Research Article

A Travel Route Recommendation System Based on Smart Phones and IoT Environment

Chenzhong Bin ^{1,2} **Tianlong Gu**,³ **Yanpeng Sun**,⁴ **Liang Chang** ² and **Lei Sun**⁵

¹School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China

²Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China

³Guangxi Experiment Center of Information Science, Guilin University of Electronic Technology, Guilin 541004, China

⁴School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin 541004, China

⁵School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China

Correspondence should be addressed to Chenzhong Bin; binchenzhong@163.com

Received 15 March 2019; Revised 26 May 2019; Accepted 1 July 2019; Published 14 July 2019

Academic Editor: Gerardo Canfora

Copyright © 2019 Chenzhong Bin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Tourism recommendation systems play a vital role in providing useful travel information to tourists. However, existing systems rarely aim at recommending tangible itineraries for tourists within a specific POI due to their lack of onsite travel behavioral data and related route mining algorithms. To this end, a novel travel route recommendation system is proposed, which collects tourist onsite travel behavior data automatically regarding a specific POI based on smart phone and IoT technology. Then, the proposed system preprocesses the behavior data to transform raw behavior sequences into Tourist-Behavior pattern sequences. Subsequently, the system discovers frequent travel routes from the generated pattern sequences by using an original route mining algorithm, named Tourist-Behavior PrefixSpan. Finally, a route-recommending method is designed to search and rank tangible travel routes according to the querying tourist's profile and constraint. The experimental results demonstrate that the proposed system is efficient and effective in recommending POI-oriented tangible travel routes considering tourists' route constraints and personal profile while ensuring that the suggested routes have considerable route values.

1. Introduction

Tourism is a popular leisure activity with the goal of visiting some Points of Interests (abbr. POIs) based on one's personal preference and constraints. Recently, industry and academia have been studying and developing tourism recommendation systems for providing convenient travel information to tourists, including next POI suggestion [1–7], Top-k POIs recommendation [8–10], and POIs travel route recommendation [11–18]. Particularly, travel route recommendations are more practical and useful than the two former kinds of POI recommendations in practice, yet they are more challenging. Travel route recommendation aims to organize a bundle of candidate POIs as a reasonable visit sequence (i.e., an itinerary) while adhering to personal constraints of a given tourist, for example, a limited time or finance budget, user-specified start, and end locations.

Thus, many researchers have studied travel route recommendation issues and designed various algorithms for solving

these problems [19]. Most of these works are orienting to city-level or district-level itinerary recommendation scenario, that is, planning a POIs travel route within a city or region for tourists. On the other side, it is difficult for tourists to instantly choose personal interested exhibits or spots and arrange these items into an itinerary under their time budget when travelling in a never visit POI, for example, a museum or a park. This situation makes most of tourists roaming or missing some potential interested items in the POI. However, few existing systems manage to recommend tangible itineraries for tourists within a given POI due to lack of rich onsite travel behavior data and related itinerary mining algorithms.

The popularity of smart phones and the flourish Internet of Things (IoT) techniques provide various means to sense onsite travel behaviors of tourists, including not only travel spatial-temporal trajectories [20] and visit durations but also the tangible travel behaviors [21] such as taking photos, standing, or walking. It is a common sense that one's onsite

travel behaviors imply his or her objective preferences and interests to some objects. For instance, tourists will spend longer visit duration or take more pictures or stand still more times to appreciate something on a spot if they are more interested in something. Thus, gathering tourists' onsite travel behaviors and mining their personal preferences and frequent travel routes could be an effective approach for recommending tangible travel routes for new similar tourists in specific POIs.

To that end, in this work, we proposed a POI-oriented travel route recommendation system based on IoT technology and smart phones. In detail, we first adopted Bluetooth low energy (BLE) beacons [22] to periodically broadcast positioning information for nearby smart phones. Also, we developed a client App running on an Android smartphone to collect onsite travel behaviors data and corresponding personal profiles and then upload collected data to the system server. Next, on the system server side, all collected travel behavior data are classified according to their personal profiles. And then a behavior sequence preprocessing method and Tourist-Behavior pattern mining algorithms were designed to generate diverse tangible travel routes. At route recommendation stage, to ensure the personalization of our recommendations, the proposed route ranking method recommends tangible travel routes for new tourists by using their personal profiles and route constraints. As all tangible travel routes are constructed from real historical onsite travel behavior, the recommended routes have high accuracy and rationality in terms of visit arrangements. Also, since the recommended routes are retrieved from the corresponding candidate route subset according to the querying tourist's profile, the visit objects of the final route can suit the personal interests of the tourist better.

Our main contributions are summarized as follows. (a) An onsite travel behavior data collecting method, which is based on tourists' smartphones and Bluetooth low energy (BLE) beacons, is designed to automatically sense onsite travel behavior under indoor and outdoor tourism scenarios. (b) Tourist-Behavior PrefixSpan algorithm is proposed to generate diverse frequent travel routes effectively based on historical Tourist-Behavior pattern sequences. (c) Travel route ranking method is proposed to recommend a list of tangible travel routes according to the querying tourist's profile and constraints so as to ensure the route value and rationality of the final travel routes. (d) Experimental results demonstrate the effectiveness of our system in recommending personalized tangible travel routes for tourists in a given POI based on historical onsite travel behavior.

The rest of the paper is organized as follows. Section 2 discusses the related works regarding travel route recommendation systems and tourism recommendation within IoT environment. Section 3 presents the research methodologies of the proposed system, where the framework of the system is described in Section 3.1; onsite travel behavior data collecting method is explained in Section 3.2. The Tourist-Behavior pattern sequences mining method and tangible travel route recommendation procedure are thoroughly explained in Sections 3.3 and 3.4, respectively. Section 4 analyzes the experimental results to validate the feasibility and performance of

the proposed system. The conclusion and future works are presented in Section 5.

2. Related Works

2.1. Travel Route Recommendation Systems. Due to the practical values of travel route recommendation systems, lots of researchers have been placing a great emphasis on solving the route planning in tourism scenarios [19, 23] in recent years. One category of these works uses the Orienteering problem [24] and its variants to approach the route planning problem. These methods formulate the problem from different perspectives, resulting in diverse problem models, which consider different problem variables and constraints. The route-generation process actually is a near-optimal solution using metaheuristic searching algorithm [25]. Accordingly, to enhance the personalization of the recommended routes, most of these works resort to acquiring more detailed user feedbacks or profiles to assist in fine-tuning the final results. In [26], the system solicits walking travel related attributes from tourists to insert concrete walking routes into POI itineraries, thereby supporting more experiential exploration of tourist destinations. Zhang et al. [27, 28] studied tour recommendation with the goal of recommending personalized itineraries based on the interest preferences of users and available touring time, while considering opening hours of POIs and uncertainty in travelling time. Other studies consider more practical factors that raise novel optimization challenges incorporating forms of situational awareness such as multiple modes of transport [29], considering traffic conditions [30–32], POI crowdedness [33, 34], and queuing times [35].

Although the optimization-based route planning systems can recommend a reasonable travel route adhering to one's preferences and constraints, the interactive preferences input of the planning process is time costly to tourists. It is impractical for tourists to spend a long time in inputting a complex user profile when entering a specific POI. Furthermore, the results of these systems are lack of diversity and less personalization due to the near-optimal solution searching methodology. Therefore, lots of other works focus on generating personalized travel routes by mining User Generated Contents (UGC), that is, data-driven approaches to route planning. The UGC adopted in previous researches include GPS trajectory datasets [36], check-in datasets [37, 38], and geo-tagged photos [39].

Chen et al. [31] adopted historical check-in data and GPS trajectories to construct a POI network and used a heuristic method to generate a favorite POIs list for a specific user in an interactive manner. Subsequently, the system requests users to specify their favorite POIs during the route-generating stage. PersTour system [13, 40] uses geo-tagged photos to determine POI and construct POI travel sequences and leverage an Orienteering problem solving model to recommend POI itineraries by both considering POI popularity and tourist personal interests. Majid et al. [41] inferred the location of POIs and their semantic meaning using clustering approaches on geo-tagged photos and used a pattern mining

algorithm to discover popular travel sequences under the context of the tour recommendation, that is, time, day, and weather. Besides, the rapid growth of online tourism websites provides massive POI reviews and travelogues. Thus, some recent works [14, 42–44] focused on generating personalized mining travelogues and POI related contents.

The data-driven based route planning systems can recommend rather personalized and reasonable travel routes; however, the limitation of these systems is that they aim at recommending city-level or district-level orienting POIs itineraries, that is, planning POIs travel routes for tourists within a city or region. They fail to generate tangible travel routes for tourists within a specific POI due to lack of rich onsite travel behavior and related itinerary mining algorithms.

2.2. Personalized Recommendation in IoT Environment. The IoT concept was first coined by Kevin Ashton in 1999 [45] in supply chain management applications based on radio frequency identification devices (abbr. RFID). At present, IoT is referring to a bundle of technologies that aim at sensing, handling, and transmitting state information of physical environments, which is broadly applied in smart cities [46, 47], smart business [48, 49], and smart tourism scenarios [50]. The goal of these smart systems is to recommend a set of personalized and valuable items or services for various users. To this end, researchers focus on recording and analyzing user behavior to learn user preferences more precisely by deploying IoT technologies.

Specifically in smart tourism applications, some studies focus on using IoT technologies and mobile devices to improve tourism experiences in an interactive way. Kuusik et al. [51] designed a smart museum system that integrates PDAs and RFID technologies to provide users with cultural contents by sensing the interactive behavior between PDAs and RFID tags, which were installed near each artwork. In [52], an indoor location-aware system was designed for a smart museum to enhance visitors' cultural experiences. The proposed system obtains visitors' localization information through a Bluetooth low energy (BLE) infrastructure installed in the museum and uses several location-aware services hosted in the system to interact with visitors according to their locations.

And some other works aim at solving the next visiting spot recommendation problem within a specific POI. Massimo et al. [5] leveraged Inverse Reinforcement Learning method to learning user preferences by observing tourists' onsite behavior in an IoT-equipped smart museum so as to predict next exhibit sequentially for tourists. Hashemi et al. [6, 7] solved the challenging next POI recommending problem by logging and mining users' onsite physical and online interaction behavior data within an IoT-augmented museum. However, the above works fail to generate personalized and tangible travel routes for tourists.

To that end, some researchers strived to solve this challenging problem by mining historical travel trajectories in an IoT-augmented environment. Tsai et al. [15] adopted RFID infrastructure to record visitors' check-in sequences of recreation facilities in a theme park and then proposed a

statistical method to find behavioral similar historical visitors so as to suggest a travel route for the current querying visitor. Luo et al. [16] studied a new path finding system that discovers the most frequent path during user-specified time periods in large-scale historical trajectory data. Tsai et al. [17] proposed a touring path suggesting system for visitors to comprehend exhibits in exhibitions or museums. The system takes previous popular visiting trajectories as the suggestion foundation and provides a time-interval sequential patterns mining algorithm improved from [18] to generate personalized travel routes. However, as the above systems only resorted to dedicated IoT devices to record the check-in behavior, they hardly learn more tangible user preferences towards each interest object from the single dimensional behavior. Meanwhile, current smartphones generally equip a camera and diverse sensors, which could be used to sense multiple dimensional onsite behaviors of tourists so as to explore high-level tourist preferences and recommend personalized tangible travel route. Although some previous researches have investigated the human activity recognition based on smartphone sensors [21, 53], there is no study on learning user preferences directly by smartphones. To the best of our knowledge, our proposed system is the first work of leveraging smartphones and IoT environment to recommend tangible travel route within POIs based on onsite travel behavior sensing and mining methods.

3. Research Methodologies

3.1. System Overview. In this work, we use the phrase *scenic area* to denote a park or a museum containing a series of sightseeing spots or exhibits, namely, *interesting spots*. At each interesting spot, entrance and exit of a scenic area need to be preinstalled a Bluetooth low energy (BLE) beacon to locate tourists in an indoor or outdoor scenario. An illustrative example of the system is shown in Figure 1. Concretely, our system adopts iBeacon [22] devices to indicate specific spots by broadcasting their own device tags, that is, positioning information. When a tourist is approaching an interesting spot with a smart phone, the phone will use the locating information to judge whether the tourist has arrived at this spot. If so, the phone will record the onsite travel behavior data at this spot. At the end of the travel, the phone uploads a complete behavior sequence and a user-specified profile to the system server. Subsequently, the server preprocesses these data, transforming them into Tourist-Behavior (TB) pattern sequences, and uses the TB pattern mining algorithm to generate candidate tangible travel routes. At the recommendation stage, the system server will recommend personalized tangible travel routes for a new tourist by using the route ranking method according to the tourist's personal profile and constraints. The recommended travel route, which contains a spot visit sequence and their respective visit durations, will help them to finish a valuable tour in the area in a comfortable way.

Figure 2 illustrates the workflow of the proposed system. Specifically, stage 1 is performed by a client App running on tourists' smart phones, which is responsible for collecting

FIGURE 1: An illustrative example for the travel route recommendation.

FIGURE 2: The workflow of the proposed system.

tourists' personal profiles and their behavior data, while stages 2 and 3 are performed on the system server side. In offline running stage 2, behavior sequence preprocessing method and Tourist-Behavior (TB) PrefixSpan algorithm are proposed to generate a series of TB pattern sequences, that is, candidate tangible travel routes. In online running stage 3, the system server recommends tangible travel routes for various tourists based on their profiles and route constraints.

3.2. Onsite Travel Behavior Collecting. Since tourists with different personal attributes may have different personal interests, stamina, walking speeds, and so forth to ensure the personalization of our recommendations, we classify and store the collected behavior sequences according to

corresponding personal profiles in our system. At the beginning of the behavior data collection process, we request each tourist to input three common and typical personal attributes, including gender, age group, and education level, as a simple profile. Then, the client App uploads an onsite travel behavior sequence and a corresponding profile together to the system server. Subsequently, at the recommendation stage, the system server uses a personal profile of the querying tourist to retrieve generated routes from the corresponding route subset for matching tourists' different interests.

3.2.1. Positioning Mechanism. The positioning mechanism is implemented based on iBeacon devices and smart phones. The iBeacon protocol is characterized as low energy

Sid	Onsite behavior sequence
01	$\langle 0, Z_e, 3, 0, 0 \rangle, \langle \mathbf{6, A, 26, 5, 4} \rangle, \langle 35, B, 45, 0, 4 \rangle, \langle 55, D, 70, 2, 2 \rangle, \langle 78, F, 90, 4, 4 \rangle, \langle \mathbf{99, E, 101, 0, 0} \rangle, \langle 108, G, 128, 3, 4 \rangle, \langle 133, Z_p, 135, 0, 0 \rangle$
02	$\langle 0, Z_e, 2, 0, 0 \rangle, \langle 4, A, 21, 4, 5 \rangle, \langle 31, B, 40, 1, 2 \rangle, \langle 50, C, 60, 2, 1 \rangle, \langle 72, F, 86, 2, 2 \rangle, \langle 92, G, 106, 1, 3 \rangle, \langle 112, Z_p, 114, 0, 0 \rangle$
03	$\langle 0, Z_e, 3, 0, 0 \rangle, \langle 9, A, 22, 1, 1 \rangle, \langle 34, B, 38, 0, 0 \rangle, \langle 46, C, 58, 3, 4 \rangle, \langle 70, F, 86, 4, 4 \rangle, \langle 90, E, 108, 6, 4 \rangle, \langle 112, G, 124, 2, 3 \rangle, \langle 130, Z_p, 132, 0, 0 \rangle$

Definition 2. A Tourist-Behavior (TB) pattern λ_i is defined as a triple $\langle b_i, NP_i, D_i \rangle$, where b_i is the location identity of spot i ; NP_i is the normalized popularity value about spot i ; D_i is the discrete visit duration at spot i . Note that the pattern λ_i is said to match the pattern λ_j if and only if $b_i = b_j$, $NP_i = NP_j$, and $D_i = D_j$.

Definition 3. Let $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_x\}$ be the set of TB patterns and let δ_i be the discrete interspot travel time in a travel behavior sequence. A sequence $\alpha = (V_1, \delta_1, V_2, \delta_2, \dots, \delta_{k-1}, V_k)$ is a TB sequence if $V_s \in \Lambda$ for $1 \leq s \leq h$ and $\delta_s = \text{DiscT}(\Delta t)$ for $1 \leq s \leq h-1$.

First, the preprocessing method cleans up the passing-by behavior data and calculates the interspot travel time and the visit duration in each travel behavior sequence. Hence, the method needs to delete the behavior data if the visit duration is shorter than a time threshold T_v , except for the entrance and exit behavior data. Let δ_i be the discrete interspot travel time for the tourist to travel from spot i to spot $i+1$, and let D_i be the visit duration at spot i ; Δt stands for δ_i or D_i ; and T_d is the metric of the discrete time. Consequently, the discrete time integer of δ_i and D_i can be derived from the following equation:

$$\text{DiscT}(\Delta t) = \left\lceil \frac{\Delta t}{T_d} \right\rceil \quad (2)$$

Second, the method calculates popularity values of each interesting spot in each travel behavior sequence. As each travel behavior sequence is collected from an individual tourist, two popularity values of the same spot in two sequences are probably different due to two different tourists' onsite behaviors. The prior knowledge of the method is that tourists will spend longer visit duration, take more pictures, or stand still more times to appreciate something at a spot if they are more interested in the spot. The popularity value of spot i in a specific sequence can be calculated by the following equation:

$$\begin{aligned} \text{Pop}_i = & w_1 \times \frac{\text{stout}_i - \text{stin}_i}{\sum_{i=1}^k (\text{stout}_i - \text{stin}_i)} + w_2 \times \frac{P_i}{\sum_{i=1}^k P_i} + w_3 \\ & \times \frac{s_i}{\sum_{i=1}^k s_i} \end{aligned} \quad (3)$$

where w_1 , w_2 , and w_3 are weights used to calculate Pop_i and $w_1 + w_2 + w_3 = 1$; for each travel behavior sequence, the total visit duration is derived from $\sum_{i=1}^k (\text{stout}_i - \text{stin}_i)$; $\sum_{i=1}^k P_i$ and $\sum_{i=1}^k s_i$ denote the total number of times of taking pictures and standing still within the sequence, respectively. Besides, to make the popularity values of spots in different sequences comparable, we normalize all popularity values in each sequence. In detail, to calculate a normalized popularity value NP_i of spot i , all spots in each sequence are ranked as a descending list according to their respective Pop_i . To calculate NP_i , the list is divided into n segments where n denotes the popularity normalization coefficient. For example, in (4), the normalization coefficient is to be 4; all spots ranking in the top $1/n$ in a specific sequence are assigned with a normalized popularity value of Ln , indicating that the querying tourist is

most likely to be interested in these spots.

$$NP_i = \begin{cases} Ln, & \text{if } \text{Pop}_i \text{ ranks in the top } \frac{1}{n} \\ \vdots & \\ L2, & \text{if } \text{Pop}_i \text{ ranks in } \left[\frac{n-2}{n}, \frac{n-1}{n} \right) \\ L1, & \text{if } \text{Pop}_i \text{ ranks in the lowest } \frac{1}{n} \end{cases} \quad (4)$$

After the preprocessing step, all travel behavior sequences are transformed into TB pattern sequences and stored in the Tourist-Behavior sequence database (abbr. TBD), according to their respective profiles. Specifically, our system divides tourist age into three groups, below 20 years, from 20 to 55 years, and over 55 years, and classifies education into three levels: preundergraduate, undergraduate, and graduate. By multiplying with two gender attributes, there are $3 \times 3 \times 2 = 18$ TBDs in total in our system with the above three profile attributes.

Example 4. Let us take the travel behavior sequences shown in Table 1 as an example to explain the travel behavior sequence preprocessing method. Suppose that T_v is set at 5 minutes; T_d is set at 10 minutes; and w_1 , w_2 , and w_3 are set as 0.4, 0.3, and 0.3, respectively. At first, the behavior data $\langle 99, E, 101, 0, 0 \rangle$ is deleted as a passing-by behavior data in *sid 01*, because its visit duration is shorter than T_v . Further, the interspot travel time and the visit duration are discretized by (2). Next, the popularity of each spot is computed; for example, Pop_A with respect to tourist #1 is calculated as $0.4 \times 20/82 + 0.3 \times 5/14 + 0.3 \times 4/18 = 0.271$; the visit duration D_A is 20 minutes; the total visit duration is 82 minutes; the number of taking pictures and standing still is 14 and 18, respectively. The corresponding TB pattern sequences are shown in Table 2.

3.3.2. Tourist-Behavior Sequential Travel Routes Generating. As the onsite travel behaviors are complex and contain noisy behavior data, for example, one making a phone call or taking a sit for a break during a visit, we need a method to discover popular travel routes and to filter noise travel behaviors. Therefore, we design the TB PrefixSpan algorithm to discover all frequent TB patterns with the corresponding interspot travel time and to construct various Tourist-Behavior (TB) sequential travel routes from a TBD. An improvement of the TB PrefixSpan algorithm compared to [54] is that due to the fact that TB pattern sequences separately contain discrete interspot travel time and spot visit durations, the TB PrefixSpan algorithm can delete visit durations of non-frequent TB patterns yet preserve intervals to ensure the accurate time arrangement of new TB sequential patterns. Before describing the TB PrefixSpan algorithm, the following definitions are given.

Definition 5. Assume two TB pattern sequences $\alpha = (V_{\alpha 1}, \delta_{\alpha 1}, V_{\alpha 2}, \delta_{\alpha 2}, \dots, \delta_{\alpha(k-1)}, V_{\alpha k})$ and $\beta = (V_{\beta 1}, \delta_{\beta 1}, V_{\beta 2}, \delta_{\beta 2}, \dots, \delta_{\beta(h-1)}, V_{\beta h})$ ($h \leq k$); β is said to be contained in α or a TB subsequence of α ; that is, $\beta \subseteq \alpha$, if there exist sequence

TABLE 2: An example of Tourist-Behavior pattern sequence.

Sid	Tourist-Behavior sequence
01	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 2 \rangle, 1, \langle B, L2, 1 \rangle, 1, \langle D, L2, 2 \rangle, 1, \langle F, L3, 2 \rangle, 2, \langle G, L3, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
02	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 2 \rangle, 1, \langle B, L2, 1 \rangle, 1, \langle C, L2, 1 \rangle, 2, \langle F, L3, 2 \rangle, 1, \langle G, L3, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
03	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L2, 2 \rangle, 3, \langle C, L3, 2 \rangle, 3, \langle F, L3, 2 \rangle, 1, \langle E, L4, 2 \rangle, 1, \langle G, L2, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle$)

TABLE 3: An example of TB sequences database.

Sid	TB pattern sequences
01	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 3 \rangle, 1, \langle C, L3, 2 \rangle, 2, \langle E, L3, 2 \rangle, 1, \langle F, L2, 2 \rangle, 1, \langle G, L4, 2 \rangle, 2, \langle H, L4, 3 \rangle, 1, \langle L, L2, 1 \rangle, 1, \langle X, L2, 3 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
02	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 3 \rangle, 1, \langle B, L3, 3 \rangle, 1, \langle C, L3, 1 \rangle, 2, \langle F, L2, 3 \rangle, 1, \langle H, L2, 2 \rangle, 1, \langle L, L4, 2 \rangle, 1, \langle R, L3, 3 \rangle, 1, \langle X, L3, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
03	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L3, 3 \rangle, 1, \langle B, L3, 3 \rangle, 2, \langle E, L2, 3 \rangle, 1, \langle G, L4, 2 \rangle, 2, \langle L, L4, 2 \rangle, 1, \langle D, L3, 2 \rangle, 1, \langle X, L2, 3 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
04	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L2, 2 \rangle, 2, \langle D, L3, 2 \rangle, 1, \langle E, L3, 2 \rangle, 2, \langle H, L4, 3 \rangle, 2, \langle L, L2, 1 \rangle, 2, \langle R, L3, 3 \rangle, 1, \langle X, L2, 3 \rangle, 1, \langle Zt, Ll, 1 \rangle$)
05	($\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 3 \rangle, 2, \langle C, L3, 1 \rangle, 2, \langle F, L2, 3 \rangle, 1, \langle H, L2, 2 \rangle, 2, \langle R, L3, 1 \rangle, 1, \langle X, L3, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle$)

indices $1 \leq j_1 < j_2 < \dots < j_h \leq k$ such that (1) $V_{\beta_1} = V_{\alpha_{j_1}}$ and $V_{\beta_2} = V_{\alpha_{j_2}}, \dots, V_{\beta_h} = V_{\alpha_{j_h}}$; (2) $\delta_{\beta_1} = \delta_{\alpha_{j_1}}$ and $\delta_{\beta_2} = \delta_{\alpha_{j_2}}, \dots, \delta_{\beta_h} = \delta_{\alpha_{j_h}}$.

Definition 6. A TB pattern γ is called a frequent TB pattern if the number of sequences in a TBD, which contains γ as the subsequence, is greater than or equal to the user-specified minimum support, called *min_sup* or *min_sup_count*. That is, γ is called a frequent TB pattern in a TBD if $\text{sup_count}_{\text{TBD}}(\gamma) \geq |\text{TBD}| \times \text{min_sup}$ or $\text{sup_count}_{\text{TBD}}(\gamma) \geq \text{min_sup_count}$, where $\text{sup_count}_{\text{TBD}}(\gamma) = |\{\beta_i \in \text{TBD} \wedge \gamma \subseteq \beta_i, 1 \leq i \leq |\text{TBD}|\}|$.

Definition 7. Assume a TB pattern sequence $\alpha = (V_{\alpha_1}, \delta_{\alpha_1}, V_{\alpha_2}, \delta_{\alpha_2}, \dots, \delta_{\alpha_{(k-1)}}, V_{\alpha_k})$, and α is called a TB sequential travel route if all TB patterns in α are frequent TB patterns; further α can be referred to as a k -length TB sequential travel route.

Definition 8. Given two TB sequential travel routes $\alpha = (V_{\alpha_1}, \delta_{\alpha_1}, V_{\alpha_2}, \delta_{\alpha_2}, \dots, \delta_{\alpha_{(k-1)}}, V_{\alpha_k})$ and $\beta = (V_{\beta_1}, \delta_{\beta_1}, V_{\beta_2}, \delta_{\beta_2}, \dots, \delta_{\beta_{(h-1)}}, V_{\beta_h})$ ($h \leq k$), β is a TB prefix of α if and only if (1) $V_{\beta_i} = V_{\alpha_i}$ for $1 \leq i \leq h$; (2) $\delta_{\beta_i} = \delta_{\alpha_i}$ for $1 \leq i \leq h - 1$.

Definition 9. Given two TB sequential travel routes $\alpha = (V_{\alpha_1}, \delta_{\alpha_1}, V_{\alpha_2}, \delta_{\alpha_2}, \dots, \delta_{\alpha_{(k-1)}}, V_{\alpha_k})$ and $\beta = (V_{\beta_1}, \delta_{\beta_1}, V_{\beta_2}, \delta_{\beta_2}, \dots, \delta_{\beta_{(h-1)}}, V_{\beta_h})$ ($h \leq k$), β is a subsequence of α . Let $1 \leq j_1 < j_2 < \dots < j_h \leq k$ be the indices of frequent TB patterns contained in α which match in β . A subsequence $\alpha' = (V_{\alpha'_1}, \delta_{\alpha'_1}, V_{\alpha'_2}, \delta_{\alpha'_2}, \dots, \delta_{\alpha'_{(g-1)}}, V_{\alpha'_g})$ of α , where $g = h + k - j_h$, is named a projection of α with respect to β if and only if (1) β is a TB prefix of α' and (2) the last $k - j_h$ TB patterns of α' are the same as the last $k - j_h$ TB patterns of α .

Definition 10. Let $\alpha' = (V_{\alpha'_1}, \delta_{\alpha'_1}, V_{\alpha'_2}, \delta_{\alpha'_2}, \dots, \delta_{\alpha'_{(m-1)}}, V_{\alpha'_m})$ be the projection of α with respect to a TB prefix $\beta = (V_{\beta_1}, \delta_{\beta_1}, V_{\beta_2}, \delta_{\beta_2}, \dots, \delta_{\beta_{(h-1)}}, V_{\beta_h})$ ($h \leq m$). Then $\theta = (V_{\alpha'_{(h+1)}}, \delta_{\alpha'_{(h+1)}}, V_{\alpha'_{(h+2)}}, \delta_{\alpha'_{(h+2)}}, \dots, \delta_{\alpha'_{(m-1)}}, V_{\alpha'_{(m)}})$ is the TB postfix of α with respect to prefix β .

The pseudocode of the TB PrefixSpan algorithm is shown in Figure 3. The α -projection database consists of postfixes of

TB pattern sequences in a TBD with respect to the TB prefix α , which is denoted as $\text{TBD}|_{\alpha}$. As the original PrefixSpan algorithm does not include the relationship among two TB patterns and their interval, a TB_Table is designed to store this type of relation where a row corresponds to a TB pattern and a column corresponds to a δ value. For instance, $\text{TB_Table}|_{\lambda_i}$ stores the support count of subsequences with respect to the current TB prefix α which has the last TB pattern λ_i . The table cell $\text{TB_Table}|_{\lambda_i}(\delta_N, \lambda_k)$ records the number of subsequences in $\text{TBD}|_{\alpha}$ containing the TB pattern subsequence $(\lambda_i, \delta_N, \lambda_k)$. Note that δ_N is an accumulated time from spot i to spot k ; that is, $\delta_N = \delta_i + \delta_{i+1} + \dots + \delta_{k-1}$.

Specifically, the algorithm initially recognizes each frequent TB pattern to construct their corresponding α -projection databases. For each $\text{TBD}|_{\alpha}$ database, the algorithm constructs the corresponding TB_Table to identify all frequent table cells. Then, for each frequent cell the element (δ_N, λ_j) is appended to the end of α to construct a new TB prefix α' and then the α' -projection database $\text{TBD}|_{\alpha'}$ is built. Recursively constructing all of the frequent TB pattern sequences in the $\text{TBD}|_{\alpha'}$ discovers all TB sequential travel routes in the TBD, which are stored in the TB sequential travel route database (abbr. TBSTR).

Example 11. Let us take five TB pattern sequences in Table 3 as an example to explain the TB sequential travel route mining process, where the *min_sup_count* is set at 2. The TB PrefixSpan algorithm can be also deemed as a Tree Traversal algorithm; each node of the growing tree corresponds to a $\text{TB_Table}|_{\lambda_i}$. As shown in Figure 4, we use red solid arrows to illustrate steps of constructing a TB sequential travel route $(\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 3 \rangle, 2, \langle F, L2, 3 \rangle, 4, \langle X, L3, 2 \rangle, 1, \langle Zt, Ll, 1 \rangle)$. In the beginning, the algorithm constructs $\text{TB_Table}|_{\langle Ze, Ll, 1 \rangle}$ which recognizes 16 1-length sequential routes listed in Table 4. Suppose that the algorithm is mining the current 3-length route $(\langle Ze, Ll, 1 \rangle, 1, \langle A, L4, 3 \rangle, 2, \langle F, L2, 3 \rangle)$. Thus $\text{TB_Table}|_{\langle F, L2, 3 \rangle}$ needs to be constructed subsequently. As shown in Table 5, the $\langle F, L2, 3 \rangle$ -projection database has two TB postfix *sid* 02 and 05. The results of $\text{TB_Table}|_{\langle F, L2, 3 \rangle}$ are shown in Table 6. There are 3 frequent TB patterns labeled in bold in the table, that is, $\langle H, L2, 2 \rangle$, $\langle X, L3, 2 \rangle$,

FIGURE 3: The pseudocode of the TB PrefixSpan algorithm.

and $\langle Zt, Ll, l \rangle$. The algorithm joins these 3 patterns behind the current route, respectively, to construct 3 new 4-length routes, and then constructs 3 corresponding TB.Tables of these patterns. The algorithm recursively traverses the tree to discover all potential TB sequential travel routes.

3.4. Travel Route Ranking and Recommending. As mentioned in Section 3.3.1, all TB pattern sequences are divided into subsets according to the personal profile after the sequence preprocess step. To make the recommended routes match the querying tourist's personal interests and characteristics better, we design a route ranking method to search valuable

TABLE 4: An example of 1-length frequent TB pattern.

TB pattern	Sup_count	TB pattern	Sup_count
$\langle Ze, Ll, l \rangle$	5	$\langle H, L2, 2 \rangle$	2
$\langle A, L4, 3 \rangle$	3	$\langle H, L2, 3 \rangle$	2
$\langle B, L3, 3 \rangle$	2	$\langle L, L2, 1 \rangle$	2
$\langle C, L3, 1 \rangle$	2	$\langle L, L4, 2 \rangle$	2
$\langle D, L3, 2 \rangle$	2	$\langle X, L2, 3 \rangle$	3
$\langle E, L3, 2 \rangle$	2	$\langle X, L3, 2 \rangle$	2
$\langle F, L2, 3 \rangle$	2	$\langle R, L3, 3 \rangle$	2
$\langle G, L4, 2 \rangle$	2	$\langle Zt, Ll, l \rangle$	5

FIGURE 4: An example of TB sequential travel routes generation process.

TABLE 5: An example of the $\langle F, L2, 3 \rangle$ -projection database $TBD|_{\langle F, L2, 3 \rangle}$.

Sid	TB Pattern $\langle F, L2, 3 \rangle$ -projection database
02	$(1, \langle H, L2, 2 \rangle, 1, \langle L, L4, 2 \rangle, 1, \langle R, L3, 3 \rangle, 1, \langle X, L3, 2 \rangle, 1, \langle Zt, L1, 1 \rangle)$
05	$(1, \langle H, L2, 2 \rangle, 2, \langle R, L3, 1 \rangle, 1, \langle X, L3, 2 \rangle, 1, \langle Zt, L1, 1 \rangle)$

TABLE 6: An example of $TB_Table|_{\lambda_i}$.

TB pattern	Interspot δ_N				
	1	2	3	4	5
$\langle H, L2, 2 \rangle$	2	0	0	0	0
$\langle L, L4, 2 \rangle$	0	1	0	0	0
$\langle R, L3, 3 \rangle$	0	0	1	0	0
$\langle R, L3, 1 \rangle$	0	0	1	0	0
$\langle X, L3, 2 \rangle$	0	0	0	2	0
$\langle Zt, L1, 1 \rangle$	0	0	0	0	2

and reasonable routes from a TBSTR matched by an input personal profile.

Thus, the method first requests the querying tourist to input a personal profile and a route constraint. The route constraint includes the intended travel duration and specified travel start and end location of the POI. Next, the method uses the personal profile to retrieve candidate TB sequential travel routes in the corresponding TBSTR. After retrieving candidate TB sequential travel routes, the server filters out travel routes that do not meet the input route constraint. In detail, the server reserves the routes of which start and end points match the user-specified entrance and exit. This is for considering the situation of multiple entrances and exits existing in one scenic area. Then, it adds up the total

route duration time $T_{t\alpha}$ of the route α by using the following equation:

$$T_{t\alpha} = \left(\sum_{i=1}^{|\alpha|-1} \delta_i + \sum_{i=1}^{|\alpha|} D_i \right) \times T_d, \quad (5)$$

where $|\alpha|$ is the length of the route α ; δ_i and D_i are the i th discrete interval and spot visit duration of the route α , respectively; T_d is the metric of the discrete time. And the server selects the candidate routes that meet the following equation:

$$(1 - \varphi) \times ITD \leq T_{t\alpha} \leq ITD, \quad (6)$$

where ITD stands for an intended travel duration of the querying tourist; φ is a filtering condition parameter between $[0, 1]$ used to set a filtering range for the candidate routes.

At last, the system server recommends the most valuable Top-k tangible travel routes for the querying tourist by calculating route values of the remaining routes. A route value consists of the total normalized popularity value and the ratio of the total visit duration to the total route duration. The candidate travel routes set is denoted as $\{CR_i \mid i = 1, 2, \dots, M\}$,

FIGURE 5: A test of onsite behavior acquisition.

where M is the total number of candidates. The route value RV_α of route α is calculated by the following equation:

$$RV_\alpha = w_{pl} \times MMN \left(\sum_{i=1}^{|\alpha|} NP_i \right) + w_{rvd} \times \left(\frac{\sum_{i=1}^{|\alpha|} D_i}{\sum_{i=1}^{|\alpha|-1} \delta_i + \sum_{i=1}^{|\alpha|} D_i} \right), \quad (7)$$

where w_{pl} and w_{rvd} are the weights used to calculate RV_α and $w_{pl} + w_{rvd} = 1$; $\sum_{i=1}^{|\alpha|} NP_i$ is the total normalized popularity value of route α ; $(\sum_{i=1}^{|\alpha|} D_i) / (\sum_{i=1}^{|\alpha|-1} \delta_i + \sum_{i=1}^{|\alpha|} D_i)$ is the ratio of the total visit duration to the total route duration of route α ; $MMN(\cdot)$ is min-max normalization function for normalizing the rank value among $\{CR_i \mid i = 1, 2, \dots, M\}$.

4. Experiment and Discussion

In this section, we designed a validation experiment and several performance analysis experiments to test our system. The iBeacon devices were based on CC2541 embedded processor developed by Texas Instruments Company. The client App was developed with Android Studio 2.3.3, which runs on Android smartphone system version 6.0.1 upwards. The system server runs on a workstation with Intel Xeon 3.5GHz and 16 GB RAM, and related application was developed by python version 2.7.13 running on Ubuntu 14.04 with Tomcat 6.0.

4.1. Validation Experiment. The primary goals of our empirical experiment are to (1) examine how well the recommended routes match a tourist's actual interests, (2) demonstrate the route value of the top recommended routes, and (3) analyze the rationality of the top recommended routes. In this section, we introduce the experimental settings of the

validation experiment and present the results of a test of onsite behavior data collection. Last, we present the results of the validation experiment to validate the ability of our system in recommending personalized tangible travel routes for tourists in a given POI based on historical onsite travel behavior.

At first, we deployed our system in a small experimental exhibition hall with 20 exhibits, where each exhibit is preinstalled with an iBeacon device. And we invited 20 male and 20 female undergraduate students as volunteers to visit the experimental hall so as to collect their onsite travel behavior data. The layout of the experimental hall is shown as in Figure 6, in which topical-related posters are exhibited with a similar length. In detail, from B1 to B5 are scientific topical posters, from B6 to B12 are sports-related, and from B13 to B20 are daily life related contents. And we set the minimum support count of TB PrefixSpan at 2 and set discrete time metric T_d at 1 minute; the popularity normalization coefficient is to be 5; the filtering condition parameter ϕ is set at 0.2; the weights w_{pl} and w_{rvd} in (7) were equally set at 0.5.

Figure 5 illustrates a test of onsite behavior acquisition. Figure 5(a) shows the experimental environment in which one volunteer carrying a smartphone is visiting an exhibit that is labeled by an iBeacon device. Figure 5(b) shows the software interface of the client App, which is selecting the nearest iBeacon device as the recognized spot, that is, Minor 3 device, to collect following onsite behavior data. Figure 5(c) presents the original behavior sequence gathered on the Minor 3 spot.

To examine how well the recommended routes match a tourist's actual interests, we requested the volunteers to fill a rating questionnaire regarding the exhibiting posters after visiting the exhibition hall so as to directly learn interests of female and male volunteers. Table 9 lists the most favorite 10 posters for female and male volunteers, respectively, which reflects that female volunteers prefer daily life topical

FIGURE 6: The result of top 1 tangible travel route for two querying tourists.

TABLE 7: Two querying tourists' route constraints and personal profiles.

Tourist	Time constraint	Personal profile
#1	45 min	Young/Female/Undergraduate
#2	30 min	Young/Male/Undergraduate

posters, while male volunteers prefer sports-related contents. Next, we assumed two querying tourists' route constraints and personal profiles that are listed in Table 7 to request route recommendations from our system. The top 3 valuable tangible travel routes recommended to two tourists are listed in Table 8. It can be easily observed from Table 8 that all candidate routes comply with corresponding tourist's time constraints. More importantly, by comparing the posters of Tables 8 and 9, we find that about 80% of top 10 favorite posters regarding the two corresponding tourists are included in both top recommended routes. For instance, the first route recommended to the female tourist suggests she spends a relatively longer time at B2, B15, and B16 posters which are the top favorite posters to female listed in Table 9, while the first route recommended to the male tourist recommends B4, B6, and B7 posters. This observation proves that our system can learn different personal preferences from real onsite travel behaviors.

To demonstrate the route value of the top recommended routes, our route ranking method ranks top 3 valuable tangible travel routes which are listed in Table 8. It can be easily observed in Table 8 that both top 1 routes recommended to young female and male tourists have the biggest route value. For example, compared to the rest two routes, the top 1 route recommended to young male tourist possesses the largest total normalized popularity value (i.e., $L38$), the longest total visit duration (i.e., 35 minutes), and the largest number of visit spots (i.e., 8 spots).

Regarding the rationality of the top recommended routes, as the recommended tangible routes are generated from onsite travel behaviors of young female or male tourists, the visit arrangements of the recommended route (e.g., the visit sequence, the interspot travel time, and the spot visit

FIGURE 7: The length of TB sequential travel routes under various data sizes.

duration) can completely comply with the layout of the exhibition hall. As a result, the recommended routes possess rather visit rationality. For instance, the interspot travel time of a tangible route recommended to young females is derived from the average walking speed of young females; and the visit duration of each spot is calculated from the historical onsite travel behaviors of young females, for example, the average reading speed of young females. Figure 6 illustrates the visit sequence of two top 1 tangible routes for two querying tourists. The results indicate that the recommended route can help two querying tourists to finish their respective time-limited visits comfortably.

4.2. Algorithm Performance Analysis. In this section, we study the performance of our TB PrefixSpan algorithm under different parameters settings. Obviously, longer travel routes, which contain more interesting spots, can meet longer travel duration query needs. Meanwhile, larger number of generated travel routes can provide more diverse personal recommendations for tourists. Thus, the length and the quantity of generated TB sequential travel routes reflect the effectiveness and quality of the recommendations in this work. Furthermore, to test the scalability of our algorithm, we construct a synthetic data set consisting of 12,000 randomly generated travel behavior sequences. All of the experimental travel behavior sequences are randomly selected from the synthetic data set. Without any other notice, the following experimental parameters settings are the same as the validation experiment.

4.2.1. Data Size of Travel Behavior Sequences. To comprehend how the number of travel behavior sequences (data size) affects the TB sequential travel route generation, the data size is changed from 500 to 2,500 and the min_sup is set at 4. Figure 7 demonstrates the average length and the longest length of the generated routes under different data sizes. As the data size is getting bigger, the length of TB sequential travel routes is getting longer; that is, the quality of routes is getting better. Therefore, the more historical onsite

TABLE 8: Top 3 candidate tangible travel routes generated for two tourists.

Tourist	Candidate tangible travel routes	Total route duration	Total visit duration	Total exhibits	Total NP _i
#1	<E0,L1,->,I,<B1,L4,4>,I,<B2,L5,5>,0,<B3,L4,3>,I,<B6,L4,2>,2,<B14,L5,4>,I,<B15,L5,6>,0,<B16,L5,6>,I,<B17,L4,5>,I,<E1,L1,->	43min	35min	8	L38
	<E0,L1,->,I,<B2,L5,5>,I,<B3,L4,3>,I,<B12,L4,2>,2,<B14,L5,4>,I,<B15,L5,6>,I,<B17,L4,5>,I,<E1,L1,->	40min	31min	7	L34
	<E0,L1,->,I,<B1,L4,4>,I,<B4,L4,4>,0,<B5,L3,2>,3,<B14,L5,4>,I,<B15,L5,6>,0,<B16,L5,6>,I,<B20,L3,2>,I,<E1,L1,->	36min	28min	7	L31
#2	<E0,L1,->,I,<B4,L5,4>,0,<B5,L4,3>,I,<B6,L5,4>,0,<B7,L5,5>,0,<B12,L5,5>,0,<B14,L4,3>,0,<B17,L4,2>,I,<E1,L1,->	28min	25min	7	L34
	<E0,L1,->,I,<B4,L5,4>,0,<B6,L5,4>,0,<B7,L5,5>,0,<B12,L5,4>,0,<B14,L4,3>,I,<B15,L4,2>,2,<E1,L1,->	26min	22min	6	L30
	<E0,L1,->,I,<B4,L5,4>,0,<B6,L5,4>,0,<B9,L5,6>,I,<B8,L4,3>,I,<B17,L4,2>,I,<E1,L1,->	23min	19min	5	L25

TABLE 9: The top 10 favorite posters of two types of volunteers.

Young Female		Young Male	
Poster No.	Poster Topic	Poster No.	Poster Topic
B2	Scientific	B4	Scientific
B14	Daily life	B6	Sport
B1	Scientific	B7	Sport
B3	Scientific	B12	Sport
B15	Daily life	B9	Sport
B16	Daily life	B5	Scientific
B17	Daily life	B8	Sport
B19	Daily life	B14	Daily life
B12	Sport	B2	Scientific
B11	Sport	B17	Daily life

FIGURE 8: The length of TB sequential travel routes under different minimum supports.

travel behavior the system gets, the higher quality of the recommendation can be generated.

4.2.2. Minimum Support of TB PrefixSpan. To discover how the minimum support parameter min_sup of TB PrefixSpan affects the quality of the TB sequential travel route, the different min_sup varying from 0.02% to 0.1% are tested with the synthetic data set.

Figure 8 presents the average length and the longest length of the generated routes under different min_sup . As the minimum support increases, the length of generated routes decreases. If the min_sup is set at 0.02%, the average length of routes is 5.14 and the longest route contains 11 visit spots. However, if the min_sup is set at 0.1%, the average length of routes declines to 3.98, and the longest route only contains 6 visit spots.

Figure 9 shows the execution time of the TB PrefixSpan algorithm under different minimum supports. As the min_sup increases from 0.02% to 0.1%, the execution time of the TB sequential travel route generation process declines from 950.59s to 200.5s. Based on the observation from Figures 8 and 9, the min_sup is suggested as 0.04% or below to ensure the quality of the routes. As the algorithm is performed

FIGURE 9: The execution time of the TB PrefixSpan algorithm under different minimum supports.

FIGURE 10: The length of TB sequential travel routes with different T_d .

in an offline stage on the system server side, the running time of the algorithm will not affect the reaction speed of the online recommendation process.

4.2.3. Information Granularity of Tourist-Behavior Pattern. According to Definition 2, each TB pattern includes a location identity, a normalized popularity value, and visit duration. Thus, the information granularity of a TB pattern, which is affected by the metric of the discrete time and the popularity normalization coefficient, consequently decides the quality of the generated travel routes. To observe the relationship between the information granularity and the route quality, that is, the average length and longest length of the generated TB sequential travel routes, the following experiments are conducted.

Regarding the metric of the discrete time T_d , a series of values ranging from 5 min to 20 min are tested by 2,000 sequences randomly selected from the data set. With different T_d settings, the average length and the longest length of the generated routes are shown in Figure 10, and the number of generated routes is shown in Figure 11. When the metric of T_d is increasing, both the length and the number of generated routes are increasing as well. The reason is that if T_d is

FIGURE 11: The number of TB sequential travel routes with different T_d .

FIGURE 12: The length of TB sequential travel routes with different normalization coefficients.

getting bigger, more TB patterns will be recognized as a same frequent TB pattern and lead to generating longer and bigger count of routes. When T_d increases, however, the time precision of the candidate routes is declined. For example, assume that the visit duration at a spot is 21 min. If T_d is set as 5 min, then the discrete time is integer 5; the time error is 4 min at the route recommendation phase. If T_d is set as 20 min, then the discrete time is integer 2; the time error increases to 19 min. Further, the accumulating time error of the candidate travel route is unacceptable under an improper T_d value. Therefore, to balance the relation between the quality of TB sequential patterns and the time error of the travel route, T_d is suggested at 10 min in this work.

Regarding the popularity normalization coefficient, we test the coefficient ranging from 4 to 10 segments with 2,000 randomly selected sequences. Figures 12 and 13 illustrate the quality of the generated travel routes with different normalization coefficients. As shown in Figures 12 and 13, when the coefficient increases, the length and the number of generated routes both decrease. This is because that the larger coefficient is set, the more TB patterns can be generated in a travel behavior sequence. That is, more normalized popularity values will decrease the support count of the corresponding TB pattern. However, more normalized popularity values

FIGURE 13: The number of generated travel routes with different normalization coefficients.

can describe a tourist's preference more precisely, which can enhance the personalization of the recommendation. Therefore, to make a proper balance between the quality and the personalization of the generated travel routes, setting the normalization coefficient as 5 is appropriate for this work. Finally, the observations of Figures 11 and 13 demonstrate that the proposed algorithm is effective in generating diverse travel routes.

5. Conclusions

The main goal of our work is to design a travel route recommendation system that recommends personalized tangible travel routes for various tourists within a given POI. First of all, we designed a novel method based on smartphone and IoT infrastructure to collect onsite travel behaviors of tourists in a specific POI automatically. To learn tourists' preferences to each interesting object or spot, we developed an Android App to record multiple onsite travel behaviors on each spot including visit duration, taking pictures, and standing. Next, we designed a travel behavior sequence pre-processing method and a Tourist-Behavior sequential route mining algorithm to generate potential frequent tangible travel routes. Furthermore, the route ranking method uses the querying tourist's personal profile and route constraint to recommend personalized tangible travel routes. Finally, experimental results demonstrate that the proposed system is efficient and effective in recommending tangible travel routes based on collected onsite travel behavior data.

Our future works include (1) deploying our system in a real-world POI, (2) using more types of smartphone sensors to gather more types of onsite travel behaviors to learn tourists' preference precisely, (3) harnessing real-time congestion information at each spot of a scenic area to generate more reasonable travel routes and further improve tourists' travel experience, and (4) using the current location and historical travel sequence of the querying tourist to generate a real-time route recommendation when the tourist requests route recommendations at an arbitrary location within a POI for improving the flexibility of our system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (nos. U1501252, 61572146, and U1711263), the Natural Science Foundation of Guangxi Province (nos. 2016GXNSFDA380006 and AC16380122), the Guangxi Innovation-Driven Development Project (no. AA17202024), and the Guangxi Universities Young and Middle-aged Teacher Basic Ability Enhancement Project (no. 2018KY0203).

References

- [1] R. Baraglia, C. I. Muntean, F. M. Nardini, and F. Silvestri, "LearnNext: learning to predict tourists movements," in *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management*, pp. 751–756, 2013.
- [2] D. Lian, V. W. Zheng, and X. Xie, "Collaborative filtering meets next check-in location prediction," in *Proceedings of the 22nd International Conference on World Wide Web*, pp. 231–232, 2013.
- [3] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: a recurrent model with spatial and temporal contexts," in *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pp. 194–200, 2016.
- [4] Y. Su, X. Li, W. Tang, J. Xiang, and Y. He, "Next check-in location prediction via footprints and friendship on location-based social networks," in *Proceedings of the 19th IEEE International Conference on Mobile Data Management*, pp. 251–256, 2018.
- [5] D. Massimo, M. Elahi, and F. Ricci, "Learning user preferences by observing user-items interactions in an IoT augmented space," in *Proceedings of the 25th ACM International Conference on User Modeling, Adaptation, and Personalization*, pp. 35–40, 2017.
- [6] S. H. Hashemi and J. Kamps, "Exploiting behavioral user models for point of interest recommendation in smart museums," *New Review of Hypermedia and Multimedia*, vol. 24, no. 3, pp. 228–261, 2018.
- [7] S. H. Hashemi and J. Kamps, "Where to go next? exploiting behavioral user models in smart environments," in *Proceedings of the 25th ACM International Conference on User Modeling, Adaptation, and Personalization*, pp. 50–58, 2017.
- [8] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy, "Rank-geoFM: a ranking based geographical factorization method for point of interest recommendation," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 433–442, 2015.
- [9] J. Wang, Y. Feng, E. Naghizade, L. Rashidi, K. H. Lim, and K. Lee, "Happiness is a choice: sentiment and activity-aware location recommendation," in *Proceedings of the Companion of the The Web Conference*, pp. 1401–1405, Lyon, France, 2018.
- [10] L. Yao, Q. Z. Sheng, Y. Qin, X. Wang, A. Shemshadi, and Q. He, "Context-aware point-of-interest recommendation using Tensor Factorization with social regularization," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1007–1010, 2015.
- [11] G. Cai, K. Lee, and I. Lee, "Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos," *Expert Systems with Applications*, vol. 94, pp. 32–40, 2018.
- [12] P. Bolzoni, S. Helmer, K. Wellenzohn, J. Gamper, and P. Andritsos, "Efficient itinerary planning with category constraints," in *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 203–212, 2014.
- [13] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized tour recommendation based on user interests and points of interest visit durations," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 1778–1784, 2015.
- [14] C. Bin, T. Gu, Y. Sun, L. Chang, W. Sun, and L. Sun, "Personalized POIs travel route recommendation system based on tourism big data," in *Proceedings of the Pacific Rim International Conferences on Artificial Intelligence (PRICAI)*, pp. 290–299, 2018.
- [15] C.-Y. Tsai and S.-H. Chung, "A personalized route recommendation service for theme parks using RFID information and tourist behavior," *Decision Support Systems*, vol. 52, no. 2, pp. 514–527, 2012.
- [16] W. Luo, H. Tan, L. Chen, and L. M. Ni, "Finding time period-based most frequent path in big trajectory data," in *Proceedings of the SIGMOD Conference*, pp. 713–724, 2013.
- [17] C. Tsai, J. J. Liou, C. Chen, and C. Hsiao, "Generating touring path suggestions using time-interval sequential pattern mining," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3593–3602, 2012.
- [18] J. Pei, J. Han, B. Mortazavi-Asl et al., "PrefixSpan: mining sequential patterns efficiently by prefix-projected pattern growth," in *Proceedings of the 17th International Conference on Data Engineering*, pp. 215–224, 2001.
- [19] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, "Tour recommendation and trip planning using location-based social media: a survey," *Knowledge and Information Systems*, pp. 1–29, 2018.
- [20] C. Yun and M. Chen, "Mining mobile sequential patterns in a mobile commerce environment," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 37, no. 2, pp. 278–295, 2007.
- [21] Y. Chen and C. Shen, "Performance analysis of smartphone-sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095–3110, 2017.
- [22] iBeacon for Developers, 2019, <https://developer.apple.com/ibeacon/>.
- [23] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: a survey," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7370–7389, 2014.
- [24] T. Tsiligrirides, "Heuristic methods applied to orienteering," *Journal of the Operational Research Society*, vol. 35, no. 9, pp. 797–809, 1984.
- [25] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "A survey on algorithmic approaches for solving tourist trip design problems," *Journal of Heuristics*, vol. 20, no. 3, pp. 291–328, 2014.
- [26] D. Gavalas, V. Kasapakis, C. Konstantopoulos, G. Pantziou, and N. Vathis, "Scenic route planning for tourists," *Personal and Ubiquitous Computing*, vol. 21, no. 1, pp. 137–155, 2017.

- [27] C. Zhang, H. Liang, and K. Wang, "Trip recommendation meets real-world constraints: poi availability, diversity, and traveling time uncertainty," *ACM Transactions on Information and System Security*, vol. 35, no. 1, pp. 1–5, 2016.
- [28] C. Zhang, H. Liang, K. Wang, and J. Sun, "Personalized trip recommendation with POI availability and uncertain traveling time," in *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*, pp. 911–920, 2015.
- [29] D. Gavalas, V. Kasapakis, C. Konstantopoulos, G. E. Pantziou, N. Vathis, and C. D. Zaroliagis, "The eCOMPASS multimodal tourist tour planner," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7303–7316, 2015.
- [30] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, "Predictive trip planning-smart routing in smart cities," in *Proceedings of the 2014 Joint Workshops on International Conference on Extending Database Technology, EDBT 2014 and International Conference on Database Theory*, pp. 331–338, 2014.
- [31] C. Chen, D. Zhang, B. Guo, X. Ma, G. Pan, and Z. Wu, "TripPlanner: personalized trip planning leveraging heterogeneous crowdsourced digital footprints," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1259–1273, 2015.
- [32] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, "Dynamic route planning with real-time traffic predictions," *Information Systems*, vol. 64, pp. 258–265, 2017.
- [33] X. Wang, C. Leckie, J. Chan, K. H. Lim, and T. Vaithianathan, "Improving personalized trip recommendation by avoiding crowds," in *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*, pp. 25–34, 2016.
- [34] T. Aoiike, B. Ho, T. Hara, J. Ota, and Y. Kurata, "Utilising crowd information of tourist spots in an interactive tour recommender system," in *Information and Communication Technologies in Tourism*, pp. 27–39, Springer, 2019.
- [35] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, "Personalized itinerary recommendation with queuing time awareness," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 325–334, 2017.
- [36] Y. Zheng, Y. Chen, X. Xie, and W.-Y. Ma, "GeoLife2.0: a location-based social networking service," *Mobile Data Management*, pp. 357–358, 2009.
- [37] 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*, pp. 1082–1090, ACM, 2011.
- [38] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring temporal effects for location recommendation on location-based social networks," in *Proceedings of the 7th ACM Conference on Recommender Systems*, pp. 93–100, 2013.
- [39] B. Thomee, D. A. Shamma, G. Friedland et al., "YFCC100M: the new data in multimedia research," *Communications of the ACM*, vol. 59, no. 2, pp. 64–73, 2016.
- [40] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," *Knowledge and Information Systems*, vol. 54, no. 2, pp. 375–406, 2018.
- [41] A. Majid, L. Chen, H. T. Mirza, I. Hussain, and G. Chen, "A system for mining interesting tourist locations and travel sequences from public geo-tagged photos," *Data & Knowledge Engineering*, vol. 95, pp. 66–86, 2015.
- [42] T. Guo, B. Guo, J. Zhang, Z. Yu, and X. Zhou, "Crowdtravel: leveraging heterogeneous crowdsourced data for scenic spot profiling and recommendation," in *Proceedings of the Pacific Rim Conference on Multimedia*, pp. 617–628, 2016.
- [43] S. Jiang, X. Qian, T. Mei, and Y. Fu, "Personalized travel sequence recommendation on multi-source big social media," *IEEE Transactions on Big Data*, vol. 2, no. 1, pp. 43–56, 2016.
- [44] G. Hu, J. Shao, F. Shen, Z. Huang, and H. Tao Shen, "Unifying multi-source social media data for personalized travel route planning," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 893–896, 2017.
- [45] K. Ashton, "That internet of things thing," *RFID Journal*, vol. 22, no. 7, pp. 97–114, 2009.
- [46] A. Zanella, N. Bui, A. P. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, 2014.
- [47] S. H. Ahmed and S. Rani, "A hybrid approach, smart street use case and future aspects for internet of things in smart cities," *Future Generation Computer Systems*, vol. 79, pp. 941–951, 2018.
- [48] C.-Y. Tsai, M.-H. Li, and R. J. Kuo, "A shopping behavior prediction system: considering moving patterns and product characteristics," *Computers & Industrial Engineering*, vol. 106, pp. 192–204, 2017.
- [49] H. Tang, S. S. Liao, and S. X. Sun, "A prediction framework based on contextual data to support mobile personalized marketing," *Decision Support Systems*, vol. 56, pp. 234–246, 2013.
- [50] D. Cavada, M. Elahi, D. Massimo et al., "Tangible tourism with the internet of things," in *Information and Communication Technologies in Tourism*, pp. 349–361, Springer, Cham, Switzerland, 2018.
- [51] A. Kuusik, S. Roche, and F. Weis, "SMARTMUSEUM: cultural content recommendation system for mobile users," in *Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology*, pp. 477–482, IEEE, Seoul, South Korea, 2009.
- [52] S. Alletto, R. Cucchiara, G. Del Fiore et al., "An indoor location-aware system for an IoT-based smart museum," *IEEE Internet of Things Journal*, vol. 3, no. 2, pp. 244–253, 2016.
- [53] H. Guo, L. Chen, G. Chen, and M. Lv, "Smartphone-based activity recognition independent of device orientation and placement," *International Journal of Communication Systems*, vol. 29, no. 16, pp. 2403–2415, 2016.
- [54] C. Tsai and B. Lai, "A location-item-time sequential pattern mining algorithm for route recommendation," *Knowledge-Based Systems*, vol. 73, pp. 97–110, 2015.
- [55] C.-H. Lee, C.-R. Lin, and M.-S. Chen, "Sliding-window filtering: an efficient algorithm for incremental mining," in *Proceedings of the 2001 ACM CIKM: 10th International Conference on Information and Knowledge Management*, pp. 263–270, 2001.