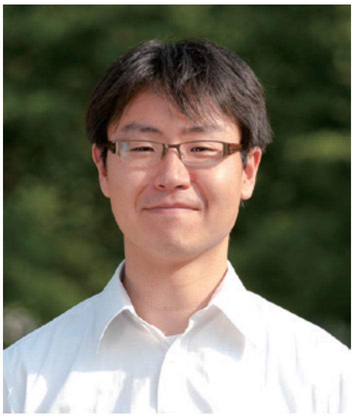
A Technique to Analyze Passenger Flow at Transport Disturbance Using Accumulated Passenger Data

Taketoshi KUNIMATSU
Assistant Senior Researcher, Transport Operation Systems, Signalling and Transport Information Technology Division

Whenever train disturbances occur, it is necessary to perform a series of adjustments such as implementation of shuttle services, suspension of trains and changing the sequence of train operation to minimize the degradation of passenger services. These actions are facilitated by correctly assessing and predicting passenger flow, but in such abnormal situations, some passengers may cancel their travels, detour to different transport facilities or take other complicated actions that were impossible in the past to assess or predict in quantitative terms. In recent years however, it has become possible to obtain recorded data in various categories related to daily train operation on a real-time basis. These data are passengers' origin-destination data (OD data) collected by the automatic ticket barrier, actual train operation time data by traffic control systems, or the number of onboard passengers by means of load compensating devices. RTRI researchers have examined techniques to quantitatively analyze and predict the passenger flow by utilizing such data when train disturbances occur.

To analyze the relation between passenger flow and train rescheduling at a transport disturbance, we devised a technique to visualize the actual data collected in the past (Fig. 1). On the rescheduled train operation diagram, this technique colors the lines according to the number of onboard passengers measured by the load compensating devices and marks with a symbol "○" the trains with substantial changes in the number of passengers compared to normal days when transport is in order. This makes it possible to easily assess the behavior of the passengers including those who shifted to the lines of other transport operators running in parallel and those passengers who were onboard the temporarily offered shuttle service trains.

As a next step, we constructed models to estimate the traffic volume in a section, or the number of passengers passing the section between two adjacent stations. We implemented multiple regression analysis using not only (1) the actual data in transport disturbances for approximately the past 10 months, and (2) the information about the place and time of the transport disturbance and the time length of the suspension, but also (3) the information about a) whether the time to resume train operation predicted and offered to users came true, and b) temporal shuttle operation in limited sections before full-scale resumption, in order to improve the precision of prediction (Fig.2). As a result, we were able to confirm that the value of the multiple correlation coefficient, an index to represent the precision of the model, took a comparatively satisfactory value of 0.7 to 0.8. Furthermore, we compared the results obtained by this model and the actual section traffic volume of two cases that were not used to construct the model (Fig. 3). Through this comparison, we were able to verify that the model is precise enough to use for making appropriate rescheduling plans before or after the resumption of train operation.



Thus, we have introduced a technique to analyze the accumulated data on the actual operation time and the number of passengers onboard the relevant trains, visualizing/predicting the passenger flow at the transport disturbance and utilizing the results thereof for decisions and discussions on train reschedulings. We will apply this technique to various lines and verify its reliability and usefulness in the future.

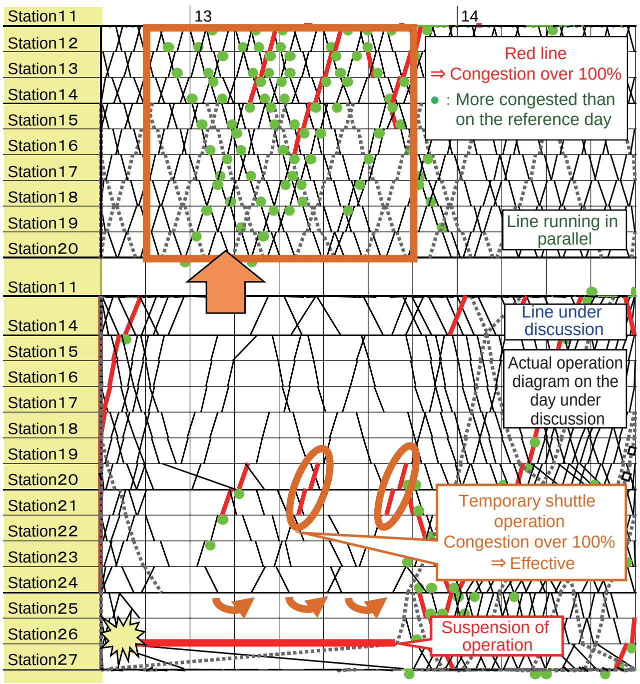


Fig. 1 Visualization of actual data

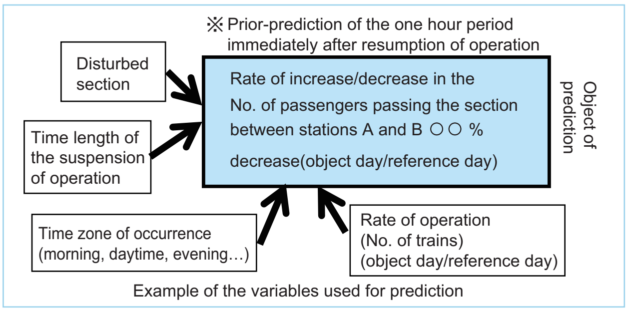


Fig. 2 A model to estimate the section traffic volume

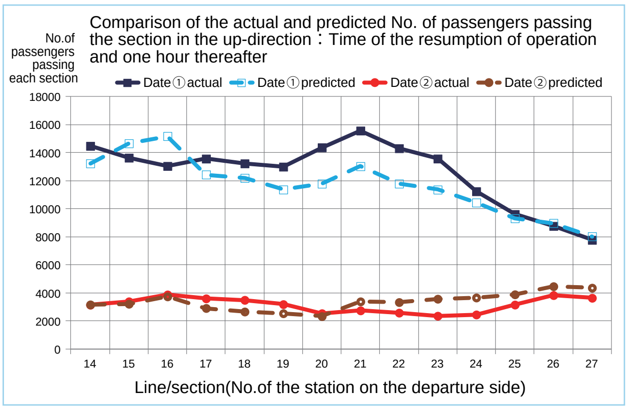


Fig. 3 Results of verification of the prediction model