Passenger trips analysis determined by processing validation data of the electronic tickets in public transport

A I Fadeev, S Alhusseini

Siberian Federal University, 79, Svobodny Avenue, Krasnoyarsk, 660041, Russia

E-mail: fai@ak1967.ru

Abstract. Based on the set of trip-legs obtained by the algorithm described in our earlier work, trip-legs were sorted by their start or end time for each ticket, we managed to identify the set of passenger correspondence which consists of one or several trip-leg based on the analysis of the transfer conditions. Advanced and in-depth study of passenger correspondence is pivotal for solving transport-planning tasks and avoiding any mistake happened at some point at trip identification. In this work, we will process two matters: distribution transfers by time and identify short activities trips. This work shows that for a two-hour period of free transfer, the share of free trips will be about 17% of the total number of trips. In details, 12% of passengers per day will take one free trip and the share of passengers who will perform two or more free transfers is about 5% of trips. For a one-hour free transfer period the share of free transfer will be about 10%. Passengers with one free transfer will make about 8% of trips. Also, after identifying the short activities trips and dividing these trips into separate trips, the share of trip without transfer increased by 7%. The transfer factor decreased by 0.04.

1. Introduction

According to electronic tickets validation data it is possible to determine transit demand and form a set of passenger trips when it is necessary. The passenger trip (transit rider) from the origin stop to the destination stop is determined by analysing the sequence of the electronic tickets validation data, which were recorded and stored by automated fare collection system. Passenger trip is formed from passenger trip segments (trip-leg), the passenger's correspondence can consist of one or several trip-legs.

Based on the set of trip-legs received from the algorithm described in our earlier work [1, 2], which were sorted by their start or end time for each ticket, we managed to identify set of passenger trips consisting of one or several trip-leg based on the analysis of the transfer conditions [4-8] (figure 1):
- Walk access link;
- Transfer penalty.

In our previous work, we used a transfer conditions for discrimination between trip segment and trip, where exceeding one of these two limits is sufficient to consider trip-leg as trip. The first condition was walk access link [3] where we assumed that if the distance between alighting stop and the next boarding stop exceed walk access link boundary, then the alighting stop would consider as destination stop for the previous trip and the next boarding as origins for next trip. Advanced and in-depth study of passenger trips determined by processing validation data of the electronic tickets in public transport is pivotal for solving transport-planning tasks and avoiding any
mistake happened at some point at trip identification. In this work, we will process two matters: identify short activities trips and distribution transfers by time.

2. **Identify short activities trips**

We observed some passenger trips, the purpose of which is to distinguish short actions at transfer points (often the duration of such actions falls within the transfer penalty) so our algorithm fails to identify these trips and combines them in one trip. We will call such trips as short activities trips. An example of short activities trips is given in figure 2 and table 1. With a considerable length of trip consisting of several trip-legs, the origin and destination stops are located within walk access link.

![Figure 1. The algorithm for determining passenger trips](image)

The short activities trips are required to be divided into separate trips in order to correctly determine transit demand. In work [4], the following criteria are proposed for identifying short activities trips:

- The time between the alighting for previous trip segment and the boarding for the current trip segment (transfer penalty) exceeds 30 minutes;
- two successive trip segments are made in the same route;
- the length of the trip (compound trip segments) is more than two times the Euclidean distance between the origin and destination stops.

The first criterion (transplantation time) is accepted and already was included in our algorithm; the second criterion proposed in [4] do not fully reflect the real situation. In many cases, the passenger’s trip cannot consider as short activities trip, even if its compound trip-legs length is significantly greater than the Euclidean distance between the origin and destination stops. An example of such trip is given in figure 3: the trip was made from stop 75 to stop 103 with a transfer at stop 66 to stop 95.
significant excess of the length of the trip over the Euclidean distance between the origin and destination stops was due to the configuration of the route network of the city.

Figure 2. Scheme of short activities trip: start at point 549 and end point 547 which are within walk access link (Schemes of trip and routes obtained using the Yandex.Maps)

Table 1. Trips-legs list for short activities trip

| Id trip-leg | Rout and dir. | Id boarding stop | Boarding time (hh:mm) | Id alighting stop | Alighting time (hh:mm) | Length, (km) | Transfer penalty, (h) |
|-------------|---------------|-----------------|------------------------|------------------|------------------------|--------------|----------------------|
| 13141301    | 53B           | 549             | 11:59                  | 615              | 12:11                  | 4,4          | 0,00                 |
| 13141300    | 49B           | 616             | 12:20                  | 881              | 13:24                  | 19,4         | 0,14                 |
| 13141318    | 20B           | 592             | 13:42                  | 418              | 14:46                  | 23,7         | 0,31                 |
| 13141299    | 2B            | 418             | 14:53                  | 175              | 15:28                  | 10,0         | 0,11                 |
| 13141332    | 2A            | 10              | 15:58                  | 212              | 16:21                  | 8,9          | 0,50                 |
| 13141319    | 2B            | 213             | 16:33                  | 641              | 17:47                  | 24,7         | 0,20                 |
| 13141283    | 83B           | 640             | 17:52                  | 1006             | 18:12                  | 6,9          | 0,09                 |
| 13141298    | 85A           | 1006            | 18:14                  | 384              | 19:29                  | 26,2         | 0,04                 |
| 13141331    | 85B           | 384             | 19:38                  | 547              | 20:57                  | 27,7         | 0,15                 |

Figure 3. Illustration of trip, which its length 4.9 times longer than the Euclidean distance between the origin and destination stops

In addition, the third criterion, which based on the identification of successive trips along the same route, is not cover all cases. In figure 4 an illustration of successive trip segments made in different
routes but on the same network section. Obviously, these trip segments should be considered as two trips, despite the fact that the transfer penalty (at stops 198, 199) was less than 30 minutes (14.4 minutes) and formally they were made in different routes.

**Figure 4.** Illustration of successive trips made in different routes on the same network section

**Figure 5.** Illustration of successive trip segments, which should be considered as separate trips according to the criterion of walk access link between the origin and destination stops

In this paper, the following criteria are proposed for identifying short activities trips. The passenger trip is considered as a short activities trip, which is subjected to fragmentation if at least one of the following conditions is met:

- Transfer penalty exceeds 30 minutes;
- the origin and destination stops of trip are located within walk access link (see figure 5);
- there are more than three trip-legs in the passenger trip (2 transfers);
- the length of trip is more than 2 times the shortest length along the route network between the origin and destination stops of the trip, this coefficient value was established on the basis of a selective analysis of trips with transfers along the route network of the city of Krasnoyarsk for April 2019;
- trip segments were made on the same network section in different directions (see figure 4): the final (alighting) stop of the later segment is within walk access link from the stopping-point of the previous segment or the starting (boarding) stop of the previous segment is within walk access link from the stopping-point of the later segment (see figure 6).

**Figure 6.** Illustration of passenger trip over a network section in different directions

A) the alighting stop of the later segment is within walk access link from the intermediate stop of the previous segment

B) the boarding stop of the previous segment is within walk access link from the intermediate stop of the later segment

In table 2 shows the distribution of short activities trips determined in accordance with the proposed indicators (criteria). The table shows that 91% of passenger trips is without transfer (consists of one segment), 8% of trips is made with one transfer and the remaining about 1% of trips with two or more transfers. The largest number of passenger trip segments is 8 (7 transfers), in April 2019 there were 2 such trips.
The criterion of transfer penalty (0.5 hours) was used in the algorithm of formation passenger trips (see figure 1), for this reason, in table 2 data on the criterion of transfer penalty is not given.

According to the criterion of walk access link, about 14% of trips with transfers (consisting of 2 or more segments) were revealed as short activities trips. For these trips, the origin and destination stops were located within walk access link (less than 500 m), while the total length of trips is a few kilometers (tens of kilometers).

Table 2. Short activities trips structure identified by proposed Criteria

| Segment | TOTAL | Walk access link | Trans. over the same net. sec. | actual to shortest possible length | TOTAL |
|---------|-------|------------------|-------------------------------|---------------------------------|-------|
|         | Amount | Weight (%)       | Amount | Weight (%)       | Amount | Weight (%)       | Amount | Weight (%)       |
| 1       | 2154165 | 91,4             | 25557  | 13,5             | 71897  | 38,0             | 54487  | 28,8             | 79716  | 42,1             |
| 2       | 189363  | 8,0              | 12893  | 6,8              | 62081  | 32,5             | 11121  | 5,9              | 65365  | 36,4             |
| 3       | 10226   | 0,4              | 1669   | 8,0              | 3038   | 15,2             | 1333   | 7,0              | 5204   | 29,0             |
| 4       | 1689    | 0,1              | 615    | 36,4             | 1642   | 97,2             | 1446   | 85,6             | 1689   | 100,0            |
| 5       | 146     | 0,0              | 54     | 37,0             | 146    | 100,0            | 146    | 100,0            | 146    | 100,0            |
| 6       | 27      | 0,0              | 4      | 14,8             | 27     | 100,0            | 27     | 100,0            | 27     | 100,0            |
| 7       | 5       | 0,0              | 1      | 20,0             | 5      | 100,0            | 5      | 100,0            | 5      | 100,0            |
| 8       | 2       | 0,0              | 2      | 100,0            | 2      | 100,0            | 2      | 100,0            | 2      | 100,0            |
| *       | 201458  | 27900            | 13,8   | 81922            | 40,7   | 61816            | 30,7   | 90162            | 44,8   |                 |
| **      | 2355623 | 8,6              | 27900  | 1,2              | 81922  | 3,5              | 61816  | 2,6              | 90162  | 3,8              |

Note. * - trips with transfers; ** all trips.

More than 40% of the total trips with transfers were classified as short activities trips by the criterion of transference over the same network section and more than 30% by the criterion of the actual length of trip to the shortest possible length (in accordance with the routes network configuration) ratio. When determining the possible trip length, trips with one segment (without transfer) or with two segments (with one transfer) was considered. Trips with a large number of segments are classified as short activities trips according to the criterion of the number of transfers.

3. Distribution transfers by time
We will study the distribution of transfers and determine their preparation during a specific period by using validation data of the electronic tickets. This study can be used to impose a reduced fare on the second trip segment, which occur after transfer during a specific time period or even provide a period of time in which the transfers are free, which gives our study the importance as it determines the number of trips that will be discounted or even free.

In table 3 is given the distribution of electronic ticket holders by the number of carried out trips per day.
Table 3. Distribution of electronic ticket holders by the number of carried out trips per day

| Num. of trips | Num. of Passenger-day | Weight (%) | Num. of trips | Num. of Passenger-day | Weight (%) |
|---------------|-----------------------|------------|---------------|-----------------------|------------|
| 1             | 876131                | 34.5       | 11            | 54                    | 0.0        |
| 2             | 1134825               | 44.6       | 12            | 25                    | 0.0        |
| 3             | 315165                | 12.4       | 13            | 11                    | 0.0        |
| 4             | 154731                | 6.1        | 14            | 9                     | 0.0        |
| 5             | 40250                 | 1.6        | 15            | 3                     | 0.0        |
| 6             | 14041                 | 0.6        | 17            | 1                     | 0.0        |
| 7             | 4204                  | 0.2        | 19            | 2                     | 0.0        |
| 8             | 1544                  | 0.1        | 24            | 1                     | 0.0        |
| 9             | 478                   | 0.0        | 31            | 1                     | 0.0        |
| 10            | 145                   | 0.0        | Total         | 2541621               | 100.0      |

The table shows that 34% of passengers using an electronic ticket make one trip a day, about 47% make two trips, 12% make three trips, etc. There is a small amount of passengers, which are carried out more than 5 trips per day. The proportion of such passengers is about 0.8%.

Consider the distribution of free or discounted fare trips under various conditions of free transfer’s period. We will consider two options for the condition of free transfers: within two hours (a) and within one-hour (b). We highlighted the trips with transfers and denote trips, for which fare should be debited or free of charge based on the conditions:

\[
\begin{align*}
\tau & = \min(t_i), i = 1, n; \\
\tau & = \min(t_i), i > \max(t_j) + \tau, i = 1, n; j = 1, n.
\end{align*}
\]

Where: \( n \) - is the number of trips for ticket per day; 
\( \tau \) - Period of free transfer, hour.

In accordance to (3.1), each first trip of the day is considered paid and then each trip, the start time of which exceeds the interval allotted for a free transfer also is considered paid.

4. Conclusion

According to table 2 on the public network of the Krasnoyarsk city, more than 8% of passenger trips currently are tagged as trips with transfers. Of these, about half (3.8%) are short activities trips, i.e. trips carried out for specific (short) purposes at transfer points. Moreover, in many cases, the conditions of match trip segments in time and walk access link are fulfilled. From table 2 shows that the sum of shares for each criteria exceeds the total share of short activities trips. This is because some trips satisfy more than one indicator, for example, the criterion of walk access link and the criterion of transference over the same network section are linked.

In table 4 a set of trips and trip segment (trip-leg) are given with and without taking into account the existence of short activities trips in the structure of transit demand. The table shows that initially in the structure of passenger trips, the share of trips without transfer amounted to about 84% of the total number of trips. The transfer factor (i.e., the ratio of the all trip segments to trips number) amounted to 1.09 (about 9% of passengers carry out transport needs with transfers).

After identifying the short activities trips and dividing these trips into separate trips, the share of trip without transfer increased to 91%. The transfer factor decreased to 1.05, i.e. only about 5% of passengers’ trips consist of two or more segments.

In table 5 provides a set of free trips for two interval options: two hours and one hour. The table shows that for a two-hour period of free transfer, the share of free trips will be about 17% of the total number of trips. In details, 12% of passengers per day will take one free trip and the share of passengers who will perform two or more free transfers is about 5% of trips.
For a one-hour period of free transfer, the share of free transfer will be about 10%. Passengers with one free transfer will make about 8% of trips. Obviously, the number of free transfer can be limited, for example, at the level of two free transfer per day. A smaller number of free transfer is illogical, as the passenger performs two work trips per day. However, the restriction to two free transfers per day is inefficient, since in this case the number of free transfers is reduced by only one percent.

Table 4. Set of trips and trip segment (April 2019, Krasnoyarsk city)

| Segment num. | Total trips | Trips after identifying the short activities trips |
|--------------|-------------|--------------------------------------------------|
|              |             | Trips | Trip segments | Share (%) | Trips | Trip segments | Share (%) |
| 1            | 2154165     | 2154165 | 83,8         | 2347027   | 2347027 | 91,3         |
| 2            | 189363      | 378726  | 14,7         | 109647    | 219294  | 8,5          |
| 3            | 10226       | 30678   | 1,2          | 1649      | 4947    | 0,2          |
| 4            | 1689        | 6756    | 0,3          | 0         | 0       | 0,0          |
| 5            | 146         | 730     | 0,0          | 0         | 0       | 0,0          |
| 6            | 27          | 162     | 0,0          | 0         | 0       | 0,0          |
| 7            | 5           | 35      | 0,0          | 0         | 0       | 0,0          |
| 8            | 2           | 16      | 0,0          | 0         | 0       | 0,0          |
| Total        | 2355623     | 2571268 | 1,09         | 2458323   | 2571268 | 1,05         |

Table 5. Distribution of trips with free transfers

| Transfer number | Two-hour period of free transfer | One-hour period of free transfer |
|-----------------|----------------------------------|---------------------------------|
|                 | Passenger-day Free trips Share (%) | Passenger-day Free trips Share (%) |
| 1               | 342639 342639 11,84                | 220660 220660 7,63               |
| 2               | 56682   113364 3,92                 | 30241 60482 2,09                 |
| 3               | 8698    26094 0,90                  | 2308 6924 0,24                   |
| 4               | 1183    4732 0,16                   | 232 928 0,03                     |
| 5               | 195     975 0,03                    | 24 120 0,00                      |
| 6               | 35      210 0,01                    | 4 24 0,00                        |
| 7               | 7       49 0,00                     | 1 7 0,00                         |
| 9               | 1       9 0,00                      |                                 |
| Total           | 488072 16,87                       | 289145 9,99                      |
| All trips       | 2893013                            | 2893013                          |

5. Acknowledgments
The authors are grateful to Municipal Public Institution of the Krasnoyarsk City (MSE) "Krasnoyarskcitytrans" for providing the data for this research.

References
[1] Fadeev A, Alhusseini S and Belova E N, 2018, Monitoring Public Transport Demand Using Data From Automated Fare Collection System Advances in Engineering Research, vol. 158, pp. 5–12
[2] Fadeev A and Alhusseini S 2019 Using Automated Fare Collection System Data To Determine Transport Demand Advances in Engineering Research, vol. 188
[3] Munizaga, M.A. and Palma, C., 2012. Estimation of a disaggregate multimodal public transport origin-destination matrix from passive Smart card data from Santiago, Chile. *Transportation Research* **24**, 9–18.

[4] Barry J J, Freimer R and Slavin H L 2009 Use of entry-only automatic fare collection data to estimate linked transit trips in New York City *Transp. Res. Rec. J. Transp. Res. Board* **2112** 53–61

[5] Barry J J, Newhouser R, Rahbee A and Sayeda S 2002 Origin and destination estimation in New York City with automated fare system data *Transportation Research* **1817** 183–7

[6] Zhao J, Rahbee A and Wilson N 2007 Estimating a rail passenger trip origin-destination matrix using automatic data collection systems *Computer Aided Civil and Infrastructure Engineering* **22** 376–87

[7] Trépanier M, Tranchant N and Chapleau R 2007 Individual trip destination estimation in a transit smart card automated fare collection system *Journal of Intelligent Transportation Systems* **11** 1–14

[8] Devillaine F, Munizaga M A and Trépanier M 2012 Detection activities of public transport users by analyzing smart card data *Transportation Research Record* **2276** 48–55

[9] Munizaga M A, Devillaine F, Navarrete C and Silva D 2014, Validating travel behavior estimated from smartcard data *Transportation Research Part C Emerging Technologies* **44**:70–79