

A method to estimate passenger flow with stored data at ticket gates

S. Myojo

Railway Technical Research Institute, Japan

Abstract

In order to improve railway diagrams, it is necessary to grasp the flow of railway passengers in detail. However, it is impossible to accomplish this goal only by counting the number of on-board passengers. We have developed a method to estimate the passenger flow based on OD (Origin & Destination) matrix data collected from ticket gates in hundreds of stations and timetables of thousands of trains. By analyzing the estimated results, it is possible to obtain the histograms of the number of passengers aboard each train for each destination, the number of passengers boarding and disembarking at each station, the number of transfer passengers at each station and other relevant factors. We have also developed a proto-type system based on this method assuming that all passengers choose shorter time paths. The estimated number of on-board passengers with this proto-type system has a strong correlation with the number reported by train conductors.

Keywords: passenger flow, ticket gate, OD.

1 Introduction

In order to improve railway diagrams or passenger announcement, it is necessary to grasp railway passenger flow in detail including the number of on-down passengers on each train and on-board number per each destination. In Japan, the Ministry of Land, Infrastructure, and Transportation holds the transportation census in order to perceive a passenger flow of public transportation in three major metropolitan areas. Japanese government utilise the census data for planning railway network and public transport operators do for scheduling trains or planning capital investment. The census provides very detailed data of passenger flow because passengers corporate with the survey fill their choice of



transportation mode on the census day. However, there is a problem that Japanese government holds it only once every five years because it is a huge-scale survey.

On the other hand, rolling stock incorporated recently is equipped with a device to measure the total weight of on-board passengers and calculate the congestion rate. This function will make it possible to obtain a total of passengers aboard all trains. However, it is time-consuming until every car is equipped with such function, and this function does not help to obtain on-down number and passengers' destination. Therefore, it is necessary to realize a method to grasp on-board number of each train on any day under the condition where not every car is equipped with that function and/or to grasp as detailed data as the Transportation Census.

We have developed a method to estimate passenger flow based on OD matrix data stored by ticket gates at stations. Although OD data do not provide directly any information about lines and trains whom passengers took, the method speculates their choice of paths including their lines and trains. By analyzing estimated result, it is possible to obtain the number of passengers aboard each train for each destination, the number of passengers boarding and disembarking at each station, the number of transfer passengers at each station and other relevant factors.

We also have developed a proto-type system based on this method assuming that all passengers select shorter time paths. It takes approximately one hour for this system to estimate one day's passenger flow of a railway network that contains more than 300 stations, and the estimated numbers of on-board passengers with this proto-type system has a close correlation with the number of passengers reported by conductors.

2 OD data during each period

OD signifies Origin and Destination. The term “OD data” means each number of passengers of each OD pair (Figure 1). Ticket gates count the number of passengers of each OD pair in each partitioned period of one day.

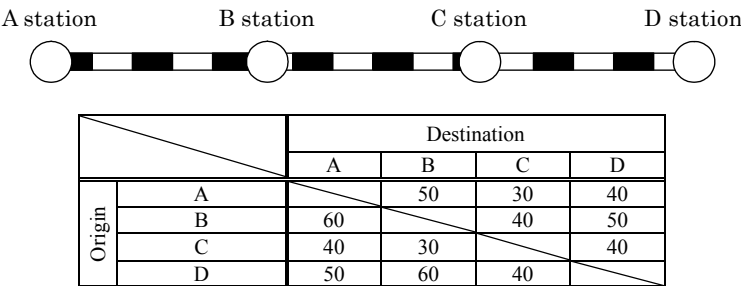


Figure 1: An example of OD data.



OD data has the following advantages:

- It is possible to obtain the data on any date because gate machines work everyday.
- The data is almost complete enumeration because gate machines count almost all passengers.

In spite of such advantages, OD data has a disadvantage that it does not provide directly any information about lines and trains that passengers took. Therefore, we have developed an estimation method assuming their routes and trains.

3 Proposed method to estimate a passenger flow

3.1 Basic idea of the method

The proposed method estimates passenger flow with following steps.

- (1) Select a pair of OD from OD data gathered from ticket gates.
- (2) Search all paths that those who travel the OD selected in procedure (1).
- (3) Calculate the number of passengers on the paths searched in procedure (2) during each period.
- (4) Store the number of passengers calculated in procedure (3).
- (5) Repeat procedures (1), (2), (3) and (4) until all pairs have been finished.
- (6) Calculate the total number of passengers that pass each link.

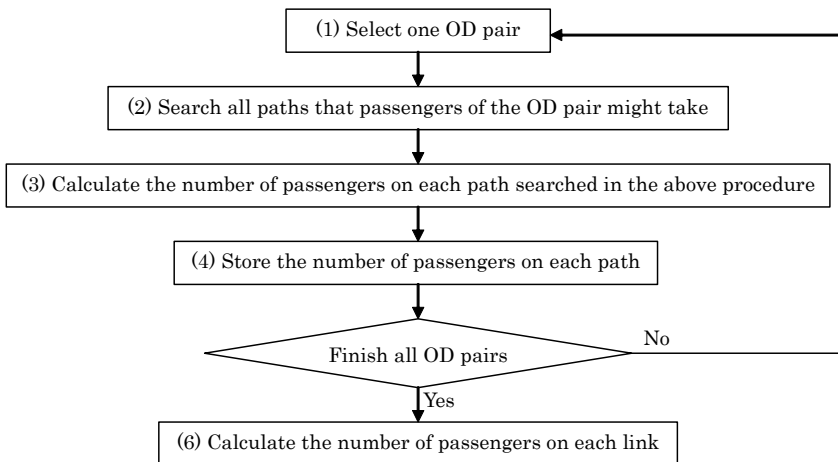


Figure 2: Procedures to estimate passenger flow.

3.2 Procedure to search paths

As described above, it is necessary to determine passengers' choice of their paths in order to estimate a passenger flow using OD data during each period. The paths in this method contain passengers' trains as well as their train lines. In

other words, a path not containing a different train from another one is not same with the one though they contain same railway line(s) at all. Therefore, estimated result can provide detailed data about passenger flow.

It is possible to search paths containing trains by transposing behaviours such as taking a train, disembarking, change of trains and other action to a directed graph. Figure 3 shows an example of components that make up a directed graph. In this graph, there are nodes representing departure from a station on a train and arrival at a station, and links representing travelling on a train, waiting for departure and change of trains. The weight (or length) of each link is time spent for the behaviour of travelling aboard a train, waiting for departure and change of trains when searching shorter time paths. When searching paths with fewer changes of trains or containing train(s) of fewer congestion rates, we add transfer resistance and congestion rate to each link in terms of time.

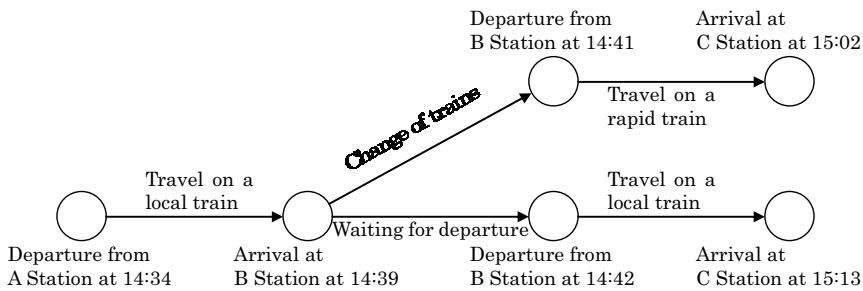


Figure 3: Components to make up a directed graph for path finding.

3.3 Calculation of the number of passengers on the searched paths

In this procedure, the estimation method distributes the number of passengers who travelled the selected OD pair to searched paths during each period. Equal allocation is the simplest method. With this method, the number of passenger on

each path is $\frac{Q_{ij}}{n}$ where Q_{ij} is the number of passengers of the OD pair i during the period j and n is the number of paths. When difference between paths is not small, equal distribution is not suitable but logit model is. With distribution method using a logit model, the estimation method calculates the number q_k of passengers on the path k with the formula (1) where V_k is the utility value of the path k .

$$q_k = \frac{e^{V_k}}{\sum_{l=1}^n e^{V_l}} \cdot Q_{ij} \quad (1)$$

3.4 Calculation of the number of passengers on each link

The estimation method calculates the number x_i of passengers on the link i with the formula (2) where q_{ij} is the number of passenger of the path j who moved on the link i and M is the number of paths of all OD pairs. The number q_{ij} is equal to the number of passengers of the path j if the path j contains the link i , and the number q_{ij} is zero if the path j does not contain the link i .

$$x_i = \sum_{j=1}^M q_{ij} \quad (2)$$

The whole of x_i are the result of estimation.

4 Various data obtained from the estimated result

4.1 A proto-type system

Based on the method, we have developed the simplest proto-type system that searches shorter time paths and allocates the number of passengers equally. We have applied this system to estimate passenger flow in the real railway network in the following condition:

Number of stations: 392

Number of trains: 4,785

Number of OD pairs: 30,604

The specification of the directed graph to search paths in this railway network is below:

Number of nodes: 216,183

Number of links: 506,233

The system can estimate within an hour with PC whose CPU is 3.6GHz Pentium® 4 processor.

4.2 Number of on-board passengers

By sorting the numbers of the passengers on links, as shown by thick arrows in Figure 4, contained in a train into near order from the starting station of the train, it is possible to obtain the change of number of on-board passengers of the train. Figure 5 shows an example of a histogram of the estimated number of passengers of a real train. The train started from (A) station and stopped at every station from (B) to (J). The shaded bars represent the number of passengers while travelling and white bars those while waiting.

4.3 Number of passengers boarding and alighting at each stop

By comparing the number of passengers before a stop and at the stop, it is possible to obtain the number of passengers who alighted at the stop.



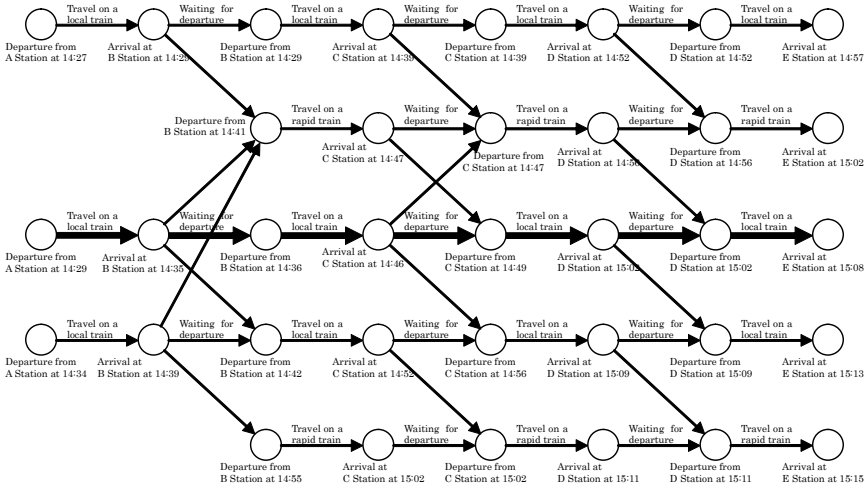


Figure 4: Links that are part of a train.

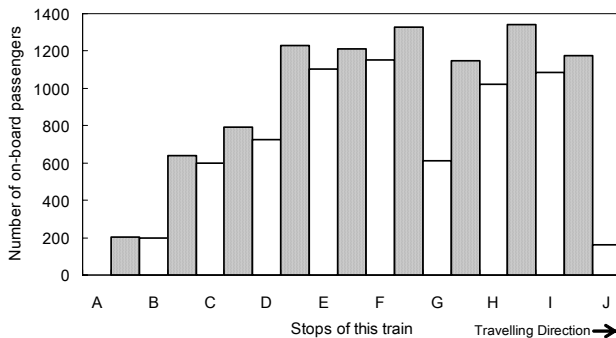


Figure 5: Numbers of on-board passengers.

By comparing that before departure and after departure, it is also possible to obtain the number of passengers who boarded at the stop. Moreover, it is possible to distinguish the number of passengers transferring to other trains from that of all passengers who alighted at a stop because the result of estimation provides the number of passengers on links representing transferring to and from the train shown by double-lined arrows in Figure 6. It is also possible to distinguish the number of passengers transferring from other trains from that of passengers who boarded on the train.

Figure 7 shows an example of a train. Bars of positive value represent the number of passengers who boarded train, and those of negative value the number of alighting passengers. It is very difficult to obtain such data by only counting the number of on-board passengers.

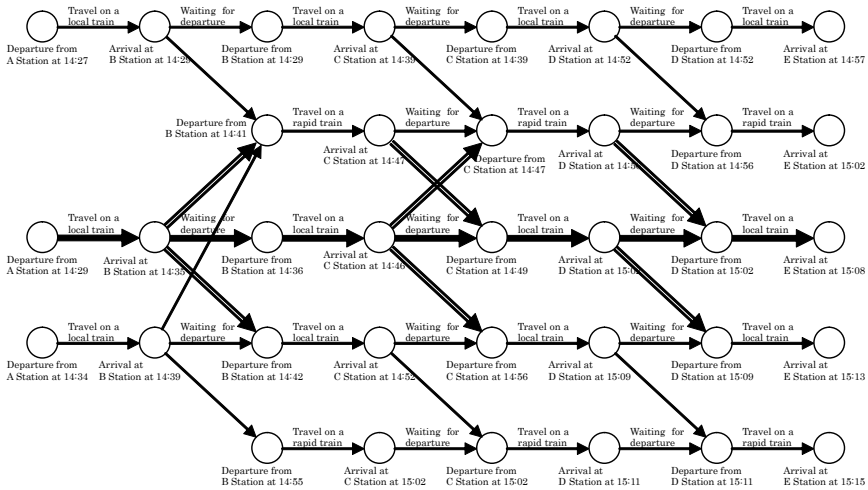


Figure 6: Links representing transfers.

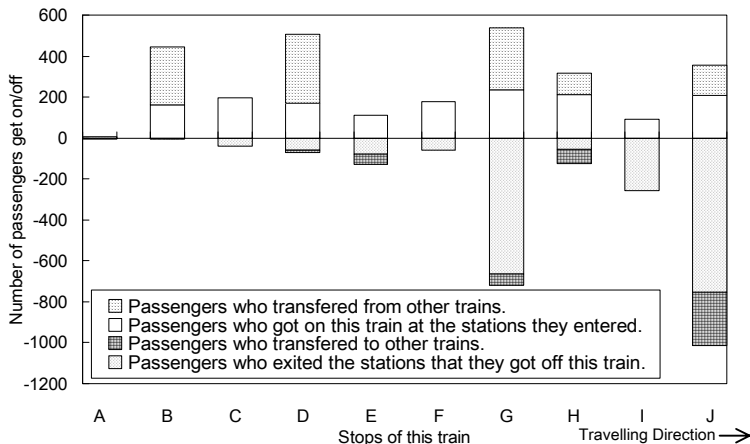


Figure 7: Number of passengers boarding and disembarking at each stop.

4.4 Number of on-board passengers for each destination

By integration of the number of passengers on each link contained in paths concerning with specific destinations, it is possible to obtain the passenger flow data about specified destination. That is, it is possible to obtain the number of on-board passenger for every destination. Figure 8 shows an example.

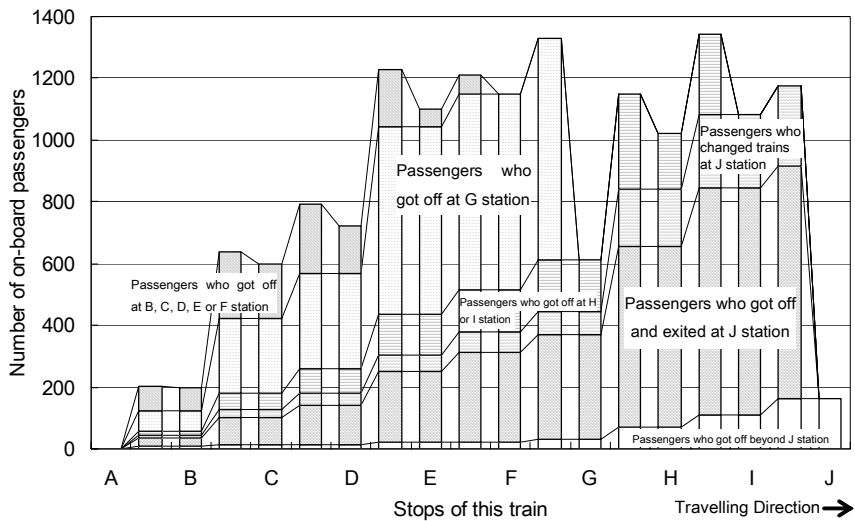


Figure 8: Number of on-board passengers for each destination.

4.5 Traffic value between two stations

Figure 9 indicates the number of passengers on links between two stations as compiled with thick arrows. It is possible to obtain a change in traffic value between the stations, figure 10.

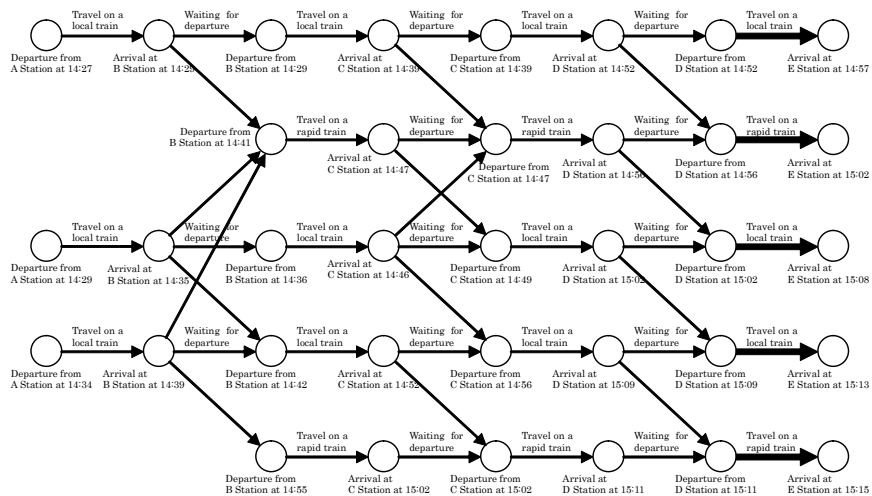


Figure 9: Links between two stations.

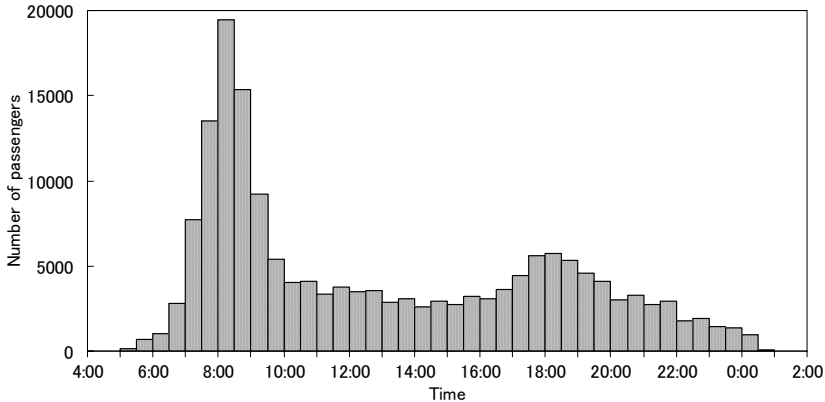


Figure 10: Histogram of traffic value between two stations.

5 Verification of estimated result

By comparing the reported number of passengers by train conductors, we have verified the number of on-board passengers estimated by the proto-type system. Train conductors of nearly a half of all trains report the maximum number aboard between two or more stations by visual counting. Figure 11 and 12 show the comparative results. Figure 11 shows the result of comparison on a line where only local trains run. Figure 12 shows that on another line where rapid trains make connections with local trains. In both lines, especially on a line of local trains only, correlation coefficient is high.

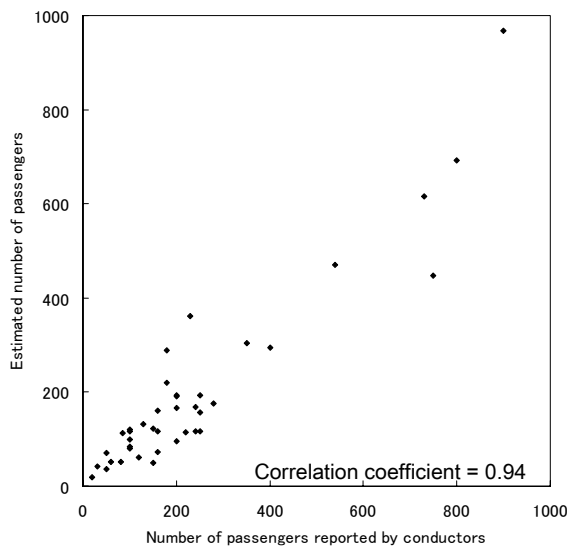


Figure 11: Comparison on a line where only local trains run.

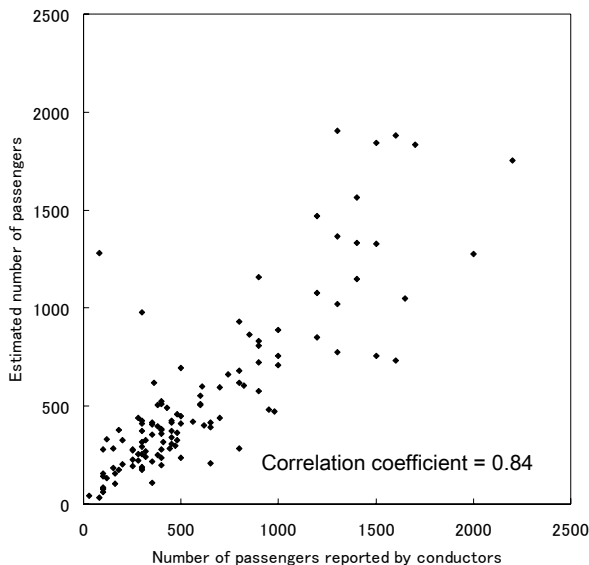


Figure 12: Comparison on a line where rapid trains make connection with local trains.

6 Conclusion

It is possible to estimate the passenger flow and to obtain detailed data containing the number of boarding passengers and disembarking passengers, the number of on-board passengers for each destination using OD data during each period stored in ticket gates. Moreover, the number of on-board passengers estimated by the simplest proto-type system has a very close correlation with the reported number of passengers by train conductors. In a line where only local trains are under operations, the simplest system can be practical.

References

- [1] Suzuki S, Goto K, Estimation of Passenger Flow from Vehicle Air Spring Pressure and Automatic Ticket Checking Data, RTRI Report No.7, pp. 39 44, 2004 (in Japanese)
- [2] Ieda H, Akamatsu T, Takagi J & Hatanaka H, Commuter's Benefit Evaluation of Train Scheduling by Network User Equilibrium Model, Infrastructure Planning Review, pp. 177 184, 1988 (in Japanese)
- [3] Nagasaki Y, Asuka M & Koyama K, Fast Estimation Method of Railway Passengers Flow, 2005 (in Japanese)