Optimized Bus Management System through Analysing the Density of Bus Commuters and Delay

Prof. Mrs. S. R. Vispute
Department of Computer Engineering,
Pimpri Chinchwad College of Engineering,
University of Pune, India,
visputesushma@gmail.com

Nikita V. Shahane
Department of Computer Engineering,
Pimpri Chinchwad College of Engineering,
University of Pune, India,
nshahane13@gmail.com

Abstract — Transport demand in most Indian cities has increased significantly due to increase in population. Public transport especially buses are getting crowded day-by-day due to heavy demand of transport facility. This has led to unregulated planning of the frequency of buses to be allocated different areas. Either the buses lines up at one time or buses get delayed for a long time. In some areas there are very few buses allocated at a very crowded route .Also there are buses lined up on the route where hardly commuters travel. This limitation of management has incurred loss in the bus management system. Hence we have proposed and implemented a system in which the number of passengers in a bus stop can be calculated and the bus service can be regulated depending on the passengers arrival. Also due to our system there has been efficient distribution of the buses on the required routes and this benefits the bus management system by reducing the loss they are facing nowadays. It also calculates the time delay a bus may take considering different parameters that may affect it to be late . It also verifies the cost generated against the actual cost collected as well benefits the commuters indirectly. All this would be done using a web application available to employees at different positions.

Keywords — *Optimization, Data Analysis, Bus Frequency Analyser, Schedule Optimiser.*

*****

I. INTRODUCTION

Data Mining is an interdisciplinary subfield of computer science. It is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating[10].

Optimization is a process or methodology of making something as fully perfect, functional or effective as possible[10].

Data Analysis is a process of cleaning, transforming, inspecting and modelling data with the goal of discovering useful information, supporting decision-making and suggesting conclusions [10].

Our system introduces about the bus management system which implemented the basic functions of the intelligent public transport management system, such as monitoring the time of bus arrival, departing from the bus station and reporting stations name automatically. This system ensures punctuality of vehicles to run, improve the automation level of reporting stations and quality of public transport service. The management system has low cost and thus it is more feasible.

The Bus Frequency Analyser module focuses on efficient distribution of the buses on a particular route. Thus, it varies the bus frequency depending on the crowd of the route or its utilization. The Bus Schedule Optimizer module calculates the

delay based on different parameters considered, while the Cost Estimator module verifies the cost.

II. EXISTING SYSTEM AND RELATED WORK

The public transport management system which is currently used indicates limitations in the management of bus services. Almost all of the work is done manually. The Depot employees take minimum three to four days for the survey of bus management system. To schedule the bus services in new area, the employees of Depot personally visit that area and survey it for few days so as to check whether is it really needed to start the bus service in that area[11].

First the Depot employees do survey by physically visiting that region, after that this request is sent to Depot manager. The Depot Manger then forward this request to the head office or Main Depot. For all this process and to come to conclusion it takes maximum one month for decision making [11].

Bus delay could be caused by a series of factors, for instance, traffic congestion, overload of passengers, traffic accident and other unpredictable situations. Due to delay in buses the commuters face problems as they have to reach the destination on time, the drivers of the buses face problems because of delay in the first trip they can't go on time for the next trip. As time management is an important factor in this digital world due to delay in one thing rest all work gets delayed [11].

Therefore, choosing crucial elements to efficiently evaluate the bus delay is a complex problem in bus delay researches and management system [11].

There is no proper analysis done resulting the government in loss as well as improper services to the commuters. Without a well deployed monitoring system, it becomes very difficult to plan for the optimization and the growth [11].

All the above limitations of the existing system can be sorted out by the system which uses data analysis algorithms to monitor commuters' density and regulate the frequency of the buses. Also delay algorithm can be used to find the delay and update the schedule accordingly [11].

Table 1. Related Work Done by Different Authors

Title of Authors Journal Finding Details Paper Details	Authors	Journal Details	Finding Details
Regulating Bus Management System Using Cloud Platform	Ranjith Ramesh, Yokesh Ezhilarasu, Prasanna Ravichandra n, and Soma Prathibha	International Journal of e- Education, e-Business, e-Management and e-Learning, Vol.2, December 2012	Here, author has optimized the bus management system by commuter density calculated by passenger feedback using GPS and gossip protocol[1].
A bus Management system Based on ZigBee and GSM/GPRS	LV ZHIAN HU HAN	InternationalConference on Computer Application and System Modeling, 2010.	This paper introduces a nt System on Computer system design Based on Application about bus ZigBee and and System management GSM/GPRS Modeling, system based 2010. on ZigBee and GSM,GPRS, which implemented the basic functions of the intelligent public transport management system, such as monitoring the time of bus arrival , departing from the bus station.[2]
Intelligent Bus Monitoring and Manageme nt System	M. A. HANNAN, A. M. MUSTAPHA, A. HUSSAIN and H. BASRI	Proceedings of the World Congress on Engineering and Computer Science, October 2012.	Here using RFID , GPRS system the movement of the bus is monitored.[3]

The proposed system will overcome all the limitations of the existing system. The cost of the product also reduces as there are no any hardware tools except a personal computer and an internet connection. The overhead of the commuters' as well as the bus operating authorities will be reduced. This system will result into efficient management of resources and regulating the bus services properly also providing the public service in efficient way.

III. Proposed System

Bus Management System is suggested for a small office management tools. Most of express bus company wouldn't invest too much on the electronic management system for their company. Therefore a standalone system is suitable for them because it won't cost them so much. This system provides a simple interface for the user. Moreover, databases typically maintain only the current value for any given data element. As a result, in a typical transaction processing system, inviolate business records are difficult, it's not impossible to locate and retrieve. The use of this system is only for the Manager and his personal assistant or anyone that have been authorized by the manager. Bus Management System includes the buses information, driver. Search function are also added so that manager can search for bus information that has been stored in the databases.

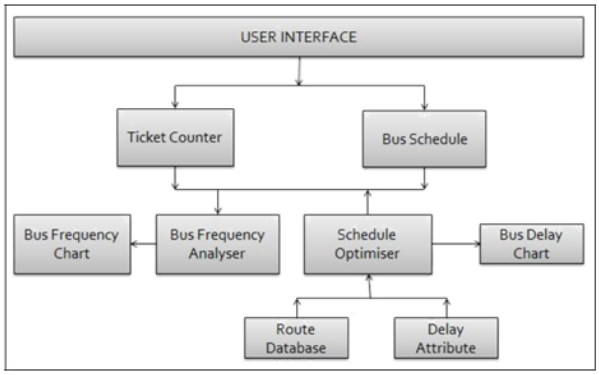


Fig 1. Block Diagram of Proposed System

Our system is divided into two modules:

A. *Bus Frequency Analyser:*

This module represents the analysis of the bus data for regulating the bus frequency. Basically, in this the frequency should be altered according to the requirement.

The analysis is done on the basis of following sub modules:

i. Ticket Count
ii. Route and Schedule Details
iii. Bus Capacity

The above sub-modules are used as input to analyse the data and give a graphical format output.

Bus Frequency Analyser system works in the following manner:

- *Starter*: Starter is the person who sits at the start and end of the route. Starter at the source of the route enters the bus details, ticket bunch handed over to the conductor and departure time of the bus. Destination starter validates the information provided by the source starter. He enters the remaining tickets and the arrival time of that particular bus.
- *System Server*: Here, whole processing of the system takes place. Calculations of the tickets sold, bus fare etc. is done here.
- *Depot Admin*: This person is a decision maker. He directly views the graph of tickets allocated versus fare of tickets. If any new bus is required for that route, admin sends notification to the system server. The system server then adds the buses on the routes as per requirement.

This system will give efficient distribution of the flow of the buses on the required route and hence reduce loss as well as give proper public service.

Fig 3(a,b,c) is the graph generated from the route details and the ticket counts. We have taken six types of tickets i.e Rs.5, Rs.10, Rs.15, Rs.20, Rs.25 and Rs30.

This graph represents the distribution of the allocated tickets Vs the remaining tickets This gives an idea about the utilised tickets on that route i.e the commuter population on that route.

The analysis is done by taking the average of the last three type of ticket i.e Rs.20, Rs.25 and Rs.30 and is compared with the overall ticket average and the bus capacity size, which is taken 45 according to the survey.

Along with this graph a notification is send to the admin whether to increase or decrease the frequency count of the bus.

The revenue generated per trip is being calculated by the cost estimator for verification.

B. Schedule Optimiser:

The operations carried out in Delay Analyser are detection of bus delays, likely cause of the delay and prediction of likely arrival time of buses at the bus-stops.

In this module we are considering the following attributes:

i. Route type
ii. Distance
iii. Actual time taken
iv. Passenger crowd

The above attributes are used to calculate the bus delays. This information will be useful for both the commuters and bus operating authorities.

Here, we are calculating the delay in the bus, analysing it and optimizing the schedule accordingly. For calculating delay we have used KNIME Tool and Naïve Bayes algorithm. In KNIME tool we take two XLS readers, one for training data and the other for test data. The training data is given input to the Naïve Bayes Learner and this input is further given to Naïve Bayes Predictor, also test data is given input for predictor. Predictor predicts the delay by analysing the test data. Naïve Bayes algorithm classifies the delay in three categories: No Delay, Moderate Delay and More Delay. The output is visible in the interactive table. Find below the Fig.3 for proper flow of the Naïve Bayes algorithm in KNIME tool. If the delay observed frequently is moderate and more, a notification is sent to system server for updating the bus timings on that particular route wherever needed.

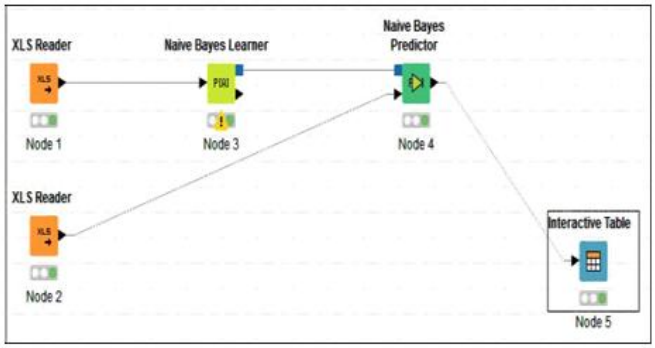


Fig 2. Design for Naïve Bayes Classifier in KNIME Tool

The above figure is the design of the Naïve Bayes classifier in KNIME tool where the bus delays are classified.

The commuters can be benefited from the information such as current bus location, expected time of arrival of a bus at bus-stop, availability of seating space in the buses, etc. [5].

The information gathered can be used to infer likely causes behind the observed bus delays, the bottleneck traffic junctions, heavy traffic time slots, and make appropriate scheduling recommendations [5].

IV. Algorithm And Tool Used

KNIME Analytic Tool:

KNIME is an open source data analytics, reporting and integration platform developed and supported by KNIME.com AG. Through the use of a graphical Network , KNIME enables users to create Data Flows, execute selected analysis steps and review the results, models and interactive views. Witten in Java and built on Eclipse, KNIME leverages Eclipse's module extension capability through the use of plug-ins. Available plug-ins support the integration, with methods for test Mining, image mining and time series analysis. KNIME also integrates various other open source projects, including Machine learning Algorithms from Weka, R and JFreeChart. It supports wrappers to call other code and provides nodes, so users can run Java, Python, Perl and other code fragments. KNIME leverages the Eclipse plug-in capability -- as a result, connector extender nodes for a wide range of systems and platforms continue to be added.

KNIME Analytics Platform provides over 1000 data analytic routines, either natively or through R and WEKA, for such topics as:

- Univariate and Multivariate
- Data Mining
- Time Series
- Image Processing
- Web Analytics
- Text Mining
- Network Analysis
- Social Media Analysis

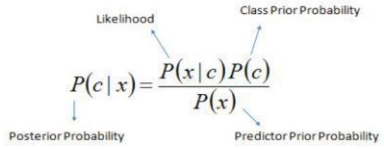
KNIME analytic work flows can be run not only through the interactive user interface but also in a batch execution mode, enabling the data analysis process to be easily integrated into your local job management and executed on a periodic basis

Naïve Bayes:

In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can

 ISSN: 2321-8169
278 – 282

be done by evaluating a closed-form expression which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.



$$P(c|X) = P(x_1|c)\times P(x_2|c)\times\cdots\times P(x_n|c)\times P(c)$$

Fig : Bayes Rule

V. Results and Discussions

The dataset used for the analysis is for five different routes i.e 300 data. It consists of the trip time of these route, number of busses allocated on each route. For calculating the delay, database used are the route details and different attributes such as expected arrival time, actual arrival time, crowd type and the route type. Using Naïve Bayes classifier these attributes are analysed.

A. Bus Frequency Analyser:

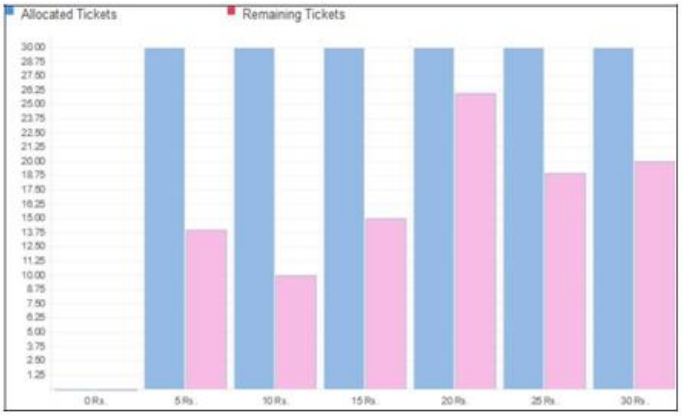


Fig 3a. Bar graph for distribution of ticket at route Nigdi-Akurdi (Ticket Count Vs Cost, Time:06:30am to 07:30am)

In this graph, if we see last three types of tickets then the utilisation of the bus is very less and hence no extra bus is required. As there are two busses allocated on this route from 06:30am to 07:30am we can also reduce the count of bus to one.

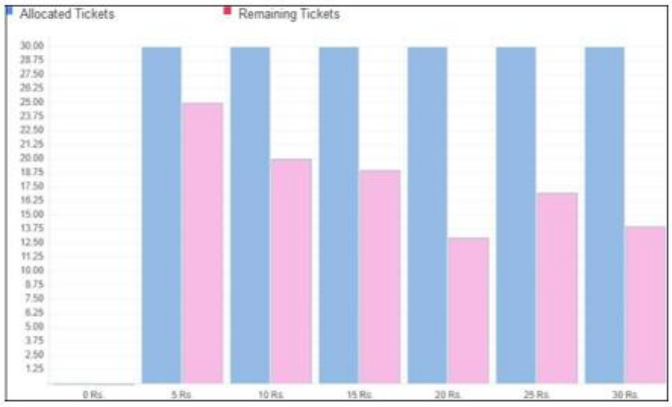


Fig 3b. Bar graph for distribution of ticket at route Nigdi-Akurdi (Ticket Count Vs Cost, Time:09:00am to 10:00am)

In this graph, if we see last three types of tickets then the utilisation of the bus is very much and hence extra busses are required. As there are three busses allocated on this route from 09:00am to 10:00am we have to increase count of the bus by one.

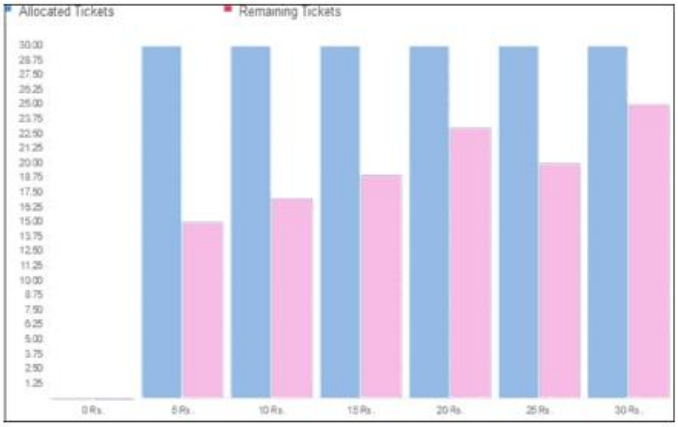


Fig 3c. Bar graph for distribution of ticket at route Nigdi-Akurdi (Ticket Count Vs Cost, Time:02:30pm to 03:30pm)

In this graph, the utilisation of the bus is very less and hence no extra bus is required. As there are two busses allocated on this route from 02:30pm to 03:30pm we have to reduce count of the bus.

Table 2. Bus Frequency Analyser Result

Route ID	Trip Time	Bus capacity	Ticket count	Change in Frequency
12	6.30am	45	34	No extra bus required
12	9am	45	88	Extra Bus required
12	2.30pm	45	20	No extra bus required

The above table represent the result for route Nigdi-Akurdi at three different time. When the ticket count exceeds the capacity to a greater extent, an extra bus needs to added on the route. This result is for route 12 at time 7.30am, 9am and 2.30 pm when there varies the density of the commuters.

B. Schedule Optimiser:

 ISSN: 2321-8169
278 – 282

Table 3. Delay Analyser Results

Row ID	Col0	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Predicti...
Row0	Route Id	Date	Day	Route Name	Scheduled Arrival Ti...	Actual Arrival Ti...	Difference	Route Type	Passenger Cro...	Class
Row1	3R	3/1/16	Tuesday	Pune Statio...	1:40	1:40	0.0	Less Crowdy	Underflow	No Delay
Row2	3R	3/1/16	Tuesday	Pune Statio...	4:40	4:40	0.0	Less Crowdy	Underflow	No Delay
Row3	3R	3/1/16	Tuesday	Pune Statio...	0:10	0:15	5.0	Moderate Cr...	Moderate	Moderate D...
Row4	3R	3/2/16	Wednesday	Pune Statio...	1:40	1:45	5.0	Moderate Cr...	Moderate	Moderate D...
Row5	3R	3/2/16	Wednesday	Pune Statio...	4:40	4:40	0.0	Less Crowdy	Moderate	No Delay
Row6	3R	3/2/16	Wednesday	Pune Statio...	0:10	0:20	10.0	Crowdy	Overflow	More Delay

The above table shows the results of the delay, where the bus delays are classified on the basis of different attributes using Naïve Bayes algorithm in KNIME tool. The accuracy obtained is 95%. Hence, Naïve Bayes classifier is best among the rest classifying algorithms for this system.

VI. Future Work

As bus management system is very vast. There are many factors in this which we can implement in future. Some of the important factors are: finding the overlapping routes and altering the bus frequencies as necessary, calculating the maintenance cost etc. By finding overlapping routes we can increase or decrease the frequency of buses on that route and get the profit for the government, also proper bus service to commuters.

VII. Conclusion

This proposed system provides proper regulation of the bus services and generate optimized schedule. Naïve Bayes algorithm has been used for classifying the delay of the bus in different ranges. The revenue generated per trip has been verified by cost estimator. This system removes the slacks in older system and provides proper analysed and optimized system.

Acknowledgement

We would like to thank Dr. Prof. J. S. Umale and Prof. Mrs. S. R. Vispute for all the help and guidance. We are very grateful to them for their kind support. Their valuable suggestions helped us a lot.

References

[1] Ranjith Ramesh, Yokesh Ezhilarasu Prasanna Ravichandran and Soma Prathibha, Member, IACSIT "Regulating Bus Management System Using Cloud Platform", International journal of e-education, e-management and e-learning, Vol. 2, No.6, December 2012.

[2] LV ZHIAN HU HAN Physics & Electronics Information Technology Department Xiangfan University Xiangfan, China "Bus Management System Based on ZigBee and GSM/GPRS".

[3] M. A. Hannan, A. M. Mustapha, A. Hussain and H. Basri, "Intelligent Bus Monitoring and Management System".

[4] Komal Satish Agarwal Computer Department MIT College Of Engineering, Pune, India "RFID Based Intelligent Bus Management and Monitoring System"

[5] Ankit Kesharwani, Vaishali Sadaphal, Maitreya Natu "Empowering Bus Transportation System Using Wireless Sensor Networks" Indian Institue Of Technology, Kanpur, Tata Research Development and Design centre, Pune, India.

[6] Hamilton, P and Suresh, S (2013). "Intelligent Agent based RFID System for On Demand Bus Scheduling and Ticketing", International Journal of Future Computer and Communication, Vol.2(5), pp.399-406.

[7] Li Zhang, Xi Ren, "A two-factor evaluation of bus delays based on GIS-T database and simulation". University of Gavle, May 2010.

[8] "Advanced Transportation Management technologies-Transit Management system."

[9] J. Owusu, F. Afukaar, and B. E. K. Prah, "Towards Improving Road Traffic Data Collection: The Use of GPS/GIS," in FIG Regional Conference, 2006.

[10] Wikipedia, "greedy algorithm, optimization, data analysis", machine learning techniques, data mining.

[11] Nikita V Shahane, Rutuja Pande, Rutuja Narkhede, S. R. Vispute, "Survey For Regulating Bus Services using Commuters Density", International Journal of Innovative Research in Computer and Communication Engineering, Vol.3, Issue 11, November 2015.