T-Drive: Enhancing Driving Directions with Taxi Drivers' Intelligence

Miss M.A. ShehnazBegum[1] ,Miss N.Nalini[2] ,Mrs.A.Angayarkanni[3]

[1,2]Student, Sri ManakulaVinayagar Engineering College, Puducherry.
[3] Assistant Professor, Sri ManakulaVinayagar Engineering College, Puducherry.

Abstract- GPS-equipped taxis uses the sensor device to detect the road surface, and taxi drivers are usually experienced with finding the fastest route to reach the destination based on their knowledge. This takes much time and rough routes are generated. In this paper, we process the smart driving directions from the historical dataset collected from large number of taxis. In our approach, we propose a time-dependent landmark graph, where the landmark is denoted as node is a road segment frequently traversed by taxis, to model the intelligence of taxi drivers. We have used a Variance-Entropy-Based Clustering approach to estimate the distribution of travel time between source and destination in different time slots. We also calculate the distance taken to travel the location. Based on this, we design a two-stage routing algorithm to compute the practically fastest route. We build our project based on a real-world trajectory dataset generated by over 33,000 taxis in a period of 3 months and evaluate the system by conducting both synthetic experiments and in-the-field evaluations. As a result, 60-70% of the routes suggested by our method are faster than the competing methods, and 20% of the routes share the same results. On average, 50% of our routes are at least 20% faster than the competing approaches.

Keywords:T-Drive,taxitrajectories,landmark graph

INTRODUCTION

FINDING efficient driving directions has become a daily activity and been implemented as a key feature in many map services like Google and Bing Maps. T-Drive concept have been implemented to give the efficient way for travelling to the taxi drivers. A fast driving route saves not only the time of a driver but also energy consumption. We can reduce our travel expenses with the help of this taxi trajectories . Therefore, this service is important for both end users and governments aiming to ease traffic problems and protect environment. Essentially, the time that a driver traverses a route depends on the following three aspects:
1) The physical feature of a route, such as distance, capacity (lanes), and the number of traffic lights as well as direction turns;
2) The time-dependent traffic flow on the route; and
3) A user's driving behavior.

Variance-Entropy-Based Clustering approach:

Variance-Entropy-Based Clustering approach is devised to estimate the distribution of travel time between two landmarks in different time slots. It will
calculate the values collected from this collected data. These clustering schemes enables taxi drivers to better understand and make improvements in their travel.
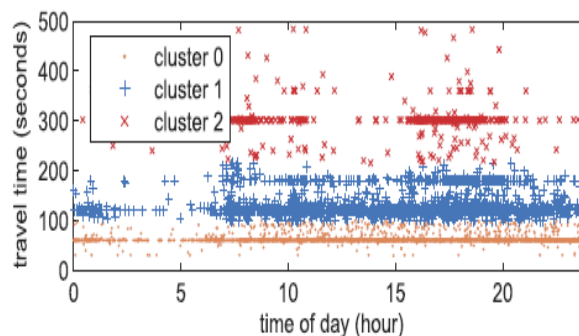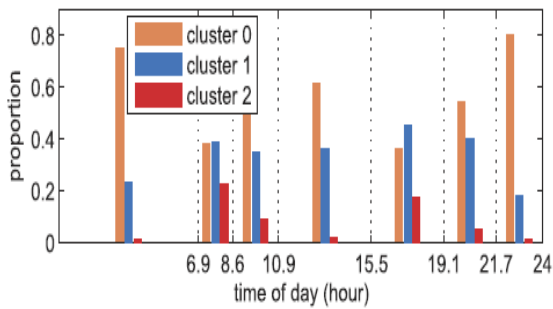

Fig.1. V-Clustering result

Fig.2.VE- Clustering result

VE-Clustering algorithm which is a two phase Clustering method, to find different time partition for different landmark edges based on the taxi trajectories. The trajectories defined as the path travelled by the taxis.

Driving Directions with Taxi Driver's Intelligence:

Driving to destination without the knowledge of the route takes much time of the drivers and also requires more energy consumption. In order to overcome these kinds of trouble we have distinguished the new way of travelling methodology for taxi drivers. Here, we have certain dataset values which are collected by 33,000 taxis in a period of 3 months.

With this data the driver can select the route along with the latitude and longitude values and the travel time is also estimated.When selecting driving directions, besides the distance of a route, they also consider other factors, such as the time-variant on road surfaces, tracks signals and direction changes contained in a route, as well as the probability of accidents. Route map will generate for the drivers to travel along the path. Route map is implemented in taxis using GPS device.The variance entropy based clustering approach is implemented to calculate the travelling distance,speed and time.These factors can be learned by experienced drivers but are too subtle and difficult to incorporate into existing routing

engines. Therefore, these historical taxi trajectories, which imply the intelligence of experienced drivers, provide us with a valuable resource to learn practically fast driving directions.
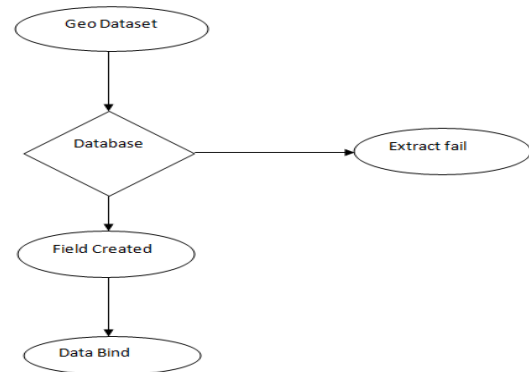
ROUTE COMPUTING

The route computing will display various paths in which the taxi drivers need to be travelled. The routes are displayed from the dataset which are already collected and stored in the database.

1. Road segmentation
2. Route Trajectories.
3. Land mark Graph.
4. Rough Route.
5. Evaluation of route time.

Road segmentation:

As a user can select any place as a source or destination, there would be no taxi trajectory exactly passing the query points. That is, we cannot answer user queries by directly mining trajectory patterns from the data. Therefore, how to model taxi drivers' intelligence that can answer a variety of queries is a challenge. We cannot guarantee there are sufficient taxis traversing on each road segment even if we have a large number of taxis. That is, we cannot accurately estimate the speed pattern of each road segment.

Route Trajectories

We build the improved system by using a real world trajectory data set generated by 33,000 taxis in a period of three months, and evaluate the system by conducting both synthetic experiments and in-the-field evaluations. The mutual influences of the GPS points in a trajectory to calculate static/dynamic score matrix for a trajectory and performs a voting-based approach among all the

candidates. As a result, each taxi trajectory is converted to a sequence of road segments.

Land mark Graph
Based on the preprocessed taxi trajectories, we detect the top-k frequently traversed road segments, which are termed as landmarks. The reason why we use "landmark" to model the taxi drivers' intelligence is that: first, the sparseness and low-sampling rate of the taxi trajectories do not support us to directly calculate the travel time for each road segment while we can estimate the traveling time between two landmarks (which have been frequently traversed by taxis).

Rough Route
Besides the traffic condition of a road, the travel time of a route also depends on drivers. Sometimes, different drivers take different amounts of time to traverse the same route at the same time slot. The reasons lie in a driver's driving habit, skills and familiarity of routes. For example, people familiar with a route can usually pass the route faster than a newcomer. Also, even on the same path, cautious people will likely drive relatively slower than those preferring to drive very fast and aggressively.

Evaluation of route time
We conduct two types of in the field studies:
1) The same driver traverses the routes suggested by our method and a baseline at different times.
2) Two drivers (with similar custom factors learned by our system) travel different routes (recommended by different methods) simultaneously. According to the results, 79.4 percent of the routes provided by our system are better than the baseline with respect to the travel time on average.
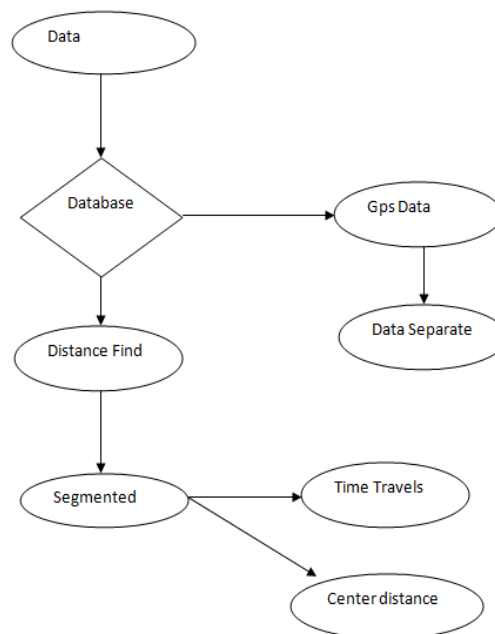
Estimation Algorithm:

Consider vehicle data. We specifically address arterial roads, which are also known as the secondary road network (highways are considered the primary road network). We consider several estimation problems: historical traffic patterns, real-time traffic conditions, and forecasting future traffic conditions. We assume that the data available for these estimation problems is a small set of sparsely traced vehicle trajectories, which represents a small fraction of the total vehicle flow through the network. We present an expectation maximization algorithm that simultaneously learns the likely paths taken by probe vehicles as well as the travel time distributions through the network.

- Intrinsic characteristics of this link (the length of the link, the number of lanes, the presence of traffic signals, etc).

- Traffic behavior characteristics such as the congestion on this particular link or on neighboring links.

- A probabilistic model of travel times through the arterial network is

- Presented along with an expectation maximization (EM) algorithm for learning the parameters of this model.

Time Dependence:

The reason why our method outperforms the RT approach is:

1) Coverage: Many road segments have neither embedded road sensors nor taxis traveling on them at a given time. At this moment, the speed constraint of a road segment is used to represent the real time track on the road segment. That is also the reason why the RT approach returns many of the same routes as the SC method.

2) Sparseness: in a time interval, the estimated travel time is still not very accurate if the number of the taxis is not large enough.

3) Open challenges: As compared to the history-based method, the RT approach is more vulnerable to noise, such as track lights, human factors (pedestrians crossing a street), and taxis looking for parking places and passengers.

It uses are,Time-dependent shortest-path problems have been studied in the case of discrete delay functions whose domain and range are the positive integers.

For the sake of management and security, these taxis have already been embedded with a GPS sensor, which enables a taxi to report on its present location to a data center in a certain frequency.

Landmark Graph Construction:

The threshold is used to eliminate the edges seldom traversed by taxis. We build two different landmark graphs for weekdays and weekends respectively. We present algorithms for finding the shortest path and minimum delay for all instances of time and under various waiting constraints and investigate properties of the derived path.

A) Matched taxi trajectories   B) Detected landmarks   C) A landmark graph
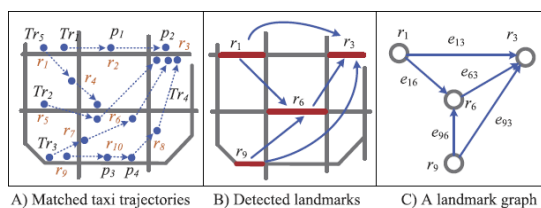
Low-Sampling-Rate:

Algorithm for processing all of the data previously collected into a form which accurately represents these trends. The landmark act to solve the rough routes collection. Pursues the local matching of a small portion of the trajectory. When matching a new position, its previous position and last matched edge are considered. Algorithms for finding the

shortest path and minimum delay under various waiting constraints are presented and the properties of the derived path are investigated. It is shown that if departure time from the source node is unrestricted, then a shortest path can be found that is simple and achieves a delay as short as the most unrestricted path.


Finding Trajectories:

The shorter a sub-path is the more unstable its travel time could be. In an extreme case, the travel time of a single road segment is terrifically impacted by traffic lights and pedestrians crossing it. So, the travel time varies in time quickly and tremendously, becoming hard to predict.Taxi drivers are provided with huge dataset of routes. In that routes, the driver should select the appropriate path he need to travel. Before travelling on the route he can estimate the travel time using this approach and the graph is provided to travel easily along the path. These trajectories are processed with sequence of road segment.
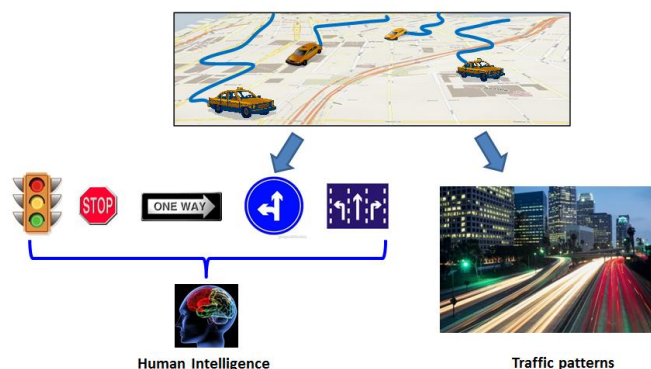

Taxi drivers will now identify which path can be travelled quickly by analyzing the various trajectories with certain devices.

Application:

Clustering approach will extract the data from the dataset collected on certain taxi's. The lantitude and longitude values are used to evaluate the travel time and distance to travel from source to destination. This application will display the current location to the taxi drivers and the travel time will also evaluate the traffic signals between source and destination. This will help the taxi drivers to estimate the correct travel time.The graph will be displayed on the evaluation.
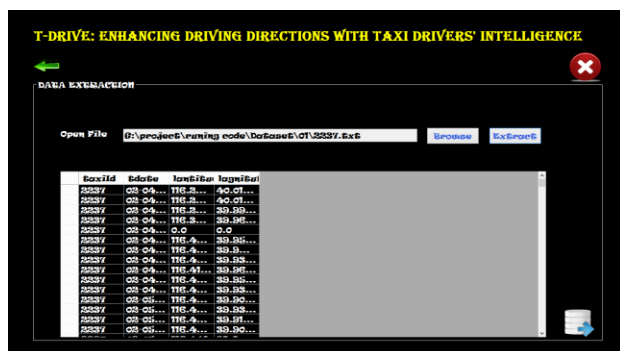
Application implementation using c#
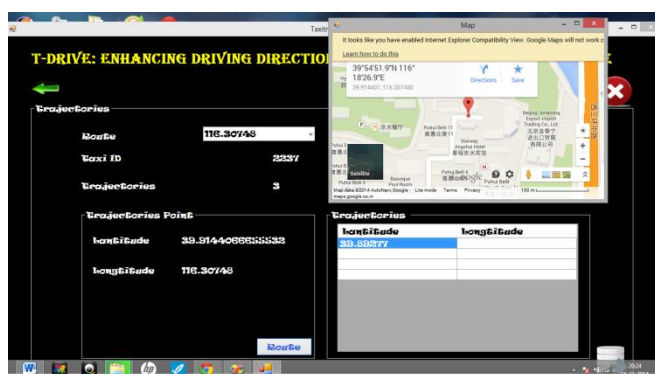
Application implementation using c#

Conclusion:

Specifically, the system mines the intelligence of experienced drivers from a large number of taxi trajectories and provide the end user with a smart route, which incorporates the physical feature of a route, the time-dependent traffic flow as well as the users' driving behaviors (of both the fleet drivers and of the end user for whom the route is being computed). We build a real system with real-world GPS trajectories generated by over 33,000 taxis in a period of three months, then evaluate the system with extensive experiments and in-the-field evaluations. The results show that our method significantly outperforms the competing methods in the aspects of effectiveness and efficiency in finding the practically fastest routes. Overall, more than 60 percent of our routes are faster than that of the existing online map services, and 50 percent of these routes are at least 20 percent faster than the latter. On average, our method can save about 16 percent of time for a trip, i.e., 5 minutes per 30-minutes driving.

References:

[1] J. Yuan, Y. Zheng, C. Zhang, W. Xie, G. Sun, H. Yan, and X. Xie, "T-Drive: Driving Directions Based on Taxi Trajectories," Proc. 18th SIGSPATIAL Int'l Conf. Advances in Geographic Information Systems (GIS), 2010.
[2] T. Hunter, R. Herring, P. Abbeel, and A. Bayen, "Path and Travel Time Inference from GPS Probe Vehicle Data," Proc. Neural Information Processing Systems (NIPS), 2009.
[3] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-Matching for Low-Sampling-Rate GPS Trajectories," Proc. Int'l Conf. Advances in Geographic Information Systems (GIS), 2009.
[4] J. Yuan, Y. Zheng, C. Zhang, and X. Xie, "An Interactive-Voting Based Map Matching Algorithm," Proc. Int'l Conf. Mobile Data Management (MDM), 2010.
[5] E. Kanoulas, Y. Du, T. Xia, and D. Zhang, "Finding Fastest Paths on a Road Network with Speed Patterns," Proc. Int'l Conf. Data Eng. (ICDE), 2006.
[6] A. Orda and R. Rom, "Shortest-Path and Minimum-Delay Algorithms in Networks with Time-Dependent Edge-Length," J. ACM, vol. 37, no.3,pp.607-625,1990.
[7]B. Liu, "Route Finding by Using Knowledge about the RoadNetwork," IEEE Trans. Systems, Man, and Cybernetics, Part A:Systems and Humans, vol. 27, no. 4, pp. 436-448, July 1997.
[8] Z. Chen, H.T. Shen, and X. Zhou, "Discovering Popular Routes from Trajectories," Proc. Int'l Conf. Data Eng. (ICDE), pp. 900-911, 2011.