PHR: A Personalized Hidden Route Recommendation System Based on Hidden Markov Model

Yundan Yang, Xiao Pan(✉), Xin Yao, Shuhai Wang, and Lihua Han

Shijiazhuang Tiedao University, Shijiazhuang 050043, Hebei, China
smallpx@163.com

Abstract. Route recommendation based on users' historical trajectories and behavior preferences is one of the important research problems. However, most of the existing work recommends a route based on the similarity among the routes in historical trajectories. As a result, hidden routes that also meet the users' requirements cannot be explored. To solve this problem, we developed a system PHR that can recommend hidden routes to users employing the Hidden Markov Model, where a route recommendation problem is transformed to a point-of-interested (POI) sequence prediction. The system can return the top-*k* results including both explicit and hidden routes considering the personalized category sequence, route length, POI popularity, and visiting probabilities. The real check-in data from Foursquare is employed in this demo. The research can be used for travel itinerary plan or routine trip plan.

Keywords: Hidden Markov Model · Route recommendation · Hidden routes · Trajectory big data

1 Introduction

With the increasing popularity of Location-based Social Network (LBSN), the dissemination and sharing of information becomes much more convenient. A large amount of user-generated content information including reviews, photos, check-in data, travel notes, GPS tracks, etc., has been accumulated. By analyzing these user-generated contents, users' preferences and behaviors can be obtained, making it possible to recommend personalized travel routes for users and better meet user needs and preferences. On the other hand, it is time-consuming and difficult to plan a qualified route since the travel information is numerous and jumbled. Most of the existing work recommends a route through comparing the similarities between the user's historical trajectories and other users' travel records. Thus, only explicit routes can be returned, while the new hidden routes, that meet user requirements, cannot be explored.

For example, Alice issues a query for requesting a path from the airport to her hotel. She also wants to pass by an art gallery, a café, and a bookstore. The route is expected to be short and with high popularity. Figure 1 shows two historical routes, which is in a blue dash line and a green solid line respectively (i.e., R_1 and R_2). We use a round

X. Wang et al. (Eds.): APWeb-WAIM 2020, LNCS 12318, pp. 535–539, 2020.
https://doi.org/10.1007/978-3-030-60290-1_42

rectangle to represent a POI whose category is shown in the rectangle. We observe that R_1 covers an art gallery and a café except a bookstore, and R_2 goes through both a café and a bookstore except an art gallery. None of the existing historical routes contains all the required POIs. However, a hidden route (*i.e.*, R' in a red line), that is combined from part of R_1 and part of R_2, can satisfy all the requirements.

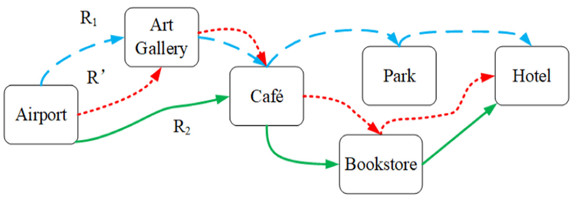

Fig. 1. Illustration for hidden routes (Color figure online)

Users' historical travelling routes can be regarded as POIs visiting sequences. We organized historical travelling routes by a Hidden Markov Model. The POIs are the unobservable hidden states, and the POI categories are observable states. Then, a route recommendation problem is transformed to a POIs sequence prediction. When a user issues a query with (origin, destination, a category sequence), an improved Viterbi algorithm is used to generate multiple POI sequence candidates. Next, the sequence candidates are ranked by the balanced scores, which consider the route length, POI popularity and the visiting probability. Finally, the top-*k* routes are displayed and returned to users. With the help of the Hidden Markov Model, both explicit and hidden routes can be found out.

2 PHR Overview

Figure 2 gives an overview of PHR. PHR is mainly composed of four modes: (1) data preprocessing; (2) Hidden Markov Model; (3) route generation; (4) route ranking and visualization. The first two modes operate offline and the last two modes is running online.

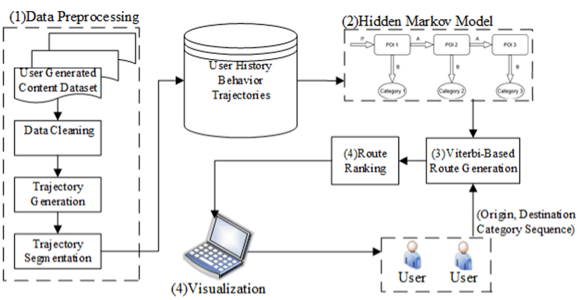

Fig. 2. PHR overview

2.1 Offline Operation Modes

Data Preprocessing. We clear the dirty data by deleting the duplicate records and the records with null values (i.e., without locations or timestamps). After data cleaning, a trajectory is generated for each user by sorting the points timestamps. Then, the trajectories are partitioned by a fixed time interval, e.g., a day, a month or a year.

Modeling. Hidden Markov Model consists of two state sets (i.e., observation state set and hidden state set) and three probability matrices λ (initial hidden state probability matrix, hidden state transition probability matrix and emission matrix). The system sets the POI categories as observation states and the POIs as the hidden states. The three probability matrices are calculated using the users' historical routes. Readers can refers to the specific computation methods in our paper [1].

2.2 Online Operation Modes

Route Generation. When a user issues a query in the form (origin, destination, a category sequence), the query is sent to the route generation mode. Given a category sequence, the route generation mode aims to find an existing or hidden POI sequences. This problem is similar to the existing sequence prediction on Hidden Markov Model [2], where Viterbi algorithm is the well-known method. Thus, we employ the improved Viterbi algorithm to generate several candidate POI sequences.

Route Ranking. The candidate POI sequences, generated in the route generation mode, have not consider the locations of the origin and the destination and users' preferences. As a result, the returned routes could be far from the origin and the destination, or the POIs in the returned routes are not popular. Thus, we fuse the distances between the first point (the last point) in the candidate route and the origin (destination), POI popularity, and the access probability of the route into a balanced score. Readers can refer to paper [1] for more computation detail as well. The candidate routes are ranked by the balance scores and the top-k routes are returned to users.

Visualization. The top-k (i.e., $k = 3$) routes are visualized and returned to users as shown in Fig. 3. User can choose the checkbox to display the route on the map. Different colored solid lines are used to indicate differently ranked routes. For instance, the blue line indicates the highest rank. Each point on the route has a label. The labels with {first, second, third} indicate the visiting order. The straight line between two consecutive points indicates the direction, which is not the navigation path.

3 Demonstration Outline

PHR is implemented using Java, SQL Server 2017 and Baidu map API. The real check-in data from New York and Los Angeles captured on the Foursquare website [3, 4] were used as experimental datasets. We generated 26,748 trajectories including 50,143 points. We use the real historical trajectories to measure the accuracy of the result, the recommendation routes' accuracy can reach more than 90% when $k = 3$ [1].

In particular, Fig. 3 shows the hidden route recommendation. Users can click the "Category Sequence" textbox to choose categories, and click the venue in the map (blue heart label) to choose origin(destination). After clicking "Submit", the top-1 route is displayed on the map by default. The top-3 routes are displayed below the map, and users can click the checkbox to show the route in the map. In addition, the route lengths are also displayed.

Besides hidden routes recommendation, PHR also supports "shortest route recommendation", "route recommendation based on Markov model" [5] and "Comparison and Statistics". This system can be extended to be used in travel itinerary plan or routine trip plan.

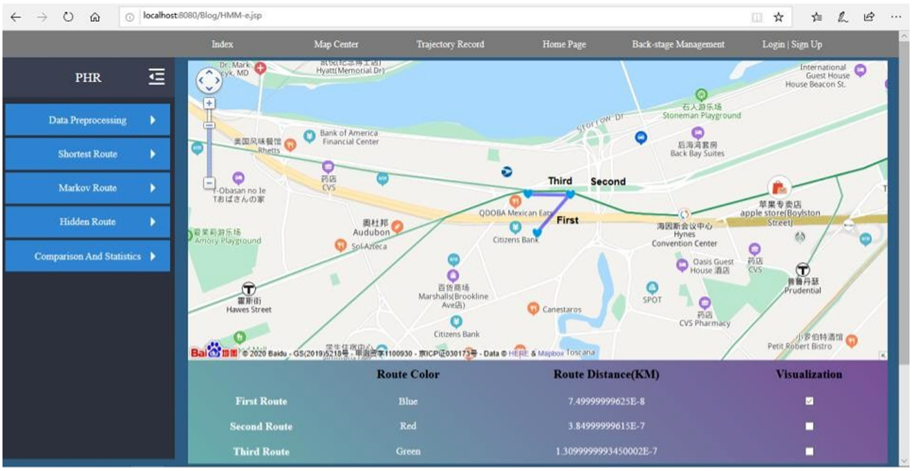

Fig. 3. Demonstration for hidden route recommendation (Color figure online)

Acknowledgment. This research was partially supported by the Natural Science Foundation of Hebei Province (F2018210109), the Key projects from the Hebei Education Department (No. ZD2018040), the Foundation of Introduction of Oversea Scholar (C201822), the Basic Research Team Project from Science and Technology Department (2019JT70803), the Fourth Outstanding Youth Foundation of Shijiazhuang Tiedao University.

References

1. Pan, X., Yang, Y.D., Yao, X., et al.: Personalized hidden route recommendation based on Hidden Markov Model. J. Zhejiang Univ. (Eng. Sci.) **54**(9), 1736–1745 (2020). (in Chinese)
2. Qiao, S.J., Shen, D.Y., et al.: A self-adaptive parameter selection trajectory prediction approach via Hidden Markov Models. IEEE Trans. Intell. Transp. Syst. **16**(1), 284–296 (2015)

3. Jie, B., Yu, Z., et al.: Location-based and preference-aware recommendation using sparse geo-social networking data. In: Proceedings of ACM GIS, pp. 199–208 (2012)
4. Wei, L.Y., Zheng, Y., Peng, W.: Constructing popular routes from uncertain trajectories. In: Proceedings of SIGKDD Annual Conference of ACM, pp. 195–203 (2012)
5. Chen, D.W., Ong, C.S., et al.: Learning points and routes to recommend trajectories. In: Author Proof Proceedings of CIKM, pp. 2227–2232 (2016)