ISSN No:-2456-2165

Automated Bus Scheduling and Travel Companion

Anandu R
Dept. of Computer Science and Engineering
St. Joseph's College of Engineering and Technology,
Palai Kottayam, India

Aishwarya Michael
Dept. of Computer Science and Engineering
St. Joseph's College of Engineering and Technology,
Palai Kottayam, India

Kishore Sebastian
Prof., Dept. of Computer Science and Engineering
St. Joseph's College of Engineering and Technology,
Palai Kottayam, India

Smitha John
Dept. of Computer Science and Engineering
St. Joseph's College of Engineering and Technology,
Palai Kottayam, India

Kennith Feryl Philips
Dept. of Computer Science and Engineering
St. Joseph's College of Engineering and Technology,
Palai Kottayam, India

Abstract:- Public transport is one of the major forms of transportation in the world. This makes it vital to ensure that public transport is efficient. Multiple buses that go through the same route may overlap due to several en-route delays and hindrances. In Order to maintain a minimum time interval between each bus on a particular route an efficient solution has to be created to combat the overcrowding in buses and inconvenient travel experience for passengers.

An efficient system can be created by scheduling the buses to maintain adequate distance between them by reorganizing the buses plying through routes and, using the schedule to create adequate distancing between buses by urging the drivers to keep to the provided schedule, variations in the bus schedule, detected using statistical inference on analytical data, we deploy/redeploy buses, which makes uses of resources at bare minimum and combats delays and keeping commuters aware of the latest schedule

Keywords:- *Bus Scheduling, Transport, DTC, KSRTC.*

I. INTRODUCTION

Public transport is one of the most popular means of transportation in various metropolitan cities across the globe. According to the Economic Survey 2005-06, buses account for nearly 60% of the total demand. While it is well understood that the public transport helps in combating air pollution and congestion caused due to single-occupancy vehicles, the usage of buses in Kerala and other cities in India has seen a nominal decline due to the uncertainty of bus timings, non availability of schedule , if we consider the situation in states such as Kerala we have a fixed schedule but it is not updated regularly based on seasonal variations and the changing demands of travel which overall has been increasing steadily over the years[1].

Multiple buses that go through the same route may overlap due to several en-route delays and hindrances. Inadequate information during manual scheduling of these buses that go in the same direction causes overcrowding in buses and inconvenient travel. Public is not aware of the present schedule and number of buses plying in a certain route. Issue of not being able to generate sufficient revenue , traffic congestion and risk of accidents as well as competition amongst drivers competing for passengers and revenue.

One of the main reasons for this decline in usage of buses is the lack of reliability of the bus routes and schedule. The timetable is often not made by the transit authorities. Moreover, it is often outdated soon due to the rapid change in infrastructure and the traffic conditions resulting in degradation in the reliability of buses. Finally, this decrease in reliability leads to unknown waiting times at the bus stops for the passengers. Due to the absence of a timetable, most bus trips are operated in an ad-hoc fashion making it extremely difficult for the passengers to trust the public transport network leading to a decrease in passenger trips. Interestingly, the various trips in a given bus-route still follow a certain pattern in a given day thanks to the pattern in the traffic conditions throughout the day.

II. BACKGROUND

Presently with respect to the Kerala State Road Transport Corporation, a proper schedule for the buses and the depots is followed at all places. This method of charting a schedule has been followed since the inception of private and trans-port buses. This traditional method has not much improved over time and still continues this not so effective scheduling scheme. In light of the current development in technology and advancements in the digital field including machine learning, deep learning, cloud resources etc, necessary changes need to be followed through in order to make processes such as bus scheduling more easier and effective.

ISSN No:-2456-2165

Currently manual operations to chart schedules and mark bus arrival and departure is done by the bus operator at bus depots . Ticketing system also is done mostly by a ticketing machine that provides a paper ticket to the passenger and some buses still follow the traditional method of paper based tickets which is given without a ticketing machine by the conductor. This is how KSRTC has implemented it bus booking so far.

III. OBJECTIVE AND SCOPE

The aim of this project is to create an bus automation infras- tructure that allows the state transport department to leverage the latest data analytics and Big data technologies to create a robust bus scheduling system that takes into account various real world parameters and creates an efficient system that finds the best schedule by understanding passenger capacity versus the the passenger count for various routes in an attempt to reduce overcrowding of buses and to reduce inconveniences to passengers. It also improve the user experience and also the general reception on the transit systems. Other notable features that the infrastructure already has/has been includedis as follows -

- **Accurate and Precise information based on GPS/GPRS navigation**: A GPS device is alreadybeing fixed in all KSRTC buses. This data can be used to track buses in real-time.
- **Number of buses in a certain route**: The count of buses that ply through a certain route depends on various factors and mainly involves the the count of the passengers that avail the bus service by booking through the app.
- **Online QR Code ticket booking of unreserved buses**: Tickets for any bus can be booked from our system. The conductor can scan the QR code and verify the ticket booked by the passenger(Unlike the current system in which reservations can only be done for specific buses).
- **Automated Scheduling of buses**: Bus timings can be altered and number of buses can be increased or de- creased according to the peak time for a specific location and for locations with less or more crowd density . For automated identification of peak time and crowd in specific locations, Regression model is used to predict the passenger count in various routes using the data collected during a period of time , which is then used to schedule an optimal bus timing.
- **Crowd Estimation**: The crowd on the bus is estimated by the number of tickets booked both using passenger application as well as the ticketing machine. To remove the conflict of too many crowd boarding on the bus and all tickets not issued/verified, we can use image processing and the data compared with the number of tickets will tell the crowd on the bus.

IV. LITERATURE SURVEY

A. Benchmark Dataset for Timetable Optimization of Bus Routes in the City of New Delhi [1]

Real-time GPS bus transit data of buses operating is being used for modeling various timetable optimization tasks as well as in other domains such as traffic management, travel time estimation, etc. with the aim to reduce the waiting timeof buses by analyzing the traffic behavior and proposing a timetable [1]

It uses an algorithm which serves as a benchmark for the dataset. The algorithm uses a constrained clustering algorithm for classification of trips. It further analyses the data statisti- cally to provide a timetable which is efficient in learning the inter and intra-month variations [1].

The system makes use of analytical data which can be used to model a time table taking into account not only the variations in the traffic during rush hours/peak times but also the seasonal variations due to natural phenomenon which leads to slowing down of traffic/creates other various discrepancies in the transit schedules [1].

Methodology

The average waiting times for passengers were studied for the routes 425 and 534, up-route and down-route separately and it was found out that the average waiting time post time schedule optimizations for the above two routes decreased and the result was found to be statistically significant with p-values¡0.05 [1].

In-order to gauge the performance of optimization algorithmthe following analysis were done:

- Comparing the waiting time with respect to the first node (stop) on the route - 425, before and after the timetable optimization to understand the performance of the clustering algorithm in defining the starting time of the schedule.
- The scheduling of the remainder of the stops can be evaluated to understand performance of the model and also learn of the intra-month variations in the route, for this purpose the data was trained on alternate days - even or odd days.
- For learning inter-month variations in a data which already contains high randomness, the algorithm was trained on first-month data and tested on the second month data.

B. A Study on Crowd Detection and Density Analysis for Safety Control [2]

Crowd Detection System consisting of Input Data, Ap- proaches, Features and Conclusion. The various approaches to handle the problem can be broadly divided into

- Detection based approaches
- Regression based approaches
- Density based approaches

ISSN No:-2456-2165

Convolutional Neural Network based methods consist of deep learning approaches for crowd detection and density analysis.[2]. CNN based methods consist of deep learning approaches for crowd detection and density analysis. Con- volutional Neural Networks (CNNs) similar to plain Neural Networks(NNs) they consist of neurons that have learnable weights and biases. Each receptive field receives a batch input and performs a convolution operation, and the result is fed into a nonlinearity function (e.g., ReLU or Sigmoid). The input image to CNN is assumed to be an RGB image; therefore, the hidden layers learn rich features which extracted in each layer that contribute to the performance of the whole network (hidden layers and classifier). This structure has benefits in terms of speed and accuracy since lots of objects need computationally expensive operations to detect and identify the features of the input.

Fully Convolutional Neural Network with skip branches produced the highest quality density map for localization tasks, with slight degradation for the counting task [2]. Deep learning has been greatly developed and applied to the crowd density estimation. It has been proved that the deep learning algorithmis more effective than the ordinary methods.

- Density estimation CNN-based models for crowd count- ing, provide accurate count but does not indicate the precise location and exact size of objects, limits further research and application, such as high-level understand- ing, localization, classification, and tracking [3].
- Unavailability of a database for some tasks such as crowd localization and behaviour analysis

C. Random Forest decision tree classification algorithm fortime time table scheduling [6][4]

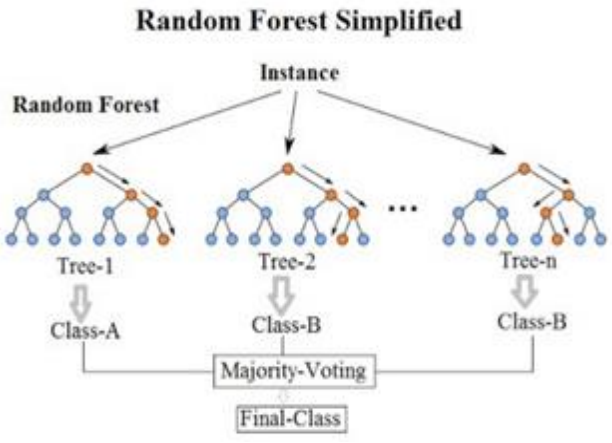

Fig. 1. Random forest voting architecture [8]

A random forest is an ensemble learning method where multiple decision trees are constructed and then they are merged to get a more accurate prediction.

Random forest runs efficiently on large data bases. It can handle thousands of input variables without variable deletion. while giving estimates of what variables are important in the classification. It has an effective method for estimating missing data and maintains accuracy when a large proportion of the data are missing. Generated forests can be saved for future use on other data. Prototypes are computed that give information about the relation between the variables and the classification. The capabilities of the above can be extended to unlabeled data, leading to unsupervised clustering, data views and outlier detection. It offers an experimental method for de- tecting variable interactions. Random forest run times are quite fast, and they are able to deal with unbalanced and missing data. They can handle binary features, categorical features, numerical features without any need for scaling. There are lots of excellent, free, and open-source implementations of the random forest algorithm. We can find a good implementation in almost all major ML libraries and tool kits [6].

Random forest, like its name implies, consists of a large number of individual decision trees that operate as an en-semble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction. A large number of relatively uncorrelated models (trees) operating as a committee will outperform anyof the individual constituent models [4].

The reason for this wonderful effect is that the trees protect each other from their individual errors (as long as they don't constantly all err in the same direction). While some trees may be wrong, many other trees will be right, so as a group the trees are able to move in the correct direction. So the prerequisites for random forest to perform well are: [4]

- There needs to be some actual signal in our featuresso that models built using those features do better than random guessing.
- The predictions (and therefore the errors) made by the individual trees need to have low correlations with each other.

The random forest is a classification algorithm consisting ofmany decision trees. It uses bagging and features randomness when building each individual tree to try to create an uncor- related forest of trees whose prediction by committee is more accurate than that of any individual tree [4].

D. Classification in Decision Tree — A Step by Step CART (Classification And Regression Tree) [5]

CART (Classification And Regression Tree) is a decision tree algorithm variation, in the previous article — The Ba- sics of Decision Trees. Decision Trees is the non-parametric supervised learning approach. CART can be applied to both regression and classification problems [5].

CART uses Gini Impurity in the process of splitting the dataset into a decision tree. Mathematically, we can write GiniImpurity as following [5]

$$I_{gini} = 1 \quad \sum_{i=1}^{j} P_i^2 \quad [5]$$

$$I_{gini} = 1 - (\text{the probability of target "No"})^2 - (\text{the probability of target "Y es"})^2$$

ISSN No:-2456-2165

Main processes when the dataset is split : [5]

- Calculate all of the Gini impurity score
- Compare the Gini impurity score, after n before using new attribute to separate data. If the node itself has the lowest score, than there is no point in separating the data
- If separating the data result in an improvement, than pick the separation with the lowest impurity score

In order to make accurate class predictions the random forest requires the below conditions to be satisfied: [4]

- We need features that have at least some predictive power. These features are what would predict the output. I After all, if we put garbage in then we will get garbage out.
- The trees of the forest and more importantly their predictions need to be uncorrelated (or at least have low correlations with each other). While the algorithm itself via feature randomness tries to engineer these low correlations for us, the features we select and the hyper-parameters we choose will impact the ultimate correlations as well [4].

E. Aanavandi [7]

Aanavandi is a web - based application that provides bus schedules to users as well as the facility to book bus tickets online [7].

- The application provides a quick access to bus schedules of private buses throughout Kerala.
- The schedule is a static timetable obtained from the government.
- It only provides bus booking facilities for long route buses and not daily buses.
- Updation of timetable does not take place frequently.

V. MODULES

The system is divided into four modules:

(1) Cloud
(2) Passenger Application
(3) Driver / Conductor Application
(4) Website for Officials

1) ***Cloud:*** Cloud computing is the delivery of computing services—including servers, storage, databases, networking, software, analytics, and intelligence—over the Internet (“the cloud”) to offer faster innovation, flexible resources, and economies of scale. You typically pay only for cloud ser-vices you use, helping lower your operating costs, run your infrastructure more efficiently and scale as your business needschange.

Using cloud infrastructure allows us to operate at mini- mal cost without having to rent server space and resources required, all we need to look after is the operating costs.Cloud computing offers tonnes of benefits to enterprises and organizations, namely:

- **Cost**: Cloud computing eliminates the capital expense of buying hardware and software and setting up and running on-site datacenters—the racks of servers, the round-the-clock electricity for power and cooling, the IT experts for managing the infrastructure. It adds up fast.
- **Speed**: Most cloud computing services are provided self service and on demand, so even vast amounts of com-puting resources can be provisioned in minutes, typically with just a few mouse clicks, giving businesses a lot of flexibility and taking the pressure off capacity planning.
- **Global Scale**: The benefits of cloud computing ser- vices include the ability to scale elastically. In cloudspeak, that means delivering the right amount of IT resources—for example, more or less computing power, storage, bandwidth—right when it is needed and from the right geographic location.
- **Productivity**: On-site data centers typically require a lot of “racking and stacking”—hardware setup, soft- ware patching, and other time-consuming IT management chores. Cloud computing removes the need for many of these tasks, so IT teams can spend time on achieving more important business goals.
- **Performance**: The biggest cloud computing services run on a worldwide network of secure data centers, which are regularly upgraded to the latest generation of fast and efficient computing hardware. This offers several benefits over a single corporate datacenter, including reduced network latency for applications and greater economies of scale.
- **Reliability**: Cloud computing makes data backup, dis-aster recovery and business continuity easier and less expensive because data can be mirrored at multiple re-dundant sites on the cloud provider’s network.
- **Security**: Many cloud providers offer a broad set of policies, technologies and controls that strengthen your security posture overall, helping protect your data, apps and infrastructure from potential threats.

For this project we decided to use Google’s Google Cloud Platform (GCP) as the cloud platform, It provides variousservices and resources that the project can utilize.

2) ***Passenger Application:*** This app also has the login and register page for passengers. Passengers can register in the appand can access to the following features:

- **Map View of Buses** - Passengers can input from where and till where to travel and the upcoming buses will be displayed on the map with their current location, bus type time to reach the bus stop, delay, and crowd density.
- **Bus Timing** - Passengers can input from where, till where to travel, bus type and the time period and the list of buses in the time period will be displayed.
- **Tracking** - This is to search for a specific bus. Bus Fleet number can be fed into the app to track that bus. This feature is added so that people can track their family members or friends or loved ones.
- **Ticket booking** - This helps the passengers to book online tickets. On making the payment the passenger gets a QR Code (Generated with a unique token) that has to be shown to the conductor on boarding the bus.

3) ***Driver/Conductor Application:*** **Conductor** In the con-ductor module, the conductor has the following features:

- **Scan Ticket QR Code** - Since passengers can book their tickets from their app, the conductor can scan the QR code and verify if it’s a valid ticket.

- **Passenger List** - The conductor can check the number of passengers traveling on the bus and also the names of people who are traveling which will be available if the ticket is booked online.
- **Emergency** - In case of emergencies like accidents and bus failures, the conductor can report the event so that the help can be provided quickly by the KSRTC officials.
- **Driver** In the driver module, the driver has the following features:
- **Map** - After the driver logs in the map loads automatically for the driver. The source and destination will automatically be fed into the app from the database. This data is stored when the driver is assigned the duty.

4) ***Website for Officials:*** This website is exclusively for the operators in the bus depots. All operators have their unique IDand a password with which they can log in.

After logging in, the Home page opens. These are the options in include in the website and displayed as a menuin the Home Page:

- **Bus tracking** - this helps operators to get real-time tracking of the bus using the bus triple code number that is unique for all buses and the starting and ending locationof the travel.
- **Bus Fare Update** - Bus fares increase very often so operators can update the bus fare for the ticket booking facilities on the user's app using this website by just updating the minimum price and price per km.
- **Bus Scheduling** - This shows the list of all available buses from one destination to the other in a particular range of time. It also includes the option of availing the passenger count on a daily basis in order to understand passenger frequency in the routes.
- **Employee details** - This provides details of a particular employee using the Employee ID and also gives the list of all conductors and drivers on duty.
- **Customer Feedback** - This shows the grievances of the passengers using the KSRTC service and their complaints or suggestions that they mentioned through the passenger app.

VI. SYSTEM DESIGN

A. Use Case Diagrams

For the proposed solution the following use case diagramsexists, for each of the modules:

Cloud

The use case diagram describes the interaction between namely the passengers, bus and the google cloud resources - Cloud dataflow, BigQuery and compute engine. The ticketing information from the passengers is taken from the ticketing machine and the online booked tickets, which are then consol- idated to obtain the overall count of the passengers travelling within the buses, along with this data packets from the buses are sent to the cloud platform in API messages formats which includes data such as location of bus, the bus number, route number and so on, which uniquely identifies each bus. The ETL pipelining functionality of the cloud dataflow resource on GCP allows batch and stream processing from multiple data sources simultaneously and summarises the information which is then used by the scheduler to make the time table for the corresponding buses.

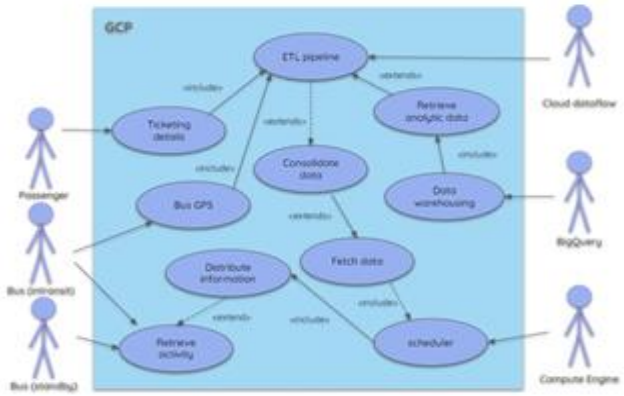

Fig. 2. Use case diagram for cloud

Passenger App

The passenger app mainly focuses on the user experience and the functionalities available to the users for carrying out various tasks such as :

- Authenticate user: Each user has their own account which they can log into and use the various facilities provided through the app.
- Requesting routing: The passenger can obtain the various routing information and the buses plying through them. The information is provided to the users through the Google Maps location services.
- Request bus tracking: A passenger can track the location of a bus for the entered fleet number, the information for which is again obtained through the tracking services in Google Maps.
- Request schedule: The passengers can obtain the schedules of all buses running through a route within a given time frame, the passengers also have the ability to filter through the buses.
- Initiate booking: Passengers can book seats using Razr Pay payment gateway, for a valid payment a unique QR code is generated which can be scanned to authenticate the seat.

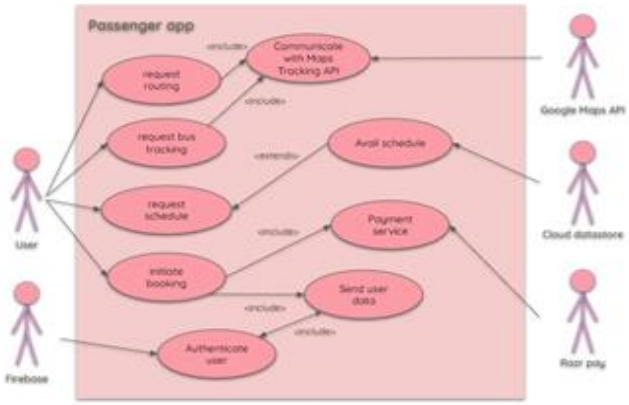

Fig. 3. Use case diagram for Passenger app

ISSN No:-2456-2165

Driver/Conductor App

The driver conductor app use case comprises the conductor, the conductor, state transport department, that are real world entities along with the GCP, Razr Pay authentication services, and the ticketing machine. The conductor can:

- Authenticate the booked tickets
- They can also report transit related issues that cause any delay or disruption in a particular route to the bus depot
- The driver obtains the routing and schedule information from the scheduler

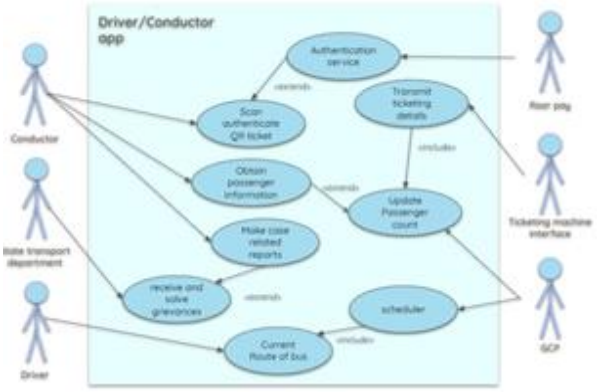

Fig. 4. Use case diagram for Driver Conductor app

Website

The website use cases include the department, operator, google maps API, cloud datastore. The operator has access various activities such as:

- Request Bus tracking: Using API calls to Google Maps Services the operator can track any bus and get its details number of passengers, current location, running duration and so on. It also relies on live information data packets being processed by the cloud dataflow.
- Request Schedule: The operator can request the schedules of all the buses on a per route basis.
- Update bus fare: The operators can easily update/ modify the bus fare rates and change the step gains.
- Request Employee details: The details regarding the employees who are currently working are available for marking attendance.

B. Sequence Diagram

The sequence diagram is depicted in figure 4.3. Here the information regarding the bus fleet number, operational data and the GPS tracking information is given to the Cloud Dataflow. The information of previous bookings and schedules over a period of time are queried from BigQuery/Firestore and this data is also given to the Cloud Dataflow/Compute Engine Instance which computes a suitable schedule based on algorithm developed, on the data and this schedule is stored back into the database service of choice and distributed back to the Transport Department, based on which suitable changes can be made to the schedule.

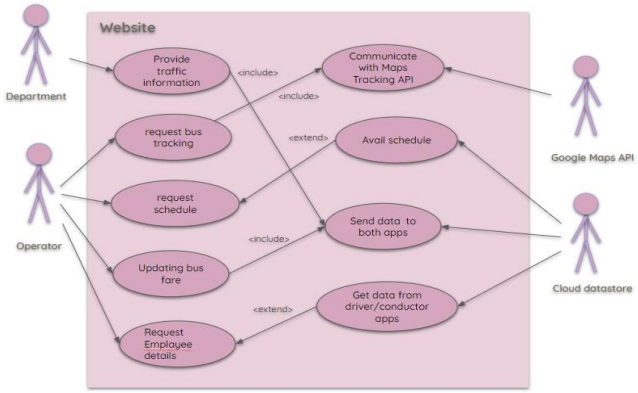

Fig. 5. Use case diagram for website

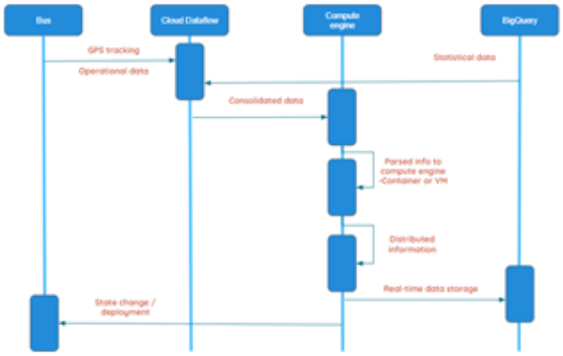

Fig. 6. Sequence diagram for clou

C. Activity Diagrams

The activities of each of the modules can be summarised as follows:

Passenger The passenger app has a login page which authenti- cates the users using their firebase authentication service. The user is able to access the following features :

- **View Bus schedule** : The user can access the bus schedule based on the boarding and destination bus stops.
- **Book Ticket** : Tickets for different types of buses can be booked online which is then redirected to the payment page wherein after successful payment of the bus ticketa QR code if generated which is used for verification by the bus conductor. In the case of a payment failure it is redirected back to the payment page.
- **Track Buses** : Buses can be tracked real - time based on their fleet number using the Google Maps API
- **View Route** : The Map can be viewed by the user.

Driver/Conductor

The driver conductor app has a login wrapper which redirects the view based on the login credentials. The conductor is able to access the following features:

- **QR code scan**: To authenticate the tickets that were booked online by the passengers
- **Track Passenger count**: Using the passenger information the conductor can infer about the number of passengers within the bus

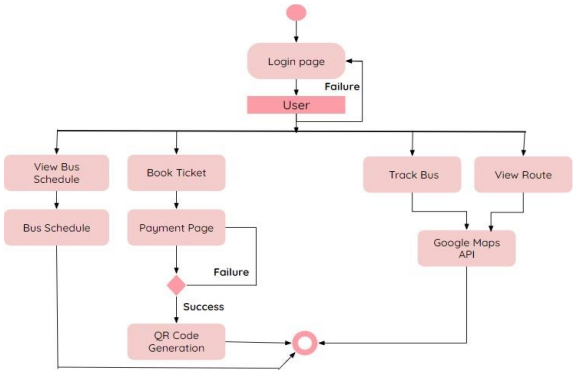

Fig. 7. Activity diagram for passenger app

- **Report emergency**: To report an emergency situation or alert the authorities on any disruptions in the route.

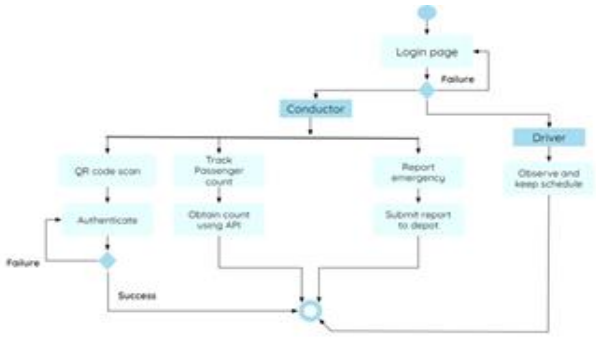

Fig. 8. Activity diagram for conductor app

The driver is able to access the following features:

- Access maps which provide the route to the driver during the journey.
- View the schedule to be followed by the bus

Website

The operator can login with valid credentials and is givenaccess to 4 major activities, they are:

- **Bus tracking**: The operator tracks all buses running.
- **Bus Scheduling**: Check all the schedules generated and currently
- **Employee Details**: The employee details are available to mark attendance and to have a general idea of number of employee currently active
- **Bus fare updation**: To easily update the bus rates and the step gains

D. ER Diagram

Entity relationship diagram depicts the basic structure of the database we use in our system. The ER diagram for our proposed system is given below which shows the relationship between the entities of our project. There are mainly four main entities. They are Bookings, Passengers, Conductor/driver and Bus. There are different attributes connecting all the three entities here. For the entity 'Passenger', the attributes are phoneno, fullname, email. For the entity 'Bus', the attributes are location, fleetno, toloc, fromloc, bustype. For the entity 'Conductor/Driver' the attributes are conductorid, fullname, phoneno. For the entity 'Booking' we have the following attributes bookid, uid, bustype, fare, to, from, phoneno, book- time.

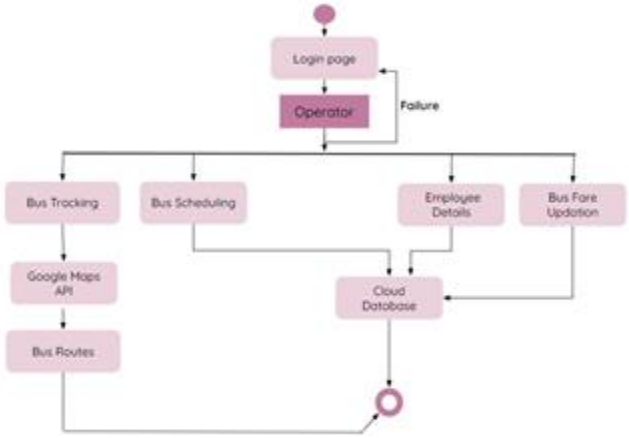

Fig. 9. Activity diagram for website

Fig. 10. ER diagram

VII. PROJECT DESIGN AND IMPLEMENTATION

Real-time GPS bus transit data of buses operating is being used for modeling various timetable optimization tasks as wellas in other domains such as traffic management, travel time estimation, etc. with the aim to analyze the traffic behavior and proposing a timetable.

The system makes use of analytical data which can be used to model a time table taking into account not only the variations in the traffic during rush hours/peak times but also the seasonal variations due to natural phenomenon which leads to slowing down of traffic/creates other various discrepancies in the transit schedules.

A. Data collection

The system collects real-time bus data while the buses arein transit and active. The data pertaining to the bus which is sampled at intervals of 10s from the bus is taken. The data collected by the system comprises of:

Variable	Description
date	Date
time	Time of day
lat and long	Location of bus
number plate	Unique bus identifier
route number	Route bus is on
Direction	Service direction

ISSN No:-2456-2165

The system also collects static data such as
- routes
- stops
- trips
- stops times

Passenger app which allows booking can be used to obtain the passenger booking information stored in firebase firestore storage service, this information is used in the algorithm to develop a schedule suggestions.

B. Modelling variables

We have generated a model that would predict the passenger count over bus routes whose bookings are done through the passenger app based on the data that is fed to it over a period of time.The names of routes that buses travel without taking the via route and the count of buses in each route are already known.

The variables collected from each route for modelling are:
- Total bus stops N are present on the route.
- To and from locations of a route
- The passenger booking data collected from app which includes:
  - Booking ID
  - Booking Time
  - Bus Type
  - Fare
  - From
  - Phone Number
  - Timing
  - To
- Number of tickets booked in a route using the help of ticketing information

C. Algorithm Used

The algorithm uses general regression model to predict the passenger count of a bus using time series data we collect overa certain period of time.

For example, assume we collect the data of bus with passenger count for Mondays of 4 separate weeks, we can predict the passenger count of the following Monday usingthe prediction model.

Working: Considering a particular route with multiple buses running, we compute the aggregate
- maxPassengerCount - Total carrying capacity of the bus times the number of buses.
- totalPassengerCount - The sum total of average passenger count of each buses.

By using these two variables we generate from the realtime data, we can infer the bus crowd density, route safety, passen- ger comfort and so on, using which we create a list of schedulesuggestions which can be used to increment or decrement the number of buses plying through that route to accomodate the larger number of passengers or save costs by running fewer buses.

D. Specific Outputs of the Project
- Increase in availability of buses.
- Passengers are more informed about the status of the buses.
- Real-Time GPS / GPRS tracking of buses ensures the safety of passengers, especially women and kids
- Understanding passenger frequency in buses according to peak time and locations.
- Cashless ticketing. (Also, paper saving)
- Maps to drivers and conductors.
- Digitalization of Information and a user-friendly and convenient method of payment.
- Systematic, efficient and reliable medium.
- An estimate of the number of passengers on each bus.
- List of all on-duty drivers and conductors.
- It helps officials to study the crowd pattern.
- Reduces traffic congestion and initiates proper manage- ment of buses.
- Helpful for the physically challenged and handicapped.
- Increase in revenue generation.
- Tackles Pollution by lesser petrol consumption.
- It encourages citizens to be updated about the latest technology.
- Encourages citizens to adhere to traffic rules.

VIII. CONCLUSION

Bus Scheduling is a system that requires real - time in- formation both from passengers as well as the bus driver or conductor side in order to make an efficient scheduler taking into consideration the passenger count and the number of tickets being booked through the app. Factors that may affect the schedule are unprecedented traffic jams and weather relateddelays or other disaster or road damage.

This report analyses different modules and the technical stack needed to make a fully functional bus scheduler. It also helped in studying different probable models that could help schedule and understand the crowd and capacity of buses. It enables to successfully implement an infrastructure system that can be used by the road transport department. It helped understand the relation between entities in our database through the Entity - Relationship Diagram. A GPS device is already being fixed in all KSRTC buses. This data can be used to track buses in real-time. Online QR Code ticket booking can be done for unreserved buses. The conductor can scanthe QR code and verify the ticket booked by the passenger. Regression modeling is used to identify suitable schedulesusing the data collected during a period of time , which is thenused to schedule an optimal bus timing. The crowd on the bus is estimated by the number of tickets booked using passenger application. As a future scope to remove the conflict of too many crowd boarding on the bus, use of image processing can be done and the data be compared with the number of tickets which will tell the crowd on the bus.

This project will be able to change the current system in bus scheduling and create a convenient and easy way to generate an optimised bus schedule beneficial to both the public as wellas the bus department.

REFERENCES

[1]. "Benchmark Dataset for Timetable Optimization of Bus Routes in the City of New Delhi" - Anubhav Jain, Avdesh Kumar, Saumya Balodi,
[2]. Pravesh Biyani *20 Oct 2019*
[3]. "A Study on Crowd Detection and Density Analysis for Safety Control"
- Mayur D. Chaudhari, Archana Ghotkar *30 April 2018*
[4]. "Deep Learning-Based Crowd Scene Analysis Survey" - Sherif Elbish-lawi, Mohamed *7 Sept 2020*
[5]. "Understanding Random Forest" - Tony Yiu *12 June 2018*
[6]. "Classification in Decision Tree — A Step by Step CART (Classification And Regression Tree)" - Arif R *5 April 2020*
[7]. "Random Forest" - Leo Breiman and Adele Cutler *8 December, 2020*
[8]. "Aanavandi" - Sujith Bhakthan, Vaisakh M.L, Sreenadh H, Antony Varghese, Rafeeq Muhammed *2008*
[9]. "Random Forest Simple Explanation" - Will Koehrsen *27 December 2017*