A Novelty-Seeking based Dining Recommender System

Fuzheng Zhang[#†], Kai Zheng[*], Nicholas Jing Yuan[†],
Xing Xie[†], Enhong Chen[#], Xiaofang Zhou[*]
[†]Microsoft Research
[#]University of Science and Technology of China
[*]School of Information Technology & Electrical Engineering, University of Queensland
{v-fuz,nicholas.yuan,xingx}@microsoft.com,
{kevinz,zxf}@itee.uq.edu.au, cheneh@ustc.edu.cn

ABSTRACT

The rapid growth of location-based services provide the potential to understand people's mobility pattern at an unprecedented level, which can also enable food-service industry to accurately predict consumer's dining behavior. In this paper, by leveraging users' historical dining pattern, socio-demographic characteristics and restaurants' attributes, we aim at generating the top-K restaurants for a user's next dining. Compared to previous studies in location prediction which mainly focus on regular mobility patterns, we present a novelty-seeking based dining recommender system, termed **NDRS**, in consideration of both exploration and exploitation. First, we apply a Conditional Random Field (CRF) with additional constraints to infer users' novelty-seeking statuses by considering both spatial-temporal-historical features and users' socio-demographic characteristics. On the one hand, when a user is predicted to be novelty-seeking, by incorporating the influence of restaurants' contextual factors such as price and service quality, we propose a context-aware collaborative filtering method to recommend restaurants she has never visited before. On the other hand, when a user is predicted to be not novelty-seeking, we then present a Hidden Markov Model (HMM) considering the temporal regularity to recommend the previously visited restaurants. To evaluate the performance of each component as well as the whole system, we conduct extensive experiments, with a large dataset we have collected covering the concerned dining related check-ins, users' demographics, and restaurants' attributes. The results reveal that our system is effective for dining recommendation.

1. INTRODUCTION

Dining out has become one of the most distinctive aesthetic features of urban life [27]. With the rapid development of smart phones and positioning technology, the emerging location based services (e.g., Foursquare, Facebook Place, JiePang) can accurately record users' location information, which provides the potential to understand users' dining behavior at an unprecedented level.

In this paper, we aim at leveraging location records, users' socio-demographic characteristics, and restaurants' attributes to recommend restaurants for users' next dining. Actually, this problem is very similar to the location prediction problem, which has long been studied in traffic analysis [39], urban planning [42] and recommendation research [5]. Most of the previous studies in these fields put an emphasis on users' historical data and the proposed prediction methods heavily rely on repetitive mobility pattern in the past. However, due to the novelty-seeking tendency, which is biologically embedded into human brains according to the uncovered genetic roots and relations to the dopamine system [9], users would also be full of enthusiasm about exploring previously unvisited places. To be specific, more than 35% location visit is made at new places (previously unvisited) each day even after half a year according to the reported results in [6]. Actually, this neophilia characteristic is extremely outstanding in dining behavior, e.g., [3] discovered that an appropriate degree of novelty-seeking as well as attendant risk could be essential ingredients in entertainment and excitement for users' dining out motivation. [17] also stated that the physiological and psychological motivators which cannot be fulfilled in a user's normal daily life are likely to be satisfied by a sense of adventure, uniqueness of the setting, experience of different cultures and the opportunity of sampling new foods.

Given above, we present a framework, termed novelty-seeking based dining recommender system (**NDRS**), to generate the top-K restaurants for the next dining. At first, to infer a user's novelty-seeking status, we present a Conditional Random Field (CRF) to model the sequential dependency of novelty-seeking statuses in consideration of spatial, temporal and historical factors that would influence the novelty-seeking decision. Furthermore, as users' socio-demographic characteristics also have a profound and lasting impact on users' novelty-seeking tendency [37], we design these characteristics as additional constraints and incorporate them into CRF seamlessly. On the one hand, when a user is predicted to be novelty-seeking (exploring new restaurants), the recommendation candidates will be generated from tens of thousands of new restaurants which locate in the city this user resides. Then a context-aware collaborative filtering method is proposed to estimate the visit aspiration of these candidates by considering both user-restaurant-time's latent relationship and restaurants' contextual features (e.g., price, taste, service quality, etc.). On the other hand, when a user is predicted to be not novelty-seeking, the recommendation candidates are generated from the restaurants he has previously visited. We then present a Hidden Markov Model (HMM) with temporal regularity observations to estimate the visit probability for each candidate.

Our evaluation consists of multiple parts. First, we conduct several experiments on novelty-seeking status inference to evaluate the performance of our CRF with constraints method. Then, we show the results of recommending new restaurants and previously visited restaurants, respectively. Finally, we evaluate the effectiveness of our whole system by comparing with several competitive baselines.

WWW 2015, May 18-22, 2015, Florence, Italy.
ACM 978-1-4503-3469-3/15/05.
http://dx.doi.org/10.1145/2736277.2741095

To the best of our knowledge, our work is the first attempt to focus on recommending restaurants from the prospective of novelty-seeking. The key contributions of this paper include the following:

- We propose a framework which designs associated recommendation strategies based on different novelty-seeking statuses.
- We apply a CRF with constraints method to infer novelty-seeking status by considering spatial-temporal-historical features and users' socio-demographic characteristics.
- We present a context-aware collaborative filtering method to recommend new restaurants, and a HMM with temporal regularity to recommend visited restaurants.

The result of this paper is organized as follows. Section 2.2 introduces the preliminary concepts and present the dining recommendation problem. Section 3 first gives an analysis of novelty-seeking tendency and then presents the overview of our system. In Section 4, we discuss how to infer novelty-seeking status. The new and visited restaurant recommendation methods are presented in Section 5 and Section 6, respectively. The experimental results are discussed in Section 7, followed by a brief review of related work in Section 8 and a conclusion of this paper in Section 9.

2. PROBLEM STATEMENT

In this section, we first clarify some terms commonly used in this article, and then explicitly present our problem.

2.1 Preliminary

Point of Interest (POI) and **Restaurant**: A POI p refers to a specific point location that someone may find useful or interesting. It is described by a latitude, a longitude, and a category (such as restaurant, gas station, etc.). Note that a restaurant v is a particular POI with "Restaurant" category.

Check-In and **Dining Check-In**: A check-in record c is a triple $\langle u, p, t\rangle$ which indicates user $c.u$ visited POI $c.p$ at the particular time $c.t$. We use the term "dining check-in" to denote the check-in made in a restaurant.

Novel Restaurant, **Regular Restaurant**, **Novel Check-In**, **Regular Check-In**: For a user u at time t, a restaurant v is novel if this user has not visited this restaurant before time t, otherwise the restaurant v is regular if this user has already visited this restaurant before time t. A check-in in a novel restaurant is called novel check-in and similar to regular check-in.

Novelty-Seeking Status: Novlety-seeking status is an indicator to represent whether a user will like to visit new places. We use a binary value to indicate whether this user would visit a novel restaurant ($s = 1$) or a regular restaurant ($s = 0$) at a particular time.

2.2 The Dining Recommendation Problem

In this study, we consider the dining recommendation problem as follows: given a collection of users' check-in history including both dinning check-ins and check-ins at other POIs, socio-demographic characteristics (e.g., gender, age, education status) and restaurants' attributes (e.g., environment, price, rating), for the next dining of a user, assume the exact time is already known, we aim to recommend top-K restaurants he would most probably visit.

3. FRAMEWORK

In this section, we first present the analysis of novelty-seeking behavior in users' dining pattern, and then present the overview of our novelty-seeking based dining recommender system (**NDRS**).

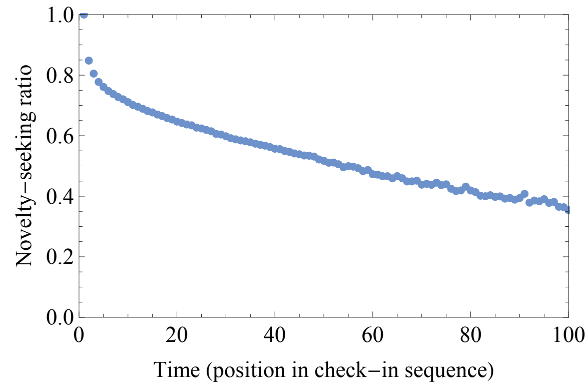

Figure 1: Novelty-seeking ratio w.r.t time

3.1 Novelty-Seeking Tendency Analysis

Before delving into the dining recommender system, a question concerned in this paper needs to be answered first: when a user dines out, will he prefers to try a novel restaurant for a certain level of freshness as well as risk-taking, or will he insist on a restaurant visited frequently in his past time? If a user frequently visits novel restaurants, the traditional statistical model based on his preference of historically visited restaurants would most probably fail. To have a deeper insight into the extent of user's exploration of novel restaurants, we define *novelty-seeking ratio* as the ratio of the number of dining check-ins with novelty-seeking status $s = 1$ to the number of total dining check-ins before a time t. As shown in Figure 1, at the beginning of dining behavior, it is obvious that almost all of them are novel. Over time, the novelty-seeking ratio declines as users will patronize the restaurants they have visited in the past. However, the novelty-seeking ratio declines slowly and it still remains about 0.4 even at the time stamp of the 100th dining visit. Such a significant part of novel dining check-ins indicates that users are keen on exploring new restaurants. On top of this, compared to traditional location recommendation techniques which mainly rely on each user's repetitive visit behavior, we should design the associated recommendation strategy in our system by taking users' novelty-seeking tendency into consideration.

3.2 System Overview

Our system is committed to generating the top-K restaurants a user would most probably visit in the next dining. In our system, by considering the influence of novelty-seeking tendency, a user's next dining decision is captured by first estimating his novelty-seeking status and then designing associated strategies to model the dining behavior for novel and regular restaurants, respectively. To infer the novelty-seeking status s of next dining, we propose a conditional random field (CRF) with additional constraints method which incorporates a CRF model with priori knowledge as constraints. Next, when the user is predicted to do exploration (novelty-seeking status $s = 1$), we propose a context-aware collaborative filtering to recommend novel restaurants. Otherwise, when the user is predicted to be backward-looking (novelty-seeking status $s = 0$), we present a Hidden Markov Model considering temporal regularity to recommend regular restaurants. Finally, the previous parts are integrated to generate the final recommendations. There are mainly three components in our system as shown in Figure 2.

- *Novelty-Seeking Status Inference*: This component presents a Conditional Random Field (CRF) with constraints model to predict the novelty-seeking status of next dining. In this method, we consider various factors that would influence the

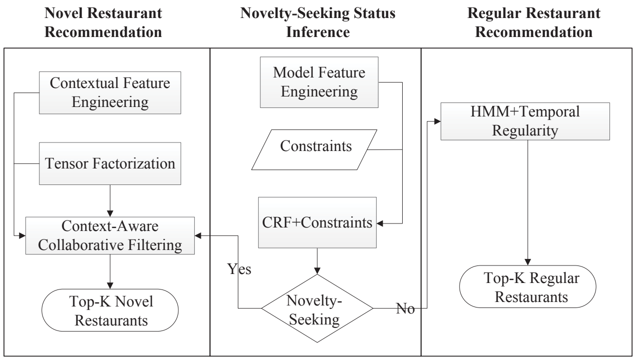

Figure 2: System overview.

novelty-seeking status, e.g., the previous novelty-seeking status, spatial-temporal limitation, users' socio-demographic characteristics, etc.

- *Novel Restaurant Recommendation*: In this component, by integrating the latent relationship in collaborative filtering and the effect of contextual features, we present a context-aware collaborative filtering method to recommend top-K novel restaurants.
- *Regular Restaurant Recommendation*: In this component, we use a Hidden Markov Model considering temporal regularity features to recommend top-K regular restaurants.

4. NOVELTY-SEEKING INFERENCE

4.1 Modeling Novelty-Seeking Behavior

According to the definition of novelty-seeking status in Section 2.1, we define our novelty-seeking status inference problem as follows: given the dining check-ins of a user, the novelty-seeking status inference problem aims to predict whether this user will visit novel restaurants (novelty-seeking status $s = 1$) or patronize regular restaurants (novelty seeking status $s = 0$) in his next dining.

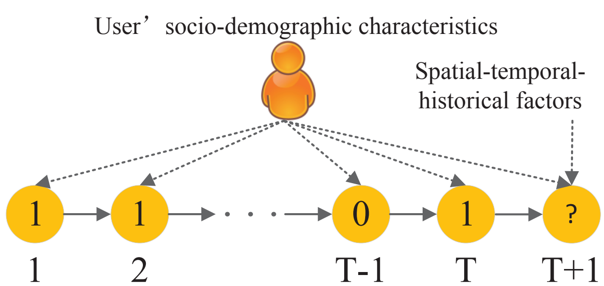

Figure 3: A user's novelty-seeking status sequence in dining check-ins, where each novelty-seeking status is influenced by the previous status, spatial-temporal-historical factors, and this user's socio-demographic characteristics.

A user's historical novelty-seeking statuses can be obtained based on the dining check-ins, and our goal is to predict the novelty-seeking status of next dining. For a user, we consider the novelty-seeking status of a check-in is influenced by the statuses of previous check-ins, since the fact that if this user has already done amount of explorations recently, he would likely have a rest and visit the company cafeteria which he is quite familiar with and vice versa. For simplicity and computational efficiency, we consider the first-order dependency as shown in Figure 3. Furthermore, various factors such as spatial-temporal limitation could also exert an effect on the novelty-seeking status. For example, the time of the dining out will influence the decision for exploration, a user will most probably have breakfast at a regular restaurant nearby his home and try some novel restaurants in the evening.

We first apply a conditional random field (CRF) to infer the next novelty-seeking status. Conditional Random Fields (CRFs) [21], which is a discriminative undirected probabilistic graphical model for parsing sequential data like natural language texts [31], has been successfully applied to sequential labeling problems in machine learning and data mining. To be specific, we first construct a linear chain CRF as shown in Figure 4. There are two kinds of nodes $\mathbf{G} = \{\mathbf{X}, \mathbf{Y}\}$, where the white nodes $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_T\}$ represent the hidden states to be inferred given the sequence of observations denoted by gray nodes $\mathbf{X} = \{X(1), X(2), \ldots, X(T)\}$. Note that in our scenario, Y_t indicates the novelty-seeking status at time t and $X(t) = (X_1(t), X_2(t), \ldots, X_m(t))$ is a vector which indicates the spatial-temporal-historical features (detailed in Section 4.2) we observe at that moment.

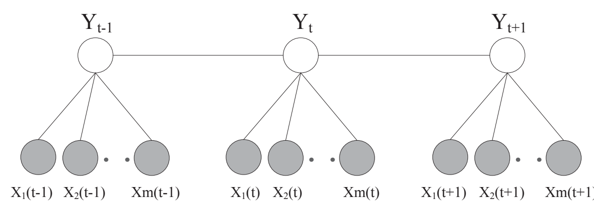

Figure 4: The graphic presentation of observations and labels in the liner chain CRF

With the labeled sequences dataset $\mathcal{D}$ with length N, CRF is typically trained by maximizing the penalized conditional log-likelihood as follows,

$$L(\lambda, \mathcal{D}) = \sum_{i=1}^{N} \log p(\mathbf{y}^{(i)}|\mathbf{x}^{(i)}) - \frac{\sum_k \lambda_k^2}{2\sigma^2}, \tag{1}$$

$$p_\lambda(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\lambda)} \exp\left(\sum_{t=1}^{T-1}\sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t+1}, \mathbf{x})\right), \tag{2}$$

$$Z(\lambda) = \sum_{\mathbf{y}} \exp\left(\sum_{t=1}^{T-1}\sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t+1}, \mathbf{x})\right). \tag{3}$$

where $f_k(y_t, y_{t+1}, \mathbf{x})$ is the feature function of the entire observation sequence and the labels at time t and $t+1$. The model assigns each feature function a numerical weight and combines them together to determine the probability of a certain value for Y_i. In addition, $\frac{\sum_k \lambda_k^2}{2\sigma^2}$ is the regularization term to avoid over-fitting.

4.2 Model Features Engineering

Based on the dining check-ins, for a user u at time t, we extract the spatial-temporal-historical features as follows,

Spatial features: As mentioned in [34], a user's novelty-seeking tendency is greatly influenced by the spatial limitation, here we consider two spatial features,

1) **Area Exploration Ratio**: this feature, which tends to reveal whether a user has fully explored the restaurants inside a area, is constructed with a procedure as like this: 1) First, we apply DBSCAN [10] (a density-based clustering algorithm) to a user's dining check-ins to detect his dining clusters. Note that each isolated

check-in is also treated as a cluster respectively. 2) Next, for each cluster, we identify a circular area which centres on the center of the cluster with a radius r (500m in our settings). In fact, there are also many other unvisited restaurants located in a discovered area 3) Finally, area exploration ratio is calculated as the ratio of the number of visited restaurants to the number of total restaurants (including both the visited and unvisited ones) located in all discovered areas. Actually, if a user prefers to explore as many restaurants as possible inside an area, it's most probably that he is enthusiastic about neophilia.

2) **Average Dining Distance** and **Dining Distance Variance**: Average dining distance measures the average distance of the previously visited restaurants, which indicates how far away a user would like to dine out. Dining distance variance calculates the variance of the distances between any two visited restaurants, which indicates whether this user would like to visit the restaurants concentrated in an area. These two quantities together can influence a user's novelty-seeking tendency to some extent in terms of the travel distance. To be more specific, if both average dining distance and dining distance variance are small, it strongly suggests that this user's dining behavior is limited to a small area (e.g., nearby to home or office), and thus the opportunity of exploring new restaurant will also be small. However, when both of these two quantities are large, it implies that this user would sometimes explore new restaurants in a faraway area.

Temporal features: The previous studies also imply that temporal status will influence a user's novelty-seeking tendency [30], here we consider three temporal features as follows,

1) **Total Time Interval**: This feature represents the time interval between this dining check-in and this user's first dining check-in, which is expressed as the number of days.

2) **Last Time Interval**: This feature represents the time interval between this check-in and the previous dining check-in, which is expressed as the number of hours.

3) **Hour of Week**: At different hours of a day, the probability of novelty-seeking might be different. In addition, the probability of novelty-seeking might also vary during different days of the week, e.g., people would prefer to explore new restaurant in weekend than in weekday. Therefore, we use a value from $\{0, \ldots, 167\}$ to represent the hour of the week.

Historical features: A user's dining history will influence his current novelty-seeking status. Accordingly, we identify two features,

1) **Visit Entropy**: This feature is calculated according to the previous visit frequency at each restaurant. To be more specific, for a user at a time, if the visit frequency of the previously visited restaurants are $f_1, f_2, \ldots, f_S$ respectively, his visit entropy is given as $H = -\sum_i \frac{f_i}{\sum_j f_j} \cdot \log_2 \frac{f_i}{\sum_j f_j}$. The visit entropy can be regarded as a kind of reference for novelty-seeking tendency, since a smaller value of this quantity implies this user's restaurant visit is more uniformly distributed on all the previously visited ones, and thus the probability of exploring new restaurant will be higher. The relationship between the entropy of a user's behavioral data and his novelty-seeking trait is also discussed in [43], which explicitly indicates that the novelty-seeking trait has a significant negative correlation with the entropy of behavioral data.

2) **Previous Novelty-Seeking Ratio**: It is calculated as the ratio of the number of novel dining check-ins to the number of total dining check-ins. This quantity measures this user's previous novelty-seeking tendency as a whole.

4.3 Constraints

However, except for these spatial-temporal-historical features and the dependency between adjacent labels contained in the single CRF, users' socio-demographic characteristics, have not been thoroughly leveraged in the model. Actually, psychologists have found that socio-demographic characteristics (i.e., age, gender, education status, etc.) have a great effect on users' novelty-seeking tendency [22]. These characteristics will influence a user's dining behavior as a whole, and thus they are more suited to be designed as a kind of priori knowledge for each user instead of the CRF's model features which mainly determine the instantaneous state. Therefore, we employ the Generalized Expectation Criterion [25] as an objective function to incorporate socio-demographic characteristics as prior knowledge. The General Expectation Criterion makes very natural an under-explored paradigm for incorporating the prior knowledge. Suppose we have already obtained the prior knowledge that the male is keen on exploring novel restaurants. To be more specific, the prior knowledge tells us desirable properties of the distribution over output variables. The distribution we want to match deriving from the prior knowledge is called target distribution, e.g., $p(s = 1|\text{male}) = 0.6$, $p(s = 0|\text{male}) = 0.4$. By carefully designing a constraint feature function denoted as $\phi(\mathbf{x}, \mathbf{y})$, the CRF model will also generate an expectation over the constraint feature function under the CRF's conditional distribution $p_\lambda(\mathbf{y}|\mathbf{x})$ as follows,

$$E_\lambda[\phi] = E_{p_\lambda(\mathbf{y}|\mathbf{x})}[\phi(\mathbf{x}, \mathbf{y})] = \sum_{\mathbf{y}} p_\lambda(\mathbf{y}|\mathbf{x})\phi(\mathbf{x}, \mathbf{y}). \tag{4}$$

The General Expectation Criterion encourages the gap (e.g., the square distance applied in this article) between the model expectation distribution $E_\lambda[\phi]$ and the target distribution is as small as possible. For the constraint feature functions, an example related to the relationship of the male and novelty-seeking status should be designed as

$$\phi_{\text{male,novel}}(\mathbf{x}, \mathbf{y}) = \sum_{t=1}^{T} \mathbb{I}(y_t = 1)\mathbb{I}(\mathbf{x}\ is\ \text{male}), \tag{5}$$

$$\phi_{\text{male,regular}}(\mathbf{x}, \mathbf{y}) = \sum_{t=1}^{T} \mathbb{I}(y_t = 0)\mathbb{I}(\mathbf{x}\ is\ \text{male}). \tag{6}$$

where $\phi_{\text{male,novel}}(\mathbf{x}, \mathbf{y})$ represents the constraint that the male would explore novel restaurant and $\phi_{\text{male,regular}}(\mathbf{x}, \mathbf{y})$ represents the constraint that the male visit regular restaurants. The indicator function $\mathbb{I}(\cdot)$ returns 1 if the statement in its parameter is true, and 0 otherwise.

To incorporate the General Expectation Criterion into the CRF model, the final objective function we intend to maximize is given by

$$L'(\lambda, \mathcal{D}) = L(\lambda, \mathcal{D}) - S(\tilde{\Phi}, E_{p_\lambda(\mathbf{y}|\mathbf{x})}[\Phi(\mathbf{x}, \mathbf{y})]). \tag{7}$$

where $\tilde{\Phi}$ is the empirical target distribution, $\Phi(\mathbf{x}, \mathbf{y})$ is the collection of constraint feature functions, and S is the score function expressing the distance (the square distance is applied in this article) between the model expectation and the target distribution. We follow the gradient-based methods mentioned in [8] to perform the optimization of Eq. (7).

For the constraint features, we consider a user's socio-demographic characteristics including gender, age and education status, where a user's education status is classified as college student and non-college student (college student are those who claim their college experience on the SinaWeibo profile while non-college students are on the opposite).

Ultimately, since both model features and constraint features needs categorical variables, each continuous variable is discretized to five categories according to the five-quantiles binning method.

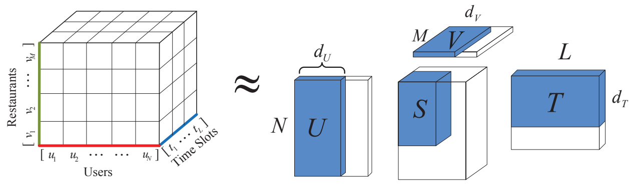

Figure 5: The illustration of tensor factorization, where the three dimensions of the tensor indicate user, restaurant and time slot. The value of each entry in the tensor indicates the visit frequency. The factorized three matrices indicate user feature, restaurant feature, and time slot feature in latent space, respectively.

5. NOVEL RESTAURANT RECOMMENDATION

Currently, if a user is predicted to be novelty-seeking ($s = 1$) in next dining, we will generate top-K recommendations from novel restaurants according to her historical dining behavior as well as other users' dining behavior. Futhermore, since the exact time of next dining is already known as mentioned in Section2.2, this scenario is similar to the user-item-time problem in a temporal-dynamic recommender system which will recommend users interested new items at a particular time. Given this, We present a context-aware collaborative filtering method to model users' preference for these novel restaurants at the given time in consideration of both latent user-restaurant-time relationship and restaurant's contextual features.

5.1 Context-Aware Collaborative Filtering

The goal of collaborative filtering method is to recommend novel restaurants to users at the next dining time according to their visit frequencies of regular restaurants. Considering user-restaurant-time as a three-dimensional tensor, denoted by $Y \in R^{N\times M\times L}$, where N is the number of users, M is the number of restaurants, and L is the number of hours in a day (for the time dimension, we divide a day into 24 hours). We apply tensor factorization, which is the state-of-the-art method used for collaborative filtering in the high dimensional situation [20]. For an entry C_{ijk} in the tensor Y, if the user i has visited the restaurant j at the time slot k in the historical data, we regard Y_{ijk} as being observed and use the frequency of this visit to denote the value. Therefore, our problem is then estimate the value of these unknown elements in the tensor.

We apply High Order Singular Value Decomposition (HOSVD) [7] to factorize the three-dimensional tensor into three matrices $U \in R^{N\times d_U}$, $V \in R^{M\times d_V}$, $T \in R^{L\times d_T}$ and one central tensor $S \in R^{d_U\times d_V\times d_T}$. The three matrices are compact representations of the three attributes in subspaces, where d_U, d_V, d_T are dimensionality parameters to balance the capability and generalization. The recosntructed value for user i, item j and time slot k is given as

$$F_{ijk} = S \times_U U_{i*} \times_V V_{j*} \times_T T_{k*}, \tag{8}$$

We regard the tensor matrix multiplication as $\times_U$, where the subscript denotes the direction, e.g., $C = S \times_P P$ is $C_{ijk} = \sum_{i=1}^{U} S_{ijk} \times P_{ij}$. The entries of the ith row of the matrix P is represented as P_{i*}.

Additionally, the single tensor factorization does not take full advantage of our data, since it only tries to find out user-restaurant-time's latent relationship in the feature space only through the observable visit frequencies. It does not consider other factors would also influence the users' dining preference for restaurants and time slots. Another important signal, the restaurant's contextual features, has not been well considered. Researchers in the hospitality management research has discovered that users' dining satisfaction and patronage is greatly influenced by the restaurant's contextual features, such as price, atmosphere and service quality [18]. Actually, an item's contextual features are often modeled into collaborative filtering to solve the uncertainty problems [2].

Assume that there are L contextual features, where the l-th feature has categorical values $1, 2, \ldots, z_l$, and the value 1 means that the l-th contextual feature was unknown, while other index values are possible contextual conditions. By incorporating these contextual features into the tensor factorization, according to [42], the reconstructed value for an entry C_{ijk} is redefined as

$$F'_{ijk} = S \times_U U_{i*} \times_V V_{j*} \times_T T_{k*} + \sum_{l=1}^{L} B_{lc_{jl}} \tag{9}$$

where c_{jl} is the value of the l-th feature for restaurant j, and $B_{lc_{jl}}$ is the parameter modeling how the l-th contextual feature with condition c_{jl} will influence the value of entry C_{ijk}. These contextual feature parameters guarantee the fact that restaurants with similar contextual features tend to have similar number of visit frequency.

In order to generate estimations for these novel restaurants, the model parameters should be learned using the observable frequency of these visited restaurants at the corresponding time slots. We define the learning procedure as an optimization problem, which is given as

$$\min_{S,U,V,T,B} L(Y, F') + \Omega(S, U, V, T, B), \tag{10}$$

where $L(Y, F')$ is the loss function given as

$$L(Y, F') = \frac{1}{||S||_1} \sum_{i,j,k} Z_{ijk} \cdot (Y_{ijk} - F'_{ijk}), \tag{11}$$

where $Z \in 0, 1^{N\times M\times L}$ is a binary tensor with nonzero entries Z_{ijk} whenever Y_{ijk} is observed. Equation (11) indicates that we consider the reconstructed accuracy for the observed entries. In addition, $\Omega(S, U, V, T, B)$ is the regularization term to avoid overfitting, which is given as

$$\begin{aligned}\Omega(S, U, V, T, B) =& \lambda \times (||S||^2_{Frob} + ||U||^2_{Frob} \\ &+ ||V||^2_{Frob} + ||T||^2_{Frob} + ||B||^2_{Frob}),\end{aligned} \tag{12}$$

Equation (10) guarantees that our context-aware collaborative filtering method can reconstruct the observations as accurately as possible and meanwhile maintain the capability of generalization. We use stochastic gradient descent [45] to solve this optimization problem.

Finally, after we obtain these parameters, for a user i, if the time slot of her next dining is k, the top-K novel restaurants will be recommended according to the preference ranking of the restaurants, which is given as

$$u_i : v_1 \succ v_2 \succ \cdots \succ v_M \longrightarrow F'_{i1k} > F'_{i2k} \cdots > F'_{iMk}, \tag{13}$$

5.2 Contextual Features Extraction

We consider four types of contextual features for restaurants as follows,

Popularity: This feature, which measures the popularity of a restaurant, is calculated as the sum of visit frequency in this restaurant for all users.

Area Popularity: This feature measures the overall popularity of an area where a restaurant locates in. It is natural to consider that when a restaurant locates in a more popular area, it has the potential to attract more consumers. For a restaurant, we first determine a circular area which centres on the center of this restaurant with a given radius (500m in our settings). The area popularity is then calculated as the sum of popularity for these POIs (including both restaurants and non-restaurants) locating in the identified area.

Area Attraction: We identify the area attraction feature according to what kind of POIs locate close to a restaurant. First, for each category C of the POIs, to determine its attraction to the restaurant category, we use the metrics defined in [16], which is given as

$$A_C = \frac{\#\text{co\_location}(C, \text{Restaurant})}{\#C} \tag{14}$$

where $\#\text{co\_location}(C, \text{Restaurant})$ refers to the frequency of co-location for category C with restaurant, while $\#C$ is the frequency of category C. Note that co-location indicates that the geographic distance between a restaurant and a POI with category C is less than 500m. The top 3 discovered POI categories are {Living Quarters, Shopping Mall, College}. Next, by aggregating nearby POIs, the area attraction of a restaurant j is given as

$$AA_j = \sum_{C} \#\text{co\_location}(C, j) \cdot A_C \tag{15}$$

where $\#\text{co\_location}(C, j)$ refers to the frequency of co-location for category C with this particular restaurant j.

Restaurant Attributes: We consider the following attributes for a restaurant: 1) *Restaurant Category*. There are total 49 restaurant categories in our dataset, such as Sichuan cuisine, Cantonese cuisine and Japanese cuisine. 2) *Price, Rating, Taste, Environment, Service Quality*. Since the restaurant in our check-in dataset is not described by these attributes, we crawl restaurant attributes from external sources and map them to the corresponding restaurant in our check-in dataset, which will be detailed in Section 7.1.

After obtaining these features, since the context-aware collaborative filtering model needs categorical feature variables, each feature with continuous value is discretized to five categories according to the five-quantiles binning method.

6. REGULAR RESTAURANT RECOMMENDATION

When a user is predicted to be not novelty-seeking ($s = 0$) in Section 4, we need to generate the top-K recommendations from a candidate pool of his regular restaurants.

In consideration of the sequential dependency widely discussed in location prediction [36] and the temporal influence on location visit [35], we apply a Hidden Markov Model (HMM) considering the temporal regularity to model a user's dining pattern. To be more specific, a user's visited restaurants in the dining check-in sequence are regarded as hidden states and the related temporal information are regarded as external observations. For the temporal information, we consider the hour of the day and distinguish weekday from weekend, thus each check-in's timestamp is mapped to a vector $t = \{h, w\}$, where $h \in \{0, 1, \ldots, 23\}$ and $w \in \{0, 1\}$.

For each user, a personalized HMM model is specified for him and we use supervised learning [24] to determine three types of parameters: emission probability, initial hidden state probability and state transition probability.

Emission probability $P(t|v)$ indicates the conditional dependency for a timestamp (observation) $t = \{h, w\}$ by a restaurant v. This probability is directly determined by maximum likelihood estimation with a Laplace smoothing as $P(h|v) = \frac{N(h,v)+1}{\sum_h N(h,v)+24}$ and $P(w|v) = \frac{N(w,v)+1}{\sum_w N(w,v)+2}$, where $N(h, v)$ is the frequency of check-ins in restaurant v at hour h and similar to $N(w, v)$. Furthermore, we observe that a user would only visit restaurants with 4.2 hour bins on average, e.g., a user might only visit restaurants at 7 a.m., 12 p.m. or 6 p.m.. To ensure nearby hours have similar conditional probabilities, the emission probability for hour h is further transformed by a Gaussian kernel smoothing function as follows,

$$P'(h|v) = \frac{\sum_{h'} \mathcal{N}(\frac{\text{dis}(h,h')}{\sigma_{h',v}}) P(h'|v)}{\sum_g \sum_{h'} \mathcal{N}(\frac{\text{dis}(g,h')}{\sigma_{h',v}}) P(h'|v)}. \tag{16}$$

where $\text{dis}(h, h') = \min(|h - h'|, 24 - |h - h'|)$ is the cyclic distance between hour h and h' in a day. $\mathcal{N}(\cdot)$ is a standard Gaussian distribution and $\sigma_{h',v}$ is set to be $N(h', v)^{-\frac{1}{5}}$ according to [33].

The initial state probability is estimated by maximum likelihood as $P(v) = \frac{N(v)}{\sum_v N(v)}$, where $N(v)$ is the check-in frequency in restaurant v. Similarly, the state transition probability can also be estimated by maximum likelihood with a Laplace smoothing as given by

$$P(v|v_p) = \frac{N(v_p, v) + 1}{2 \times \sum_{v'} N(v_p, v')} \tag{17}$$

where $N(v_p, v)$ indicates the frequency of visiting restaurant v right after restaurant v_p.

Given these parameters estimated, we generate the top-K regular restaurants according to the visit probability of a restaurant v as follows,

$$p(v|v_p, h, w) \propto p(v_p|v) \cdot p(h|v) \cdot p(w|v) \cdot p(v) \tag{18}$$

7. EXPERIMENTS

In this section, we first describe and analyze the data we collected from two specific websites. Based on this huge dataset, we then conducted extensive experiments to evaluate the performances of novelty seeking status inference, efficient novel restaurant recommendation, and regular restaurant recommendation, respectively. Finally, we validate the performance of our whole framework **NDRS**.

7.1 Data Collection and Description

We tested the performance of our system based on two publicly available websites: SinaWeibo and DianPing.

SinaWeibo, which is the largest social netowrk website in China, also provides location-based services such as check-ins. We crawled 135 million check-ins from 3 million users with the breath-first strategy using Lifespec crawling platform [41]. For each POI, we crawled its description including id, geographic coordinates and category. For each user, we also crawled his demographic information including gender, age and education status before 1st June 2014. To clean the datasets, we first filter the noisy data, e.g., repeated check-ins at the same place in quite a short interval (1hr is adopted in our setting). Next, since our system focuses on the dining check-ins of a user, we need sufficient restaurant visits. On top of this, we adopt two strategies to further filter the check-in dataset: 1) For a user, we only focus on the dining check-ins which are located in the city where she resides (note that we represent the city she resides is the place where most of her check-ins locate). The reason is our novel restaurant recommendation method only focus on the restaurants in her residence city. For example,

Table 1: Summarization of collected dataset for different cities (partially presented due to page limit).

City	Shanghai	Beijing	Chongqing	Guangzhou	Chengdu	Shenzhen	Nanjing	Tianjin	Xian	Shenyang
Users	183,239	69,412	6,787	42,694	10,788	9,462	12,760	10,137	7,807	6,015
Restaurants	34,127	27,120	6,948	18,625	12,074	9,931	8,570	7,349	7,095	5,268
Dining check-ins	2,582,914	895,120	89,749	542,563	152,227	135,302	167,826	120,802	86,651	65,841

Figure 6: A restaurant's description on DianPing

if a user resides in Beijing, it is actually difficult and meaningless to recommend restaurants in Shanghai if she does not visit Shanghai frequently. 2) We remove those users who have fewer than 10 dining check-ins, which is to ensure sufficient observable dining history. 3) For a city, if the number of users who reside in this city is smaller than 5,000 or the number of restaurants located in this city is smaller than 1000, we remove all the users who reside in this city and their related check-ins. The reason is that when we generate the Top-K novel restaurants for a user in Section 5, all the restaurant candidates are selected from the city he resides. For each city, the context-aware tensor factorization method is applied, respectively. Thus, we need enough users and restaurants in a city to ensure the user-restaurant-time tensor factorization.

After the filtering procedures, we eventually obtained a collection of 361,218 unique users from 21 cities all over China, with 4,941,060 dining check-ins. Table 1 summarizes the statistics of the final dataset for different cities.

Furthermore, since the restaurant in our SinaWeibo check-in dataset is only described by geographic coordinates and category, to obtain the restaurant attributes discussed in Section 5.2, we need to link external data sources. As shown in Figure 6, DianPing (similar to Yelp, the largest online review website in China) is such a desired website where the restaurants are described by sufficient attributes in detail. First, for each of the 21 cities from our SinaWeibo check-in dataset, we crawled all restaurants located in this city from DianPing by grabbing web pages directly, and extracted prices, ratings, tastes, environments and service qualities from the raw webpages. Then, we applied a two staged method described in [32] (including both geographic filtering and title string match) to map each restaurant from SinaWeibo to a restaurant from DianPing. Finally, we explicitly observed 500 paired results, where 91.5% of the pairs are correctly matched (the match precision is good enough for later process). In addition, we find that only 5.7% of restaurants from SinaWeibo can not be mapped to any restaurant from DianPing.

To evaluate the performance of our proposed methods, for each user, her previous 90% dining check-ins were used as training data and the remaining dining check-ins are used as testing data.

7.2 Experiments for Novelty-Seeking Inference

In this subsection, we study the performance of novelty-seeking status inference. First, for each constraint feature, we constructed the target distribution according to the empirical distribution in the training data. For example, if 60% of the dining check-ins generated by the male is novel and 40% is regular, the target distribution for constraint feature "male" is $p(s=1|\text{male})=0.6, p(s=0|\text{male})=0.4$. Next, we compared our method (CRF with model features, optimized with constraints), shortened as "**CRF(M)+C**" against the following baselines,

- **LR** (Logistic Regression): This algorithm uses both the model features and constraint features for the Logistic Regression.
- **CRF+C** (CRF with only constraints): This algorithm uses the same settings as **CRF(M)+C**, except that it does not incorporate model features.
- **CRF(M)** (CRF with only model features): This algorithm uses the same settings as **CRF(M)+C**, except that it does not incorporate constraints.
- **NSTM** (Novelty-Seeking Trait Model): This Bayesian model presented in [43] is a state-of-the-art method to model user's sequential behavior in consideration of novelty-seeking and preference. To accommodate our scenario, we set a user's novelty-seeking status as 0 or 1 (it is originally set from 1 to 5 in [43]).

Table 2: The results of novelty-seeking inference.

	Accuracy	True Positive Rate	False Negative Rate
CRF(M)+C	0.823	0.786	0.845
LR	0.641	0.593	0.667
CRF(M)	0.795	0.762	0.811
CRF+C	0.633	0.624	0.639
NSTM	0.748	0.767	0.719

If we consider novelty-seeking status $s=1$ as positive and the other as negative, the results of true positive rate, false negative rate as well as accuracy are shown in Table 2, where accuracy indicates the percentage of correct prediction for all check-ins in the testing data, true positive rate indicates for these novel check-ins ($s=1$) in the testing data, the percenrage of correct prediction, and false negative rate indicates for these regular check-ins ($s=0$), the percenrage of correct prediction. It is clear that our method significantly outperforms competitors in all three criterions. For example, by using the constraints, our method achieves an improvement over **CRF(M)**. Compared our method with **CRF+C**, it is obvious that model features are significantly crucial to the detection of novelty-seeking status. We can see that **LR** performs worst due to the fact that this method does not take the sequential dependency of novelty-seeking statuses into consideration. In addition, the reason why **NSTM** is defeated by our method **CRF(M)+C** is that **NSTM** only considers a user's dining sequence, while our method incorporates various model features as well as users' socio-demographic characteristics as constraints. Since regular restaurant recommendation is easier than novel restaurant recommendation as shown in Section 7.3 and Section 7.4, to give a more accurate recommendation as a whole, we would expect that the False Negative Rate is as large as possible (if a user want to visit regular restaurants, we expect that the error of predicting her novelty-seeking status as $s=1$ is as small as possible). In view of this, our method also outperforms these baselines evidently in the False Negative Rate criterion.

7.3 Experiments for Novel Restaurant Recommendation

In this subsection, we study the performance of our context-aware collaborative filtering model in various situations.

We compare our context-aware collaborative filtering method (**CACF** for short) against the following methods,

- **CF** (Collaborative Filtering): This algorithm uses the same settings as **CACF**, except that it does not incorporate contextual features.
- **LR** (Logistic Regression). This algorithm uses a Logistic Regression with all the contextual features in our **CACF**.
- **PPTM** (Personal Popularity Tendency Matching). This algorithm proposed in [28] gives an effective novel item recommendation by reasonably penalizing popular items while improving the recommendation accuracy, which is a state-of-the-art method in recommending novel items.

The evaluation method of novel recommendation is as follows: 1) For each city, we collect all the users who reside in this city and then use their training check-ins (the previous 90% check-ins mentioned in Section 7.1) to learn a model and then obtain a user's evaluation result on her novel check-ins in the testing data. Note that we only test the performance on novel check-ins in the testing data, therefore regular check-ins in the testing data are ignored. 2) The final result is the average value of the results from all users.

We used nDCG, a widely adopted metric in information retrieval, to evaluate the performance of novel restaurant recommendation. We first listed recommendation candidates in a descending order according to the estimation, and used $\mathrm{rel}_i = 1$ to indicate that the i-th recommended restaurant was just the one visited by the user and $\mathrm{rel}_i = 0$ otherwise. Next, we used nDCG@p to evaluate the performance given by

$$\mathrm{nDCG@p} = \frac{\mathrm{DCG_p}}{\mathrm{IDCG_p}}, \quad \text{where} \quad \mathrm{DCG_p} = \sum_{i=1}^{p} \frac{2^{\mathrm{rel}_i} - 1}{\log_2(i+1)}, \tag{19}$$

where IDCG_p is the value of DCG_p for a perfect ranking.

We used nDCG@10 in our reported results, which are shown in Figure 7, where dim means the feature dimension ($dim = d_U = d_V = d_T$) in the tensor factorization. The table shows that **CACF** outperformed the baselines in various situations. Compared **CACF** with **CF**, it shows that incorporating contextual features is effective for the recommendation. Compared **CF** with **LR**, it also implies that tensor factorization is more suitable for novel restaurant recommendation.

7.4 Experiments for Regular Restaurant Recommendation

In this subsection, we evaluate the performance of regular restaurant recommendation. We compared our temporal related Hidden Markov Model (**TM** for short) against the following baselines,

- **MC** (Markov Chain): This method models sequential behavior by learning the transition graph over restaurants. Actually, this algorithm only uses the transition probability in our **TM**.
- **TR** (Temporal Regularity): This method models the temporal regularity of restaurant visit. It only uses the emission probability in our **TM** to generate the recommendation list.

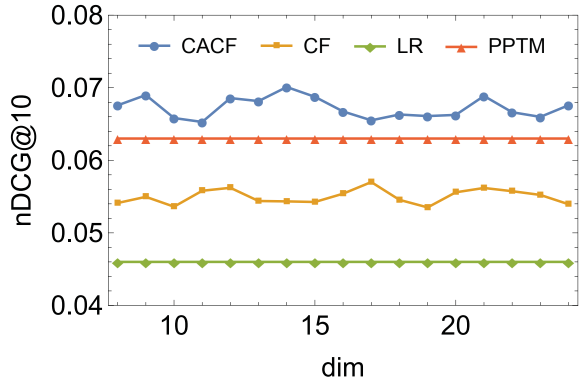

Figure 7: The results nDCG@10 of recommending w.r.t. latent dimension dim

- **OF** (Order by Frequency): This algorithm gives a recommendation list of restaurants according to this user's visit frequency in a restaurant.

The evaluation method of regular recommendation is as follows: 1) For each user, we use her training check-ins (the previous 90% check-ins mentioned in Section 7.1) to learn a model and then obtain this user's evaluation result on her regular check-ins in the testing data. Note that we only test the performance on regular check-ins in the testing data, therefore novel check-ins in the testing data are ignored. 2) The final result is the average value of the results from all users.

We also used nDCG@10 to evaluate the performance of each algorithm, and the results are shown in Table 3. It shows our **TM** outperforms the competitors greatly, which implies that the combination of Markov dependency and temporal regularity is more suitable for regular restaurant recommendation.

Table 3: The results nDCG@10 of recommending regular restaurants

TM	MC	TR	OF
0.374	0.327	0.313	0.332

7.5 Experiments for NDRS

In this subsection, we evaluate the performance of our whole system **NDRS**, which combines the results from the three components together to generate the final next dining recommendation list. We compared our system against the following baselines:

- **CACF+**: This method has the same settings as **CACF** mentioned before, except that for each user, the recommendation list is based on all the restaurants (including both novel and regular restaurants) located in the city. It does not emphasize the regularity in a user's previous dining pattern.
- **TM**: This method only use the regular restaurant recommendation as mentioned in Section 6, which does not consider recommending novel restaurants.
- **FPMC-LR**: This methods is proposed in [5], which embeds users' preference, personalized Markov chain, and localized region constraints into next POI recommendation.

Note that in this part, the testing data includes both novel and regular check-ins. We also employ nDCG@10 for the performance evaluation. As shown in Figure 8, our system **NDRS** enjoys the best performance compared to its competitors. Compared **NDRS** to

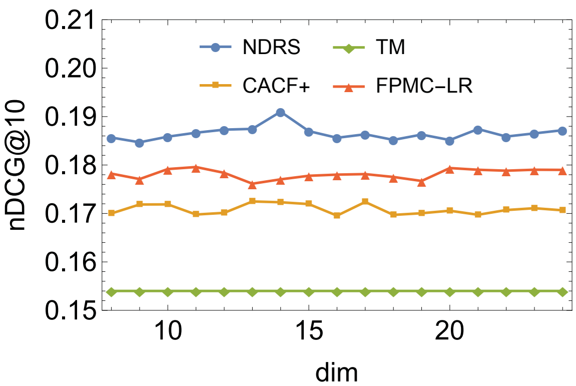

Figure 8: The results nDCG@10 of dining recommendation w.r.t. latent dimension dim

CACF+ and **TM**, it shows that integrating novel and regular restaurant recommendation together will give a better result. Compared **NDRS** to **FPMC-LR**, it also shows that using novelty-seeking status to determine whether recommending novel restaurants is effective for the dining recommendation scenario.

8. RELATED WORK

8.1 Location Recommendation

In recent years, with the rapid accumulation of spatial-temporal records in the check-in data and the prevalence of various interesting real-world applications [40], the location recommendation problem has received much attention. Ye et al. [38] exploited the social and geographical characteristics of users and location/places to generate the next location recommendation. Zheng et al. [47] used GPS data and users' comments at various locations to discover interesting locations and possible activities that can be performed for recommendation. Cheng et al. [4] first fused matrix factorization with geographical and social influence for POI recommendation.

Compared to previous works which mainly focus on regular mobility patterns, our work primarily aims at proposing a framework based on novelty-seeking status to predict whether a user will visit novel restaurants or regular restaurants.

8.2 Novelty-Seeking in Dining Behavior

Novelty seeking is also termed sensation seeking or neophilia. It has long been studied in psychology, consumer behavior and health science [11, 14, 23]. Acker and Mcreynolds [1] mentioned that novelty-seeking appears to be that through internal drive and external motivating force, the individual is then motivated to seek out novel information. There are two aspects to novelty-seeking that are likely to be correlated. The first aspect is seeking new and potentially discrepant information, which is emphasized by Fiske and Maddi [12]. The second aspect is the extent to which individuals would like to vary their choices among familiar contexts [13].

Novelty-seeking is closely related to dining behavior. Kivela et al. [17] proposed that with the dining experience, the quest to sample various food styles is one of many appealing experiences users aspire to achieve. Kivela and Chu [19] discovered that the adventure and hedonistic dishes found in a restaurant are sources of pleasure they help satisfy individuals' sensational desires particularly. Pizam et al. [29] moved forward to conclude that it is possible to predict the meal and types of food that tourists would prefer based on determining the relative level of novelty-seeking.

Compared to previous works which investigated this problem mainly by surveys or interviews, we present a computational framework to infer a user's novelty-seeking status and then use novelty-seeking status to provide insight for different recommendation strategies.

8.3 Novelty in Recommendation

It has been well acknowledged that novelty and diversity are important aspects in evaluating the performance of a recommender system [26]. Zhang et al. [46] addressed five measures to capture the novelty and redundancy of relevant documents in an adaptive information filtering system. Using these measures, the system can determine whether an item which is considered relevant contains any novel information to the user. Hurley and Zhang [15] took the view that novel items have greater utility and focus on strategies to recommend novel items. The authors formulated the trade-off between novelty and matching quality as a binary optimization problem and used an explicit control parameter to allow the tuning of this trade-off. Oh et al. [28] proposed an efficient novel-recommendation method which can help to diversify recommendations by reasonably penalizing popular items while improving the recommendation accuracy. Zhou et al. [48] presented an algorithm specifically to address the challenge of novelty and diversity, and showed it can be used to resolve the novelty-accuracy dilemma when combined in an elegant hybrid with an accuracy-focused algorithm. To better capture user's novelty tastes, Zhang and Hruley [44] proposed to partition the user profile into clusters of similar items and composed the recommendation list of items that match well with each cluster.

Compared with previous influential works that primarily aimed at recommending novel items, we step further to consider the novelty-seeking tendency of the user, which can help to understand the user's behavior at an intrinsic level.

9. CONCLUSION

This paper proposes a dining recommender system termed **NDRS**, which gives associated recommendation strategies according to different novelty-seeking statuses. Following the framework, we first design a CRF with constraints to infer novelty-seeking status. Next, a context-aware collaborative filtering method and a HMM with temporal regularity method are proposed for novel and regular restaurant recommendation, respectively. The extensive experiments we have conducted validated the effectiveness of our dining recommender system. Besides, this research sheds new light on other recommender systems such as POI and music recommendation, which can be used in more application scenarios.

References

[1] M. Acker and P. McReynolds. The" need for novelty": A comparison of six instruments. *The Psychological Record*, 1967.

[2] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23(1):103–145, 2005.

[3] T. D. Andersson and L. Mossberg. The dining experience: do restaurants satisfy customer needs? *Food Service Technology*, 4(4):171–177, 2004.

[4] C. Cheng, H. Yang, I. King, and M. R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In *AAAI*, volume 12, page 1, 2012.

[5] C. Cheng, H. Yang, M. R. Lyu, and I. King. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 2605–2611. AAAI Press, 2013.

[6] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.

[7] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[8] G. Druck, G. Mann, and A. McCallum. Learning from labeled features using generalized expectation criteria. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 595–602. ACM, 2008.

[9] R. P. Ebstein, O. Novick, R. Umansky, B. Priel, Y. Osher, D. Blaine, E. R. Bennett, L. Nemanov, M. Katz, and R. H. Belmaker. Dopamine d4 receptor (d4dr) exon iii polymorphism associated with the human personality trait of novelty seeking. *Nature genetics*, 12(1):78–80, 1996.

[10] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[11] R. Feng et al. Temporal destination revisit intention: The effects of novelty seeking and satisfaction. *Tourism management*, 28(2):580–590, 2007.

[12] D. W. Fiske and S. R. Maddi. Functions of varied experience. 1961.

[13] M. Givon. Variety seeking through brand switching. *Marketing Science*, 3(1):1–22, 1984.

[14] J. Hauser, G. J. Tellis, and A. Griffin. Research on innovation: A review and agenda for marketing science. *Marketing science*, 25(6):687–717, 2006.

[15] N. Hurley and M. Zhang. Novelty and diversity in top-n recommendation–analysis and evaluation. *ACM Transactions on Internet Technology (TOIT)*, 10(4):14, 2011.

[16] P. Jensen. Network-based predictions of retail store commercial categories and optimal locations. *Physical Review E*, 74(3):035101, 2006.

[17] J. Kivela and J. C. Crotts. Tourism and gastronomy: Gastronomy's influence on how tourists experience a destination. *Journal of Hospitality & Tourism Research*, 30(3):354–377, 2006.

[18] J. Kivela, R. Inbakaran, and J. Reece. Consumer research in the restaurant environment, part 1: A conceptual model of dining satisfaction and return patronage. *International Journal of Contemporary Hospitality Management*, 11(5):205–222, 1999.

[19] J. J. Kivelä and C. Y. H. Chu. Delivering quality service: Diagnosing favorable and unfavorable service encounters in restaurants. *Journal of Hospitality & Tourism Research*, 25(3):251–271, 2001.

[20] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.

[21] J. Lafferty, A. McCallum, and F. C. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. 2001.

[22] J. Lahti, K. Räikkönen, J. Ekelund, L. Peltonen, O. T. Raitakari, and L. Keltikangas-Järvinen. Socio-demographic characteristics moderate the association between drd4 and novelty seeking. *Personality and Individual Differences*, 40(3):533–543, 2006.

[23] M. Leyton, I. Boileau, C. Benkelfat, M. Diksic, G. Baker, and A. Dagher. Amphetamine-induced increases in extracellular dopamine, drug wanting, and novelty seeking: a pet/[11c] raclopride study in healthy men. *Neuropsychopharmacology*, 27(6):1027–1035, 2002.

[24] H. Mamitsuka. Predicting peptides that bind to mhc molecules using supervised learning of hidden markov models. *Proteins Structure Function and Genetics*, 33(4):460–474, 1998.

[25] G. S. Mann and A. McCallum. Generalized expectation criteria for semi-supervised learning of conditional random fields. 2008.

[26] S. M. McNee, J. Riedl, and J. A. Konstan. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems*, pages 1097–1101. ACM, 2006.

[27] M. Miele and J. Murdoch. The practical aesthetics of traditional cuisines: slow food in tuscany. *Sociologia Ruralis*, 42(4):312–328, 2002.

[28] J. Oh, S. Park, H. Yu, M. Song, and S.-T. Park. Novel recommendation based on personal popularity tendency. In *2013 IEEE 13th International Conference on Data Mining*, pages 507–516. IEEE, 2011.

[29] A. Pizam, G.-H. Jeong, A. Reichel, H. van Boemmel, J. M. Lusson, L. Steynberg, S. Volo, C. Kroesbacher, J. Kucerova, N. Montmany, et al. The relationship between risk-taking, sensation-seeking, and the tourist behavior of young adults: A cross-cultural study. *Journal of Travel Research*, 42(3):251–260, 2004.

[30] G. V. Rebec, J. R. Christensen, C. Guerra, and M. T. Bardo. Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-choice novelty. *Brain research*, 776(1):61–67, 1997.

[31] B. Roark, M. Saraclar, M. Collins, and M. Johnson. Discriminative language modeling with conditional random fields and the perceptron algorithm. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, page 47. Association for Computational Linguistics, 2004.

[32] T. Scheffler, R. Schirru, and P. Lehmann. Matching points of interest from different social networking sites. In *KI 2012: Advances in Artificial Intelligence*, pages 245–248. Springer, 2012.

[33] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[34] C. Song, T. Koren, P. Wang, and A.-L. Barabási. Modelling the scaling properties of human mobility. *Nature Physics*, 6(10):818–823, 2010.

[35] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.

[36] L. Song, D. Kotz, R. Jain, and X. He. Evaluating location predictors with extensive wi-fi mobility data. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 1414–1424. IEEE, 2004.

[37] P. A. Weaver, K. W. McCleary, L. Lapisto, and L. T. Damonte. The relationship of destination selection attributes to psychological, behavioral and demographic variables. *Journal of Hospitality & Leisure Marketing*, 2(2):93–109, 1994.

[38] M. Ye, P. Yin, and W.-C. Lee. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 458–461. ACM, 2010.

[39] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, and Y. Huang. T-drive: driving directions based on taxi trajectories. In *Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems*, pages 99–108. ACM, 2010.

[40] N. J. Yuan, Y. Wang, F. Zhang, X. Xie, and G. Sun. Reconstructing individual mobility from smart card transactions: A space alignment approach. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pages 877–886. IEEE, 2013.

[41] N. J. Yuan, F. Zhang, D. Lian, K. Zheng, S. Yu, and X. Xie. We know how you live: exploring the spectrum of urban lifestyles. In *Proceedings of the first ACM conference on Online social networks*, pages 3–14. ACM, 2013.

[42] F. Zhang, D. Wilkie, Y. Zheng, and X. Xie. Sensing the pulse of urban refueling behavior. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 13–22. ACM, 2013.

[43] F. Zhang, N. J. Yuan, D. Lian, and X. Xie. Mining novelty-seeking trait across heterogeneous domains. In *Proceedings of the 23rd international conference on World wide web*, pages 373–384. International World Wide Web Conferences Steering Committee, 2014.

[44] M. Zhang and N. Hurley. Novel item recommendation by user profile partitioning. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 01*, pages 508–515. IEEE Computer Society, 2009.

[45] T. Zhang. Solving large scale linear prediction problems using stochastic gradient descent algorithms. In *Proceedings of the twenty-first international conference on Machine learning*, page 116. ACM, 2004.

[46] Y. Zhang, J. Callan, and T. Minka. Novelty and redundancy detection in adaptive filtering. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 81–88. ACM, 2002.

[47] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. In *Proceedings of the 19th international conference on World wide web*, pages 1029–1038. ACM, 2010.

[48] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J. R. Wakeling, and Y.-C. Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, 107(10):4511–4515, 2010.