Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media

Traveling Path Recommendation Using Temporal Transit Patterns

Hsun-Ping Hsieh[+], Cheng-Te Li[#]

[+]Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan

[#]Research Center for Information Technology, Academia Sinica, Taipei, Taiwan.

d98944006@csie ntu.edu.tw, reliefli@gmail.com

Abstract

In this work, we aim to collectively recommend traveling paths by leveraging the check-in data through mining the moving behaviors of users. We call such traveling paths as *Temporal Transit Patterns* (TTP), which capture the representative traveling behaviors over consecutive locations, from the big check-in data. To achieve such goal, we propose a novel *Temporal Transit Pattern Mining* method (TTPM-method), which devises an unsupervised mechanism that automatically summarizes the representative travel patterns us guarantee better time efficiency lower and memory usage. Based on the mined TTP, a novel system, TTP-Rec, is then developed. TTP-Rec not only allows users to specify starting/end locations, but also provides the flexibility of the time constraint requirement (i.e., the expected duration of the trip). Considering a sequence of check-in points as a traveling path, we mine the frequent sequences with a ranking mechanism to achieve the goal. Our TTP-Rec targets at travelers who are unfamiliar to the objective area/city and have time limitation in the trip.

Introduction

Nowadays, location-based services (LBS), such as Foursquare[1] and Gowalla[2], keep track of personal geospatial journeys through check-in actions. With smart phones, users can easily perform check-in actions, and the geographical information of locations with timestamps is stored in LBS. Eventually a large-scaled user-generated location sequences (i.e., routes) data are derived. Such location sequence data can not only collectively represent the real-world human geo-activities, but also serve as a handy resource for constructing location-based recommendation systems. Since the user-moving records implicitly reveal how people travel around an area with rich spatial and temporal information, including longitude, latitude, and recording timestamp, one reasonable application leveraging such user-generated location sequence data is to recommend travel routes. While existing trip planning systems (Zheng et al. 2009) (Yuan et al. 2010) (Arase et al. 2010) consider either the shortest geodesic distance, frequent sequence or the shortest time period to plan routes, we believe it would be more useful if the *representative* routes with visiting time information can be recommended from the user check-in data.

Check-in records rapidly accumulate and update over time, so that an efficient and scalable algorithm is demanded to mine the useful travel patterns from the big check-in data. However, discovering travel patterns under efficiency and scalability concerns from large-scaled location data had not ever carefully tackled yet. In this paper, we propose to mine the Temporal Transit Patterns (TTP), which capture the representative traveling behaviors over consecutive locations, from the big check-in data. A TTP is a sequence of locations, in which each transit between locations is associated with a visiting timestamp. For example, an example TTP is <(MMOA museum, Macy's store, 0, 3), (Macy's store, H&M store, 3, 5), (H&M store, Time Square, 5, 6)>, which refers to that the recommended route consists of MMOA museum, Macy's store, H&M store and Time Square in order. And the pattern suggests that the staying time as well as the transportation time between MMOA museum and Macy's store is 3-0=3 hours, and so on. We believe finding such notable travel patterns can not only allow us perform route planning in an effective manner but also benefit transportation scheduling.

Our goal is to recommend suitable travel paths, which consists of sequences of check-in locations, for travelers with time requirement. We formulate a k-TP (i.e., top-k traveling paths) problem to achieve the proposed composer. The central idea is to represent a traveling path as a sequence of check-in points and to mine the top-k time-based frequent subsequences. Given (a) a database of user check-in records, (b) the starting and/or ending location points, and (c) the time constraint of this trip, our goal is to return a ranked list of top k frequent traveling paths satisfying the query requirements.

Existing work to discover traveling paths is limited. The approaches of mining GPS trajectories to provide route recommendation are discussed in (Yuan et al. 2010), (Zheng et al. 2009). Yuan et al. infer fastest routes from historical trajectories on the road network. Based on the users' interested activity, Zheng et al recommend the travel sequence from GPS trajectory database. Arase et al. mine

[1] https://foursquare.com/

[2] http://gowalla.com/

travelers' frequent trip patterns from photo database (Arase et al. 2010). To the best of our knowledge, we are the first to tackle the recommendation of frequent traveling paths with time constraints using location-based check-in data.

The Proposed Method

We collect a large-scaled check-in data from Gowalla, which contains 6,442,890 check-ins and 950,327 friendships from Feb. 2009 to Oct. 2010. We give the system overview of TTP-Rec in Figure 1.

Preprocessing. For each user, we transform his check-in records into sequences of locations. We also associate the time span between two consecutive locations by averaging the time duration between two check-in places for all people. A database of traveling paths is then built.

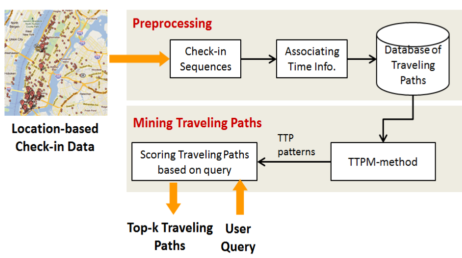

Figure 1. System Overview.

Temporal Transit Patterns Mining. We model the travel behaviors among locations into a Route Transit Graph (RTG), in which nodes represents locations, and edges denotes the transit behaviors of users between locations with certain time intervals. The time-aware transit patterns, which are required to satisfy frequent, closed, and connected requirements due to respectively physical meanings, are mined based on the RTG transaction database.

The goal is to mine frequent subsequence from the database of traveling paths. For a query containing some locations with a time constraint, our TTP-Rec system first retrieves the sequences with query locations. Then we devise a TTPM-method to find the frequent paths of locations. A Temporal Transit Pattern is defined as $<(S_1, D_1, t_{s1}, t_{e1}), (S_2, D_2, t_{s2}, t_{e2}), \ldots, (S_h, D_h, t_{sh}, t_{eh})>$, where t_{s1}=0, and all the edges in the pattern are sorted in increasing order. The temporal transit pattern should satisfy *route connected property*, which means a TTP $<(s_1, d_1, t_{s1}, t_{e1}), (s_2, d_2, t_{s2}, t_{e2}), \ldots, (s_h, d_h, t_{sh}, t_{eh})>$ should follows route connected constraint if $\forall\, s_i \in \{s_2, s_3, \ldots, s_h\}$ we can always find an edge (i, j) whose $i < j$, and $d_j = s_i$ or $s_j = s_1$. We measure the importance of a subsequence by calculating its *support* value, which is the number of sequence containing such subsequence. A subsequence P is *frequent* if its support is not less than *minsup*, where *minsup* is a user-specified minimum support threshold. In our system, we prefer to set *minsup* = 10% to ensure that the system contains enough traveling paths. The proposed algorithm consists of two stages. First, we mine all frequent patterns of length one (denotes 1-patterns), P in the database. Next, for each frequent k-pattern ($k \geq 1$) found P, we build the projected database for each frequent k-pattern found. And then we scan its projected database to find local frequent 1-patterns e which is connected with P. For each local frequent 1-pattern e, we concatenate P with e to form a frequent $(k+1)$-pattern. The concatenations are recursively performed in a depth-first search manner until no more frequent closed patterns can be found.

To return a ranked list of top-k traveling paths, we provide three ranking criteria. The first one is the support value, which determines the popularity of mined subsequences. The second one is the time constraint. If the total time of a mined traveling path is much shorter than time constraint, the remaining time could be wasted. We rank traveling paths based on the remaining time (the less, the better). The third one is the social effect. If a user has some check-in records in the past, we model user similarity according to the location type they visited in the history, and give higher rank for the routes generated from similar users. Such three ranking criteria can be combined to have better quality of traveling paths.

System Interface

The system interface of TTP-Rec is shown in Figure 2. Travelers are allowed to input initial and destination locations or choose popular attractions we suggest, as well as the expected travel time duration. The generation top-k traveling paths are shown in the right panel. Furthermore, TTP-Rec can also recommend the transportation mode between locations, which is developed by integrating with Google Map API.

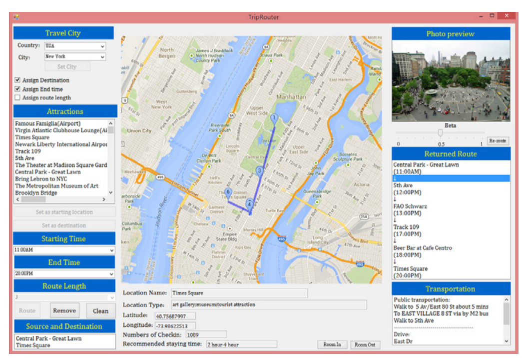

Figure 2. The interface of TTP-Rec system.

References

Arase,Y., Xie, X., Hara, T., and Nishio, S.. Mining people's trips from large scale geo-tagged photos. In ACM MM, pages 133-142, 2010.

Yuan, J., Zheng, Y., Zhang, C., Xie, W., Xie, X., Sun, G., and Huang, Y.. T-drive: driving directions based on taxi trajectories. In ACM GIS, pages 99-108, 2010.

Zheng, Y., Zhang, L., Xie X, and Ma, W.-Y. Mining interesting locations and travel sequences from GPS trajectories. In WWW, pages 791-800, 2009.