



# GPS based Location Services and Analysis of Transit Data using Data Mining

Prof A.G. Dongre<sup>1</sup>, Ankita Kadale<sup>2</sup>, Snehal Khavle<sup>3</sup>, Ajinkya Dalvi<sup>4</sup>

Professor, Computer, PVG'S COET, Pune, India<sup>1</sup>

Student, Computer, PVG'S COET, Pune, India<sup>2,3,4</sup>

**Abstract:** To reduce the wait times and uncertainties of citizens travelling through public transportation system such as bus. To give better services by providing real time bus location as well as passenger count to the user and analysis of transit service performance using classification. In this fast life, everyone is in hurry to reach their destinations. In this case waiting for the buses is not reliable. People who rely on the public transport their major concern is to know the real time location of the bus for which they are waiting for and the time it will take to reach their bus stop. This information helps people in making better travelling decisions. This project provides the user with a user friendly GUI wherein he/she can get information about the real time location of the bus and also availability of bus for a particular route along with passenger count inside the bus. The real time location of the bus is tracked using GPS system of an android phone and the passenger count is obtained using IR Sensors which is fitted inside the bus at the front and rear doors. The data from this transit system can be analysed to discover patterns and useful information. Problem faced by the public transport organizations can be solved by using data mining technique. Service performance analysis based on historical data will help organizations to improve their functioning and efficiency. The analysis of gathered data, would help organizations to predict rush on particular routes and handle the same. This may also help redesign and improve the existing transportation policies.

**Keywords:** GPS; location tracking; real time location; android; smart phone; IR Sensor; classification Knn; data mining.

## I. INTRODUCTION

Effective transportation system has effective movement of goods and people which leads to better quality of life and better social and economic growth of the society. Transportation system forms the heart of the system. With the population boom vehicle population is also rapidly increasing which is further leading to heavy traffic. Optimal solution to this problem is the use of public transport. However public transport schedules are unreliable and waiting for bus for long results is waste of time. But a system that provides complete information namely the number of buses that go to the required stop, bus numbers, bus timings, time taken for the bus to reach, the routes through which the bus goes, maps that guides the passengers and most importantly tracking the real time bus location coordinates and finding correct time the bus will take to reach its bus stop.

However the bus service is not reliable. To come over these day to day problems, our system uses the Android platform with the built in GPS receiver in the smart phone. Along with location information, whether the bus is full or partially full is also a matter of concern, for this the passenger count inside the bus is obtained using IR Sensors placing one while entering and another while leaving the bus.

The technical progress in computerized data acquisition and storage has resulted in the growth of vast databases.

With the continuous increase and accumulation, the huge amounts of the computerized data have far exceeded human ability to completely interpret and use. In order to understand and make full use of these data repositories, a few techniques have been developed over a period of time. In this proposed system the data from transit system can be analysed to discover patterns and useful information.

The analysis of gathered data, would help organizations to predict rush on particular routes and handle the same. This may also help to redesign and improve the existing transportation policies. We are using knn classification for passenger flow prediction.

## II. RELATED WORK

Real time bus tracking systems are standalone systems that displays the arrival time of the buses on LCD screens on every bus stop. The system comprises of the power source, battery, LEDs, RF transceiver, microprocessor. RF transceiver is installed over every bus that polls a signal that contains its GPS coordinates. The data will then be processed by the microprocessor. RF transceivers are installed at every bus stop to receive information regarding bus coordinates. These will be passive circuits and will get active only when transmitter enters the range of reception. Bus location is displayed on the LCD screen along with the bus number [1].

SMS (Short message service) is used over the GSM networks to transfer the bus location coordinates. The GPS receiver at the buses computes the longitude and latitude of the vehicle coordinates. This information is sent to the central server over the GSM networks using SMS and this information is stored in the database. Users can retrieve the information by sending the route number and the bus number. SMS is sent to user that contains the arrival time of the bus[2].

Web based vehicle tracking system comprises of the vehicle- mounted tracking devices, central server system and the web based application. Through the web based application users can track the bus graphically using the web based systems also enables users with different operating systems platforms to easily reach the details with the help of internet access[3].

Bus tracking system using Android application uses the inbuilt GPS service provided by the Smartphone. Smartphone is mounted on each bus and gets its GPS coordinates. These coordinates are transferred to the central server. Users can retrieve information through android application where users select the route number,

bus number and receive the arrival time of the bus with respect to the user's current location instead of the bus stop. Maps are also used to graphically plot the bus and the user on the Google- Maps[4].

The study analysed the performance of transit services using data collected from Automatic Fare collection (AFC) system in Seoul with respect to service measure such as punctuality, crowding and operational speed of buses. In order to analyse the transit service performances, they developed data mining logics and applied them to a case study [5]

Analysis relies on use of a generative model based classification. Additionally, some existing approaches consider either classification the trips of each passenger individually or disregard the identity of passenger. They consider each passenger as a single observation and conduct the classification accordingly. The approach is applied on a real dataset from metropolitan area of Rennes (France) with four weeks of smart card data containing trip made by both bus and subways.[6]

### III. PROPOSED SYSTEM

#### System Architecture

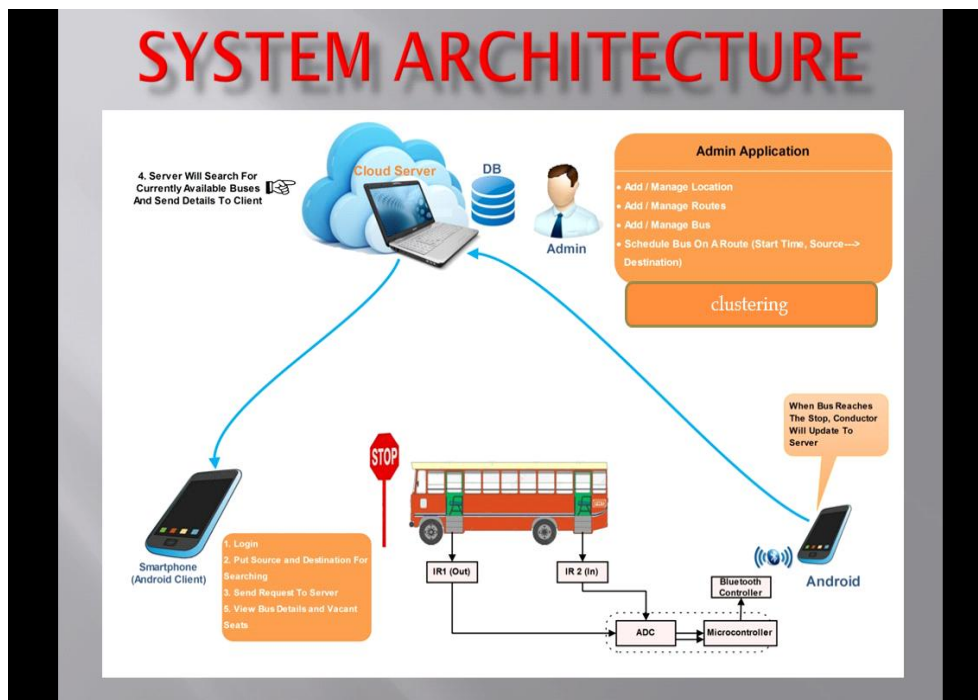


Fig.1. System Architecture

Bus tracking system using Android application uses the inbuilt GPS service provided by the Smartphone. Smartphone is made available to the conductor of the bus and therefore we get the GPS coordinates. These coordinates are transferred to the server. Users can retrieve information through android application where users select

the route number, bus number and receive the arrival time of the bus with respect to the user's current location instead of the bus stop. Also the passenger count is obtained using IR Sensors which are integrated inside the bus giving the real time information about the number of passengers travelling through the bus.



The project is divided into two parts.

- Interfacing of hardware, fetching and delivery of data.
- Data analysis and prediction.

#### A. Interfacing of Hardware, Fetching and Delivery of Data

This part consists of:

- Location tracking through GPS system.
- Passenger count through IR Sensor.
- Administration application.
- Bus Driver application.
- User application.
- Server and server database.

#### I. Location Tracking through GPS system

The driver of the bus would be provided with a android phone. Through the GPS of the phone, the location of the bus is tracked for real time. This information regarding the location is send to the server database using WIFI as the medium of data transfer. The server sends this information to the administration and also the user application regarding the location of bus.

#### II. Passenger count through IR Sensors

To keep a count of passengers entering and leaving the bus, we make use of IR Sensors. We make use of two IR Sensors, the first IR Sensor is used to count the number of passengers entering the bus and the other is used to count passengers leaving the bus. Whenever a passenger enters the bus through the entry door, the count is incremented by one and while exiting the count is deducted by one. This count through the IR Sensor is send to the driver application using Bluetooth (driver application further explained). This data regarding the passenger count is send to the server and the server then sends the data to the user application and administrator.

#### III. Administration application:

An application for the administration is provided to do the following tasks:-

- Add locations.
- Manage location information.
- Add bus numbers.
- Add routes related to buses as well as their time.
- Updating information regarding changes in bus schedule or any other.

The admin application sends the data to user application and driver through server. It receives real time bus location and the passenger count of that particular bus through sever database.

#### IV. Conductor application:

The location of the bus is tracked through GPS. The GPS system exits in the android phone of the bus driver. The GPS data is send to the server through WIFI. The data from the IR Sensor is send to the android phone of the bus driver using Bluetooth and then send to the server database.

#### V. User application:

The real time location of the bus can be viewed by the passengers through the user application. This application also has information about the number of passengers present in the current bus. Users can also view the bus routes and arrival time of buses. User would be provided with a search button wherein he/she is supposed to enter the source and destination information. After clicking on the search button user would be provided with all the available buses along with routes and timing information.

#### VI. Server and Server Database: The server is responsible for carrying and delivery of data. The

The server carries the data from applications and stores at its database. The user query is serviced using server database.

#### VII. Data analysis and Prediction:

Analytics of the gathered data forms the other part of our project. Prediction is done over passenger flow on a particular route at particular time based on the IR count, historical data as well as on other factors such as weather. Scheduling of buses will be done based on prediction of passenger flow and demands. We are using K-mean classification to identify patterns and trends and provide predictions.

**We summarize the working of the first part of our system in the following way:**

- IR Sensors will detect the passenger count and the data will be generated.
- This generated data will be transferred to android app of the driver via BLUETOOTH.
- On receiving the data, the values will automatically be updated at the driver app.
- This updated data and the location will be sent to the web server using SOAP/XML protocol.
- Once the data is at the server it will be used by the android app with the client and by the admin portal.

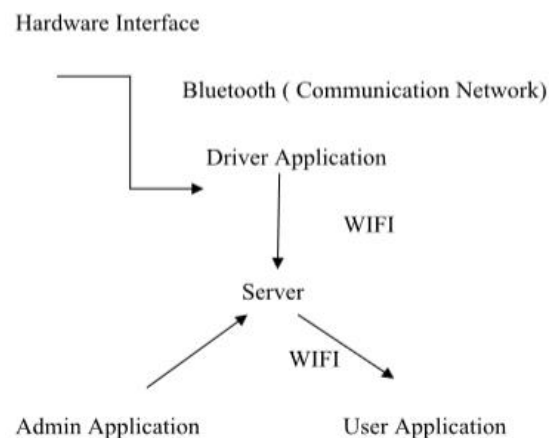


Fig.2. Flow of first part.



Server carries data from the application and stores it at its database. The user query is served using server Database.

B. Part 2- Analysis of Data

Analysis of data is a process of inspecting, cleansing, transforming and modelling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. We are using K-nearest neighbours classification to identify passenger flow prediction.

K nearest neighbours is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition as a non-parametric technique.

**Algorithm:** A case is classified by a majority vote of its neighbours, with the case being assigned to the class most common amongst its K nearest neighbours measured by a distance function. If  $K = 1$ , then the case is simply assigned to the class of its nearest neighbour.

Distance function =  $\text{Sqrt} [(X1-Y1)^2 + (X2-Y2)^2]$

Choosing the optimal value for K is best done by first inspecting the data. In general, a large K value is more precise as it reduces the overall noise but there is no guarantee. Cross-validation is another way to retrospectively determine a good K value by using an independent dataset to validate the K value. Historically, the optimal K for most datasets has been between 3-10. That produces much better results than 1NN.

4) Example:

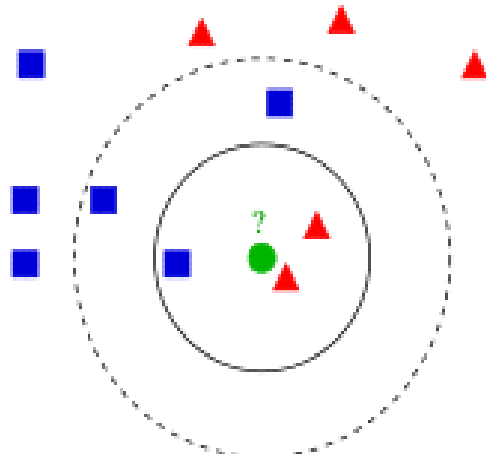


Fig 3: Example KNN

1. The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles.
2. Now the value of k can be anything. The more the value of k the more precise will be the prediction.
3. Let's suppose for the above example we take value of  $k = 3$  (solid line circle) then the green dot is assigned to the second class because there are 2 triangles and only 1 square inside the inner circle neighbouring the green dot.
4. If  $k = 5$  (dashed line circle) the green dot is assigned to the first class because there are 3 squares vs. 2 triangles inside the outer circle neighbouring the green dot.
5. This exact algorithm is used to find out the frequency of the buses there should to maintain the smooth transportation. The prediction is done on the following parameter

	A	B	C	D	E	F	G
1	Day	Source	Destination	Time	Passenger Count	Frequency	
2							
3							
4	Monday	upper	swargate	morning	600	7	2/6/2017
5				afternoon	300	6	
6				evening	550	7	
7				night	300	6	
8							
9	tuesday			morning	550	7	2/7/2017
10				afternoon	300	6	
11				evening	550	7	
12				night	300	6	

Fig 4 : Dataset

6) We put the current day, source, destination, time and a rough passenger count for that day using the algorithm it will find the frequency of busses there should be for the particular route.

IV. METHODOLOGY

• **Android SDK-**

A software development kit that enables the developers to create applications for the android platform. In this project for the development of the Client app and the Conductor

app we will be using the Android Software development kit.

• **Glassfish server –**

Glassfish server is used in the project for developing the web services. It provide very interactive Admin console.

• **JDBC - Java Database Connectivity (JDBC)** is an application programming interface (API) for the programming language Java, which defines how a client may access a database.



- **MySQL** – Database running on server.
  - **Swing** – Swing is a GUI widget toolkit for java. Swing will be used in this project for the development of GUI for Client app, Conductor app and at the admin side.
  - **GPS**- The real time vehicle tracking system has been presented through the use of the global positioning system technology module to receive the location of the vehicle to forward onto the server for displaying the real time position of the vehicle on the map.
- [8] J. Lee, K. Hong, H. Lee, J. Lim and S. Kim, “Bus information system based on smart-phone Apps”, in Proc. of KSCI Winter Conference (2012), pp. 219-222.
- [9] S. Chandurkar, S. Mugade, S. Sinha, M. Misal and P. Borekar, “Implementation of Real Time Bus Monitoring and Passenger Information System”, International Journal of Scientific and Research Publications, vol. 3, no. 5, (2013), pp. 1-5.
- [10] K. G. Zografos, K. N. Androusoyopoulos and V. Spitadakis, “Design and assessment of an online passenger information system for integrated multimodal trip planning”, Trans. Intell. Transport Syst. vol. 10, (2009), pp. 311–323.

## V. CONCLUSION

GPS Based Bus tracking System is an Application that has its client side on the Android platform. Application is free of cost and easy to install on device. The accuracy of this system relies on the GPS coordinates generated using satellites while reliability of the system depends on the GPRS facility. The system is effective where Internet is accessible. The proposed system also predicts the passenger flow using classification. This system also takes into account the traffic conditions during that day of week and time of the day.

## ACKNOWLEDGMENT

We would like to thank the reviewers for their detailed comments, suggestions and constant support throughout the reviewing process that helped us to significantly improve the quality of paper

## REFERENCES

- [1] ChhedaGaurav, GajraNiket, ChhayaManal, DeshpandeJitesh, GhargeSaylee, “Real Time Bus monitoring and Passenger Information System”, in “International Journal of Soft Computing and Engineering (IJSCE)”, ISSN: 2231-2307, Volume-1, Issue-6, January 2012 34.
- [2] Maruthi R., Jayakumari C., “SMS based Bus Tracking System using Open Source Technologies”, in International Journal of Computer Applications (0975 – 8887) “, Volume 86 – No 9, January 2014 .
- [3] Salim A., IdreesIbrahim , “Design and Implementation of Web-Based GPS-GPRS Vehicle Tracking System”, in International Journal of Computer Science and Information Technologies , Vol 3, Issue 12, 443448, December 2013.
- [3] YashaSardey, PranotiDeshmukh, PoojaMandlik, SaurabhShelar, MinalNerkar, “A Mobile Application for Bus Information System and Location Tracking using Client-Server Technology”, Presented at International Journal of Emerging Technology and Advanced Engineering Volume 4, Issue 4, April 2014.
- [4] Jin kiEom\*, Ji Young Song\*\*, dae-seop Moon\*\*\*, "Analysis of public transit performance using smart card data in seoul" ,KSCE journal of civil engineering, vol.00, No.0/000 0000, springer, 2015
- [5] Mohamed k.EiMahrsi , Etienne Come , Johanna Baro, LatifaOukhellou, "Understanding passenger pattern in public transit through smart card and socioeconomic data", UrbComp'14, august 24, 2014
- [6] Development of an Effective Online Bus Pass Generation System for Transportation System for Transportation Service in Karnataka State.
- [7] Caulfield and M. O'Mahony, “An examination of the public transport information requirements of users”, IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, (2007), pp. 21–30.