International Journal of Pure and Applied Mathematics
Volume 118 No. 24 2018
ISSN: 1314-3395 (on-line version)
url: http://www.acadpubl.eu/hub/
Special Issue
http://www.acadpubl.eu/hub/

Smart Transit System A boon for Smart Cities

S.Vignesh Babu , M.Varun
Undergraduate Scholars
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Thandalam, Chennai.
K.Devaki
Professor
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Thandalam , Chennai

May 26, 2018

Abstract

Smart transit system (STS) is a major area in Smart city initiative of the Indian Government. A report published by the Union Transport Ministry on March 22, 2016, complains that some of the route buses are overcrowded as it carries twice the number of passengers than its seating capacity, overruling the RTO rules and restrictions. Hence this project is an imminent requirement for todays fast paced environment. The aim of this project is to develop a mobile application for Smart Transit System using IoT and Data Analysis technologies to provide info to the passengers about the buss proximity, seat availability and arrival time estimate in advance and to optimize the number of buses plying through the routes for efficient fuel usage and human resources. The buss proximity is determined by Automatic Vehicle Location (AVL) device which uses inbuilt

GPS system of the smartphone or cell tower triangulation method to identify and upload the location information to the centralized server. The passenger request data for each bus is collected from the commuter interface and transmitted to the centralized database, which is further used in the analysis process, to obtain the final optimized number of buses that can be operated through that route. STS with such capabilities will be an integral part in the upcoming smart city development.

***Key Words*:** IoT; cloud computing; AVL; cell tower triangulation

1 INTRODUCTION

Cities are engines of growth for the economy of every nation, including India. With increasing urbanization, urban areas are expected to house 40% of Indias population and contribute 75% of Indias GDP by 2030 [14]. This requires the comprehensive development of physical, institutional, social and economic infrastructure. Development of Smart Cities is a step in that direction. Public transportation system is the lifeline of the residents. Therefore, over the years, a large number of initiatives have been undertaken by various State Governments and Central Ministries [14] to usher in an era of the smart transport system. The current metropolitan transport corporation(Chennai) has around 3866 buses and an approximate estimation of 48 lakh people per day travelling using this transport service. The government is seeking all possible measures to increase the profit obtained through public transport and also to cater the peoples need. The general public (i.e.) the commuters are oblivious about the buss info and has to wait for a long time uncertain of the bus status. For long, buses are not often tracked of their location. At the same time, the transport department has inadequate information for operating new buses in the most crowded route so as to maximize their profit as well as to satisfy the commuters need. Most of the time during the lean hours the bus goes empty and much of fuel and time is wasted. This model introduces a mobile application to provide clear information about the bus, to the commuters availing the transport system in advance so as to avoid inconvenience in their travel. Hence, a new method has been proposed using technologies like Raspberry Pi, AVL (Automatic Vehicle Location) using inbuilt GPS [4], cloud computing and data analytics, to provide optimization of the number of buses in operation, insights about the bus proximity for the commuters, decrease in fuel expenditure and eventually reduce the emission of harmful air pollutants and greenhouse gases from the vehicles.

2 LITERATURE SURVEY

Few works with varied functionalities have been proposed on smart transportation, some of the primary papers have been referred. In [1] this paper the author proposes a mobile application that deals mainly with the shortest route between the source and destination, provides the approximate time of travel for the same. It also provides the advanced reservation and seat availability details through short message service (SMS). The author [2] predicts the real time traffic using the data collected from the user mobile phones through the applications installed and compares it with the historical data collected from the ONE ITS service and applies an algorithm to analyse and predict the traffic in real time, and to disseminate the information about the traffic to their clients through the smartphone application.

3 METHODOLOGY

The architecture diagram of the STS using latest technologies like AVL (Automatic Vehicle Location) using inbuilt GPS, cloud computing and data analytics, Raspberry Pi is elaborated as such. The proposed system as shown in Fig 1 consists of the three Components. i) Bus proximity identification ii) Booking and AVL iii) Bus frequency optimization and data analysis interface component.

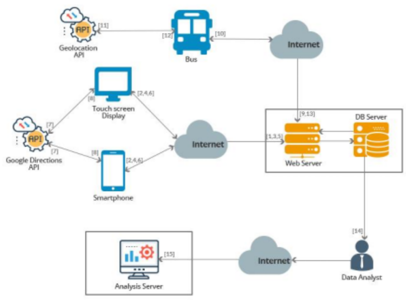

Fig 1 Architecture Diagram of STS

A. BUS PROXIMITY IDENTIFICATION:

This component can be accessed by web browser through a touch screen display supported by the Raspberry Pi device or through an app installed in the personal smartphone of the passengers. This component accepts the boarding and leaving points submitted by the user and sends it as a query to the cloud database server, and that server responds with the list of buses which has the matching stops of the boarding point and the leaving point in its route of navigation. The user can select any one the bus from the list displayed. Upon this request, the database server returns the info related to the selected bus which includes the buss proximity, numbe r of free seats and estimated arrival time of the bus to the boarding point. The number of free seats is calculated by the proposed algorithm discussed below. The buss proximity is displayed in the google map by passing the latitude and longitude values of the bus to the google directions API and the JSON file returned is parsed to obtain the proximity detail and the ETA (Estimated Arrival Time) for the bus.

ALGORITHM This system uses an algorithm as given below for calculating the seat availability in the bus. The pre-requisites for the algorithm are the boarding count and the get down count values which will be updated each time a ticket is booked in the conductor component and the total number of seat capacity of the bus value will be predefined in the database by then administrator. The proposed algorithm is as follows:

Algorithm: Algorithm for Vacant seat count identification **Input:** Boarding_stop, Destination_stop, Database D **Globals:** const Bs: Boarding_stop const Ls: Leaving_stop const TS: Total number of seats in the bus TBC: Total_boarding_count = NULL TLC: Total_leaving_count = NULL Bc: Boarding_count Lc: Leaving_count Output: Free_seats, available in the bus Method: (1) Fetch the Bc and Lc from the D (2) Calculate sum of Bc and Lc for (i=1; i¡Bs; i++) TBC = TBC + Bc[i]; TLC = TLC + Lc[i]; (3) To calculate the free seats Free_seats = TS TBC + TLC

The above algorithm can be generically applied in problems where dynamic change in the input parameter value competing for available resources occurs and calculation of the present available resource need to be calculated.

B. BOOKING AND AVL

The bus conductor can access this component through the app installed on their smartphone. The initial step is the login page where the conductor has to enter the credentials known to them. These credentials are sent as a query to the database server for validation. Upon successful validation, the booking page is displayed in the app. The conductor can enter the boarding point, leaving point and number of seats requested by the passenger and has to submit those details. The details are then stored in the database server accordingly. In the backend the automatic vehicle location database updating is done at regular intervals, the data is sent in the form of latitude and longitude of the vehicles current position. The latitude and longitude of the vehicle are found, either, by using inbuilt GPS of the smartphone device or by cell tower triangulation method using the Geolocation API of the World Wide Web ConsortiumRecommendation [4]. The data stored in the database which contains the details collected from this component is used forcalculations in the bus proximity identification component.

C. BUS FREQUENCY OPTIMIZATION AND DATA ANALYSIS INTERFACE One of the primary objective of the proposed system is to optimize the number of busses plying in a particular route [9]. It is better to know the optimum frequency of busses to prevent wastage of fuel and provide a hassle-free ride to the prospective commuters. The data analysis process consists of certain functionalities which enable the user to view and infer results and eventually utilize those results in the decision-making process. The inference is obtained by running predictive algorithms on the accumulated data. This optimization is done with the help of algorithms like Auto Regressive Integrated Moving Average and Random Forest Regression [8] [11]. The univariate ARIMA algorithm is used to predict the trend depicting the frequency of commuters boarding a particular bus route. The Random Forest algorithm on the other hand is used to estimate the number of busses to be operated in a particular route. Parameters such as time interval, date, route etc. are considered. These data are obtained by accessing the database which gets updates as and when a commuter access the system. The data being fed into the regressor along with the custom inputs provided by the admin the day, time interval and bus number enable to produce the exact optimised number of busses. On inferring this output, the admin can take necessary steps to add or reduce the number of busses for a route. The stochastic data is unpredictable and may change in different time frame, therefore Auto ARIMA is used in the application. This chooses the best fit of ARIMA version by accessing the parameters. The data is initially converted into time series data and the ARIMA function takes these data as parameters along with other parameters such as approximation and trace. The approximation is taken a default value FALSE so as to decrease the search time for the [13] best fit ARIMA model. Trace is used to display the considered list of ARIMA models, therefore it is set as FALSE. The value parameters of ARIMA, such as number of autoregressive terms (p), number of non-seasonal differences needed for stationarity (d), number of lagged forecast errors in the prediction equation (q) are automatically decided by the obtained best [8] AIC, AICc or BIC value. On application of this time series algorithm we get the trend for the commuters usage data. Ensemble method of regression helps in estimating a more accurate prediction by applying the decision tree regressor multiple times in the stochastic data which is gathered. The parameters such as day of week as X and count Y is fed into the regressor along with number of trees.

$$Fn = 1/B\Sigma_{b=1}^{B} f_b(x\prime) = y^{'} \qquad (1)$$

where $x^{\hat{'}}$ consist of the training data antecedent X- Day of the week (training data) and consequent Y count (training data), $Y^{\hat{'}}$count of searches (response),B-number of bags

The summation of the predicted values f_b($x^{\hat{'}}$) (for each bag) is divided by the total number of bags to get the average of all the outputs. In this way high accuracy in estimation of bus is obtained by applying the algorithm.

textbfFor Example:

The ticket trend for sample data bus number 17D represents the predicted ticket sales for the upcoming dates [8]. This is predicted with the help of ARIMA algorithm which handles the dataset. The univariate ARIMA algorithm based on previous dependent value to predict future independent value, allow to estimate the future trend with minimal error.

The optimization for sample data helps to estimate the number of busses flying in that route for a particular day and time interval [7]. The Random Forest regression model is applied to produce the output for the decision-making process. Based on this output the number of busses can be varied. In this way the predicted value of estimate bus count and trend of the demand is observed.

4 RESULTS AND DISCUSSION

DATASET

The dataset as shown in Fig 2 is accessed by the admin dashboard consisting of attributes such as bus number, date, day, time interval of request, accumulated count. These attributes are obtained from the user interface (i.e.) on click of the bus number, while searching, accessing the location and checking the seat availability. Each search is recorded as request at that time interval. The aggregated count along with other parameters are considered for prediction and optimization.

RESULTS

busnum	dates	day	segment	count
17D	1/1/2018	Monday	morn	141
17D	1/1/2018	Monday	noon	70
17D	1/1/2018	Monday	eve	185
17D	1/1/2018	Monday	night	100
17D	1/2/2018	Tuesday	morn	170
17D	1/2/2018	Tuesday	noon	20
17D	1/2/2018	Tuesday	eve	166
17D	1/2/2018	Tuesday	night	72
17D	1/3/2018	Wednesd	morn	171
17D	1/3/2018	Wednesd	noon	44
17D	1/3/2018	Wednesd	eve	138
17D	1/3/2018	Wednesd	night	97
17D	1/4/2018	Thursday	morn	145
17D	1/4/2018	Thursday	noon	27
17D	1/4/2018	Thursday	eve	197
17D	1/4/2018	Thursday	night	26
17D	1/5/2018	Friday	morn	210
17D	1/5/2018	Friday	noon	11
17D	1/5/2018	Friday	eve	201
17D	1/5/2018	Friday	night	31
17D	1/6/2018	Saturday	morn	122

Fig 2 test dataset sample

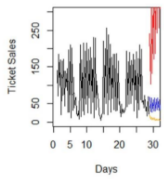

Fig 3 Prediction of trend for sample data

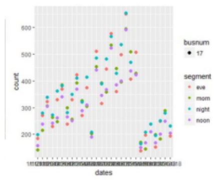

Fig 4 Graph showing the commuters trend for test period

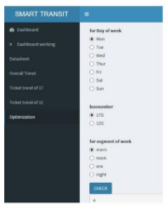

Fig 5 Dashboard for optimization of bus count

The predicted output for sample bus number 17D for the test dataset is represented in Fig 3. The black line represents the data considered and the blue line represents the trend projected. The orange line represents the maximum and minimum deviations.

The graph represented in Fig 4 gives the total requests accessed for all the busses. The time interval is represented by colours and the bus number to which each request pertains is represented by the size of the point. The x-axis represents the date of the requests and y-axis represents the total count of requests that has been accumulated for each bus.

The optimization of number of busses requires two data for control, namely day and interval as shown in Fig 5. The admin can also view the data which is being processed by browsing through the functionality provided.

5 CONCLUSION

We have developed a mobile application for Smart Transit System to help the commuters to take an informed choice about theirtravel and to optimize the number of buses using data mining algorithms, thereby optimizing the cost of operation, fuel consumption, pollution, traffic management and at the same time serving the commuters requirement. The experiments conducted for the sample data set with the proposed method produced good results and predicted the optimized number of buses. The prediction made foroptimizing the bus can be applied to a robust, scalable dataset.

References

[1] Nouf Mohammad Al Shammary and Abdul Khader Jilani Saudagar, Smart Transportation Application using Global Positioning System, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 6, 2015

[2] Hans-Arno Jacobsen, Kaiwen Zhang, Yingqi Yue, Smart Phone Application for Connected Vehicles and Smart Transportation, Middleware 2013 Posters and Demos Track, December 9-13, 2013, Beijing, China

[3] Bryan A. Weber, ARNP, PhD, Hossein Yarandi, PhD, Meredeth A. Rowe, RN, PhD, Justus P. Weber, A comparison study: paper-based versus web-based data collection and management, Applied Nursing Research 18 (2005) 182 185

[4] Geolocation API Specification 2nd Edition, W3C Recommendation 8 November 2016, https://www.w3.org/TR/geolocation-API/

[5] M. E. Mallia and K. Simpson, Wireless global positioning system fleet tracking system The university at Albany. Report No. C-11-12/ 14-27, 2014

[6] S. Motahari, H. Zang, S. Bali, and P. Reuther, Mobile applications tracking wireless user location, In Proceedings of IEEE Global Communications Conference (GLOBECOM), 37 Dec. 2012, Anaheim, CA, pp. 20062011

[7] Andy Liaw and Matthew Wiener, Classification and Regression by RandomForest, Vol. 2/3, December 2002 ResearchGate Publications

[8] Chenghao Liu1,2, Steven C. H. Hoi2, Peilin Zhao3, Jianling Sun, Online ARIMA Algorithms for Time Series Prediction, Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

[9] Albert Nagy, Jzsef Tick , Improving Transport Management with Big Data Analytics , SISY 2016 IEEE 14th International Symposium on Intelligent Systems and Informatics August 29-31, 2016, Subotica, Serbia

[10] E.Dhamo, LL.Puka, Using the R-package to forecast time series: ARIMA models and Application, December 2010 ResearchGate Publications

[11] BOX, G.E.P. and G.M. Jenkins (1970) Time series analysis: Forecasting and control, San Francisco: Holden-Day

[12] http://cran.r-project.org/web/views/TimeSeries.html

[13] http://ucanalytics.com/blogs/step-by-step-graphic-guide-to-forecasting-through-arima-modeling-in-r-manufacturingcase-study-example/

[14] https://blog.mygov.in/editorial/the-idea-behind-the-buzz-of-indian-smartcity/