Physica A 515 (2019) 258–269

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Trip destination prediction based on multi-day GPS data

Fang Zong [a], Yongda Tian [a], Yanan He [a], Jinjun Tang [b,*], Jianyu Lv [a,*]

[a] College of Transportation, Jilin University, Changchun, 130012, China

[b] School of Traffic and Transportation Engineering, Smart Transport Key Laboratory of Hunan Province, Central South University, Changsha, 410075, China

HIGHLIGHTS

- A model for destination prediction is constructed by using multi-day GPS data.
- The habit destination choice is studied in the pre-trip destination prediction model.
- The during-trip destination prediction model is developed based on Hidden Markov model.
- The results indicate that the model improves the forecasting precision.
- The model can be applied to trip distribution prediction and crowded location analysis.

ARTICLE INFO

Article history:
Received 26 May 2018
Received in revised form 31 August 2018
Available online xxxx

Keywords:
Trip destination
GPS data
Habit
Support point
Hidden Markov model
Markov chain

ABSTRACT

This study presents a model system for trip destination prediction with multi-day GPS data. The pre-trip destination prediction model and the during-trip destination prediction model are constructed and calibrated for trips on weekdays and weekends, respectively. Combining Markov chain and Multinomial logit model, the habit of multi-day destination choice is learned and utilized in developing the pre-trip model. By introducing support points, the during-trip model is developed with the Hidden Markov model. The estimation results indicate that this study improves the precision for destination prediction by calibrating the models for trips on weekdays and weekends separately and considering habit of multi-day destination choice. The findings suggest that travelers destination choice behavior follows not only the continuity between adjacent destinations, but also the inertia at the same time of day in consecutive days, and consecutive weeks as well. The quantitative comparison of the habit-based factors indicate that the inertia between adjacent destinations has the greatest effect on weekends' destination choice, and that in consecutive days has the most significant impact on weekdays' destination choice. In addition to real-time travel navigation based on the during-trip destination forecasting, the model can be applied to Advanced Traveler Information System which provides travelers with pre-trip information, such as traffic condition and commercial facilities around the destination. Through pre-trip destination prediction, study findings can also be utilized in trip distribution prediction or crowded location analysis, which presents a wide application in transportation planning and management.

* Corresponding authors.
E-mail addresses: jinjuntang@csu.edu.cn (J. Tang), lvjy16@mails.jlu.edu.cn (J. Lv).

https://doi.org/10.1016/j.physa.2018.09.090
0378-4371/

1. Introduction

Traveler's daily trip arrangement usually follows a pattern. That is, most of the regular trips are towards the known places, such as home, the workplace, a favorite cinema and a fashionable shopping center [1]. Besides, statistics of multi-day destination data indicate that there are also regulars of destination choice between different days and weeks. In this case, the daily and multi-day habit of destination choice can be studied and applied in destination prediction. Comparing to previous studies concerning destination prediction, the models concerning daily and multi-day habit need only traveler's previous destination choice recorders, which can be obtained directly from GPS device. Other information, such as travelers' socioeconomic data or geographic information, is not needed. Therefore, habit-based destination prediction is applicable to navigation system or other circumstance that traveler's socioeconomic data or geographic information is not available.

The destination prediction can be divided into pre-trip prediction and during-trip prediction, which provide travelers, planners or managers with destination prediction information before trips or during trips, respectively. The during-trip prediction is mainly applied to real-time travel navigation. In addition to route recommendation, the information about key facilities around a trip destination, such as the gas station, the shopping mall and the hospital, can also be recommended quickly to the traveler. In this way, destination prediction can not only advance the efficiency of route navigation but also enhance the attractiveness of facilities around destinations. Comparing to the during-trip destination prediction, the application area of the pre-trip prediction is much wider, which includes presenting pre-trip information through the Advanced Traveler Information System (ATIS), predicting trip distribution which is one of the major component in traffic demand forecasting, as well as being utilized in crowded location forecasting and analyzing, which can be applied to transportation management.

With the application of GPS technology, it is easy to record a traveler's multi-day movement as a sequence of time-stamped locations. Through analyzing the personal movement data collected by GPS, the movement patterns, i.e., the origin where a traveler usually departs from and the destination where a traveler usually goes to, can be discovered. The trip destination prediction model then can be developed by using the multi-day GPS data.

This paper tries to predict trip destinations by learning travelers' destination choice patterns based on multi-day GPS data. In order to predict trip destination for not only real-time travel navigation, but also pre-trip applications, the trip destination prediction models for both during-trip and pre-trip situations will be developed. Besides, the prediction models for trips on weekdays and weekends will be calibrated separately in order to enhance the forecasting precision, as statistical results show that the choice of trip destination on weekdays and weekends has different patterns.

The remainder of this paper is organized as follows. Section 2 gives a survey of the related work. Section 3 describes the process of GPS data preprocessing, and presents the basic characteristics of destination choice. In Section 4, the pre-trip destination prediction model is constructed by using the Markov chain and the Habit-based model. In Section 5, we describe a Hidden Markov model (HMM) to conduct the during-trip destination prediction. The paper is concluded with Section 6, in which we summarize our findings and discuss our study limitations and directions for future research.

2. Existing literature

To the authors' knowledge, Ashbrook and Starner's study [2] is the first one that predicted trip destination based on GPS data. They developed a Markov model to predict the next trip destination based on a set of previously visited candidate destinations. In order to enhance the predication accuracy, Ashbrook et al. [3] proposed a during-trip destination prediction model by introducing support points in the Hidden Markov model. The results indicate that it can improve the prediction accuracy by using support points in destination prediction. However, the study shows an unstable accuracy, which is from 71.81% to 94.57%. One of the reasons maybe that the distinct selection patterns of trip destinations on weekdays and weekends are not considered in this paper. Alvarez-Garcia et al. [1] developed a destination prediction system with the Hidden Markov model. It achieved total independence from a street-map data base by generating a "support-map" in which the main characteristics, such as support points and visiting frequency of the trips are taken into account.

Comparing to the studies using only GPS data, some researchers considered other information in destination predication, such as geographical data and travelers' socioeconomic information. For example, Huang et al. [4] modeled destination choice with Mixed-effects logit models. The effects of land use, road network structure, and axis of travel on home-based and non-work destination choice were investigated based on in-vehicle GPS travel data. Huang et al. [5] developed a framework that examined two-destination choice in the context of home-based trip chains. They utilized the survival models to determine the selection probability of a destination by considering the information concerning land use, transportation network and travel time, etc. Krumm et al. [6] obtained a probabilistic map of destinations by using Bayesian inference. Lin et al. [7] developed a predictive hierarchical model to learn and infer a traveler's daily movements. Simmons et al. [8] predicted the next link road and the final destination by constructing a Hidden Markov model. In these studies, a process of map matching was needed in order to adapt the GPS points received from the receptor to the map database. Besides, Kitamura et al. [9] analyzed the impact of travelers' socioeconomic data, travel characteristic, and land use on travelers' destination choice. Sivakumar et al. [10] developed a conceptual and econometric framework of non-work trip destination choice that is comprehensive in its incorporation of spatial cognition, heterogeneity in preference behavior, and spatial interaction. Auld et al. [11] developed a Multinomial logit model to predict on-work destination choice, in which the choice set is constrained by what has already been planned in the travel schedule. These studies need the supplementary investigations related to