BR²: A Travel Behavioral Approach to Personalized Route Recommendation Based on GPS Trajectories

Rodrigo Augusto de Oliveira e Silva
University of Calgary
Calgary, Alberta, Canada
radeoliv@ucalgary.ca

Ge Cui
University of Calgary
Calgary, Alberta, Canada
cuig@ucalgary.ca

Seyyed Mohammadreza Rahimi
University of Calgary
Calgary, Alberta, Canada
smrahimi@ucalgary.ca

Xin Wang
University of Calgary
Calgary, Alberta, Canada
xcwang@ucalgary.ca

ABSTRACT

Various unknown circumstances may affect users' travel behaviors between two locations on the road network, hence it is complicated to provide satisfactory personalized route recommendations. In this paper, we believe that users' travel behaviors are reflected and can be learned from their historical GPS trajectories. The Behavior-based Route Recommendation (BR²) method is proposed to compute personalized routes based exclusively on users' travel preferences. The concepts of appearance and transition behaviors are used to describe users' travel behaviors. The behaviors are extracted from users' past travels and the missing behaviors, of unvisited locations, are estimated with the Optimized Random Walk with Restart technique. Then, the temporal dependency of travel behaviors is considered by constructing a time difference interval histogram. Last, a behavior graph is generated to allow the maximum probability route computation with the shortest path algorithm, resulting in the most likely route to be taken by a user. Experiments conducted on two real GPS trajectory data sets demonstrate the efficiency and effectiveness of the proposed method.

1 INTRODUCTION

With advances in Global Positioning System (GPS) technology and the popularity of mobile devices, massive amounts of human movement data in GPS trajectories have been collected and are available for research, which provides an alternative way to explore travel route recommendation. Route recommendation systems usually suggest routes based on the optimization of cost functions of either distance or travelling time. However, it has been observed that in the real world the shortest or quickest routes are often not taken [5]. A personalized route recommendation system, on the other hand, provides route recommendations based on users' travel preferences [8]. For instance, some drivers want to reach the destination as fast as possible, while others are willing to take a slightly longer route that does not go through highways or busy roads. The identification of all factors which may affect users' travel behaviors is still very challenging and an attempt to acquire as much as these aspects as possible may prove to be infeasible, as distinct factors affect each person particularly [21].

The proper representation of GPS readings allows the analysis of the relationship between users and their behaviors. Users' travel behaviors reflected in historical GPS trajectories are the implicit feedback of their preferences, as they do not explicitly rate preferred routes. Since it is laborious to use surveys and questionnaires to acquire the explicit feedback about traveling preferences from individual users, users' travel behaviors are captured based on implicit feedback obtained from their past travels.

The consideration of users' preferences in route recommendation can be problematic when suggesting routes to untraveled locations. To alleviate this issue, some studies have been conducted in personalized route recommendation based on collaborative filtering (CF) methods, e.g. item-based CF and matrix factorization (MF), to acquire more information through users that share similar preferences [3, 4, 8]. The use of CF has been proven to provide more accurate recommendations compared to the shortest distance route method, but they usually require quadratic space and considerable time to process [3, 22]. In this context, it is a complex and computational expensive task due to the huge amount of data needed to describe user travel behaviors and the overall sparseness of known data. The rapid and accurate estimation of missing users' travel behaviors is crucial to support personalized route recommendation systems.

Time is another essential factor for the effective comprehension of users' travel behaviors, as users' behaviors tend to be similar during specific intervals of the day. For instance, a user that has a behavior of taking a route between an origin and destination at noon might have a different behavior at midnight, but the same behavior is expected to be shown at times close to noon. To accurately represent temporal dependency and to effectively incorporate it in the recommendation process are challenges that need to be addressed.

In this study, we propose a personalized Behavior-based Route Recommendation method, named BR², to extract users' travel behaviors and compute the route with maximum probability, which refers to the route that users are most likely to take based on their previous travels. In BR², the implicit information of users' travel behaviors is extracted from their historical movements. Compared with popular CF methods, the random walk with restart (RWR) has been proven to provide better estimations when considering users' implicit feedback [18]. Therefore, we utilize a RWR approach to estimate users' missing behaviors and investigate optimizations to better train the model. In addition, we make use of temporal dependency within users' travel behaviors to enhance the accuracy of the recommendation. In summary, the contributions of this study are:

- The missing behavior frequencies for each user on specific road segments are estimated by using the RWR technique, as it yields more accurate estimations due to implicit data nature. We utilize the Optimized Random Walk with Restart (ORWR) to improve the efficiency of the behavior estimation.
- The data sparsity problem is mitigated with the consideration of temporal dependency. A time difference histogram is computed based on time interval differences from GPS historical data, which weights users' preferences at similar time accordingly and provide more data for the recommendation.
- Experiments are conducted on two real GPS trajectory data sets in China, Geolife from Beijing and taxi GPS trajectories from Shenzhen. The results indicate that the proposed approach BR^2 outperforms the baseline methods.

In the remainder of this paper, Section 2 outlines the related works regarding CF techniques and route recommendation. Section 3 presents the personalized route recommendation methodology and explains the proposed method step by step. Section 4 discusses the experiment results. Finally, Section 5 exposes the conclusion and further work of this research.

2 RELATED WORK

2.1 Collaborative Filtering Techniques

Recommendation systems are popular for providing predictions that interest people in various scenarios. One of the approaches of recommendation systems is given by CF, which considers information of preference from multiple users to predict missing ratings of users on items, and some of its popular methods are the item-based CF, MF, and RWR.

The item-based CF is one of the most simplistic methods which considers the similarities between items to calculate values for others without a user rating [25]. Since every item needs to be compared to compute the similarity score, in the worst-case scenario, the total number of evaluations is equal to the combination of two for all the elements in the data set. The algorithm does not scale well for large data sets, involves intensive processing, and requires $O(n^2)$ space to store the similarities between n items [20].

Among the latent factor models, MF is one of the most used methods for recommendation problems [10]. The objective of this technique is to represent rating data into two vectors of latent factors, in which the dimensionality may vary based on the data itself, so that the dot product of the two vectors result in approximate values of known ratings. The values of the latent factor vectors can be learned from the known data by the minimization of regularized squared error with the use of the stochastic gradient descent algorithm.

Random walk is a process that describes the probabilities of series of random movements in a dimension space. In recommender systems, one of the techniques that considers personal preferences is the RWR. While the usual approach of random walk traverses a graph only based on its structure, leading to an exclusive convergence, the RWR makes uses of a probability of returning to the original node on each movement. Therefore, the technique allows personalization by having lower ranking values given to nodes farther from the origin node [14]. RWR is graph-based, having users and items as nodes and ratings as edges connecting the user-item pair. The missing ratings are estimated by traversing the graph according to the weights and the restart probability until convergence.

2.2 General Route Recommendation

General route recommendation aims to provide a route between two points in a road network, an origin and a destination, based on a given cost function. Researches have been conducted to study human movement and provide route recommendations based on historical GPS trajectories.

Chen et al. [2] study the most popular route based on users' traveling behaviors. A popularity indicator is used to discover the frequent routes in a network in order to assist users when they travel to an unfamiliar area. Wei et al. [24] propose an algorithm to construct popular routes from historical trajectories in regions of the road network.

These studies contribute to the area of route planning and travel recommendation, but neither of them is truly personalized, as they do not study users' personal route preferences on the road network.

2.3 Personalized Route Recommendation

Differently from general route recommendation, personalized route recommendation considers users' preferences in the definition or calculation of a cost function. While some studies explore users' preferences through explicit feedback, collecting information by directly querying users, others assume that preference factors cannot be modelled entirely and that users might not be fully aware of their own preferences. Our study is focused on the latter assumption.

Several researches explore personalized route recommendation through the optimization of the cost function based on multiple criteria either weighted directly by users or defined by their driving preferences information [1, 5, 7, 11, 16, 26]. Users' preferences are explicitly collected, modeled, and/or used in the calculation of candidate routes, which might not reflect their true behaviors and are invariant to time.

Another approach is focused on providing personalized tourism route recommendation based on social media data. Studies have used users' data to model route attributes [9], build data set of popular locations [6], and even mine their preferences and temporal information [31]. Nevertheless, these approaches focus on movement between points of interest or visiting locations, disregarding the evaluation of users' preferences directly on the road network.

Personalized route recommendation based on historical GPS trajectories is another branch of research. McGinty and Smyth [15] proposed a personalized route planning method that considers historical trajectories to derive implicit driving preferences without defining a preference model. The method combines and reuses routes sections from previous travels to generate new routes, but it fails to recommend routes to unfamiliar areas of the road network. As an extension to their previous work, McGinty and Smyth [8] used a type of distributed case-based reasoning strategy, in which historical trajectories of similar drivers are borrowed to recommend routes in unfamiliar map territories. However, considering the behaviors of similar drivers by directly using their trajectories in the recommendation might not precisely reflect the preferences of a driver.

Letchner et al. [12] introduced a method of personalized route planning by considering users' historical trajectories, extracted from GPS readings, in the calculation of an inefficiency ratio that represents the proportion of time extended in a trip compared to the shortest possible time. Liu et al. [13] explored personalized route recommendation for self-drive tourists not only considering the drivers' visiting preferences, but also real-time traffic information. In these approaches, the authors define a metric to represent users' implicit preferences but include other factors like distance and travel time as part of the objective function.

Cui et al. [3] extracted users' travel behaviors from their historical GPS trajectories to represent their preferences. By predicting missing travel behaviors with CF technique, a route with maximum probability is computed for specific users. In addition, Cui and Wang [4] proposed a different representation of travel behavior and improved the performance of route recommendation. However, the methodologies disregard the implicit data nature and temporal dependency is not fully explored.

A different approach is proposed by Wang et al. [23], in which neural networks are used to learn the optimal cost functions of the A* algorithm. The presented results show considerable improvement in precision, recall, and F1-score, but the training time might impede its usage in real-world applications with new data constantly being fed into the system.

3 BEHAVIOR-BASED ROUTE RECOMMENDATION

This section discusses the proposed method of the Behavior-based Route Recommendation (BR²). The following preliminaries are first defined for this research.

3.1 Preliminaries

The preliminaries of this study are defined as follows.

Definition 1 - Road network. The road network is a graph $G = (V, E)$ composed by a set of vertices V and edges E. A vertex $v \in V$ represents the boundary of road segments and an edge $e \in E$ represents a road segment, containing starting and ending vertices, denoted as $e.start$ and $e.end$, respectively, where $e.start \in V$ and $e.end \in V$.

Definition 2 - GPS reading. A GPS-reading is a 3-tuple $p = (t, lat, lng)$ in which t represents a timestamp, and lat and lng are the latitude and longitude of the location of the GPS-reading.

Definition 3 - GPS trajectory. A GPS trajectory $trj = (p_1, p_2, p_3, \cdots, p_k)$ consists of a sequence of GPS-readings, such that $p_i.t - p_{(i-1)}.t > 0$ and $1 < i \leq k$.

Definition 4 - Route. Given a road network $G = (V, E)$, a route R starting from vertex v_i and ending at vertex v_j is a sequence of connected road segments $R = (v_i, e_1, e_2, e_3, \cdots, e_l, v_j)$, where $v_i, v_j \in V$ and $e_i \in E$; e_i is the i-th road segment in R, $e_i \neq e_j$ if $i \neq j$, $e_1.start = v_i$, and $e_l.end = v_j$.

Two types of behaviors are proposed to extract the user's movement behaviors on the road network. First, to provide a global overview of how frequently users are in specific locations at a certain time on the road network, the concept of appearance behavior is defined to represent the relation between location and time of users' movement. The second, transition behavior, captures the sequential relation of appearance behaviors, not only giving a sense of location regularity but most importantly of direction. Both behaviors are formally described as follows.

Definition 5 - Appearance behavior. An appearance behavior is defined as tuple of the road segment and the time, denoted as $b = (e, t)$, where e is an edge of the road network and t is a time interval of a day, and it describes the location and time of a user's movements. For a given user u and an associated appearance behavior b, the frequency that user u has behavior b is represented by $frq(u, b)$.

Definition 6 - Transition behavior. A transition behavior t_b represents the relationship of two sequential appearance behaviors $b_i = (e_i, t)$ and $b_j = (e_j, t)$ at the same time interval t, such that $e_i.end = e_j.start$. It is denoted as an ordered tuple $t_b = (b_i \rightarrow b_j)$ or $tb_{(i \rightarrow j)}$ in short. Similarly, the frequency that user u has transition behavior tb is given by $frq(u, tb)$.

Both appearance and transition behaviors contribute to capture the implicit travel preferences of each user. This implicit travel preference information is essential in solving the problem of the personalized route recommendation, as the recommendations need to truly reflect users' priorities.

3.2 Framework Overview

The framework of the proposed BR² is illustrated in Figure 1. The first step of BR² is data preprocessing. The GPS trajectories are split into trips with defined origins and destinations, the trips are matched to the road network by applying the map matching technique, and the outliers are removed from the trajectories. As a final step of the preprocessing component, the appearance and transition behaviors are generated from the historical users' routes. Since users, in general, travel on few routes daily, covering a limited number of road segments in a study area, the missing travel behaviors for each user need to be estimated. In the second step, the RWR technique is used to estimate users' appearance and transition behaviors on each untraveled road segment. With the missing behaviors frequencies estimated, the temporal dependency is evaluated in the data set by building a time interval difference histogram. The histogram indicates how the data is distributed and suggests the number of intervals that should be considered in the route recommendation process. Then, the probabilities are calculated from the travel behaviors for a defined origin and destination considering the Markov property. In addition, the Laplace smoothing method is applied to estimate the probability of users' missing travel behaviors for the road segments that have never been visited previously by any user.

Finally, the last stage is the recommendation of the route with maximum travel behavior probability. To facilitate the route computation, a behavior graph is constructed to represent the travel behaviors and the relationship among them. Dijkstra's algorithm is used in the behavior graph to find the route maximum travel behavior probability.

3.3 Data Preprocessing

The preprocessing can be divided into three parts. First, since a GPS trajectory usually tracks users for a long period of time and contains multiple trips of the users, trajectory segmentation is first applied to divide the raw trajectory into several sub-trajectories [3]. After the trajectory segmentation, each GPS trajectory corresponds to a single travel route. Second, since the trajectory is usually noisy due to the urban canyon or measurement errors, outliers in GPS trajectories are detected and removed. In this study, if the distance between a GPS point and its nearest road segment exceeds a threshold of 180 seconds, the GPS point is considered as an outlier. Besides, given the maximum moving speed, if the distance that a GPS point moves exceeds a threshold within the time interval from its previous GPS point, it is also taken as outlier. Lastly, GPS trajectories are mapped onto

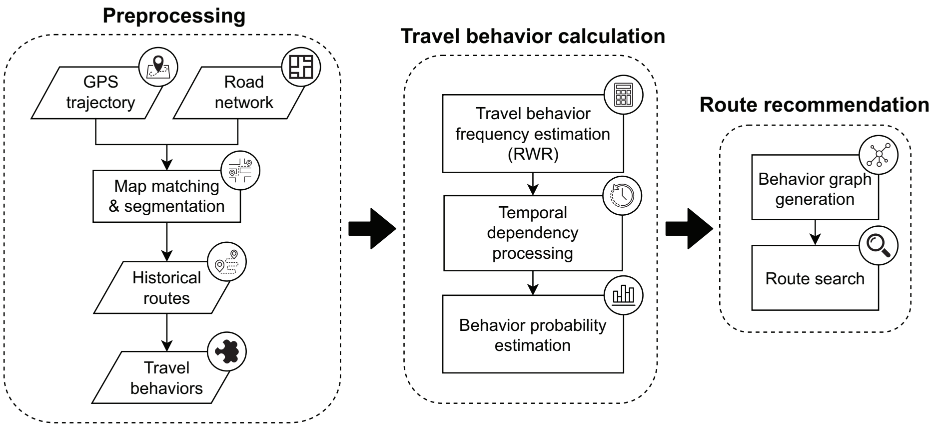


Figure 1: Behavior-based Route Recommendation framework overview.

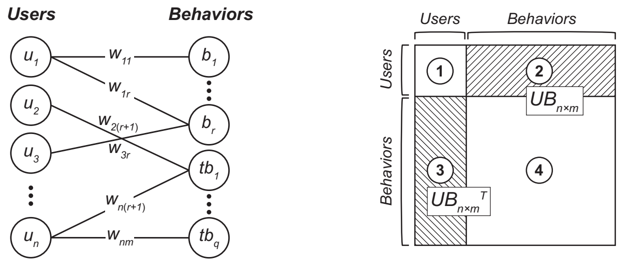
the road network to obtain users' historical routes with the map matching method proposed in [17].

3.4 Behavior Frequency Estimation

Each user in general covers a very limited number of road segments in a city, so the missing travel behaviors for each user need to be estimated. In this section, we discuss how the RWR is applied in the estimation and provide the motivation for the use of ORWR.

3.4.1 Random Walk with Restart for Behavior Estimation. In this study, the appearance and transition behaviors are first extracted from the historical users' trajectories and a user-behavior matrix $UB_{n\times m}$ is built. In user-behavior matrix, rows represent users $U = (u_1, u_2, \cdots, u_n)$, columns represent both appearance and transition behaviors $B = (b_1, b_2, \cdots, b_r, tb_1, tb_2, \cdots, tb_q)$, where $m = r + q$, and the element $UB_{i,j}$ of user i and behavior j represents the frequency of the pair (u_i, b_j) or (u_i, tb_j). The user-behavior matrix can also be represented as a user-behavior graph, in which the nodes (users and behaviors) are connected through edges with weights as the corresponding frequency of the user-behavior pair, as illustrated in Figure 2a.

The RWR technique tackles the problem based on a graph and can approach it with an adjacency matrix representation. Considering the structure of the user-behavior graph shown in Figure 2a, the idea of RWR is to traverse the graph by either moving to a neighbor node or going back to an initial node based on a given restart probability value. Therefore, a behavior node is highly related to a user when it is visited multiple times. A behavior node associated with a higher score value represents the higher probability of being visited from a user node when the graph is traversed. The score values are represented as the weights of the edges in the graph. Overall, the RWR estimates the probability values of all edges in the user-behavior graph by incrementally updating the user-behavior probability values based on past behaviors of the user and the behaviors of the similar behaving user. Since users and behaviors are interpreted as nodes in an undirected graph, the adjacency matrix of the graph is generated with the combination of both users and behaviors, resulting in a large and sparse symmetric matrix, illustrated in Figure 2b. The adjacency matrix has $(n + m)$ rows and columns, composed by four distinct parts: part 1 is a $n \times n$ matrix (users by users) composed by similarities between users, part 2 is the user-behavior matrix, part 3 is the transposed user-behavior matrix, and part 4



(a) Bipartite graph **(b) Adjacency matrix**

Figure 2: User-behavior graph representations.

is a $m \times m$ matrix (behaviors by behaviors) composed similarities between behaviors. This structure is commonly unoptimized due to the exceedingly higher number of behaviors compared to the number of users in a data set.

The adjacency matrix and restart probability c are provided as inputs to the RWR algorithm. The score matrix r, is calculated by:

$$r^{(t+1)} = (1 - c) \begin{bmatrix} 0 & \overline{UB}_{n\times m} \\ \overline{UB}_{n\times m}^T & 0 \end{bmatrix} r^{(t)} + c\, I_{(n+m)\times(n+m)}$$

where $\overline{UB}_{n\times m}$ represents the row-normalized user-behavior matrix, $\overline{UB}_{nxm}^T$ represents the transposed row-normalized user-behavior matrix, $r^{(t)}$ is the score matrix of t-th iteration – initially set as an identity matrix – and $I_{(n+m)\times(n+m)}$ is the identity matrix with $(n + m)$ rows and columns. The score matrix is calculated iteratively until convergence, in which the difference between the new and previous scores is smaller than a defined threshold ε.

$$|r^{(t+1)} - r^{(t)}| < \varepsilon$$

After convergence, the RWR algorithm returns the score matrix, which contains the normalized probability values associated with the previously unknown behavior frequencies.

Considering the total number of users n and behaviors m, the adjacency matrix representation contains $(n + m)^2$ elements. Consequently, the time complexity of the algorithm is defined by the total number of iterations for convergence k and the matrices multiplication in every iteration, resulting in a time complexity of $O(k(n + m)^3)$.

3.4.2 Optimized Random Walk with Restart for Behavior Estimation. The common RWR algorithm approach makes use of an adjacency matrix as input, processed for as many iterations as needed until the process converges. However, to fit the data into an adjacency matrix representation, both sets of users and behaviors need to be combined in rows and columns to form a symmetric matrix, as shown in Figure 2b.

This representation could be useful if similarity measures between users and between behaviors are available, describing the inner relations between themselves, as the data could be used to fill parts 1 and 4 of the adjacency matrix presented in Figure 2b. Nevertheless, since we do not explore the similarities between users and between behaviors, the matrix representation is unquestionably unoptimized and leads to unnecessary processing overhead, as the total number of elements in the matrix is exceedingly higher than the actual data in the user-behavior matrix. For instance, for n users and m behaviors the total number of elements in the adjacency matrix is $(n+m)^2$ while the actual data consists of the part 2 in Figure 2b, containing $(n * m)$ elements.

To better handle the estimation of missing behavior frequencies we use the ORWR algorithm, which considers the user-behavior bipartite graph represented by the user-behavior matrix only [19]. Using the user-behavior matrix instead of the entire adjacency matrix has a strong impact in performance, resulting in a time complexity reduction from $O(k(n+m)^3)$ to $O(knm^2)$.

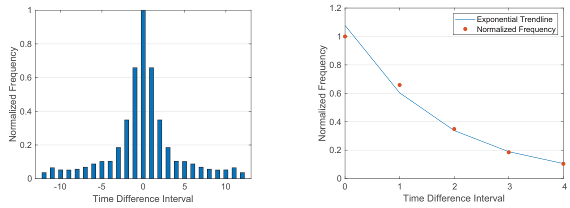
3.5 Temporal Dependency

Time is one of the important factors that influence users' actions throughout the day, as people are more prone to go to different places during specific times [27]. In addition, it is possible to identify that people present similar travel behaviors at closer time intervals. For instance, many people go to work from home by roughly the same route at similar times in the morning and it is not expected for them to show this movement behavior during the evening, many hours apart from their common travel behavior. Therefore, the temporal information should also be considered when providing personalized route recommendations.

An appropriate strategy to predict user's behaviors by considering temporal information is to study the relation between behaviors at different time intervals. Behaviors associated to a given road segment are most likely to be shown on closer time intervals. Temporal dependency is implemented in this study by considering existing behaviors of the same road segment at similar time intervals in the calculation of travel behaviors' probabilities.

To identify how behaviors at different time intervals impact the route recommendation process, a time difference histogram is generated by comparing behaviors of all users related to the same road segments in pairs. For example, if there are 24 time intervals in total, each time interval representing each hour of the day, the time difference histogram is divided into values ranging from -12 to 12, with increments of one, and each interval consists of the frequency that behaviors related to the same road segment happened at a specific time difference.

A pair of appearance behaviors b_i and b_j refer to the same road segment if $b_i.e = b_j.e$ and the correspondent absolute time difference interval is defined as $|k| = |b_i.t - b_j.t|$. Since transition behaviors consist of sequential appearance behaviors, the rationale is the same. Given a total number of appearance and transition behaviors r and q respectively, the frequency for each time difference interval value k is computed according to the



(a) (b)

Figure 3: (a) Normalized time difference histogram. (b) Five first intervals of time difference histogram approximated by an exponential function ($a = 1.0778, b = -0.58$).

following equation:

$$freq(k) = \begin{cases} \sum_{i=1}^{r} \binom{frq(b_i)}{2} + \sum_{i=1}^{q} \binom{frq(tb_i)}{2} & |k| = 0 \\ \sum_{i=1}^{r} \sum_{j=1}^{r} frq(b_i)\, frq(b_j) \ + & i \neq j \text{ and} \\ \sum_{i=1}^{q} \sum_{j=1}^{q} frq(tb_i)\, frq(tb_j) & |k| \neq 0 \end{cases}$$

Since we evaluate all behaviors against themselves, only the absolute time difference intervals are calculated and equally divided for their correspondent interval. For instance, frequencies of time difference intervals of absolute value of one hour are calculated and divided in half for the intervals of -1 and 1. Finally, the obtained frequencies are normalized in the range from zero to one. For example, Figure 3a illustrates a histogram obtained from a real-world data set Geolife by the computation of time difference of behaviors of all driving travelers.

As expected, most of the frequencies are concentrated in the first few intervals close to the zero interval, showing that most behaviors happen at similar time intervals. As an attempt to avoid overfitting and to reduce the noise from the data, a function can be used to represent the values with most importance, closer to the zero interval. In this study, an exponential function $f(x) = ae^{bx}$ is used to represent the data while conveying the notion of decay as the time interval is farther from the evaluated behavior. Figure 3b presents how closely an exponential function represents the first five intervals of the distribution.

The values from the exponential function for each time interval allow weighting the behaviors frequencies from time intervals close to the recommendation time interval. In Figure 3b, for instance, the weighting values $w1$, $w2$, $w3$, and $w4$, regarding the time difference intervals of 1 to 4, are approximately 0.6, 0.3, 0.2, and 0.1, respectively. The frequencies of similar time intervals are weighted by multiplying them with the corresponding value of the exponential function for a time difference interval. The weighted frequencies are added to the frequency of the behavior at the time interval used for recommendation. For example, for two appearance behaviors b_i and b_j of the same road segment and user u at time intervals t and $t+1$ respectively, if the recommendation is provided at time interval t and the correspondent value of the exponential function for a time difference interval of 1 is $w1$, the total frequency of behavior b_i is $frq(u, b_i) = frq(u, b_i) + w_1\, frq(u, b_j)$.

3.6 Probability Calculation for Travel Behavior

Based on behavior frequencies, probabilities are computed for a user u given a route R at a specific time t. The route with maximum probability reflects the route that the user is most inclined to take, and it is preferred above all others. The route probability is defined as:

$$P(R|u,t) = P(e_1, e_2, e_3, \cdots, e_l|u,t) = \frac{P(e_1, e_2, e_3, \cdots, e_l, t|u)}{P(t|u)}$$

where $e_1, e_2, e_3, \cdots, e_l$ represent the road network edges, and e_1 and e_l are the incident edges with the origin and destination vertices, respectively. If the user u and time t are known, $P(t|u)$ is constant, thus simplifying the problem to the maximization of the numerator.

With the assumption that the series of behavior probabilities are described as Markov property, the probability of a user behavior depends on the immediate previous behavior, if not related to the origin, simplifying the representation of the problem. Extending the previous equation, the route probability is given by:

$$P(b_1, b_2, b_3, \cdots, b_l|u) = P(b_1|u)P(tb_{1\to 2}|u)P(tb_{2\to 3}|u)\cdots P(tb_{l-1\to l}|u)$$

where $P(b_1|u)$ is the appearance behavior probability of origin and $P(tb_{i-1\to i}|u)$ is the transition behavior probability, from appearance behavior b_{i-1} to appearance behavior b_i.

Since traditional pathfinding algorithms search for the path with the minimal weight, the probability function is transformed in order to shift the problem objective from maximization to minimization. It is achieved by using the logarithm of inversed probabilities as follows:

$$L = \frac{1}{P(b_1|u)P(tb_{1\to 2}|u)P(tb_{2\to 3}|u)\cdots P(tb_{l-1\to l}|u)}$$

$$\ln L = \ln \frac{1}{P(b_1|u)P(tb_{1\to 2}|u)P(tb_{2\to 3}|u)\cdots P(tb_{l-1\to l}|u)}$$

$$\ln L = \ln \frac{1}{P(b_1|u)} + \sum_{i=1}^{l-1} \ln \frac{1}{P(tb_{i\to i+1}|u)}$$

If there is no historical data of users' travels through some road segments, the related behaviors will not exist. To prevent the designation of zero to the probabilities of nonexistent behaviors, the Laplace smoothing method is employed for both appearance and transition behaviors by considering the following:

$$P(b|u) = \begin{cases} \frac{\widehat{frq}(u,b)+\alpha}{\sum_{b_i\in S_o}\widehat{frq}(u,b_i)+\alpha d} & \widehat{frq}(u,b) > 0 \\ \frac{\alpha}{\sum_{b_i\in S_o}\widehat{frq}(u,b_i)+\alpha d} & \text{otherwise} \end{cases}$$

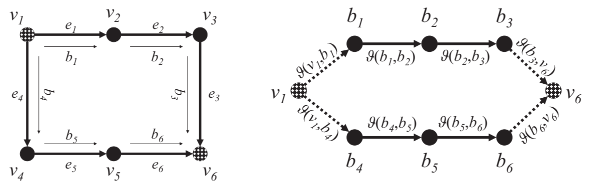
$$P(tb_{i\to j}|u) = \begin{cases} \frac{\widehat{frq}(u,tb_{i\to j})+\alpha}{\sum_{k=1}^{d}\widehat{frq}(u,tb_{i\to k})+\alpha d} & \widehat{frq}(u,tb_{i\to j}) > 0 \\ \frac{\alpha}{\sum_{k=1}^{d}\widehat{frq}(u,tb_{i\to k})+\alpha d} & \text{otherwise} \end{cases}$$

where $\widehat{frq}$ represents the estimated behavior frequency, $P(b|u)$ is the probability of appearance behavior b of user u at time t, S_o is the set of appearance behaviors starting at origin o, $P(tb_{i\to j}|u)$ is the probability of transition behavior $tb_{i\to j}$, α is the smoothing parameter, and d is number of appearance behaviors – the behaviors related to the destination road segment, in the case of transition behavior.

3.7 Route Search Based on Probability

The structure of the road map network graph does not allow the representation of travel behavior probabilities as weights in edges. Therefore, a different structure is constructed to enable graph traverse considering probabilities. It is denominated behavior graph and its structure is illustrated in Figure 4. The behavior graph represents each vertex as an appearance behavior and each edge as a transition behavior. The weight ϑ defines the correspondent behavior probability. In addition, the origin and destination vertices of a given trajectory are included in the structure and their edges represent appearance behaviors, illustrated with a different notation from other edges and vertices in the graph.

Despite adding complexity in the process by generating a new structure to represent probabilities, the personalized recommended route can be computed as the least log-inverse probability path, which can be solved by any traditional shortest path algorithm.



(a) Road network (b) Behavior graph

Figure 4: Comparison of road network and behavior graph, given a trajectory from origin v_1 to destination v_6.

Table 1: Summary of training data used in experiments.

Features	Data set	
	Geolife	Shenzhen
Users	22	274
Behaviors	26,123	677,068
Time interval	1 hour	1 hour
Elements in user-behavior matrix	574,706	185,516,632
Sparseness in user-behavior matrix	95.45%	99.63%
Segments in road network	15,493	29,411
Vertices in road network	38,485	62,113

4 EXPERIMENTAL RESULTS

In this section, the performance of the proposed BR² is evaluated through the experiments on two real data sets. The first data set, Geolife [28–30], contains trajectories extracted from drivers in the central district of Beijing and the second data set consists of taxi drivers' trajectories from Shenzhen. The training data has 80% of the trajectories while 20% was used for testing purposes. Some details of the data sets are presented in Table 1.

Four accuracy measures are considered to evaluate the performance of the model: precision and recall based on the number of road segments and based on the distance of road segments. The measures definitions are presented as follows.

$$Precision_{segments} = \frac{\#\ of\ correct\ road\ segments}{\#\ of\ road\ segments\ on\ recommended\ route}$$

$$Recall_{segments} = \frac{\#\ of\ correct\ road\ segments}{\#\ of\ road\ segments\ on\ true\ route}$$

$$Precision_{distance} = \frac{distance\ of\ correct\ road\ segments}{distance\ of\ recommended\ route}$$

Table 2: Recommendation performance of baseline methods and BR² with and without temporal dependency (BR² - TD).

	Precision		Recall	
	Segments	Distance	Segments	Distance
		Geolife		
SD	0.246	0.266	0.215	0.239
MaP2R	0.533	0.537	0.543	0.569
BR² - TD	0.576	0.573	0.562	0.587
BR²	**0.631**	**0.637**	**0.613**	**0.642**
		Shenzhen		
SD	0.289	0.311	0.247	0.267
MaP2R	0.555	0.562	0.520	0.562
BR² - TD	0.628	0.632	0.580	0.613
BR²	**0.688**	**0.692**	**0.626**	**0.654**

$$Recall_{distance} = \frac{distance\ of\ correct\ road\ segments}{distance\ of\ true\ route}$$

The different aspects of the route recommendation performance are assessed by the conduction of experiments. First, we compare the performance of BR² with other baseline methods. Second, we evaluate the route recommendation performance based on behavior estimation, considering the ORWR in the proposed BR² method and the previously discussed CF techniques. Last, the influence of temporal dependency on the performance of route recommendation is assessed.

4.1 Overall Recommendation Performance

In this experiment, the recommendation performance of BR² is compared against MaP2R [4] and the shortest distance (SD) route method. With respect to the parameters associated with MaP2R, the number of latent factors in MF was set as 30 and the Laplace smoothing parameter as 1.0^{-5} for Geolife data set and 1.0^{-7} for Shenzhen. The BR² used the same values as MaP2R for the Laplace smoothing parameters, the ORWR considered a convergence value of 5.0^{-10} and restart probability of 0.5 for Geolife and 0.1 for Shenzhen data sets, and three time difference intervals were used for temporal dependency. The obtained precision and recall values of the three recommendation methods are presented in Table 2.

The obtained results show that BR² outperforms the other methods in all accuracy measurements. For Geolife data set, BR² performs on average 38.9% better than the shortest distance route method and 8.5% better than MaP2R. Similarly, for Shenzhen data set, BR² shows an average enhancement of 38.6% and 11.5%. The better performance of BR² is due to a more effective estimation of users' travel behaviors using RWR and to the consideration of temporal dependency between travel behaviors.

4.2 Performance Based on Behavior Estimation

The effective estimation of users' travel behaviors is critical for personalized route recommendation. In this experiment, we compare the route recommendation performance of ORWR against two popular CF methods, i.e. the item-based CF (represented as IBF) and MF, in the estimation of the frequencies of missing behaviors. We do not consider the temporal dependency in the experiment.

In this analysis, the number of latent factors in matrix factorization was set as 30, the Laplace smoothing parameter as 1.0^{-5} for Geolife data set and 1.0^{-7} for Shenzhen, the convergence value in ORWR as 5.0^{-10}, and the restart probability as 0.5 for Geolife and 0.1 for Shenzhen. The results are shown in Table 3. Due to the large size of the Shenzhen data set, the experiment with the item-based CF could not be conducted, as a result of scalability issue [20].

Table 3: Recommendation performance with different CF methods.

	Precision		Recall	
	Segments	Distance	Segments	Distance
		Geolife		
IBF	0.507	0.517	0.522	0.549
MF	0.533	0.537	0.543	0.569
ORWR	**0.576**	**0.573**	**0.562**	**0.587**
		Shenzhen		
IBF	-	-	-	-
MF	0.555	0.562	0.520	0.562
ORWR	**0.628**	**0.632**	**0.580**	**0.613**

As seen in the results, for both data sets, the route recommendation with the ORWR obtained the highest precision and recall values for the number of road segments and distance compared to IBF and MF due to its capability to better handle implicit data. In addition, the higher amount of data in Shenzhen justifies a larger gap in performance between the ORWR and the other methods.

4.3 Temporal Dependency Impact

This experiment evaluates the influence of the temporal dependency (TD) in the accuracy of recommendations. The exponential functions, $y = 1.067e^{-0.57|x|}$ for Geolife data set and $y = 1.0057e^{-0.038|x|}$ for Shenzhen data set, were built to represent the generated data distribution histogram, as exemplified in Figure 3b. The weight for the behaviors with different time intervals was determined according to the function. These functions were obtained considering behaviors of the closest three time intervals, as it yielded the best accuracy gain in the recommendation. The results presented in Table 2 show the difference in precision and recall of BR² when temporal dependency was not considered.

As seen in the results, the overall accuracy of the model considerably decreased when temporal dependency was not considered, negatively impacting the accuracy, on average, by 5.6% and 5.1% for Geolife and Shenzhen data sets, respectively. The impact of temporal dependency depends on the behavior distribution and the amount of similar behaviors considered in the recommendation process.

5 CONCLUSION AND FUTURE WORK

In this study, we proposed the Behavior-based Route Recommendation (BR²) to provide personalized routes based on users' preferences. The methodology uses the appearance and transition behaviors to capture users' travel preferences, adopts the ORWR to estimate missing behavior frequencies, evaluates temporal dependency through a time difference interval histogram, and searches for the route with maximum probability on the behavior graph. Experiments show the effectiveness of the proposed method.

The lack of data of users with few behaviors was alleviated with temporal dependency but the problem with routes without any previous data was not focused in this research. It is believed that with the inclusion of spatial correlation and other data such as road type, number of lanes, speed limit, and traffic density can lead to a more concise and robust recommendation. Future work also includes the comparison of the optimized model with other personalized route recommendation methods that may provide a treasure trove of improvements to be incorporated to the model.

Acknowledgments. The research is supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant to Xin Wang.

REFERENCES

[1] P. Campigotto, C. Rudloff, M. Leodolter, and D. Bauer. 2017. Personalized and Situation-Aware Multimodal Route Recommendations: The FAVOUR Algorithm. *IEEE Transactions on Intelligent Transportation Systems* 18, 1 (Jan 2017), 92–102. https://doi.org/10.1109/TITS.2016.2565643

[2] Z. Chen, H. T. Shen, and X. Zhou. 2011. Discovering popular routes from trajectories. In *2011 IEEE 27th International Conference on Data Engineering*. 900–911. https://doi.org/10.1109/ICDE.2011.5767890

[3] Ge Cui, Jun Luo, and Xin Wang. 2018. Personalized travel route recommendation using collaborative filtering based on GPS trajectories. *International Journal of Digital Earth* 11, 3 (2018), 284–307. http://www.tandfonline.com/doi/abs/10.1080/17538947.2017.1326535

[4] Ge Cui and Xin Wang. 2017. MaP2R: A Personalized Maximum Probability Route Recommendation Method Using GPS Trajectories. In *Advances in Knowledge Discovery and Data Mining*, Jinho Kim, Kyuseok Shim, Longbing Cao, Jae-Gil Lee, Xuemin Lin, and Yang-Sae Moon (Eds.). Springer International Publishing, Cham, 168–180.

[5] J. Dai, B. Yang, C. Guo, and Z. Ding. 2015. Personalized route recommendation using big trajectory data. In *2015 IEEE 31st International Conference on Data Engineering*. 543–554. https://doi.org/10.1109/ICDE.2015.7113313

[6] Siyuan Du, Hua Zhang, Hualin Xu, Jirui Yang, and Oscar Tu. 2019. To make the travel healthier: a new tourism personalized route recommendation algorithm. *Journal of Ambient Intelligence and Humanized Computing* 10, 9 (01 Sep 2019), 3551–3562. https://doi.org/10.1007/s12652-018-1081-z

[7] Stefan Funke and Sabine Storandt. 2015. Personalized route planning in road networks. In *Proceedings of the 23rd SIGSPATIAL International Conference on advances in geographic information systems (SIGSPATIAL '15)*, Vol. 03-06-. ACM, 1–10.

[8] Lorraine Mc Ginty and Barry Smyth. 2001. Collaborative Case-Based Reasoning: Applications in Personalised Route Planning. In *Case-Based Reasoning Research and Development*, David W. Aha and Ian Watson (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 362–376.

[9] Gang Hu, Yi Qin, and Jie Shao. 2018. Personalized travel route recommendation from multi-source social media data. *Multimedia Tools and Applications* (18 Oct 2018). https://doi.org/10.1007/s11042-018-6776-9

[10] Y. Koren, R. Bell, and C. Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (Aug 2009), 30–37. https://doi.org/10.1109/MC.2009.263

[11] H. Kriegel, M. Renz, and M. Schubert. 2010. Route skyline queries: A multi-preference path planning approach. In *2010 IEEE 26th International Conference on Data Engineering (ICDE 2010)*. 261–272. https://doi.org/10.1109/ICDE.2010.5447845

[12] Julia Letchner, John Krumm, and Eric Horvitz. 2006. Trip Router with Individualized Preferences (TRIP): Incorporating Personalization into Route Planning. https://www.microsoft.com/en-us/research/publication/trip-router-individualized-preferences-trip-incorporating-personalization-route-planning/

[13] Long Liu, Jin Xu, Stephen Shaoyi Liao, and Huaping Chen. 2014. A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication. *Expert Systems With Applications* 41, 7 (2014), 3409–3417.

[14] B. Lv, W. Yu, L. Wang, and J.A. McCann. 2014. Efficient processing node proximity via random walk with restart. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8709, 542–549.

[15] Lorraine McGinty and Barry Smyth. 2000. Personalised Route Planning: A Case-Based Approach. In *Advances in Case-Based Reasoning*, Enrico Blanzieri and Luigi Portinale (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 431–443.

[16] S. Nadi and M.R. Delavar. 2011. Multi-criteria, personalized route planning using quantifier-guided ordered weighted averaging operators. *International Journal of Applied Earth Observation and Geoinformation* 13, 3 (2011), 322 – 335. https://doi.org/10.1016/j.jag.2011.01.003

[17] Paul Newson and John Krumm. 2009. Hidden Markov Map Matching Through Noise and Sparseness. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '09)*. ACM, New York, NY, USA, 336–343. https://doi.org/10.1145/1653771.1653818

[18] H. Park, J. Jung, and U. Kang. 2017. A comparative study of matrix factorization and random walk with restart in recommender systems. In *2017 IEEE International Conference on Big Data (Big Data)*. 756–765. https://doi.org/10.1109/BigData.2017.8257991

[19] Seyyed Mohammadreza Rahimi, Rodrigo Augusto de Oliveira e Silva, Behrouz Far, and Xin Wang. 2019. Optimized Random Walk with Restart for Recommendation Systems. In *Advances in Artificial Intelligence*, Marie-Jean Meurs and Frank Rudzicz (Eds.). Springer International Publishing, Cham, 320–332.

[20] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based Collaborative Filtering Recommendation Algorithms. In *Proceedings of the 10th International Conference on World Wide Web (WWW '01)*. ACM, New York, NY, USA, 285–295. https://doi.org/10.1145/371920.372071

[21] A. M. Tawfik, H. A. Rakha, and S. D. Miller. 2010. Driver route choice behavior: Experiences, perceptions, and choices. In *2010 IEEE Intelligent Vehicles Symposium*. 1195–1200. https://doi.org/10.1109/IVS.2010.5547968

[22] Hanghang Tong, Christos Faloutsos, Christos Faloutsos, and Jia-Yu Pan. 2006. Fast Random Walk with Restart and Its Applications. In *Proceedings of the Sixth International Conference on Data Mining (ICDM '06)*. IEEE Computer Society, Washington, DC, USA, 613–622. https://doi.org/10.1109/ICDM.2006.70

[23] Jingyuan Wang, Ning Wu, Wayne Xin Zhao, Fanzhang Peng, and Xin Lin. 2019. Empowering A* Search Algorithms with Neural Networks for Personalized Route Recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*. ACM, New York, NY, USA, 539–547. https://doi.org/10.1145/3292500.3330824

[24] Ling-Yin Wei, Yu Zheng, and Wen-Chih Peng. 2012. Constructing Popular Routes from Uncertain Trajectories. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12)*. ACM, New York, NY, USA, 195–203. https://doi.org/10.1145/2339530.2339562

[25] S. Wei, N. Ye, S. Zhang, X. Huang, and J. Zhu. 2012. Item-Based Collaborative Filtering Recommendation Algorithm Combining Item Category with Interestingness Measure. In *2012 International Conference on Computer Science and Service System*. 2038–2041. https://doi.org/10.1109/CSSS.2012.507

[26] P. Yawalkar and S. Ranu. 2019. Route Recommendations on Road Networks for Arbitrary User Preference Functions. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. 602–613. https://doi.org/10.1109/ICDE.2019.00060

[27] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. 2013. Time-aware Point-of-interest Recommendation. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13)*. ACM, New York, NY, USA, 363–372. https://doi.org/10.1145/2484028.2484030

[28] Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. 2008. Understanding Mobility Based on GPS Data. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*. ACM, New York, NY, USA, 312–321. https://doi.org/10.1145/1409635.1409677

[29] Yu Zheng, Xing Xie, and Wei-Ying Ma. 2010. GeoLife: A Collaborative Social Networking Service among User, location and trajectory. *IEEE Data(base) Engineering Bulletin* (June 2010). https://www.microsoft.com/en-us/research/publication/geolife-a-collaborative-social-networking-service-among-user-location-and-trajectory/

[30] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*. ACM, New York, NY, USA, 791–800. https://doi.org/10.1145/1526709.1526816

[31] X. Zhu, R. Hao, H. Chi, and X. Du. 2017. FineRoute: Personalized and Time-Aware Route Recommendation Based on Check-Ins. *IEEE Transactions on Vehicular Technology* 66, 11 (Nov 2017), 10461–10469. https://doi.org/10.1109/TVT.2017.2764999