An Analysis of Bus Ticket Sales in East Bangalore

Yogalakshmi Jayabal  
j.yogalakshmi@iiitb.org

S. Rajagopalan  
raj@iiitb.ac.in

International Institute of Information Technology, Bangalore

Abstract

This paper investigates different aspects of demand modelling for bus transport systems based on the data obtained from Electronic Ticketing Machine (ETM). Nowadays, ETM’s have been introduced by many Public Transit Agencies as part of improving their operations and services. The data used in this study is the ticket sales data from the Bangalore Metropolitan Transport Corporation (BMTC)\(^1\). BMTC approximately makes 69000 vehicle trips with a traffic revenue of Rs5.17 crores everyday. The ETM data of BMTC has approximately 40 million transactions per month. This ETM data can be utilized effectively to understand passenger movement, identification of peak and off-peak hours of the day, popular Origin-Destinations, operator’s efficiency in terms of revenue generated, load-profiles at 1. route-level, 2.corridor-level, 3. Origin-Destination (OD) wise etc across Bangalore. This paper focuses on generation of Origin-Destination matrices from this ETM data to understand the user behaviour between different OD pairs, duration of peaks and off-peaks for the OD pairs across the different times of the day. This OD data will help in understanding the spatio-temporal bus ridership demand in Bangalore. The work presented in this paper provides details on the methodology for generating the OD matrix and additional inferences that are possible from the ETM data. This work also presents a number of analysis tasks that were executed, to derive information from ETM data for travel demand modelling and operational planning of public transit agencies. A major finding is that while nearly two thirds of ticket sales happen during peak period, peak periods themselves were a small fraction of the overall operating hours.

1 Introduction

Urbanization has resulted in greater demand for movement of people and goods which mandates good mobility within the city. Public Transport plays an important role in mobility in any city. Transportation Planners are often required to analyze various parameters to ensure effective services. Bangalore Metropolitan Transport Corporation (BMTC) is the public bus transit operator in Bangalore in India. There are around 6600 buses with around 2500 routes that are operated in the city. These buses are equipped with automatic vehicle location system (GPS) and electronic ticketing machines. To attract more people, public bus transport-the fleet operator

\(^1\)https://www.mybmtc.karnataka.gov.in/info-1/BMTC-Glance/en
should provide quality service to passengers. It is important to estimate the demand for public transit which in turn affects the operational policies and strategies of the public transit agency. Appropriate estimation of the peak and off-peak time, peak and off-peak loads leads to better understanding and modelling of the travel demands and operations.

The ETM is a handheld device that records the transaction when a passenger requests a ticket. The introduction of the Automated data collection source like Electronic Ticketing Machine (ETM) plays a vital role in the absence of smart cards or travel cards in Bangalore. Hence, building tools to explore this ETM data and asking the right analytic questions provides us with a better understanding of the passenger movement and therefore system’s behaviour. Bangalore has two types of ticketing system: 1. trip based tickets, and 2. pass-tickets. The pass-tickets can be one of the following: 1. Student pass, 2. Day pass, 3. Monthly pass, or 4. Senior Citizen. Every transaction in BMTC-ETM captures data like Ticket_id, waybill_Id, waybill_no, schedule_no, trip_no, etim_no, route_no, route_id etc. Using these, various key performance indicators like Total number of passengers [route-level, Daily], High boarding/alighting stops, In-vehicle passenger volume or Occupancy, Occupancy ratio, Average revenue per shift etc can be computed to know the effectiveness of the services provided.

2 Related Literature

There are many factors like speed of the bus, schedule adherence, passenger demand, travel time etc that affect the effectiveness of public travel. One of them is passenger demand. Passenger demand modelling and estimation is one of the important task in transport operations. The conventional method of data collection like household surveys, travel surveys to understand the demand are both expensive and time consuming and hence they are infrequent[Cui07]. Also, surveys would be conducted on few sample routes or links or zone and hence the comprehensive view of existing demand of a city may not be understood completely. In contrast, there is a need for frequent analysis and updating of real time traffic scenarios to improve the public transport operations. Hence, the Automated Data Collection (ADC) systems have gained importance. The Automated Data Collection (ADC) systems include Automatic Passenger Counting (APC) [Fur06], Automatic Vehicle Location (AVL) [Fur06] and Automatic Fare Collection (AFC) [Nun17]. Smart card data [Ort15], cellphone data [Dem17] and social media data are some of the other data sources that are being used now-a-days to understand the travel demand. There are a lot of literature available that explore ADC data to understand the system.

Yu et al [Sha16] have proposed to forecast bus passenger trip flow for transit route design and optimization. They have used Artificial Neural Network (ANN) to forecast the bus passenger trip flow and have validated with a dataset from China. The ANN model is based on the influence factors like each traffic zone land use (the proportion of residential, commercial and industrial traffic), accessibility to bus stations, area and distance between zones etc. They have used the OD pairs as a base from a survey that was carried out to forecast the passenger flow. Kinene [Kin09] employed Random Forest machine learning algorithm to predict the hourly demand for buses along all routes in ‘Orebro city in Sweden. They have considered factors like day of the week, weather season, time of the day, customer types etc for predicting the hourly demand for buses. Kinene also suggests that these information can be used to decide the frequency on a given route considering these factors. Cui [Cui07] in his thesis has developed an algorithm to estimate bus passenger OD-matrix using the data from Automatic Vehicle location (AVL) and Automated Fare Collection (AFC). Initially, a single route OD-matrix is estimated from a seed matrix that is derived from AFC data. Then Iterative Proportional Fitting and Maximum Likelihood Estimation (MLE) techniques are used to estimate OD matrix for single routes. Then network level OD-matrix are estimated. Ji et al [Yji17] have proposed Hierarchical Bayesian model to estimate the trip-level OD flows and a period-level OD flow from the samples OD flow data collected by the WIFIsensors and the fareboxes. They have used bus load and average journey length to reflect indirectly on the accuracy of their proposed OD estimation method. Li et al [Dli11] also have proposed an OD estimation matrix for each route using the data collected from the farebox. They have presented an OD estimation model based on trajectory search algorithms to track passenger trips, using the pre-processed smart card data. They have used one day smart card data from Jinan city. They also suggest that the estimated OD pairs can be used to evaluate route network and optimize bus scheduling. Janine [Jan08] has also proposed to construct Automated Bus Origin-Destination matrix using farebox and AVL data.

Most of the works in literature for travel demand analysis are based on the Automatic Fare Collection (AFC) or Farebox. The data that we have analyzed is from Electronic Ticket Machines (ETM) which has few more details than that of Farebox. A few works are available in literature that analyze data from ETM machines. Cyril et all [Cyr17] have analyzed ETM data of Kerala State Road Transport Corporation for 6 depots in Trivandrum.
city for modelling intercity public transport demand to predict the number of trips on a given day. Kalanidhi et al. [Kal13] have used ETM data along with OD pattern of travels taken from Chennai City Traffic Study of Chennai Metropolitan Development Authority, passenger opinion survey and GPS data to study the accessibility of urban transportation networks and assessing its influence on the public transport ridership. Wang et al. [Wan11] have proposed a methodology to infer bus passenger travel behaviour, ODpair inference using the smart card transactions and AVL data in London.

In this study the objective is to analyze the ODpairs to understand the passenger distribution and hence to obtain the temporal and spatial variation in ridership and hence passenger travel characteristics. This paper focuses on generating the passenger movement from the ETM data and some of the key performance indicators like Total number of passengers [route-level, day-level], load profiles of routes, identification of peak and off-peak hours based on the number of tickets sold.

3 East Bangalore - A case study

This section presents the area of study and provides details on the data collected and the methodology used to generate the ODpairs. Bangalore is the fifth largest urban city in India with a population of about 8.5 million as of 2011 with an area of 709 sq km. The below map shows the boundary of Bangalore and the portion highlighted in red is the study region which is East-Bangalore2 BMTC is the government agency that operates public transport bus service in Bangalore. It has different types of services like 1. General, 2. Samartha, 3. Suvarna, 4. BIG 10, 5. Big Circle, 6. Atal Sarige, 7. Vajra, 8. Vayu vajra, 9. Marcopolo and Corona AC, 10. Metro Feeder and 11. Hop On Hop Off. These BMTC buses are operated from 48 depots3 within the city and are numbered from 1 to 48. In some BMTC services, the tickets are issued using a Electronic ticket machine (ETM) and in few other services, the manual (pre-printed) tickets are issued. This study analyzes both ETM data and manual ticket data sold in buses operated from four depots 6, 25, 28 and 41, which cater to the East Bangalore population. In the introduction, it was mentioned that two types of tickets - trip based tickets and pass tickets are available in BMTC. This study focuses only on trip based tickets as the information about the travel made by pass ticket holders is not available. It is assumed that the analysis results could be a representative of the total public transit passengers. BMTC is the government agency that operates public transport bus service in Bangalore. It has different types of services like 1. General, 2. Samartha, 3. Suvarna, 4. BIG 10, 5. Big Circle, 6. Atal Sarige, 7. Vajra, 8. Vayu vajra, 9. Marcopolo and Corona AC, 10. Metro Feeder and 11. Hop On Hop Off. These BMTC buses are operated from 48 depots4 within the city and are numbered from 1 to 48. In some BMTC services, the tickets are issued using a Electronic ticket machine (ETM) and in few other services, the manual (pre-printed) tickets are issued. This study analyzes both ETM data and manual ticket data sold in buses operated from four depots 6, 25, 28 and 41, which cater to the East Bangalore population. In the introduction,  

---

2 East-Bangalore was identified as study region since ticket sales data was predominantly available for this region from the 4 depot’s data and is not exclusive of east bangalore region.

3 https://www.mybmtc.karnataka.gov.in/info-1/Depots/en

it was mentioned that two types of tickets - trip based tickets and pass tickets are available in BMTC. This study focuses only on trip based tickets as the information about the travel made by pass ticket holders is not available. It is assumed that the analysis results could be a representative of the total public transit passengers.

### 3.1 Electronic Ticketing Machine

The Electronic ticket machine (ETM) is a handheld device which weighs about 800gms. They are GPRS\(^5\) enabled ETM which transmits ticket data to ITS server every 5 minutes. The figure 2 shows both ETM and manual ticket. When a ticket is issued using the ETM, there are as many details as 50 parameters, that are sent to the data server that is placed in BMTC data center. The parameters that we have analyzed are given in Table 1.

### 3.2 Data Collection and Pre-processing

The ETM data along with data in manual tickets for the month of December 2018 and July 2019 from depots 6, 25, 28 and 41 was provided to us for analysis. Each data file size was between 250MB to 800MB. Each data file had 59 parameters including: ticket_id, waybil_Id, waybill_no, schedule_no, trip_no, etim_no, route_no, route_id, transaction_no, ticket_no, ticket_type_short_code, ticket_sub_type_short_code, ticket_from_stop_id,

---

3.3 Pre-Processing of data files

One data file for each depot (6, 25, 28, 41) was provided consisting of ticket sales of all the routes that operates from the depot. This data file of each depot is processed to check for any inconsistent data type values, spurious rows etc. The data processing steps followed are:

1. From each depot data, generate separate files for every route.
2. Simultaneously, extract only the required parameters of Table 1 for every route.
3. The route-level data files for each depot and month (December 2018 and July 2019) are extracted separately. This extracted route file data size is of the order of few KBs and becomes the base data for further analysis.

A snapshot of the generated route-level ticket sales data of route 139 is shown in the figure 3.

4 Data Analysis

The route-level data files extracted for each depot forms the base data for all our analysis tasks. The following analysis were carried out on these data:

1. Total Number of Passengers route wise and day wise.
2. Hourly occupancy of passengers route wise, day wise and vehicle wise.
3. Load profile - Occupancy trip wise and by stop wise
4. Identification of the location and time of peaks and valleys in the distribution of ticket sales month wise and hence check for any patterns.
5. Distribution of users based on identified Origin-Destination pairs.
4.1 Total number of passengers

The total number of passengers route-level trip-wise, schedule-wise and day-wise computed using a Python script. The sample output for some of the routes are as shown in table 2:

Table 2: Sample output of computed Total number of Passengers

<table>
<thead>
<tr>
<th>Route_no</th>
<th>Vehicle_no</th>
<th>Ticket_date</th>
<th>Total_no_of_passengers</th>
<th>Total_ticket_amount</th>
<th>Shift_no</th>
<th>Trip_no</th>
<th>Scheduled_Start_time</th>
<th>Scheduled_End_time</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBS-13K</td>
<td>KA01FA1881</td>
<td>12/12/2018</td>
<td>3</td>
<td>28</td>
<td>2</td>
<td>1</td>
<td>2018-12-12 12:09:27</td>
<td>2018-12-12 13:06:39</td>
</tr>
<tr>
<td>SBS-13K</td>
<td>KA01FA1881</td>
<td>12/13/2018</td>
<td>4</td>
<td>34</td>
<td>2</td>
<td>1</td>
<td>2018-12-13 11:47:51</td>
<td>2018-12-13</td>
</tr>
<tr>
<td>SBS-13K</td>
<td>KA01FA1881</td>
<td>12/18/2018</td>
<td>6</td>
<td>50</td>
<td>2</td>
<td>1</td>
<td>2018-12-22 12:04:44</td>
<td>2018-12-22 12:32:31</td>
</tr>
<tr>
<td>500-QG</td>
<td>KA57F1926</td>
<td>12/29/2018</td>
<td>29</td>
<td>471</td>
<td>2</td>
<td>1</td>
<td>2018-12-29 07:25:45</td>
<td>2018-12-29 08:01:52</td>
</tr>
<tr>
<td>500-QG</td>
<td>KA57F1926</td>
<td>12/2/2018</td>
<td>17</td>
<td>281</td>
<td>2</td>
<td>1</td>
<td>2018-12-02 15:46:47</td>
<td>2018-12-02 17:06:30</td>
</tr>
</tbody>
</table>

4.2 Hourly Occupancy

The term occupancy of a bus is defined as the following. It is given by:

\[
\text{Occupancy} = x + y, \quad (1)
\]

\[ x = \text{Number of people who are inside the bus when it arrives at a stop}, \]
\[ y = \text{Number of people boarding the bus at that stop} - \text{Number of people alighting at that stop} \]

The occupancy at a route level helps to understand the passenger demand in the route at different times of the day. It also helps to understand the peak and off-peak times of the given route. The figure 4 gives the hourly occupancy of route: V-500D between December 3rd – 7th and figure 5 given the hourly occupancy of route:SBS-1K between December 3rd – 7th.

It could be observed that the route: V-500D have 2 clear peaks, one in morning between 8.30 A.M to 10.30 A.M and one in evening between 17.30 P.M to 20.30 P.M. Whereas, the route:SBS-1K has a sharp morning peak, but the evening peak relatively blunt compared to the morning peak. These give the times when the routes are most used. Another important factor to observe is that the peak(morning/evening) load is 2 – 3 times that of the load in off-peak times. This pattern is consistent across all days in the week as shown. The similar pattern is also observed across all weeks of the month. The figure 6 shows the hourly occupancy in V-500D for 4 weeks in December 2018. We could compute Utilization by examining whether the occupancy is < 100% or > 100%. This piece of information would be a valuable feedback to be considered by the operations and planning team of public transit agency while scheduling.

4.3 Load Profile

The load profiles of a route gives a much detailed information such as the trip-wise, stop-wise and time-wise occupancy. These also allow us to infer the trip times of various trips made throughout the day and how they vary in peak and off-peak hours of the day. The figure 7 show the different trips made by the route:335C in December 2nd – 7th. It can be observed that the trips that start between 8.00 A.M and 10.30 A.M take slightly little longer time to complete the trip compared to other trips made in the day.
Figure 4: Hourly occupancy on V-500D

Figure 5: Hourly Occupancy on SBS-1K
Figure 6: Hourly occupancy on V-500D on 4 weeks in December 2018

Figure 7: Load Profile of 335C between 2-7th in December 2018
4.4 Generate Origin-Destination pairs

The Origin-Destination pairs from the route-level ticket sales data have to be generated to understand the spatio-temporal passenger distribution across Bangalore. This also helps in identification of the peaks and valleys in the distribution. The steps followed to generate the OD pairs from the route-level ticket sales data are given below.

1. From the route-level ticket sales data, the distinct Origin-Destination (OD) pairs for every 15 minutes are extracted along with the number of passengers and ticket amount.

2. The OD pairs (same OD pairs could occur in multiple routes) for every 15 minutes for each week (only for weekdays) are generated in separate files.

3. The week-wise files from step no 2 are generated for the months of December 2018 and July 2019 separately. There are four weeks in December 2018: Week 1: December 3rd to 7th, Week 2: December 10th to 14th, Week 3: December 17th to 21th and Week 4: December 24th to 28th. There are five weeks in July 2019: Week 1: July 1st to 5th, Week 2: July 8th to 12th, Week 3: July 15th to 19th, Week 4: July 22nd to 26th and Week 5: July 29th to 31st.

4. Once the week-wise OD pairs for each depot are obtained, the same OD pairs across four depots in the same week and in the same time interval (i.e. 24 hours of the day are divided into 15 minutes) are combined.

5. From step no 4, one file for each week of December 2018 and July 2019 are output.

6. The passengers count of same OD pair across time intervals (i.e. every 15 mins) are summed up to get the total number of passengers for that OD pair in that week.

7. The OD pair file from step 6 got for each week is sorted in descending order according to the total number of passengers.

8. The sorted OD pair file is parsed to extract the top 100 OD pairs.

9. The top 100 OD pairs from step 8 are analyzed for duration of peaks and the total number of tickets sold in the peak duration.

Top 5 of the generated OD pairs for the 2 weeks of December are shown in table 3. From table 3, it can be observed that most of the OD pairs from Week 1 of December are occurring in other week of December as well. This informs that passenger movement across weeks remain similar. The next step is to examine the ticket sales in these top 100 OD pairs for the peak/off-peak times of ticket sales. The peak and off-peak times are identified using a Python script. Though in many of the routes, there are only 2 peaks (morning and evening peak) observed, in many other routes multiple peaks are observed. Also, since the maximum passenger count
<table>
<thead>
<tr>
<th>From_bus_stop_id</th>
<th>From_stop_name</th>
<th>To_bus_stop_id</th>
<th>To_stop_name</th>
<th>Passenger_count</th>
<th>Total_ticket_amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>134</td>
<td>Kundalahalli Gate</td>
<td>1629</td>
<td>Marathahalli</td>
<td>6401</td>
<td>32316</td>
</tr>
<tr>
<td>6930</td>
<td>AECS Layout Cross</td>
<td>2050</td>
<td>Sathya Sai Hospital</td>
<td>5341</td>
<td>74001</td>
</tr>
<tr>
<td>134</td>
<td>Kundalahalli Gate</td>
<td>154</td>
<td>NAL Manipal Hospital</td>
<td>4313</td>
<td>77013</td>
</tr>
<tr>
<td>2280</td>
<td>Hope Farm (Towards Varthuru)</td>
<td>140</td>
<td>White Field Post Office</td>
<td>4210</td>
<td>21143</td>
</tr>
<tr>
<td>7030</td>
<td>Bellanduru</td>
<td>2619</td>
<td>Marathahalli Bridge</td>
<td>4171</td>
<td>79290</td>
</tr>
</tbody>
</table>

Week - 1: December 3rd to 7th

<table>
<thead>
<tr>
<th>From_bus_stop_id</th>
<th>From_stop_name</th>
<th>To_bus_stop_id</th>
<th>To_stop_name</th>
<th>Passenger_count</th>
<th>Total_ticket_amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>134</td>
<td>Kundalahalli Gate</td>
<td>1629</td>
<td>Marathahalli</td>
<td>6346</td>
<td>32093</td>
</tr>
<tr>
<td>6930</td>
<td>AECS Layout Cross</td>
<td>2050</td>
<td>Sathya Sai Hospital</td>
<td>5285</td>
<td>73504</td>
</tr>
<tr>
<td>134</td>
<td>Kundalahalli Gate</td>
<td>8456</td>
<td>Kempegowda Bus Station</td>
<td>4439</td>
<td>113625</td>
</tr>
<tr>
<td>2280</td>
<td>Hope Farm (Towards Varthuru)</td>
<td>140</td>
<td>White Field Post Office</td>
<td>4233</td>
<td>21219</td>
</tr>
<tr>
<td>134</td>
<td>Kundalahalli Gate</td>
<td>154</td>
<td>NAL Manipal Hospital</td>
<td>4105</td>
<td>72389</td>
</tr>
</tbody>
</table>

Week - 2: December 10th to 14th
between different ODpairs varies (as shown in figure 8), there is a need to systemically identify the peaks and off-peaks between different ODpairs. The steps of the algorithm to identify the peak and off-peak times of the day is given in 1.

**Algorithm 1: Identification of peak time and duration**

| Data: Top 100 ODpairs data file, Weekdf = ODpair data file of a week |
| Result: ODpair, peaktime, peak_duration, number_of_passengers_in_peak_duration |

Abbreviations: pxc = passenger_count;

1. foreach odpair ∈ top100odpairs do
2. peak_pxc = Weekdf . max_passenger_count ;
3. foreach row ∈ Weekdf do
4. Iterate through the ODpair week data file containing passenger count in 15 mins time interval, to identify and retrieve all the time intervals at which the passenger count is atleast 70% of the peak_pxc.
5. Store for each ODpair = peaktime, peak_duration, no_of_passengers_in_peakduration ;
6. end
7. return ODpair, peaktime, peak_duration, no_of_passengers_in_peakduration;

The algorithm 1 is executed for four weeks ODpair data files of December and five weeks ODpair data files of July. The algorithm 1 provides 3 outputs for each week. They are for each ODpair, the peak passenger count, peak duration, time at which peak occurred. Additionally, the total travel time for every ODpair is computed in every week. Using these outputs the following two ratios are computed for every ODpair for every week.

\[
\text{peak}_{-}\text{pxc}_{-}\text{ratio} = \frac{\text{Average\_weekly\_peak\_passenger\_count}}{\text{Average\_weekly\_passenger\_count}} \quad (2)
\]

\[
\text{peak\_time\_ratio} = \frac{\text{Total\_peak\_duration\_of\_week}}{\text{Total\_travel\_time\_of\_week}} \quad (3)
\]

The week 4 of December 24th to 28th being a holiday week and Week 5 of July 29th to 31st having only 3 days are ignored for peak behaviour analysis. The sample output of peak_pxc_ratio for 10 ODpairs for all the weeks considered for analysis in December 2018 and July 2019 are shown in table 4. From the table 4 the following observations can be made.

1. The percentage of ticket sales in these ODpairs across weeks in both months are similar.
2. The variance in the peak ticket sales percentage is also less than or equal to 5%.

The travel time was then computed to examine the duration for which these peak ticket sales occurred. The peak_time_ratio as in eqn:3 was computed. The peak_time_ratio across weeks also remains similar. They are as shown in table 5. These peak time ratio are very low indicating that the time for which the peak occurs is very small. This behaviour was observed across weeks in both the months. This also is an evidence that the peak ticket sales are really high compared to the off-peak ticket sales. The top 10 ODpairs for which the peak ticket sales was observed is presented in table 6.

5 **Discussion**

The Table 6 shows that more than 30% of ticket sales occurs in the peak times. The ticket sales in some of the ODpairs goes as high as 60%. The column Mean under peak_pxc_ratio shows the mean of peak_pxc_ratio of 7 weeks (3 weeks in December and 4 weeks in July). Similarly the Mean under peak_time_ratio shows the mean of peak_time_ratio of 7 weeks. The peak duration are very less compared to the total trips time. This behaviour needs to be considered while scheduling. Jara Díaz et al [Ser17] have provided an analytical explanation that in urban cities — the number of buses and vehicle size is determined by the characteristics of demand during peak period and adjusting frequencies for other off-peak period whose characteristics are very different from that of the peak duration. They have shown numerically that minimizing social costs(operator and user) for the whole day results in a larger fleet of smaller size buses than if only peak period is considered for determining the fleet size and capacity.
### Table 4: Sample Computed peak\(_{pxc}\) ratio for some ODpairs

<table>
<thead>
<tr>
<th>odpair_id</th>
<th>odpair_stopnames</th>
<th>December 2018</th>
<th>July 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>peak(_{pxc}) ratio (as in eqn:2)</td>
<td>W1</td>
<td>W2</td>
</tr>
<tr>
<td>134_1629</td>
<td>Kundalahalli Gate, Marathahalli</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td>6930_2050</td>
<td>AECS Layout Cross, Sathya Sai Hospital</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>7030_2619</td>
<td>Bellanduru, Marathahalli Bridge</td>
<td>0.22</td>
<td>0.23</td>
</tr>
<tr>
<td>2280_140</td>
<td>Hope Farm (Towards Varthuru), White Field Post Office</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>1218_1234</td>
<td>Sony World 80ft Road, Koramangala, Dhoopanahalli</td>
<td>0.23</td>
<td>0.23</td>
</tr>
</tbody>
</table>

### Table 5: Sample Computed pxc\(_{time}\) ratio for some ODpairs

<table>
<thead>
<tr>
<th>odpair_id</th>
<th>odpair_stopnames</th>
<th>December 2018</th>
<th>July 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pxc(_{time}) ratio (as in eqn:3)</td>
<td>W1</td>
<td>W2</td>
</tr>
<tr>
<td>134_1629</td>
<td>Kundalahalli Gate, Marathahalli</td>
<td>0.0404</td>
<td>0.0577</td>
</tr>
<tr>
<td>6930_2050</td>
<td>AECS Layout Cross, Sathya Sai Hospital</td>
<td>0.029</td>
<td>0.0193</td>
</tr>
<tr>
<td>7030_2619</td>
<td>Bellanduru, Marathahalli Bridge</td>
<td>0.0307</td>
<td>0.0168</td>
</tr>
<tr>
<td>2280_140</td>
<td>Hope Farm (Towards Varthuru)</td>
<td>0.0166</td>
<td>0.0161</td>
</tr>
<tr>
<td>1218_1234</td>
<td>Sony World 80ft Road, Koramangala, Dhoopanahalli</td>
<td>0.0228</td>
<td>0.0058</td>
</tr>
</tbody>
</table>

### Table 6: Top 10 bus stops with high peak\(_{pxc}\) ratio in peak time and the variance in peak time is very small

<table>
<thead>
<tr>
<th>odpair_id</th>
<th>odpair_stopnames</th>
<th>December 2018</th>
<th>July 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>peak(_{pxc}) ratio</td>
<td>mean</td>
<td>variance</td>
</tr>
<tr>
<td>2092_6914</td>
<td>Kadugodi Bus Station, Pattandur Agrahara Gate</td>
<td>0.66</td>
<td>0.0277</td>
</tr>
<tr>
<td>2092_140</td>
<td>Kadugodi Bus Station, White Field Post Office</td>
<td>0.46</td>
<td>0.0127</td>
</tr>
<tr>
<td>9010_9288</td>
<td>Police Station, Indiranagar, Military Bridge</td>
<td>0.43</td>
<td>0.0104</td>
</tr>
<tr>
<td>2092_154</td>
<td>Kadugodi Bus Station, NAL Manipal Hospital</td>
<td>0.42</td>
<td>0.0068</td>
</tr>
<tr>
<td>2619_2581</td>
<td>Marathahalli Bridge, Dodda Nekkundi (Towards Hebbala)</td>
<td>0.39</td>
<td>0.0055</td>
</tr>
<tr>
<td>5557_2595</td>
<td>Kadabisanahalli, Bellandur City Light Apartment</td>
<td>0.33</td>
<td>0.0064</td>
</tr>
<tr>
<td>7030_5228</td>
<td>Bellanduru, Kadabisanahalli</td>
<td>0.32</td>
<td>0.0088</td>
</tr>
<tr>
<td>403_6919</td>
<td>Pattandur Agrahara Gate, Hope Farm (Towards Hoskote)</td>
<td>0.32</td>
<td>0.0081</td>
</tr>
<tr>
<td>2055_6929</td>
<td>White Field TTMC (Vydehi Hospital), AECS Layout Cross</td>
<td>0.31</td>
<td>0.0024</td>
</tr>
<tr>
<td>134_2595</td>
<td>Kundalahalli Gate, Bellandur City Light Apartment</td>
<td>0.31</td>
<td>0.0009</td>
</tr>
</tbody>
</table>
The analysis tasks based on the ticket sales data as shown in this paper also show that the peak behaviour is very different from the off-peaks in the system. Hence, the process of planning and scheduling needs to consider both the peak and the off-peaks in the urban transit system.

6 Conclusion

The use of automatic data collection techniques various advantages. This study investigates the potential of ETM data and in general ticket sales data for the purposes of operations and planning. The ticket sales data can provide insights into quantitative measures for operational performance. This paper has shown a methodology for generating ODmatrices from ticket sales data along with various other analytical tasks. This paper also shows the effectiveness of ticket sales data for understanding various important performance indicators of the public transit agency. Future works involve coming up with schedule modelling based on Jara Díaz study.

6.0.1 Acknowledgements

The authors thank BMTC for sharing their data to us for analysis. This research received funding from the Netherlands Organisation for Scientific Research (NWO) in the framework of the Indo Dutch Science Industry Collaboration programme [NWO, Den Haag, PO Box 93138,NL-2509 AC The Hague,The Netherlands]. We are thankful to NWO, Royal Shell and Prof. Sebastian Meijer, the Principal Investigator of this project.

References


