Proposed Advance Taxi Recommender System Based On a Spatiotemporal Factor Analysis Model

Santosh Thakkar, Supriya Bhosale, Namrata Gawade , Prof. Sonia Mehta

Department of Computer Engineering, Alard College of Engineering and Management, Pune

ABSTRACT

Nowadays, Taxi industry is growing very rapidly .Everyone uses taxies for long or short distance journey instead of travelling by own car. In this paper, an online recommendation model is presented to help the driver to decide about the best stand to head in each moment, minimizing the waiting time. The proposed approach uses time series advance techniques to predict the solution of real-time problem faced by taxi driver every day. We combine some information with the live current network status to produce our output. Also we are implementing taxi management system based on GPS which has become an important tool for efficient taxi business. It can be used not only for the sake of management, but also to provide useful information for taxi drivers to earn more profit by mining the historical GPS trajectories. In this paper, we propose a taxi recommender system for next cruising location which could be a value added module of the management system.
Keywords: Location-based services, GPS, recommender systems, taxicabs, trajectories.

1. INTRODUCTION

Due to introduction of advance technology in market all the systems have become smart and advance. Every system will now be available on android market. And one of the current trends in application is taxi application. Now a day's people use mobile applications to book a taxi as per there requirements. Also due to the dramatic cost down of Global Positioning System (GPS) devices, taxi fleet management system based on GPS has become very popular for taxi companies. By using this system, a taxi company is able to keep track time-stamped of its taxi cabs. Furthermore, additional information, such as the status of a taxi, include waiting at a stand, cruising, occupied, off shift, can also be tracked. The purpose of this work is to recommend a good location for the taxi driver to cruise to such that he can earn more profit than cruise based on his own experience. Several factors shall be considered for guiding a taxi driver cruising to a more profitable location. First, distance between current location and the recommended location should be short to save time and energy. Secondly, when the taxi arrives at the recommended location, the waiting time for next passengers should also be short. Thirdly, if the taxi driver is able to pick up passengers at the recommended location, the fare for the trip is preferred to be large.

2. PREVIOUS WORK

2.1 Dispatching System
Taxi dispatching systems are attracting growing attention from researchers with the development of intelligent transportation systems and the popularization of GPS sensors. Most existing dispatching systems assign a task to taxi drivers based on nearest neighbour principle in terms of distance or time. Use the naive Bayesian classifier with developed error-based learning approach to infer the number of vacant taxis at a given time and location which can be used to enhance the dispatching system. Yamamoto et al. propose a fuzzy clustering and adaptive routing approach to improve dispatching system by assigning vacant taxis adaptively to the locations with high expectation of potential customers. Different from the centralized dispatching, our recommendation system provides suggestions to taxi drivers/passengers allowing them to make their own decisions. Typically, for a dispatching system, the customers need to book a taxi by telephone/internet in advance, and it is usually not free of charge. Most passengers hail a taxi along the road or stand where the taxis are available instead of booking a taxi. Besides, our method aims to maximize the profit for a particular taxi driver instead of balancing the income of all the taxi drivers, which is usually a goal of a dispatching system. In addition, our approach can be combined with a dispatching system so as to complement each other.

2.2 Location Recommendation for Taxi Drivers
A number of recent works provide recommendations for taxi drivers present a novel model to recommend a taxi driver with a sequence of pick-up points so as to maximize a taxi driver's profit. This work formulates the target problem by a mobile sequential recommendation (MSR) problem. Combined with taxi driving fraud detection method and some business insights such as tip distributions, introduce a taxi business intelligence system in study the passenger-finding

International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org
Volume 4, Issue 4, April 2015   **ISSN 2319 - 4847**

strategies (hunting/waiting) of taxi drivers in Hangzhou. In this work, L1-Norm SVM is used to select features for classifying the passenger-finding strategies in terms of performance. Our approach is different from the above methods in the following aspects: 1) We provide recommendations to both taxi drivers and passengers, which mobilizes them and reduces the disequilibrium of the demand and supply. 2) Instead of a grid/cell-based partition of the map, our recommendation is provided on road-segment level, which enables more accurate and meaningful understanding of the taxi drivers' behaviours as well as a more practical recommendation for both the taxi drivers and the passengers. 3) We focus on the off-peak hours to help the driver make the first step decision whenever and wherever they want to decide a destination to go. In practice, the "first step" recommendation would be more effective since usually the drivers are not willing to remember a sequence of places. 4) We develop an algorithm to distinguish the parked status from traffic jams and propose a solution to detect the parking places in an urban area. 5) We target the challenges when building the system based on sparse data and facilitate the on-line recommendation with a partition-and-group framework.
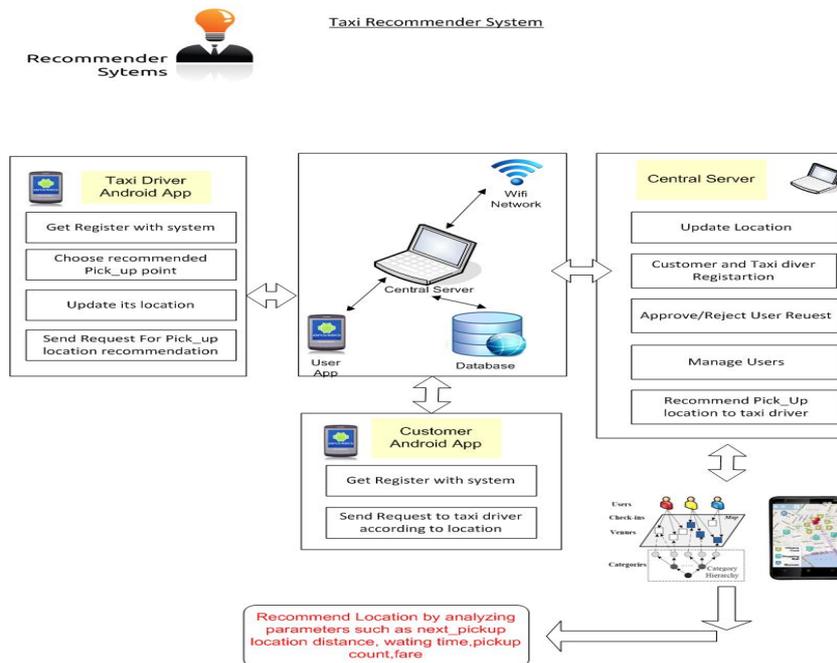
3.SYSTEM ARCHITECTURE

Fig 3.1 System Architecture

1. **Customer App**
Initially, Customer will send a request for Taxi using Android App. The request is then send to the server.
2. **Server**
 The request is processed by the server and check for the taxi driver which is near to the customer. It is forwarded to the nearer taxi driver using GPS (Global Positioning System).

3. **Driver App**
 Nearer Taxi Drivers get the pickup request from the server and the nearest driver which is available accepts the request. System architecture represents the flow of the software system. In this system, the customer initially sends the request for the taxi to the server. The Server acknowledges the location of the customer using Global Positioning System. It searches for the taxi driver available at that instant location using the taxi driver's current location. The server then accordingly forwards the customer request to the taxi drivers nearby the location. The taxi drivers nearer to that location get the customer request for the taxi. One of the Driver as per their location and convenience accept the request.

4.ALGORITHM USED

1] Grid _based Clustering Algorithm
The grid-based clustering approach differs from the innovational clustering algorithms in that it is concerned not with the data points but with the value space that surrounds the data points. In general, a typical grid-based clustering algorithm consists of the following five basic steps:
1. Creating the grid structure, i.e., partitioning the data space into a finite number of cells.
2. Calculating the cell density for each cell.
3. Sorting of the cells according to their densities.

International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org

Volume 4, Issue 4, April 2015   **ISSN 2319 - 4847**

4. Identifying cluster centers.

5. Traversal of neighbor cells.

2] Temporal and Spatial Analysis

The algorithm details are as follows:

1. Temporal Analysis

a. Divide historical GPS data into two categories (i.e., weekdays and weekends) so that the properties of taxi demand can be reasonably observed.

b. Divide a day into 24 time slots of equal length (i.e., the length of a time slot is an hour)

c.   Analysis of average waiting time in every place recorded in the historical GPS data.

2. Spatial Analysis

a. Data clustering using grid clustering is done on a set of

GPS points on the   historical GPS trajectories for dividing

the map into fixed square areas.

b. Four main factors are calculated which are average waiting time (denoted by T) in each cluster, the distance between two clusters (denoted by D), the average revenue from each cluster (denoted by R), the transition probability which captures the relation of the passenger get-off location and the next passenger get-on location (denoted by P)

i) The average waiting time of a place is obtained by dividing the total waiting time by the number of times that passengers get on a taxi in that place.

ii) Distance between two clusters is calculated from fixed size of each cluster

iii) Average revenue of each cluster is calculated by dividing the total revenue of each cluster by the number of occupied trips.

iv) The transition probability $p(L_{off} j L_{on})$ represents the probability that   passengers get on a taxi at cluster L_i and get off the taxi at cluster L_j ,is calculated as:

$p_{ij}(on\text{-}off) = Num(L_j\text{-}off\, jL_i\text{-}on)/Num(L_i\text{-}on)$

where $Num(L_i\text{-}on)$ is the number of occupied trips with passengers who get on a taxi at cluster L_i and $Num(L_j\text{-}off\, jL_i\text{-}on)$ is the number of occupied trips with passengers who get off a taxi at cluster L_j , after getting into the taxi at cluster L_i.

v) The transition probability $p(L_{on}jL_{off})$, which represents the probability of the passengers getting off a taxi at cluster L_i and the taxi driver going to pick up the next passenger at cluster L_j ,is calculated as:

$P_{ij}(off\text{-}on) = Num(L_j\text{-}onjL_i\text{-}off)/Num(L_i\text{-}off)$

where $Num(L_i\text{-}off)$ is the number of cruising trips with last passengers who get off a taxi at cluster L_i and $Num(L_j\text{-}onjL_i\text{-}off)$ is the number of occupied trips with passengers who get on a taxi at cluster L_j given that their previous cruising trips started at cluster L_i.

vi) Expected revenue can also be calculated as follows:

The expected revenue of cluster L_i for a taxi driver is given by:

$$r_i = \Sigma 8j\ p_{ij}(on\text{-}off)\_r_{ij}$$

 where r_{ij} is the average total revenue from cluster L_i to cluster L_j . r_i is obtained by multiplying the distance from cluster L_i to cluster L_j by the unit fare per mile.

 vii) Building of two location to location graph models on-off and off-on graph by observing the routes of taxi driver.

I. Build of an On-Off graph model as follows:

A weighted directed graph $G_{on\text{-}off}$ is used to present the moving trajectories of the occupied trips of all taxi drivers,

$$G_{on\text{-}off} = (V; E_{on\text{-}off}; W_{on\text{-}off})$$

 where V is a set of nodes that represents clusters, $E_{on\text{-}off}$ is a set of edges that represents occupied trips, and a weight set $W_{on\ off}$ of edges represents the number of transitions between clusters. The edge with a weight of three from point A to B represents that there were three occupied trips from cluster A to B in the historical taxi trajectories.

II. Build an Off-On L-L graph model, similar to the On-Off model, as follows.

A weighted directed graph $G_{off\ on}$ represents the moving trajectories of the cruising trips of all taxi drivers,

$$G_{off\text{-}on} = (V; E_{off\text{-}on}; W_{off\text{-}on})$$

where V is a set of nodes that represents the cluster set $E_{off\text{-}on}$, which is a set of edges that represents cruising trips, and the weight set $W_{of\text{-}on}$ of edges represents the number of transitions between clusters.

viii) Finding score(S_k) of each cluster :

According to the four factors described above, S_k represents the scores of cluster k, which finally finds the best candidate location to seek for the next passenger.

$S_k = W_R$ (4) $\_ GR_i(k) + W_D \_ GD_i(k) + W_T \_ GT_i(k) + W_P \_ GP_i(k)$.

5.SYSTEM RESULTS

Following are some screenshots of our system. Each step show how and where we apply above algorithm in our system.

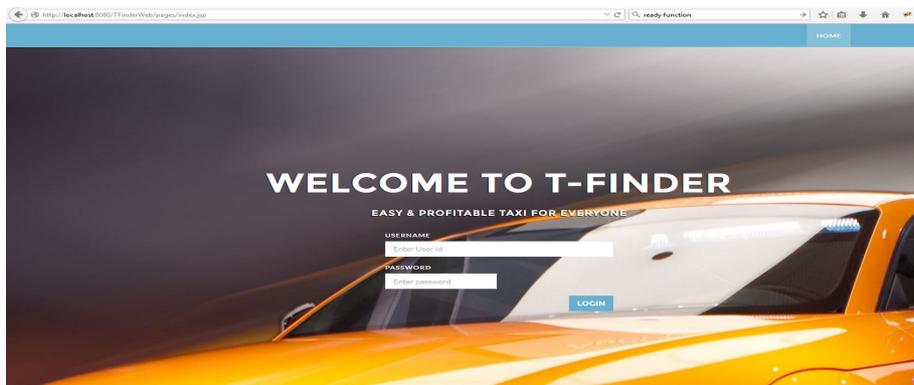**Step 1:** Here the first page on which user will login in with their credentials.

Fig 5.1 Login Page

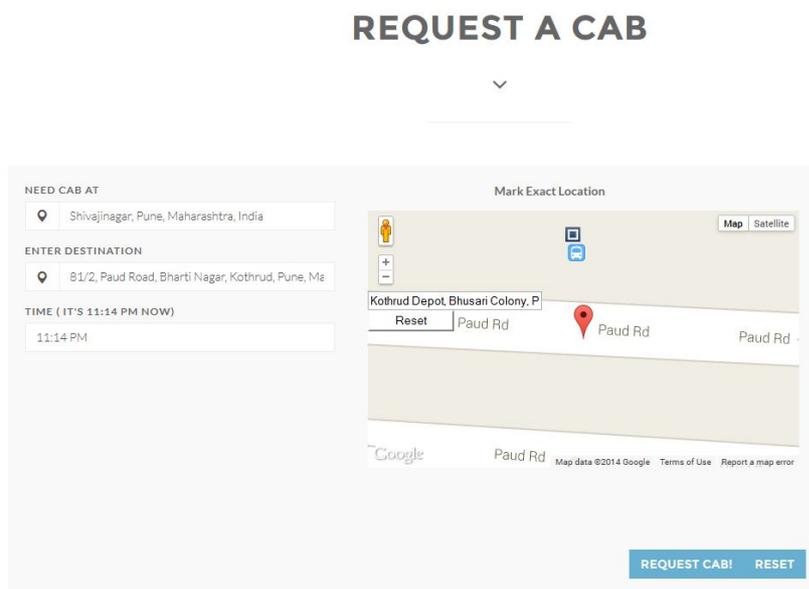**Step 2:** After successful login user will request for taxi as shown below.

Fig 5.2 Cab request page

Step 3: From this step it shows that how all the above mention algorithm are used in system.

Fig 5.3 Page showing cab request status

International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org
Volume 4, Issue 4, April 2015   **ISSN 2319 - 4847**

Step 4: Below screen shows the result of T-finder Cluster for average trips per day. By using this results taxi driver will decide which trip is more reliable for him. This result will help driver to make more profit as well as reduce waiting time.
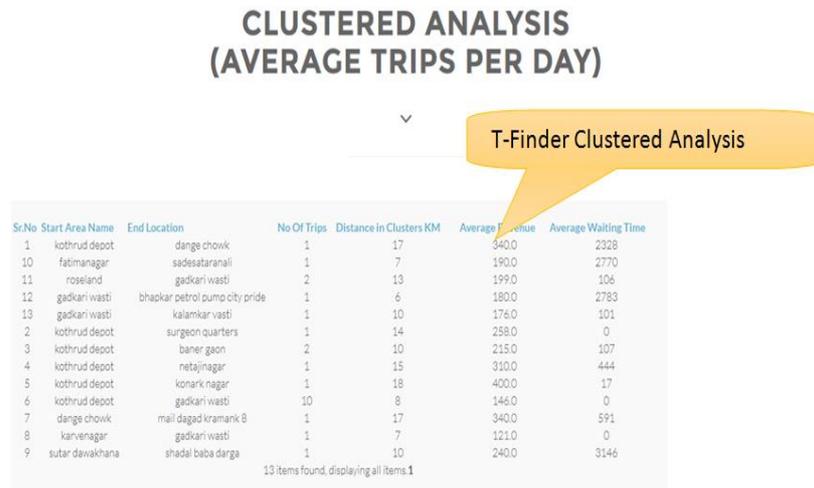

Fig 5.4 Page showing cluster analysis

Step 5: following screen shows the areawise priority cluster. This will help driver to decide which trip is nearer as well as more reliable for him.

Fig 5.5 Cluster analysis part 2

Step 6: following screen output show the revenue wise cluster.

These cluser are divided into three category ex. High, medium, low. So this will help driver to decide which trip is in which cluster gives him more revenue.
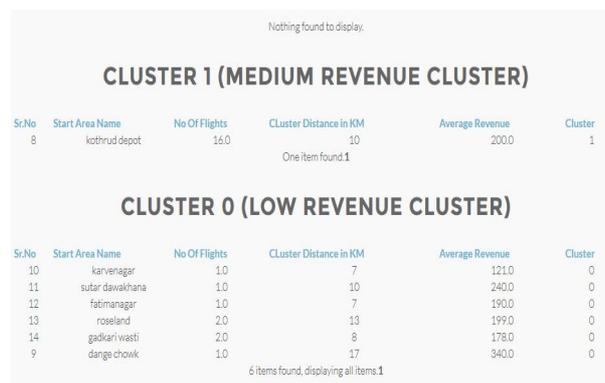

Fig 5.6 Cluster Results

6.CONCULSION

To save the time for finding a taxicab and reduce unnecessary traffic flows as well as energy consumptions caused by cruising taxicabs, we proposed a taxi recommender system based on the pick-up behaviors of high-profit taxi drivers and the mobility patterns of passengers learned from a large number of taxi trajectories. We investigate four factors on recommending taxi drivers the next cruising location. The four factors are obtained by analyzing the historical data trajectories according to spatio-temporal relation and location-to-location graph models. We evaluated the stability and reliability of our recommender system using real world data sets.

REFERENCES

[1] L. Moreira-Matias, R. Fernandes, J. Gama, M. Ferreira, J. o. Mendes-Moreira, and L. Damas, "An online recommendation system for the taxi stand choice problem (poster)." in VNC, 2012, pp. 173–180.

[2] An Effective Taxi Recommender System Based on a Spatiotemporal Factor Analysis Model Yu-Ling Hsueh, Ren-Hung Hwang, and Yu-Ting Chen Dept. of Computer Science & Information Engineering, National Chung Cheng University, Taiwan fhsueh@cs.ccu.edu.tw, rhhwang@cs.ccu.edu.tw, wendy10445@gmail.comg 2014 International Conference on Computing, Networking and Communications, Mobile Computing & Vehicle Communications Symposium.

[3] D. Zhongwei and J. Minhe, "Spatiotemporal structure of taxi services in Shanghai: Using exploratory spatial data analysis," presented at the 19th International Conference on Geoinformatics, 2011.

[4] N. Y. C. Taxi and L. Commission, "Taxi of Tomorrow Survey Results,"ed: New York, 2011.

[5] H. Chang, Y. Tai, and J. Hsu, "Context-aware taxi demand hotspots prediction," International Journal of Business Intelligence and Data Mining, vol. 5, pp. 3-18, 2010.

[6] W. Min and L. Wynter, "Real-time road traffic prediction with spatiotemporal correlations," Transportation Research Part C: Emerging Technologies, vol. 19, pp. 606-616, 2011.

[7] B. Li, D. Zhang, L. Sun, C. Chen, S. Li, G. Qi, and Q. Yang. Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset. In Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on, pages 63 –68, march 2011.

[8] T. Takayama, K. Matsumoto, A. Kumagai, N. Sato, and Y. Murata, "Waiting/cruising location recommendation based on mining on occupied taxi data," International Journal of Systems Applications, Engineering and Development, vol. 5, no. 2, pp. 224–236, 2011.