The Fifth International Conference on Railway Technology
RAILWAYS 2022
22–25 August 2022
Montpellier, France

Proceedings of the Fifth International Conference on
Railway Technology:
Research, Development and Maintenance
Edited by J. Pombo
Civil-Comp Conferences, Volume 1, Paper 23.6
Civil-Comp Press, Edinburgh, United Kingdom, 2022, doi: 10.4203/ccc.1.23.6


# Estimation of station congestion using real-time data

**Toru Sahara[1], Sei Sakairi[1], Yugo Shibata[1], Yuhei Soda[1]**

**[1]JR East Research and Development Center, East Japan Railway Company Tokyo, Japan**

**Abstract**

We believe that it is necessary to quickly grasp the congestion of the station and deal with it. In this research, we studied a method of calculating the congestion situation in a station by collecting data from train weight sensors and automatic ticket gates. Since the data used this time is data that can be acquired in real time, the situation can be immediately reflected even in the event of an abnormality. We believe that this method can accurately calculate congestion even in abnormal situations.

We have developed a method that can calculate the congestion status of a station using the information on the automatic ticket gates and the number of passengers of the train. We are developing a prototype system “Station Congestion Visualization function” and testing it. In the future, we believe that opening it to the public will avoid congestion at stations. We are making these developments to improve passenger convenience. We believe it will help further enhance the value of the railway.

**Keywords:** station, big data, simulation.

## 1 Introduction

It is very important for railway companies to understand congestion. Reduction in service quality or even dangerous condition can be leaded by congestion. To prevent or deal with this, railway companies are making it possible to grasp congestion. JR East carries 17.5 million customers a day. As a result, heavy congestion occurs every day. Further, when a large event or transportation failure occurs, more severe congestion occurs. In order to ascertain the number of passengers on the train, we have developed a "Train Congestion Visualization System" using a train weight sensor. Our trains are equipped with weight sensors. The acquired weight data is collected by the server via the train radio. By combining this information with the location information of the train, we have developed a system that visualizes where the running train is and how crowded it is. The train dispatcher who manages the train operation can use this system. By using this system, dispatchers can now conduct traffic control with more consideration for passengers. We commercialized the system in April 2016. It now supports 18 routes in the Tokyo metropolitan area. Information on some functions is also made available to general customers. You can see the train congestion information by using our app "JR East App". On the other hand, when a major event or transportation failure occurs, congestion not only at the train but also at the station may become a problem. We believe that it is necessary to quickly grasp the congestion of the station and deal with it. In this research, we studied a method of calculating the congestion situation in a station by collecting data from train weight sensors and automatic ticket gates. Since the data used this time is data that can be acquired in real time, the situation can be immediately reflected even in the event of an abnormality. We believe that this method can accurately calculate congestion even in abnormal situations.

## 2 Methods

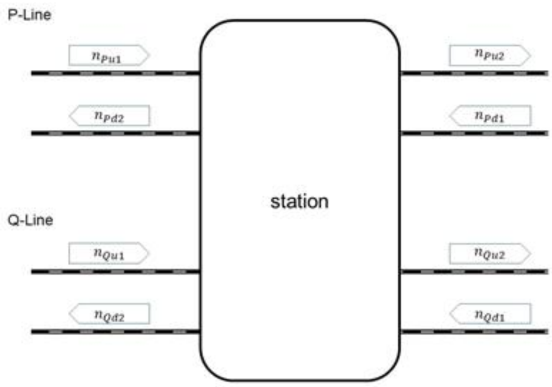
We have developed a method that can calculate the congestion status of a station using the information on the automatic ticket gates at the station in addition to the information on the number of passengers of the train used in “Train Congestion Visualization System”. In JR East, information on automatic ticket gates can be collected in real time at major stations (36 stations) of Tokyo metropolitan area. When a passenger enters or leaves an automatic ticket gate, information is collected at the station server, and the information is collected at a collection base. As a result, although there is a delay of about 15 to 30 minutes, information on automatic ticket gates at stations can be collected. In order to calculate the station congestion from this information, we first calculated the number of passengers and where they were for a certain period of time, and created a database. The number of passengers on the day and the number of people using the automatic ticket gate are converted and corrected. In this way, we calculate the station ticket gates and platform congestion for that day. By performing the correction, it is possible to accurately calculate even when an event occurs. We have devised three estimation methods and assumed 4 cases of occurrence of station congestion and examined which of the three estimation methods was the most appropriate.

### 2.1 Estimation methods

We devised the following three estimation methods, compared and evaluated those with higher accuracy. Table 1 and Figure 1 explains three estimation methods.

Table 1 : Three estimation methods

| | |
|---|---|
| Method 1 | The number of passenger boarding the train and the number of passenger changing trains are calculated as increasing as the number of visitors to the station increases. The number of trains from the station is normalized so that the totals match. |
| Method 2 | The number of passenger getting on the train increases as the number of visitors at the station increases. The number of trains from the station is normalized so that the totals match. In addition, the number of passenger who want to transfer and board the train increases according to the number of trains before the transfer in the previous time zone and further according to the number of trains after the transfer. |
| Method 3 | The number of passenger getting on the train increases as the number of visitors at the station increases. The number of trains from the station is normalized so that the totals match. In addition, the number of passenger who want to transfer and board the train increases according to the number of trains before the transfer in the previous time zone. |




$n$ : monthly average number of passengers on the train

$n'$ : the number of passengers on the train

$N_{Pu}$ : monthly average number of passengers heading for "Pu"

$N_{iPu}$ : monthly average number of passengers visiting the station and riding "Pu"

$N_{oPu}$ : monthly average number of passengers getting off "Pu" and leaving the station

$N_{trQuPu}$ : monthly average number of passengers transferring from "Qu" to "Pu"

$N'_{iPu}$ : the number of passengers visiting the station and riding "Pu"

Estimation method 1

$$N'_{iPu} = N'_i \times \left(\frac{N_{iPu}}{N_i} \times \frac{n'_{Pu2}}{n_{Pu2}} \times k_{on} \times \frac{N_{iPu} + N_{trQuPu} + N_{trQdPu}}{N_{iPu}}\right)$$

Estimation method 2

$$N'_{Pu} = N_{iPu} \times \frac{N'_i}{N_i} \times k_{on} + N_{trQuPu} \times \frac{n'_{Qu1}}{n_{Qu1}} \times \frac{n'_{Pu2}}{n_{Pu2}} + N_{trQdPu} \times \frac{n'_{Qd1}}{n_{Qd1}} \times \frac{n'_{Pu2}}{n_{Pu2}}$$

Estimation method 3

$$N'_{Pu} = N_{iPu} \times \frac{N'_i}{N_i} \times k_{on} + N_{trQuPu} \times \frac{n'_{Qu1}}{n_{Qu1}} + N_{trQdPu} \times \frac{n'_{Qd1}}{n_{Qd1}}$$

$$k_{on} = \left(\sum_{j=P_u,P_d,Q_u,Q_d} \frac{N_{ij}}{N_i} \times \frac{n'_{j2}}{n_{j2}}\right)^{-1}$$

Figure1 : How to estimate in each three methods

2.2 Assumptions

And we assumed four cases of possible congestion and decided to verify the estimation method described in the previous section. Table 2 shows conditions of four cases. As a side note, following conditions are in common.

- Only P-Line and Q-Line stop this station
- The average numbers of passengers on P-Line and Q-Line are all 500. (nPu1 = 500, nPu2 = 500, …)
- The average numbers of transferring passengers are all 100. (NtrQuPu = 100, NtrQdPu = 100, …)
- The average numbers of passengers visiting the station and riding each train are all 100. (NiPu = 100, NiPu = 100, …)
- The average numbers of passengers getting off each train and leaving the station are all 100. (NiPu = 100, NiPu = 100, …)

Table 2 : Conditions of four cases

| | |
|---|---|
| Case 1 | • 600 passengers, 100 more than usual, arrived at the station on P-Line U-direction. (n'Pu1 = 600)<br>• 200 passengers, 100 more than usual, got off P-Line U-direction and leaving the station. (N'oPu = 200) |
| Case 2 | • 600 passengers, 100 more than usual, arrived at the station on P-Line U-direction. (n'Pu1 = 600)<br>• 600 passengers, 100 more than usual, got on P-Line U-direction from the station. (n'Pu2 = 600) |
| Case 3 | • 600 passengers, 300 more than usual, arrived at the station on P-Line U-direction. (n'Pu1 = 800)<br>• Of the increased passengers of 300, 200 transferred to Q-Line U-direction and 100 got off at the station. (n'Qu2 = 700, N'oPu = 300) |
| Case 4 | • 800 passengers, 400 more than usual, visited the station and got on the train. (N'i = 800)<br>• Of the increased passengers of 400, 100 passengers got on the train in each of 4 directions. (n'Pu2 = 600, n'Pd2 = 600, n'Qu2 = 600, n'Qd2 = 600) |

# 3 Results

We calculated the number of people getting on and off each train using three estimation methods, and adopted the method with the smallest average error as the estimation method for this tool. Table 3 shows the average error of each case. We can say that estimation Method 3 has the highest accuracy in 4 cases. In addition, this method requires less calculation load than other estimation methods.

Table 3 : Assumption and results

| | Method1 | Method2 | Method3 |
|---|---|---|---|
| Case1 | 8.0% | 5.9% | 2.9% |
| Case2 | 7.1% | 6.7% | 5.7% |
| Case3 | 11.3% | 10.1% | 6.8% |
| Case4 | 25.0% | 11.7% | 5.0% |

We have built a prototype system that visualizes station congestion at major stations in the Tokyo metropolitan area using Method 3. This system targets the major stations in the Tokyo metropolitan area (26 stations) where congestion can be easily calculated. We have developed this system as an additional function of the Train Congestion Visualization System introduced earlier. With this system, we color the stations on the map red, yellow, and green. Each color indicates the congestion level of each station: red shows highly congested, green shows vacant and yellow is middle level. Furthermore, when a station is selected, the congestion of platforms and ticket gates at that station can be viewed in detail. Figure 2 shows screenshot of this system.


Figure2 : Screenshot of Station Congestion Visualization function

## 4 Conclusions and Contributions

We use information on the number of passengers, information on train delays and congestion to provide better transport services without personal identification. We have developed “Train Congestion Visualization System” that visualizes train congestion and delays. We use this system in our train control rooms. For example, when transportation problems occur on any lines, dispatchers can make arrangements for traffic control based on each train congestion and delays by using this system. They also uses this system to review arrangements carried out for a previous situation. They strive to improve their skills of arrangements to provide good service for trains.

We are also developing “Station Congestion Visualization function” that uses this information to calculate station congestion. We are developing a prototype system and testing it. In the future, we believe that opening it to the public will avoid congestion at stations. Some people came to avoid crowded places including train congestion because they were afraid of infecting COVID-19. If they know where are crowded in advance, they may choose empty route although this is not the shortest distance to a destination. We are making these developments to improve passenger convenience. We believe it will help further enhance of our railway services.

## References

[1] Fujiwara, M., Ning, R., 2015. “Passenger flow simulator for systematic optimization of station layout and train timetable”, in: Proceedings of the 6th International Conference on Railway Operations Modeling and Analysis (RailTokyo 2015), Tokyo, Japan.

[2] John J, Fruin., 1974. “Pedestrian - planning and design”

[3] Johansson, F., Peterson, A., Tapani, A. 2014. “Local performance measures of pedestrian traffic”, Public Transport 6, pp. 159–183.