items in $I_B | U_B$ of their random interests, where the activeness of users is less than threshold δ and the stableness of user–item interest is also less than threshold θ. Users do not have declining repeat purchase behavior. Hence, the proposed ReRec-IASI based on declining repeat purchase cycle of items will be invalid for item recommendation in IAUSI. For this reason, this research considers the recommendation algorithm for the IAUSI users by combining the item–KNN algorithm and total hot-sale index, where it assumes that items with higher total hot-sale index in IAUSI may be preferred by users. In particular, the total hot-sale index of item i_s, denoted by $\varphi_{i_s}^t$, refers to an index that is the largest total sales quantity before time t of item i_s, after the range standardized calculation. The calculation of $\varphi_{i_s}^t$ is as in Equation (21), where $C^{i_s t}$ is the largest total sales before time t of item i_s, $max \sum_{i_{s'} \in I} C^{i_{s'} t}$ is the largest $C^{i_{s'} t}$ among all the $C^{i_s t}$ of items in $I_B | U_B$, and $min \sum_{i_{s'} \in I} C^{i_{s'} t}$ is the smallest $C^{i_{s'} t}$ among all the $C^{i_s t}$ of items in $I_B | U_B$. The bigger the largest total sales quantity, the greater the total hot-sale index. $\varphi_{i_s}^t$ is a decimal between 0 and 1.

$$\varphi_{i_s}^t = \frac{C^{i_s t} - min \sum_{i_{s'} \in I_B | U_B} C^{i_{s'} t}}{max \sum_{i_{s'} \in I_B | U_B} C^{i_{s'} t} - min \sum_{i_{s'} \in I_B | U_B} C^{i_{s'} t}} \tag{21}$$

Similar to the ReRec-IASI approach, this research considers the ReRec-IASUI method by adding an incentive factor which is a hot-sale index to the item–KNN recommendation

algorithm. However, as users in IAUSI have unstable interest, the research recognizes the similarity by reversing it from 1, and then multiplying by the cumulative purchase amount of other users for item i_s. The improved similarity can show not only that the recommended items were purchased by similar users, but also that the recommended items are diverse. This is in line with the characteristics of unstable purchase interest of users in IAUSI. Moreover, combined with the total hot-sale index, the improved similarity will further increase the hit rate of recommended items. The $\widetilde{R}_{u_k i_s}^{t+1}$ function of ReRec-IAUSI can be formed as Equation (22).

$$\widetilde{R}_{u_k i_s}^{t+1} = \sum\nolimits_{i_{s'} \in I_B | U_B} q_{u_k i_{s'}}^t * (1 - \text{sim}(i_s, i_{s'})_t) + \varphi_{i_s}^t \tag{22}$$

where $q_{u_k i_{s'}}^t$ is the cumulative purchase of item $i_{s'}$ by user u_k at time t. $\text{sim}(i_s, i_{s'})_t$ is the similarity between item i_s and item $i_{s'}$ at time t. $\varphi_{i_s}^t$ is the hot-sale index at time t of item i_s.

5. Experiments

5.1. The Dataset

The dataset used in this paper comes from a community e-commerce platform T-app, with 11,350 purchase records from June 2017 to August 2019. It contains 1064 users and 137 kinds of items. The characteristics of each record include user ID, item ID, purchase time, purchase quantity, price, payment method and other attributes. Specifically, the data from June 2017 to April 2019 (10,343 records) are used as the training set, and the data from April 2019 to August 2019 (1007 records) are used as the test set. The user–item recommendation models are trained on the training set, and are tested on the test set.

The purchase behavior of users on the T-app platform has obvious characteristics of repurchase. For instance, by analyzing the data of a time phase, it is found that among 955 users who have made purchases, 58.74% have repeat purchases. In Figure 3, it can be seen that the total repurchases of 23% of repurchase users is larger than 15. The average repurchase time of repurchase users is 3.61. Among the repurchase users, 10.33% repurchase the same item more than six times. In an extreme case, it is found that one user has repurchased the same item up to 43 times under the investigated time duration. In Figure 4, it can be seen that, among all the types of item (105 types), 78.10% (82 types) have been repurchased by users, and in 17% of the repurchased items, the total number of times repurchased by users is more than 120.
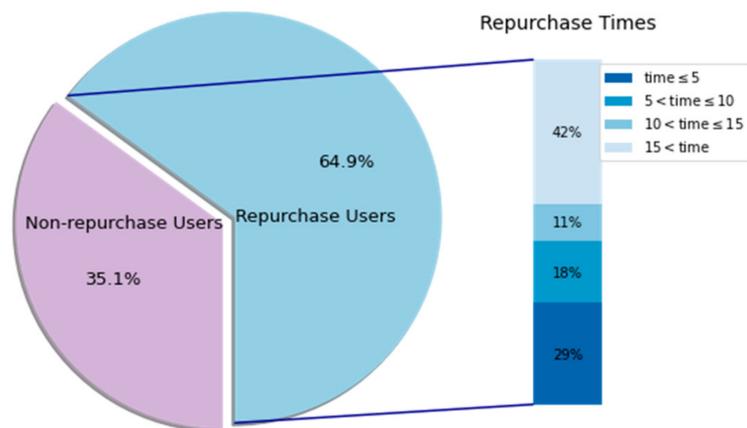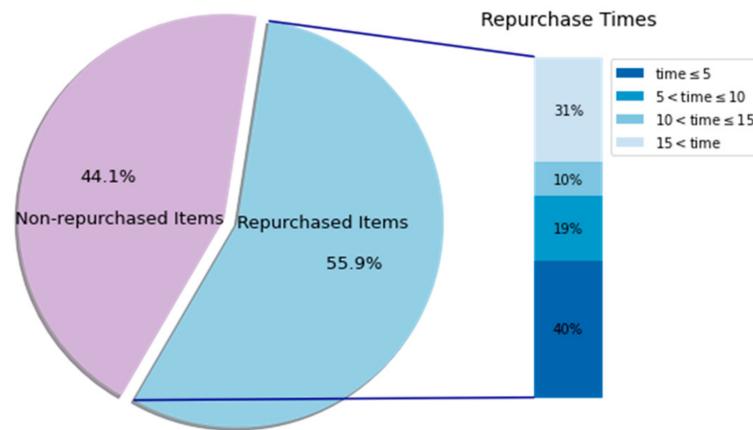

Figure 3. Proportions of repurchase users.

Figure 4. Proportions of repurchased items.

5.2. Experimental Setup

In the traditional collaborative filtering recommendation, the user–item score matrix is usually used as the original data for the recommendation calculation. This paper adopts offline experiments for verification, and the user's cumulative purchase is used as the score. First, according to the user classification model, the user-item is classified into four categories: active users with stable interest, active users with unstable interest, inactive users with stable interest and inactive users with unstable interest. Then, the recommendation calculation is carried out for each category, and the improved recommendation algorithms for active users and inactive users are evaluated respectively. The results are then compared with that of the traditional CF, SVD, SVD++, and NMF algorithms.

The repurchase cycle refers to the time interval between the *n*th and the (*n* + 1)th purchase of item i_s by user u_k. For an item, the repurchase cycle of different users at the same time period is different, and that of the same user at different time periods is also different. So, if the items' repurchase cycle is calculated by each user by time, it could be highly random and prone to overfitting. Therefore, for the active users' stable purchase behavior, the average repurchase cycle of the top three users in purchase quantity of a certain item is used as the repurchase cycle. For the items included in $I_A|U_B$, as the overall interest of users is inactive, the repurchase cycle of the user who purchases the largest quantity of an item is regarded as the repurchase cycle of this item. Examples of repurchase cycle for some items included in $I_A|U_A$ are shown in Table 2 and for some items included in $I_A|U_B$ in Table 3.

Table 2. Repurchase cycle of typical items included in $I_A|U_A$.

Item ID	Repurchase Cycle (Days)	Name
2	14.07	ZY
38	24.65	TB-Mo
61	21.25	TB-Th
68	14.84	ZQB-F
69	20.37	ZYB-We
73	15.51	HB-We

Table 3. Repurchase cycle of typical items included in $I_A | U_B$.

Item ID	Repurchase Cycle (Days)	Name
2	10	ZY
38	16	TB-Mo
61	24	TB-Th
68	12	ZQB-Fr
69	14	ZYB-We
73	11	HB-We

In the experiments, each type of user behavior model can produce a corresponding item recommendation list. After sorting in descending order according to the purchase possibility, the recommended items can be selected according to the top N method. $N_{(U_A)}$ is the number of active users and $N_{(U_B)}$ is the number of inactive users, and they can be expressed as Equations (23) and (24), respectively.

$$N_{(U_A)} = N_{(I_A | U_A)} + N_{(I_B | U_A)} \tag{23}$$

$$N_{(U_B)} = N_{(I_A | U_B)} + N_{(I_B | U_B)} \tag{24}$$

Here, $N_{(I_A | U_A)}$ is the recommended item quantity from items included in $I_A | U_A$. $N_{(I_B | U_A)}$, $N_{(I_A | U_B)}$, $N_{(I_B | U_B)}$ are similar in meaning to $N_{(I_A | U_A)}$. So, it is easy to discover that the recommendation list of active users is composed of $N_{(I_A | U_A)}$ stable interests and $N_{(I_B | U_A)}$ unstable interests. Similarly, the recommendation list of inactive users is composed of $N_{(I_A | U_B)}$ stable interests and $N_{(I_B | U_B)}$ unstable interests.

Considering the actual situation of T-app, its operators should select the best combination of items in different user–item classifications for recommendation. Hence, here this research uses the grid search method to test the models. Firstly, let the total number of recommendation items be less than the number of all items, N_{max}, for each type of user. Both the number of stable items and unstable items should be less than N_{max}. That is to say, it has constraints (25) and (26). In the test experiment, N_{max} is set as 25. Secondly, with constraints (25), $N_{(U_A)}$ has multiple combinations of $N_{(I_A | U_A)}$ and $N_{(I_B | U_A)}$, and it is the same as $N_{(U_B)}$. For instance, when the total number of recommended items N_{max} is 5, $\left(N_{(I_A | U_A)}, N_{(I_B | U_A)} \right)$ can be able to (0,5), (1,4), (2,3), (3,2), (4,1), (5,0). It can select the optimal combination among the six combinations as the recommended combination when $N_{(U_A)} = 5$.

$$\begin{cases} 0 \leq N_{(I_A | U_A)} \leq N_{max} \\ 0 \leq N_{(I_B | U_A)} \leq N_{max} \\ 0 \leq N_{(I_A | U_A)} + N_{(I_B | U_A)} \leq N_{max} \end{cases} \tag{25}$$

$$\begin{cases} 0 \leq N_{(I_A | U_B)} \leq N_{max} \\ 0 \leq N_{(I_B | U_B)} \leq N_{max} \\ 0 \leq N_{(I_A | U_B)} + N_{(I_B | U_B)} \leq N_{max} \end{cases} \tag{26}$$

5.3. Evaluation Metrics

Three evaluating indicators are used to gauge the algorithm performance, precision (Pre), recall (Rec) and F-measure, defined in Equations (27)–(29). Precision is defined as the ratio of items that users like to all recommended items in the recommended list. Recall is defined as the ratio of the items that users like in the recommended list to all the items that users like in the system. Generally, precision and recall must be used at the same time to fully evaluate the quality of the algorithm. Some researchers have proposed an indicator called F-measure that comprehensively integrates the precision and the recall. Therefore,

the evaluation indicators used in this paper are precision, recall, and F-measure to measure the precision of item recommendation. The three expressions are shown as (27)–(29).

$$Pre = \frac{TP}{TP + FP} \tag{27}$$

$$Rec = \frac{TP}{TP + FN} \tag{28}$$

$$F - measure = \frac{2 \times Pre \times Rec}{Pre + Rec} \tag{29}$$

Here, TP is the number of items that have been recommended and purchased; FP is the number of items that are recommended but not purchased; and FN is the number of items that have not been recommended but purchased.

5.4. Experimental Results

Figure 5 shows the comparison results of the proposed ReRec algorithm on active users (i.e., the combination of ReRec-ASI and ReRec-AUSI) compared with four baseline methods, traditional User CF, SVD, SVD++ and NMF algorithms. It sets $w = 1$, and $N_{(U_A)} \in [5, 25]$. It can be seen from Figure 5 that, on the purchase prediction of active users, the proposed ReRec algorithm performs better than the traditional User CF, SVD, SVD++ and NMF algorithms in terms of the three evaluation indicators, precision, recall and F-measure. This indicates that the proposed ReRec algorithm for active users in this paper improves the hit rate of item recommendation and ensures the precision of recommendation results.
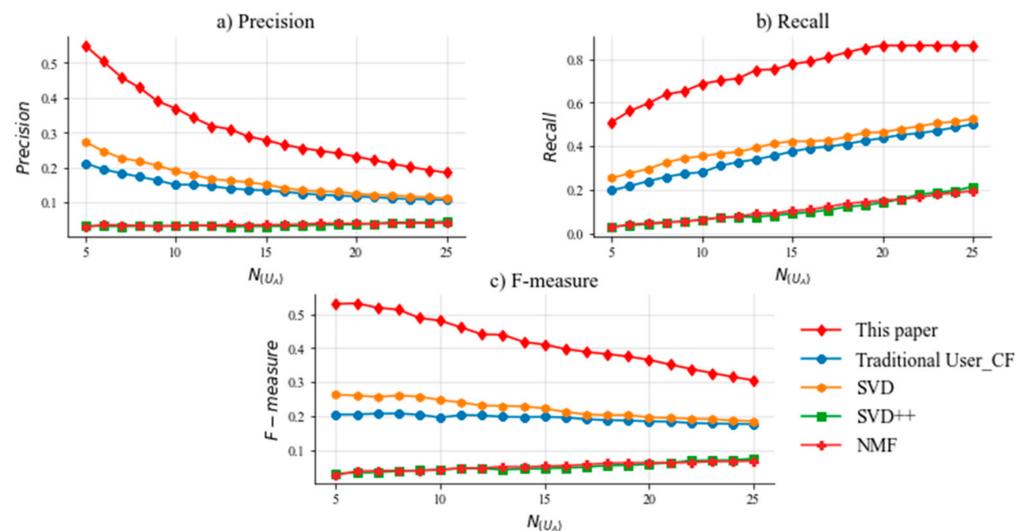

Figure 5. The comparison results of the proposed ReRec algorithm (the combination of ReRec-ASI and ReRec-AUSI) and four baselines on active users.

Figure 6 shows the comparison results of the proposed ReRec approach with the baselines on inactive users (the combination of ReRec-IASI and ReRec-IAUSI). It sets the parameters α and σ as 0.75 and 0.2284, respectively. The total number of recommendation items of $N_{(U_B)}$ is the same as $N_{(U_A)}$. It can be seen that, in the purchase prediction of inactive users, when $N_{(U_B)} \in [6, 25]$, the improved Item CF algorithm proposed in this paper is superior to the evaluation indicators of traditional Item CF, SVD, SVD++ and NMF algorithms in terms of precision, recall and F-measure. Because the number of item type in the test data is relatively smaller than the number of users, the purchase prediction performance for inactive users is not as good as that for active users. However, the purchase prediction of inactive users based on the improved Item CF algorithm still improves the hit

rate of item recommendation within a certain range, and also ensures a higher precision of recommendation results.
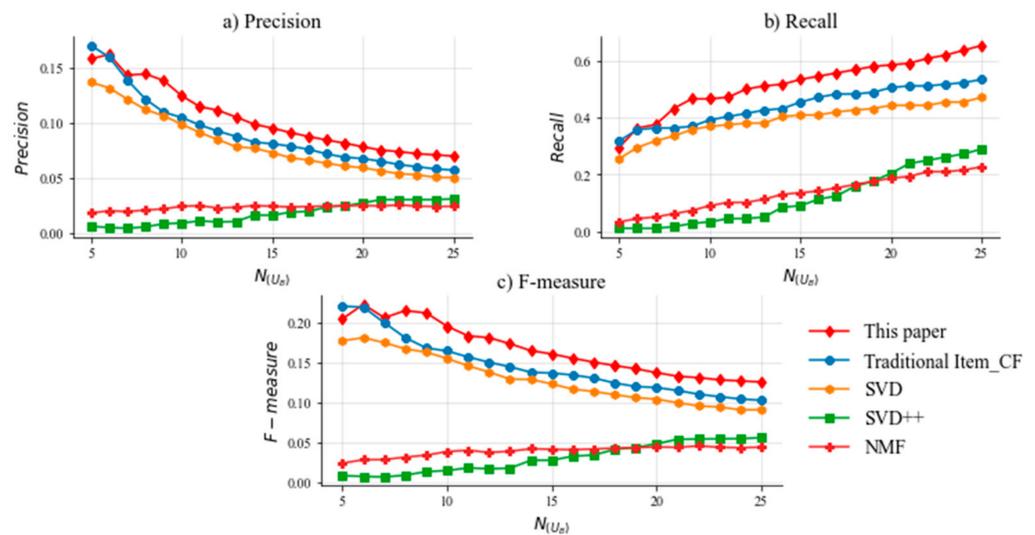

Figure 6. The comparison results of the proposed ReRec approach (the combination of ReRec-IASI and ReRec-IAUSI) and four baselines on inactive users.

The poor performance of the baselines can be explained because all ratings in the user item rating matrix are regarded as equal, ignoring the heterogeneity of users' interests, i.e., user's personalized interest and users' public interest. The SVD method, which is derived from linear algebra, has a solid mathematical foundation in matrix approximation. However, it lacks a user's preference model and an item's preference model of the user's interest in the item. In the SVD++ method, a bias model and the latent vectors of the user and the item are used to model the user's interest in the item. Using stochastic gradient descent to update the bias vector and latent vector of each observed rating in the user item rating matrix can result in a large amount of computation. The advantage of the NFM model is that the elements of latent users and item vectors can be non-negative, while its disadvantage is that the precision of rating prediction is reduced.

In summary, none of the baselines improve the recommendation algorithms according to different types of user behavior on the temporal horizon. Although some scholars have added the user's personalized behavior into the item recommendation algorithm, they more often than not ignore user loyalty in recommendations that may drive the users' repeat purchase. It holds that the users' loyalty to the shopping platform and items has a non-negligible impact on the successful recommendation of items. Following this line of thought, this research proposes the ReRec algorithm based on user behavior classification and item repurchase cycle. The proposed ReRec algorithm can predict the possibility of repeat purchase in order to recommend the top N items to users and improve the user experience of the recommendation system.

5.5. Sensitivity Analysis of Parameter w

In the proposed ReRec approach for active users, incentive factor w is an important parameter. In order to analyze the influence of w on the recommendation process, it conducts sensitivity analysis on the parameter w. Figure 7 illustrates the F-measures with $N_{(I_A|U_A)}$ and $N_{(I_B|U_A)}$, when other conditions are fixed and w varies. The following conclusions can be drawn from Figure 7. When $w \in [0, 5]$, for $N_{(I_A|U_A)} \leq 10$ and $N_{(I_B|U_A)} \leq 15$, the F-measures with various combination of $N_{(I_A|U_A)}$ and $N_{(I_B|U_A)}$ are better than that of other conditions.
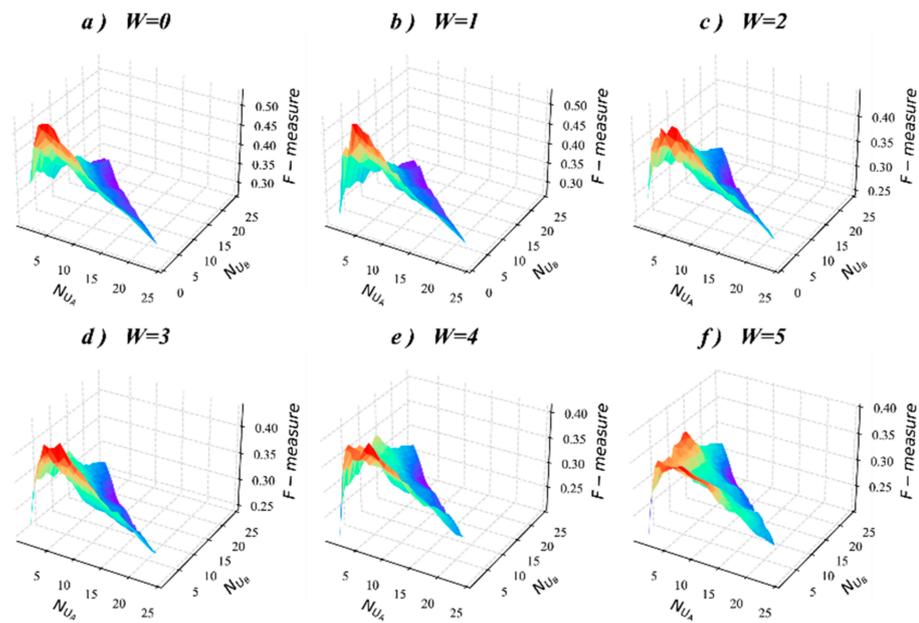
Figure 7. The F-measures of different values of w.

Figure 8 illustrates the F-measures with incentive factor w given $N_{(U_A)} = 3$. It can be seen that, with the value of w increasing in [0,5], the value of F-measure first increases and then decreases. When $w > 5$, the values of F-measure are kept stable. Therefore, the research further analyzes the evaluation indicators with $w \in [0,5]$.
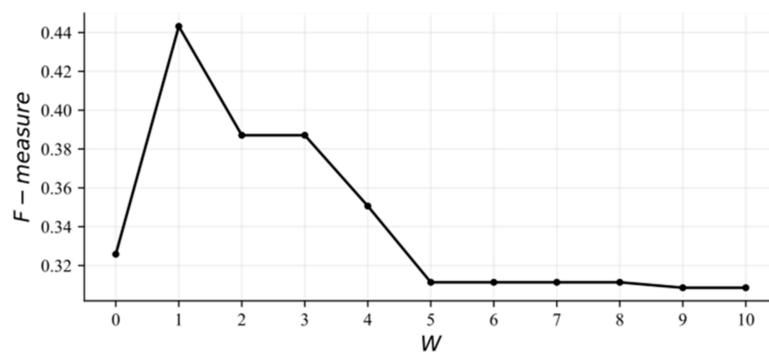

Figure 8. The F-measures for incentive factor w.

Figure 9 illustrates the evaluating indicators (precision, recall and F-measure) with the total recommended quantity $N_{(U_A)}$ when $w \in [0,5]$. It can be seen that when $N_{(U_A)}$ increases in the range $[0,5]$, the variation trend of precision is relatively unstable. In comparison, the recall and F-measure go up firstly and then go down. While $N_{(U_A)}$ increases in the range $[5,25]$, the precisions gradually decrease, while the recall increases. As a result, the F-measures decrease. It is evident that when $N_{(U_A)} > 7$, the performances of three evaluating indicators at $w = 1$ are better than that at other values of w. Therefore, the ReRec algorithm should be used with the setting as $w = 1$.
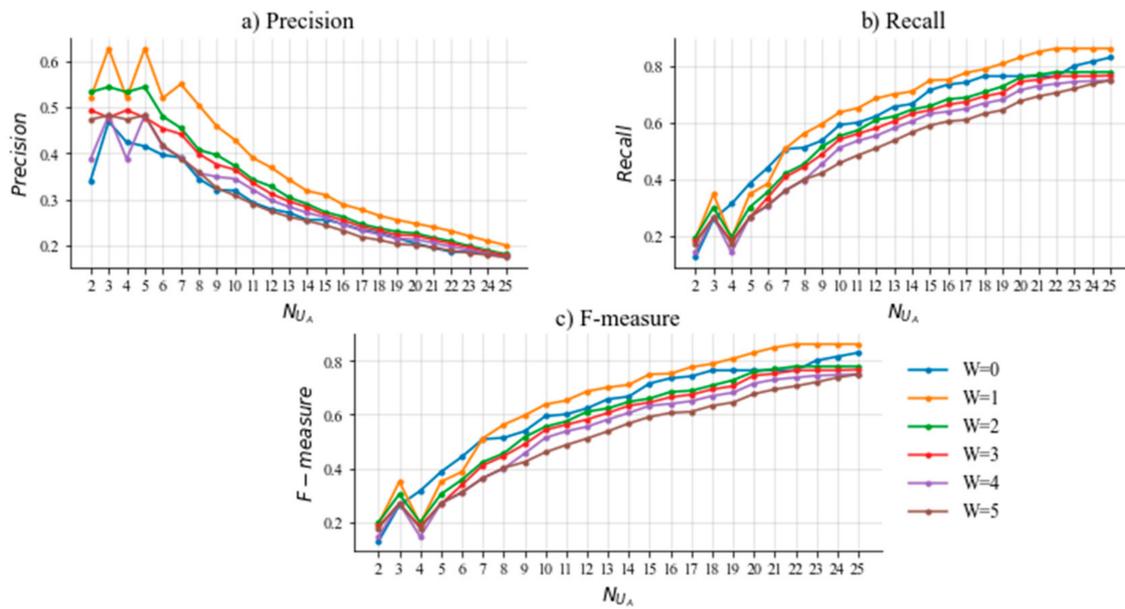
Figure 9. The change of three evaluating indicators at given *w*.

In the proposed ReRec approach for inactive users, the grid search method is adopted to carry out ReRec-IASI and ReRec-IAUSI. The precision, recall and F-measures with $N_{(I_A|U_B)}$ and $N_{(I_B|U_B)}$ are shown in Figure 10. It can be seen that when $N_{(I_A|U_B)}$ is fixed, the precisions of the recommendation results are decreasing along with the increase of $N_{(I_B|U_B)}$. When $N_{(U_B)}$ is small, the precisions and F-measures are large. The recalls of recommendation results are large when $N_{(I_A|U_B)}$ and $N_{(I_B|U_B)}$ and are approximately equal to each other.
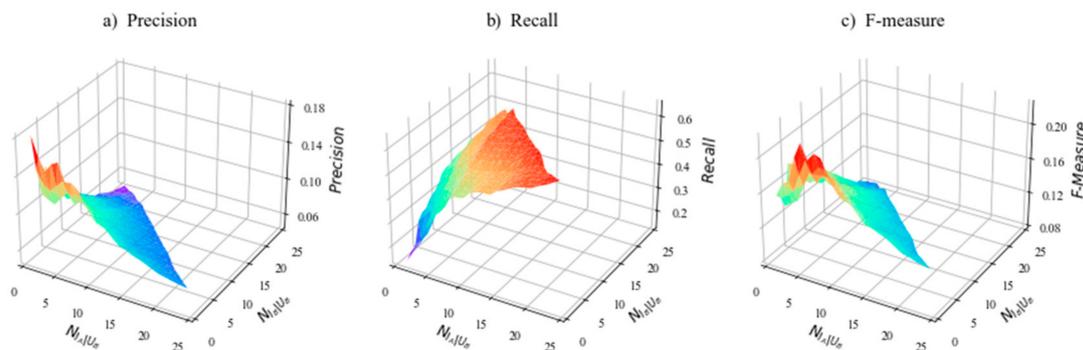

Figure 10. The changes of three evaluation indicators with the combination of inactive users.

5.6. Discussion of Important Results

The proposed methods are trained in the training set and evaluated in the test set. Three evaluating indicators are used to gauge the algorithm performance as precision, recall and F-measure (Equations (27)–(29)). We conduct our experiments on a real-life community e-commerce platform. Results show that the proposed ReRec method provides better performance compared to the existing methods (namely traditional CF, SVD, SVD++, NMF). The discussion of the important results of the proposed methods is analyzed as follows.

Four types of user-item interactions are obtained before applying ReRec: active users with stable interest (ASI), inactive users with stable interest (IASI), active users with unstable interest (AUSI), and inactive users with unstable interest (IAUSI). For active users, the hit rate of item recommendation shows a marked improvement, while for inactive users, the hit ratio increased slightly. Compared with inactive users, active users use the platform more

frequently, so it is easier to detect their buying interest. The reason for the poor performance of the baselines may be that none of the baselines improves the recommendation algorithms according to different types of user behavior on the temporal horizon.

The performance of ReRec is analyzed based on varying the values of incentive factor w. With the value of w increasing in [0,5], the value of F-measure first increases and then decreases. When $w > 5$, the values of the F-measure are kept stable and low. When $w = 1$, the ReRec algorithm shows the highest precision. It is evident that when $N_{(U_A)} > 7$, the performances of three evaluating indicators at $w = 1$ are better than that at other values of w. Therefore, the ReRec algorithm should be used with the setting as $w = 1$.

Finally, the practical contribution is summarized. The result of recommendation is stable, which can provide support for business management decision-making in enterprises, and the effectiveness of the algorithm is verified. For instance, precise marketing strategies based on customer heterogeneity can be implemented, thus reducing the operating costs of community e-commerce platforms.

6. Concluding Remarks

To fill in the research gap from the perspective of repeated purchase behavior and improve the process of the generation of a recommendation list, this research proposed a novel approach called ReRec (Repeat purchase Recommender) to recommending items to users in a divide and conquer manner. The proposed method includes ReRec-ASI, ReRec-AUSI, ReRec-IASI and ReRec-IAUSI. Experiments are conducted on a real dataset collected from a community e-commerce platform. Compared with well-known existing methods (e.g., SVD, SVD++) the ReRec method improves the recommend performance by at least 13.6% (measured by F-measure). Specifically, for active users, with $w = 1$ and $N_{(U_A)} \in [5, 25]$, the ReRec-ASI, ReRec-AUSI shows a significant improvement (at least 50%) in recommendation. With α and σ as 0.75 and 0.2284, respectively, the proposed ReRec-IASI and ReRec-IAUSI are also superior to (by at least 13.6%) the evaluation indicators of traditional Item CF when $N_{(U_B)} \in [6, 25]$.

Although the proposed ReRec approach performs well in this study, there are still some gaps to be explored in the future:

Firstly, the size of test set needs to be expanded, because this paper only uses four months of consumption data to test the algorithm at present. The amount of data that can be used now on T-app is limited. In the future, there will be more consumption data available this large-scale data can be used to verify the algorithm.

Secondly, when it has consumption data over a long time, such as consumption data for several years, it would attempt to improve the recommendation algorithms with centralized consumption behaviors such as seasonal consumption and holiday consumption, which is an interesting problem in recommendation.

Author Contributions: Conceptualization, J.W. and W.Z.; methodology, J.W., Y.L. and W.Z.; software, Y.L. and L.Y.; validation, Y.L., L.Y. and X.N.; formal analysis, J.W., Y.L., L.S. and W.Z.; investigation, Y.L. and L.S.; resources, J.W. and L.S.; data curation, Y.L., L.Y. and X.N.; writing—original draft preparation, Y.L., L.Y. and X.N.; writing—review and editing, J.W. and W.Z.; visualization, Y.L. and L.Y.; supervision, J.W. and W.Z.; project administration, J.W., L.S. and W.Z.; funding acquisition, J.W. and W.Z. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zsolnai, L. Green business or community economy? *Int. J. Soc. Econ.* **2002**, *29*, 652–662. [CrossRef]
2. Lao, J.; Zhong, Y.; Tan, Z. Study of community E-commerce model based on intelligent building. *J. Intell.* **2007**, *26*, 39–41.
3. Kim, H.K.; Oh, H.Y.; Gu, J.C.; Kim, J.K. Commenders: A recommendation procedure for online book communities. *Electron. Commer. Res. Appl.* **2011**, *10*, 501–509. [CrossRef]
4. Zhang, W.; Du, Y.; Yang, Y.; Yoshida, T. DeRec: A data-driven approach to accurate recommendation with deep learning and weighted loss function. *Electron. Commer. Res. Appl.* **2018**, *31*, 12–23. [CrossRef]
5. Iwanaga, J.; Nishimura, N.; Sukegawa, N.; Takano, Y. Improving collaborative filtering recommendations by estimating user preferences from clickstream data. *Electron. Commer. Res. Appl.* **2019**, *37*, 100877. [CrossRef]
6. Ghasemi, N.; Momtazi, S. Neural text similarity of user reviews for improving collaborative filtering recommender systems. *Electron. Commer. Res. Appl.* **2020**, *45*, 101019. [CrossRef]
7. Riyahi, M.; Sohrabi, M.K. Providing effective recommendations in discussion groups using a new hybrid recommender system based on implicit ratings and semantic similarity. *Electron. Commer. Res. Appl.* **2020**, *40*, 100938. [CrossRef]
8. Verstrepen, K.; Bhaduriy, K.; Cule, B.; Goethals, B. Collaborative filtering for binary, positive-only data. In Proceedings of the 23rd ACM SIGKDD Conference, Halifax, NS, Canada, 13–17 August 2017; pp. 1–21.
9. Chen, J.; Wei, L.; Zhang, L. Dynamic evolutionary clustering approach based on time weight and latent attributes for collaborative filtering recommendation. *Chaos Solitons Fractals* **2018**, *114*, 8–18. [CrossRef]
10. Verbert, K.; Manouselis, N.; Ochoa, X.; Wolpers, M.; Drachsler, H.; Bosnic, I.; Duval, E. Context-aware recommender sys-tems for learning: A survey and future challenges. *IEEE Trans. Learn. Technol.* **2012**, *5*, 318–335. [CrossRef]
11. Mezni, H.; Benslimane, D.; Bellatreche, L. Context-aware service recommendation based on knowledge graph em-bedding. *IEEE Trans. Knowl. Data Eng.* **2021**, *99*, 1–14. [CrossRef]
12. Nguyen, V.-D.; Sriboonchitta, S.; Huynh, V.-N. Using community preference for overcoming sparsity and cold-start problems in collaborative filtering system offering soft ratings. *Electron. Commer. Res. Appl.* **2017**, *26*, 101–108. [CrossRef]
13. Resnick, P.; Iacovou, N.; Suchak, M.; Bergstrom, P. Group Lens: An open architecture for collaborative filtering of net news. In Proceedings of the ACM 1994Conference on Computer Supported Cooperative Work, Chapel Hill, NC, USA, 22–26 October 1994; pp. 175–186.
14. Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; pp. 285–295.
15. Wang, C.; Zheng, Y.; Jiang, J.; Ren, K. Toward Privacy-Preserving Personalized Recommendation Services. *Engineering* **2018**, *4*, 21–28. [CrossRef]
16. Brand, M. Fast online SVD revisions for lightweight recommender systems. In Proceedings of the Third SIAM International Conference on Data Mining, San Francisco, CA, USA, 1–3 May 2003; pp. 37–46.
17. Koren, Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NV, USA, 24–27 August 2008; pp. 426–434.
18. Kim, J.; Park, H. Toward Faster Nonnegative Matrix Factorization: A New Algorithm and Comparisons. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 353–362. [CrossRef]
19. Gorgoglione, M.; Panniello, U.; Tuzhilin, A. The effect of context-aware recommendations on customer purchasing behavior and trust. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 85–92. [CrossRef]
20. Zimdars, A.; Chickering, D.M.; Meek, C. Using temporal data for making recommendations. In Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence, Seattle, WA, USA, 2–5 August 2001; pp. 580–588.
21. Campos, P.G.; Díez, F.; Bellogín, A. Temporal rating habits: A valuable tool for rating discrimination. In Proceedings of the 2nd Challenge on Context-Aware Movie Recommendation, Chicago, IL, USA, 27 October 2011; pp. 29–35.
22. Liang, X.; Yang, Q. Time-dependent models in collaborative filtering based recommender system. In Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Milano, Italy, 15–18 September 2009; pp. 450–457.
23. Qin, G.; Du, X. An efficient collaborative filtering algorithm with user hierarchy. *Comput. Sci.* **2004**, *10*, 138–140.
24. Xing, C.; Gao, F.; Zhan, S.; Zhou, L. A collaborative filtering recommendation algorithm incorporated with user interest change. *J. Comput. Res. Dev.* **2007**, *02*, 296–301. [CrossRef]
25. Zhang, Y.; Liu, Y. A Collaborative Filtering Algorithm Based on Time Period Partition. In Proceedings of the Third International Symposium on Intelligent Information Technology & Security Informatics, Washington, DC, USA, 2–4 April 2010; pp. 777–780.
26. Chen, J.; Lu, Y.; Shang, F.; Zhu, T. A novel recommendation scheme with multifactorial weighted matrix decomposition strategies via forgetting rule. *Eng. Appl. Artif. Intell.* **2021**, *101*, 104191. [CrossRef]
27. Wu, F.; Yu, L.; Feng, M. A collaborative filtering algorithm based on time effect. *Comput. Eng. Sci.* **2017**, *39*, 2095–2101.
28. Fader, P.S.; Hardie, B.G.; Lee, K.L. RFM and CLV: Using iso-value curves for customer base analysis. *J. Mark. Res.* **2005**, *42*, 415–430. [CrossRef]
29. Hermann, E. Memory: A Contribution to Experimental Psychology. *Ann. Neurosci.* **2013**, *20*, 155–156.