Dynamic Passenger Assignment during Disruptions in Railway Systems

Zhu, Yongqiu; Goverde, Rob

Publication date
2017

Document Version
Accepted author manuscript

Published in
Proceedings of the 5th IEEE International Conference on models and technologies for intelligent transportation systems, 26 - 28 June 2017, Napoli (Italy)

Proceedings of the 5th IEEE International Conference on models and technologies for intelligent transportation systems, 26 - 28 June 2017, Napoli (Italy)

Dynamic Passenger Assignment during Disruptions in Railway Systems

Yongqiu Zhu
Department of Transport and Planning
Delft University of Technology
Delft, The Netherlands
Email: Y.Zhu-5@tudelft.nl

Rob M.P. Goverde
Department of Transport and Planning
Delft University of Technology
Delft, The Netherlands
Email: R.M.P.Goverde@tudelft.nl

***Abstract*—Passenger-oriented rescheduling problems receive increasing attention. However, the passenger assignment models used for evaluating the rescheduling solutions are usually simplified by many assumptions. To estimate passenger inconvenience more accurately, this paper establishes a dynamic passenger assignment model during disruptions, in which the time-dependent demand, disruption-induced service variations and vehicle capacities are all taken into account. Event-based simulation is adopted to implement the model of the dynamic loading and unloading procedures of passengers. Based on the model, individual travels can be tracked, thus making the estimation of individual passenger delay possible. By aggregating individual inconvenience, the performance of a given rescheduling solution/contingency plan can be evaluated. Furthermore, recommendations such as adding train units can also be proposed, as illustrated in the case study.**

***Index Terms*—Disruptions, railway systems, dynamic assignment, event-based simulation**

I. Introduction

Disruptions that result in track blockages or station closures have been annoying railway systems for years. Many efforts have been made either to prevent their occurrences or to mitigate the consequences once they occurred. For the latter, a widely-adopted method is to use pre-defined contingency plans that are further modified by traffic controllers according to the specific conditions [1]. However, such plans together with the manual modifications are only proposed from the perspective of operators. Passengers who should have been put first are neglected. Therefore, it is of great concern to take passengers into account when dealing with disruptions.

Recently, some contributions to passenger-oriented rescheduling problems towards either disturbances or disruptions have been made by [2]–[6]. In these papers, models and algorithms are proposed to generate rescheduling solutions automatically with the purpose of minimizing passengers' inconvenience. However, as their main focus is on optimization, the passenger assignment models that are used for evaluation in these papers are simplified by many assumptions. A more realistic passenger assignment model is proposed by [7] to evaluate the mitigation effects of real-time information that are offered to passengers during disruptions.

In this paper, we pay more attention to what information should be offered to traffic controllers to help them take better dispatching measures from the perspective of passengers. For this purpose, a dynamic passenger assignment model with more realistic considerations is established, which could be integrated in rescheduling models.

The terminology on dynamic transit assignment models is perceived differently across researchers [8]. Usually, three aspects are partly or completely included to represent the characteristic *dynamic*. They are time-dependent OD matrices, service variations and vehicle capacity constraints. In [9], both passenger demand and train headways that vary with time are considered in a schedule-based transit assignment model, whereas vehicle capacity is assumed to be infinite. As such, the model is more suitable for strategic planning of proposed transit systems with the focus on emerging phenomena of macroscopic congestion. In [10], [11], under given time-dependent passenger demand, schedule-based transit assignment models are proposed by taking finite vehicle capacity into account, with the assumption of trains operating precisely on schedule. Likewise, a passenger equilibrium flow model with the inclusion of vehicle capacity constraints is established in [12], while the supply side is assumed to be constant. Clearly, service variations are overlooked in these models, however in the real-world, they cannot be fully avoided. In [13]–[15], service variations are considered and described as irregularities of train dwell and running times that are thought to be relevant to the passenger loadings of the corresponding trains. Their focus is on the average effects of minor service perturbations on passengers' reactions rather than on the real-time management of major service disruptions on passengers' reactions [16].

In case of major disruptions, service variations are extremely different from the ones on normal days. Train services could be delayed, short-turned or completely cancelled. Besides, the service levels are different in the three phases that constitute a disruption: the first transition phase from the planned timetable to the disruption timetable, the second phase where the disruption timetable is performed, and the third recovery phase from the disruption timetable to the planned timetable [1]. Passengers who start their journey in different phases might be affected differently, and even the on-board passengers who start their journeys before the first phase could be affected. Although the dynamic transit assignment in [7] pays attention to the case of major disruptions, it doesn't explicitly show which disruption phases are considered. However, to quantify passengers' inconvenience accurately, it is necessary to establish a passenger assignment model that takes all three phases of a disruption into account, which is one of the contributions of this paper. In addition, time-dependent OD matrices and vehicle capacities are also included.

Note that instead of looking at passengers' pre-trip travel decisions, the dynamic passenger assignment model proposed in this paper mainly focuses on passengers' en-route travel decisions. This means that passengers are assumed to have planned paths in mind before they actually arrive at the origin stations, however, possibly they have to re-plan their paths due to service variations during disruptions. Such an assumption is justified, since nowadays passengers can rely on various travel-planner applications in mobile phones or the official websites of operator companies to find their preferred paths. This is particularly true for passengers who have a clear travel purpose (e.g. commuters). Thus, once disruptions occur, passengers would make en-route travel decisions by comparing the alternative paths during disruptions with their planned paths.

In what follows, we first explain the network modelling approach in Section II. Then, in Section III, the proposed dynamic passenger assignment framework is shown, followed by the introduction of the main parts in the framework. Finally in Section IV, a case study of a complete track blockage in part of the Dutch railway network is performed.

II. Network formulation

A space-time network is a typical approach to formulate transit services [10]–[14], especially for schedule-based services. In view of the real-time environment of disruptions, it is most suitable to formulate the train services in a space-time network. Generally, the formulation of pedestrian movements like boarding/alighting nodes and arcs are included in the space-time network. This results in an increasing amount of nodes and arcs, thus slowing down the computations. Therefore, we exclude the formulation of pedestrian movements, and propose another network formulation. An example is given in Fig. 1.

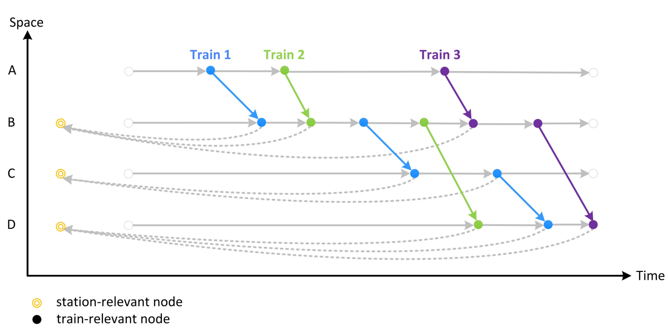

Fig. 1. Network formulation

In Fig. 1, nodes are either train-relevant or station-relevant. The train-relevant nodes consist of arrival nodes and departure nodes that are assigned with information of a time instant, type (arrival or departure), corresponding train number and station, etc. The station-relevant nodes are also called virtual nodes here, which are convenient for path generation. For each station, a virtual node is created and each arrival node at the station is linked to the virtual node as a predecessor. Note that virtual nodes are assigned with information of type (virtual) and corresponding station only. This means they don't have a timing attribute in Fig. 1.

When we search paths for a particular OD, a pair of source and sink node has to be determined. Here, the chosen source node is a departure node at the origin station. For the sink node, instead of trying every possible arrival node at the destination station, only the virtual node at the destination station is chosen. More details about path generation can be found in Section III.

Likewise, arcs are either train-relevant or station-relevant. The train-relevant arcs start from departure nodes and point to the space-adjacent arrival nodes that share the same train numbers (i.e. one single train). The station-relevant arcs are divided into two categories. One includes the arcs that connect two time-adjacent nodes in the same station. Such arcs enable passengers to wait to board trains at origins, transfer from one train to another, or dwell at the station in a train. Along with the aforementioned train-relevant arcs, the weights of these arcs are the time differences between the predecessor node and the successor node. One other category are the arcs from arrival nodes to virtual nodes in the same station. These arcs enable passengers to leave the railway system once they arrive at the final destination. We assign equal weight to each of such arcs, since they are not used to distinguish paths.

This formulation is in fact a compact transformation of the timetable into a network. The formulation of pedestrian movement is discarded, which brings the advantage of short computation times in searching paths. However, it also leads to an inability to distinguish transfers when searching the paths, which makes penalizing transfers impossible. To overcome this issue, we generate all reasonable paths first, and perform afterwards an analysis on the transfers. This is further explained in Section III.

III. Dynamic assignment model

The framework of dynamic passenger assignment in this paper consist of three parts, as shown in Fig. 2. Part I assigns each passenger with a planned path. Part II decides when a passenger would take a re-plan action if his planned path is delayed or cancelled. Since our assignment model is event-based, it is necessary to pre-determine the trigger event that determines when a passenger may re-plan his travel (i.e. en-route travel decision). Part III simulates the passengers' loading procedure to confirm the finally realized paths, by taking unsuccessful boarding due to full trains into account. In the following, the three parts are explained successively.

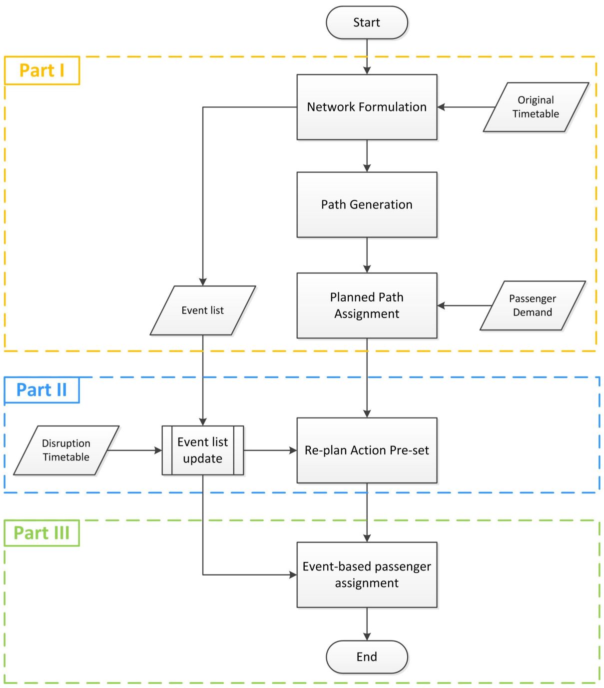

Fig. 2. Framework of the dynamic passenger assignment model

A. Passenger planned path assignment (Part I)

In part I, three modules are included: the network formulation, path generation, and planned path assignment. First of all, the original timetable is transformed into a space-time network as shown in Fig. 1. The resulting departure nodes and arrival nodes are sorted in time-ascending order to constitute the event list based on which part III can proceed. In other words, events are represented by departure nodes and arrival nodes. Next, we generate paths for each OD pair in the concerned time period. We assign each passenger with a planned path from the path set by assuming that a passenger always chooses the least transfer path with the earliest departure time regarding his arrival time at the origin station. Here, a passenger's planned path is characterized by a sequence of events in time-ascending order.

In this paper, for each departure node at each station, the paths from the departure node to any other station within a reasonable arrival time horizon are generated (if any). Waiting times at origins are excluded in the paths, but are considered in the procedure of assigning paths by comparing passenger arrival times with path departure times. In other words, the total travel times of paths consist of only in-vehicle times and transfer waiting times if any. The method of generating paths for each departure node is introduced in an example below.

In Fig. 3, suppose the paths from the departure node of train 1 at station A to station D need to be generated. Then, the departure node is chosen as the source node, and the virtual node at station D is chosen as the sink node. Additionally, to exclude waiting times at the origin, we cut the station arc starting from the source node. After that, a shortest path algorithm is performed, and the shortest path that takes the minimal time is generated.

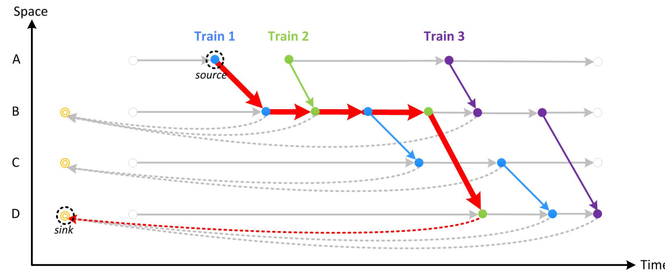

Fig. 3. The shortest path from the departure node of Train 1 at station A to station D (bold red)

Next, we find the nodes at station D, that are after the destination arrival node of the shortest path but are within a certain time horizon. Here, the time horizon is determined by the inverse of the frequency of services that are relevant to the source node, since passengers would not skip a service for another one that is actually the same type but operates later. Besides, to avoid meaningless waiting times at destinations, the arcs between these nodes are eliminated (as Fig. 4 shows). Then, a k-shortest path algorithm is performed to find the paths that have reasonable total travel times from the source node to the sink node. Here, the value of k in the k-shortest path algorithm is set as the number of destination arrival nodes that are within the chosen time horizon. In fact, the value of k has impact on the quality of the path set, while *quality* is interpreted as whether the paths that passengers would plan to choose are all contained in the path set.

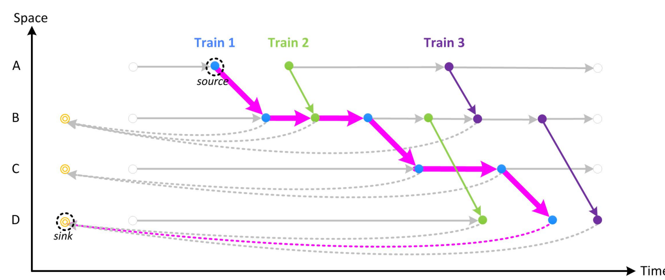

Fig. 4. The second shortest path from the departure node of Train 1 at station A to station D (bold pink)

The way we set the value of k in this paper is actually based on the assumption that between the departure node and each arrival node, only one path is available. It could happen that two paths have the same total travel time, but the needed transfer counts are different. Even so, the least transfer path can be found under the given k since such paths mostly cost less total travel times compared to the ones that need transfers.

B. Passenger re-plan actions pre-set (Part II)

In part II, according to the given disruption timetable, the event list is updated. Next, a comparison between each passenger's planned path and the updated event list is made to check whether the planned path would be delayed/cancelled or not. If yes, an event is assigned to the passenger, which triggers that the passenger re-plans his travel.

In order to assign affected passengers with proper trigger events, the following information is needed: reasons of re-planning paths (cancellation or delay), the left transfer counts of planned paths, planned boarding times at origin stations and transfer stations (if any), start time of disruption, and forced off stations of on-board passengers whose planned paths are cancelled. Based on this information, decision trees regarding where to take re-plan actions for passengers who encounter delay and cancellation are described in Fig. 5 and Fig. 6, respectively. Here, a forced off station is the place where a train service is short-turned or completely cancelled. Note that, as the left transfer counts may be more than one, thus once the boarding time of first left transfer is before disruption start, we cannot directly make the "no re-plan" decision, but need to check the boarding time of next left transfer.

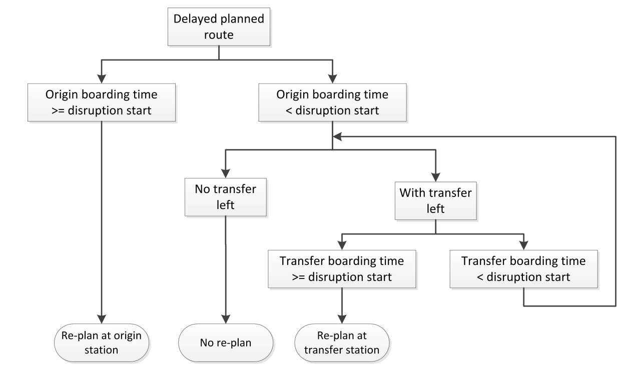

Fig. 5. Re-plan decision tree for passengers with delayed planned paths

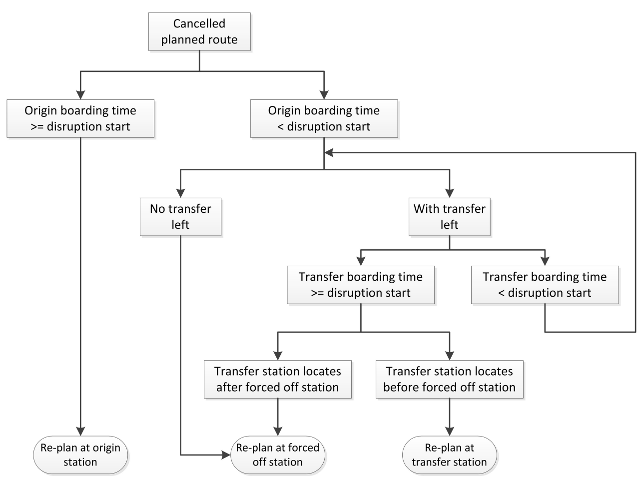

Fig. 6. Re-plan decision tree for passengers with cancelled planned paths

For the passengers who re-plan at origin stations, the re-plan actions are triggered when the passengers' arrivals are simulated. For the passengers who re-plan at transfer/forced off stations, the trigger events are the planned arrival events that correspond to the transfer or forced off stations.

C. Passenger realized path confirmation (Part III)

Unlike in part II where we can only see the passengers who are affected due to delayed or cancelled trains, in part III we can check which passengers are affected by full trains. In this part, passengers' arrivals and the loading/unloading procedures are implemented by event-based simulation. The events are represented by the arrival and departure nodes that are generated during the network formulation. All events are contained in the event list E. Before simulation, E is sorted in ascending order regarding the time instant of each event: $Time(e)$. Besides, the previous system time $clock_{\mathrm{p}}$ is set to 0. Below, the simulation steps are described.

Step 1: Choose the first element of E as the current event e_{c} and set the current system time $clock_{\mathrm{c}}$ to $Time(e_{\mathrm{c}})$.

Step 2: Simulate passenger arrivals. Find all passengers P_1 of each arrival at origin station s_{o} between $clock_{\mathrm{p}}$ and $clock_{\mathrm{c}}$. For each $p \in P_1$, increase the number of passengers in s_{o} (i.e. $V(s_{\mathrm{o}})$) with one unit. Next, if e_{c} is the re-planning trigger event of p: $Re(p)$ of p, then implement the re-plan action of p. After iterating over P_1, skip to step 4, if e_{c} is an arrival. Otherwise, continue.

Step 3: Simulate loading procedures. Find all passengers P_2 who wish to board the departing train t_{d} at departure station s_{d}. For each $p \in P_2$, if the available capacity $C(t_{\mathrm{d}})$ in t_{d} is bigger than 0, p boards t_d. Then, $C(t_{\mathrm{d}})$ and $V(s_{\mathrm{d}})$ both decrease with one unit. However, if $C(t_{\mathrm{d}}) = 0$, the re-plan action of p is implemented. After iterating over P_2, skip to step 5.

Step 4: Simulate unloading procedures. Find all passengers P_3 who alight from the arrival train t_{a}. For each $p \in P_3$, $C(t_{\mathrm{a}})$ increases with one unit. Next, if the arrival station s_{a} is the expected destination, p is removed from the simulation. Otherwise, $V(s_{\mathrm{a}})$ increases with one unit, and the re-plan action could be implemented for p, if e_{c} is $Re(p)$.

Step 5: Set $clock_{\mathrm{p}}$ to $clock_{\mathrm{c}}$. Then, remove e_{c} from E. If E is empty, the simulation ends. Otherwise, return to step 1.

The re-plan action mentioned above is implemented as follows: first, search the shortest path that results in minimal delay with the delay lower than the maximal acceptable delay: u. If such a shortest path exists, the passenger chooses it for the following travel. Otherwise, the passenger leaves the railway, thus being removed from the simulation. Consequently, the number of passengers staying in the corresponding station decreases with one unit.

IV. Case Study

The case study is performed on part of the Dutch railway network where six lines operate on normal days (see Fig. 7). However, when a complete track blockage occurs between Hm and Dn, lines 1900 and 9600 cannot run as usual. In practice, the adopted contingency plan for this disruption case is as follows: for line 1900, the operation between Eindhoven (Ehv) and Hrt is cancelled, thus trains are short-turned at Hm, while for line 9600, the operation between Hm and Dn is cancelled, thus trains are short-turned at Hrt. Apparently, such a contingency plan only concerns which changes should be made on services to accommodate the reduced infrastructure capacity. Whereas during disruptions the vehicle capacity could be insufficient, thus resulting in passengers' unsuccessful boardings and possibly more delay at the destination. To avoid this issue, traffic controllers need to decide whether it is necessary to add more units to particular trains. In the following case, we use the proposed dynamic passenger assignment model to help traffic controllers make such decisions under different scenarios.

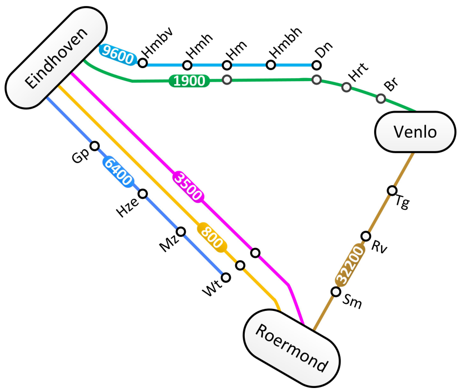

Fig. 7. The concerned line network

Here, the scenario is differentiated by two factors, the available capacity of each train when reaching the concerned network and the passenger maximal acceptable delay on the destination. The first factor represents the supply that can be given. The second factor implicitly represents the demand, since the lower the maximal acceptable delay is, the less the passengers would continue within the railway. In fact, the second factor is influenced by the alternatives outside the railway. If the outside alternatives are much faster compared to the alternatives provided by the railway, it is possible that passengers would drop the railway and turn to the outside alternatives to continue their travels. In such a case, adding more train units could be unnecessary, since lots of passengers might be lost.

In the following, scenarios are constructed based on different value settings of the two factors to explore the need for adding train units.

A. Scenario construction

The value of maximal acceptable delay u is set from 0 to 95 (in minutes) with intervals of 5 min. The reason to set u up to 95 will be explained later. Under each u, two scenarios are constructed. In one scenario, the available capacity of each train is set to be infinite, which enables passengers to board any train they wish to board. In the other scenario, the available capacity of each train is finite. For each train, we set its initial available capacity as the maximal passenger demand across all of its running processes within the concerned network on normal days. In other words, we assume that on normal days the available capacity of each train when reaching the concerned network happens to ensure successful boarding within the concerned network.

In each scenario, a disruption of a complete track blockage between Hm and Dn is set between 11:00 to 12:00, and the passenger demand from 10:00 to 13:00 is generated. The reason to generate passenger demand one hour before the disruption start and one hour after the disruption end is that in addition to the passengers who start their travels during the disruption period, the passengers who start their travels either before the disruption start or after the disruption end could also be affected.

B. Indicators chosen

For each scenario, whether more train units would be added is decided according to the influence of the scenario on the affected passengers. Here, the affected passengers contain the passengers who leave the railways (left passengers) and the passengers who stay in the railways to complete their travels but are delayed at the destinations (delayed passengers). The indicators of the amount of left/delayed/affected passengers and the total delay of left/delayed/affected passengers are chosen to reflect the influence of each scenario.

After implementing each scenario with the proposed dynamic passenger assignment model, the values of the indicators are calculated by aggregating individual travel information. In particular, the total delay of left passengers is computed as the amount of left passengers multiplied by the maximal acceptable delay corresponded to the scenario, since travels outside the railways are not tracked in the assignment model. Here, the assumption is made that if no paths within the railways are acceptable, we consider the worst case of the resulting delay of the outside path. The calculated indicators of each scenario are shown in Fig. 8 and Fig. 9.

C. Result analysis

From Fig. 8 we can see that with increasing maximal acceptable delay (i.e. u), the amount of left passengers decreases under all scenarios with either infinite or finite vehicle capacity. When u reaches 95 min, no passengers leave anymore, which is exactly the reason of setting the value of u up to 95. Besides, the decrease of the amount of left passengers is relatively sharp when u is at 20, 35, or 65 min. The reason could be relevant to the set disruption length (60 min) and the cycle time of services (30 min). In addition, the gap between the yellow line and the black line reveals the influence of full trains on the amount of left passengers, which is however not distinct. Nevertheless, when looking at the gap between the cyan line and the magenta line, we find that the influence of full trains on the amount of delayed passengers is rather distinct, particularly at higher u. By looking at the gap between the red line and the blue line, we find that the amount of affected passengers due to full trains is rare for u under 15, but grows with increasing u between 20 and 55. When u reaches 60, it sharply increases and continues such rapid growing till $u = 65$. Accordingly, we can conclude that in our case, adding more train units could help to reduce the amount of passengers affected by full trains if the maximal acceptable delay is 20 min or higher.

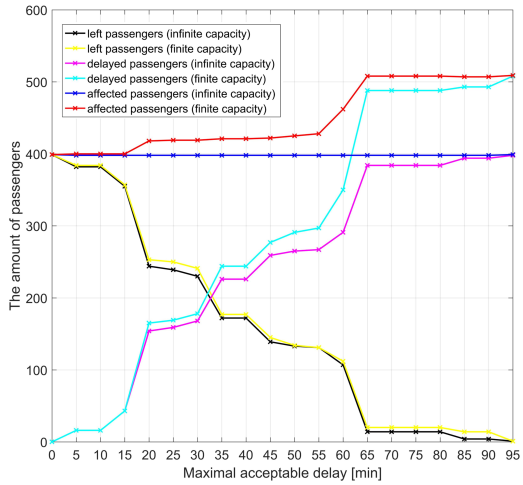

Fig. 8. The amount of left/delayed/affected passengers in each scenario

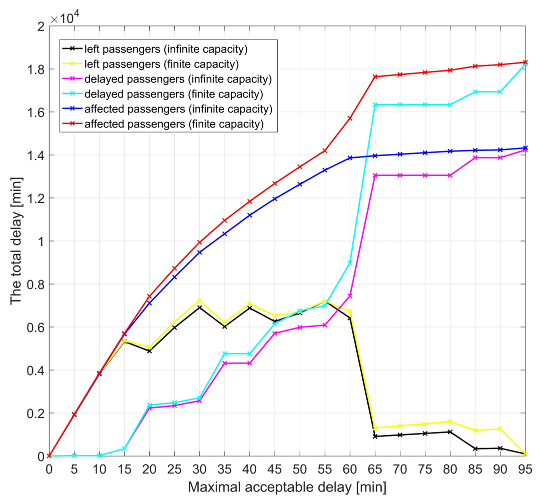

Fig. 9. The total delay of left/delayed/affected passengers in each scenario

Instead of looking at how many passengers are affected only, it is also necessary to check how serious passengers are affected. In Fig. 9, by looking at the gap between the cyan line and the magenta line, we find that due to full trains, the total delay of delayed passengers grows with increasing u from $u = 20$. Similar trend can be found by looking at the gap between the red line and the blue line, which represents the total delay of affected passengers due to full trains. Therefore, we can conclude that in our case, if the maximal acceptable delay is 20 minutes or higher, adding more train units could help to reduce the total delay of passengers affected by full trains.

Overall, we conclude that under a given contingency plan, the need to add train units or even how much should be added depends on the maximal acceptable delay that is actually determined by the outside alternatives.

V. Conclusions

In this paper, a dynamic passenger assignment model during disruptions in railway systems is proposed. In the case study, we showed one possible application of the model of determining the need for adding train units. However, more applications could be performed based on the proposed model. For example, the resulting passenger delay of cancelling a particular train can be calculated, which can be used as the train cancellation weight in the objective function of an optimization model for timetable rescheduling. Besides, considering the fluctuation of day-to-day passenger demand and the frequency of disruptions, reasonable vehicle capacity allocations for improving the service resilience during disruptions can be proposed.

Acknowledgement

This work is partly supported by the program of China Scholarship Council (No. 201507000056).

References

[1] N. Ghaemi, O. Cats, and R. M. P. Goverde, "Railway disruption management challenges and possible solution directions," *Public Transport*, pp. 1–22, 2017.
[2] A. Schöbel, "A model for the delay management problem based on mixed-integer-programming," *Electronic notes in theoretical computer science*, vol. 50, no. 1, pp. 1–10, 2001.
[3] S. Kanai, K. Shiina, S. Harada, and N. Tomii, "An optimal delay management algorithm from passengers viewpoints considering the whole railway network," *Journal of Rail Transport Planning & Management*, vol. 1, no. 1, pp. 25–37, 2011.
[4] T. Dollevoet, D. Huisman, M. Schmidt, and A. Schöbel, "Delay management with rerouting of passengers," *Transportation Science*, vol. 46, no. 1, pp. 74–89, 2012.
[5] L. Kroon, G. Maróti, and L. Nielsen, "Rescheduling of railway rolling stock with dynamic passenger flows," *Transportation Science*, vol. 49, no. 2, pp. 165–184, 2014.
[6] F. Corman, A. DAriano, A. D. Marra, D. Pacciarelli, and M. Samà, "Integrating train scheduling and delay management in real-time railway traffic control," *Transportation Research Part E: Logistics and Transportation Review*, 2016.
[7] O. Cats and E. Jenelius, "Dynamic vulnerability analysis of public transport networks: Mitigation effects of real-time information," *Networks and Spatial Economics*, vol. 14, no. 3-4, pp. 435–463, 2014.
[8] Y. Liu, J. Bunker, and L. Ferreira, "Transit users route-choice modelling in transit assignment: A review," *Transport Reviews*, vol. 30, no. 6, pp. 753–769, 2010.
[9] C. Tong and S. Wong, "A stochastic transit assignment model using a dynamic schedule-based network," *Transportation Research Part B: Methodological*, vol. 33, no. 2, pp. 107–121, 1999.
[10] M. Poon, S. Wong, and C. Tong, "A dynamic schedule-based model for congested transit networks," *Transportation Research Part B: Methodological*, vol. 38, no. 4, pp. 343–368, 2004.
[11] Y. Hamdouch and S. Lawphongpanich, "Schedule-based transit assignment model with travel strategies and capacity constraints," *Transportation Research Part B: Methodological*, vol. 42, no. 7, pp. 663–684, 2008.
[12] S. Nguyen, S. Pallottino, and F. Malucelli, "A modeling framework for passenger assignment on a transport network with timetables," *Transportation Science*, vol. 35, no. 3, pp. 238–249, 2001.
[13] A. Nuzzolo, F. Russo, and U. Crisalli, "A doubly dynamic schedule-based assignment model for transit networks," *Transportation Science*, vol. 35, no. 3, pp. 268–285, 2001.
[14] Y. Hamdouch, W. Szeto, and Y. Jiang, "A new schedule-based transit assignment model with travel strategies and supply uncertainties," *Transportation Research Part B: Methodological*, vol. 67, pp. 35–67, 2014.
[15] O. Cats, J. West, and J. Eliasson, "A dynamic stochastic model for evaluating congestion and crowding effects in transit systems," *Transportation Research Part B: Methodological*, vol. 89, pp. 43–57, 2016.
[16] G. Gentile and K. Noekel, *Modelling public transport passenger flows in the era of intelligent transport systems*. Switzerland: Springer International Publishing, 2016.