Network Effects in Railways

Alex Landex, al@ctt.dtu.dk
Otto Anker Nielsen, oan@ctt.dtu.dk
Centre for Traffic and Transport, Technical University of Denmark

1 Abstract

Railway operation is often affected by network effects as a change in one part of the network can influence other parts of the network. This influence can even be far away from where the original change was made. The network effects occur because the train routes (often) are quite long and that the railway system has a high degree of interdependencies as trains cannot cross/overtake each other everywhere in the network.

First the article describes network effects in general (section 2). In section 3 the network effects for trains – and how they can be measured by queuing time is described. When the trains are affected by network effects also the passengers are affected. Therefore, section 4 describes the network effects for passengers and how they can be measured using passenger delay models. Before the concluding remarks in section 6, section 5 discusses how the operation can be improved by examining network effects in the planning process.

Keywords: Railway, Network Effect, Passenger delay, Queuing time, Correspondence

2 Introduction

When railway capacity and delays are investigated, the analyses are often restricted to a single railway line or section of the network. However, a change in one part of the network can influence other parts of the network. This influence can even be far away from where the original change was made. These influences are denoted as network effects and occur because train routes (often) are quite long and that the railway system has a high degree of interdependencies as trains cannot cross each other or overtake each other everywhere in the network.

Network effects are dependent on the given infrastructure and timetable and can result in longer travel times for trains and passengers. The passengers can be further affected of the network effects because not all the wanted correspondences to/from other trains can be kept due to too many interdependencies – or network effects. Furthermore, the network effects can result in reduced capacity as some trains or train routes can make it impossible to operate other planned/desired trains or train routes. This is shown in figure 1 where it is not possible to operate more trains on the single track line section because of the many trains operated on the double track railway line.

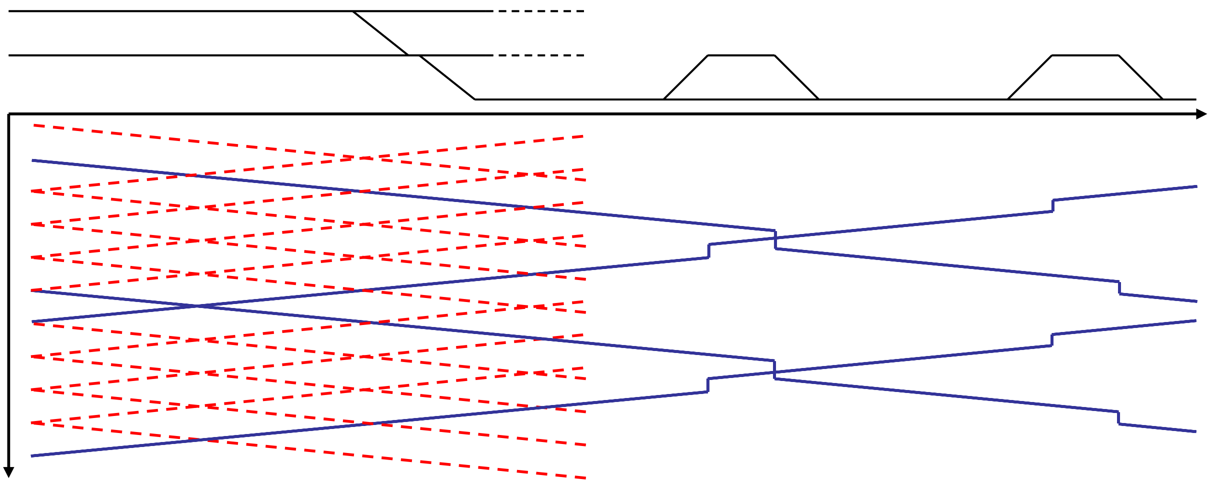


Figure 1: Limitations in the degree of freedom in the timetable

Identifying network effects of changes on the main lines, a nationwide candidate timetable for one standard hour must be worked out. However, the nationwide timetable in Denmark depends on the train services to/from Germany and Sweden. To evaluate all the network effects it is therefore not enough to create a nationwide candidate timetable. It is necessary to include the trains to/from Germany and Sweden and thereby also the nationwide timetables of Germany and Sweden and so forth.

When the analysis area is large, the risk of network effects is high too. This is because when a large analysis area is examined it can result in bigger changes in the infrastructure and/or timetables. Major changes in the infrastructure and/or timetables may influence many trains in the analysis area, and these trains may influence other trains outside the analysis area.

However, even smaller analysis areas may generate network effects. This is due to the way of planning the timetable in Denmark and many other countries. All train services can be placed in a hierarchy, cf. figure 2, where the train services placed in the top of the hierarchy is planned and timetabled before trains further down in the hierarchy.

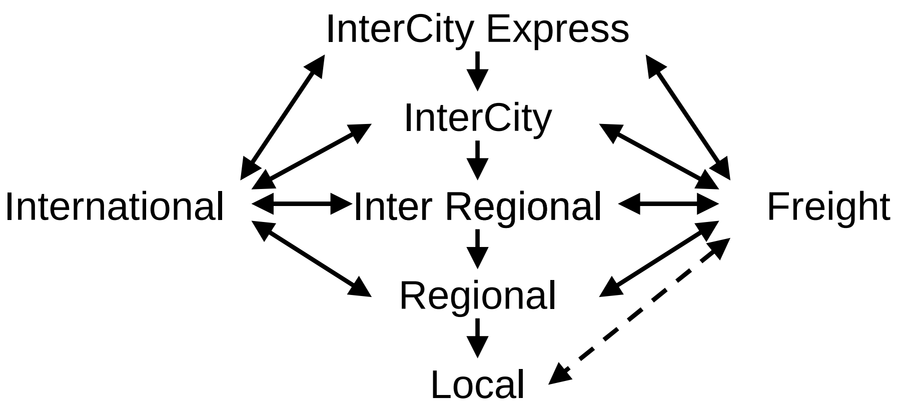


Figure 2: The hierarchy of the train service. Inspired by (Hansen, Landex & Kaas 2006) and (Landex, Kaas & Hansen 2006)

Even small changes in the timetable of a train in the upper level of the hierarchy may influence other trains further down in the hierarchy, because these trains are planned according to the train high up in the hierarchy. Since trains high up in the hierarchy often travel long distances, the changes for other train services can occur far away from the analysis area.

Although the risk of network effects is known many kinds of analysis/projects, the effects of changed timetables and/or infrastructure are only studied locally. It can be due to lack of resources, or because the network effects are uncertain (or insignificant), or because one wish only to evaluate the project locally, isolated from the remaining network.

An example to illustrate the network effects is the Danish railway line between Aalborg and Frederikshavn, cf. figure 3. It is a single track line with a one-hour service. The travel time in one direction is 63 minutes and 66 minutes in the other direction (Hansen 2004). The speed on the line is now examined increased from 120 km/h to 180 km/h.

This project can be evaluated locally. However, the traffic in the northern part of Denmark is not timetabled independently of the remaining network. The trains are part of the nationwide Intercity system (cf. figure 3) and are therefore adapted to the arrival and departure times of the IC-trains at Aalborg (as well as the crossing possibilities in the northern part of Denmark).

If the crossing in the candidate timetable for the upgrading project is moved to obtain benefits locally, e.g. 10 minutes for one of the directions, it would result in nationwide changes. This is because most regional trains have connection(s) to and from IC-trains. A change in the northern part of Denmark will therefore influence the regional trains between Copenhagen and Nykøbing F (in the southern part of Denmark) because of the connection at Ringsted cf. figure 3. This change may very well result in time benefits (or losses) at other lines of the network.

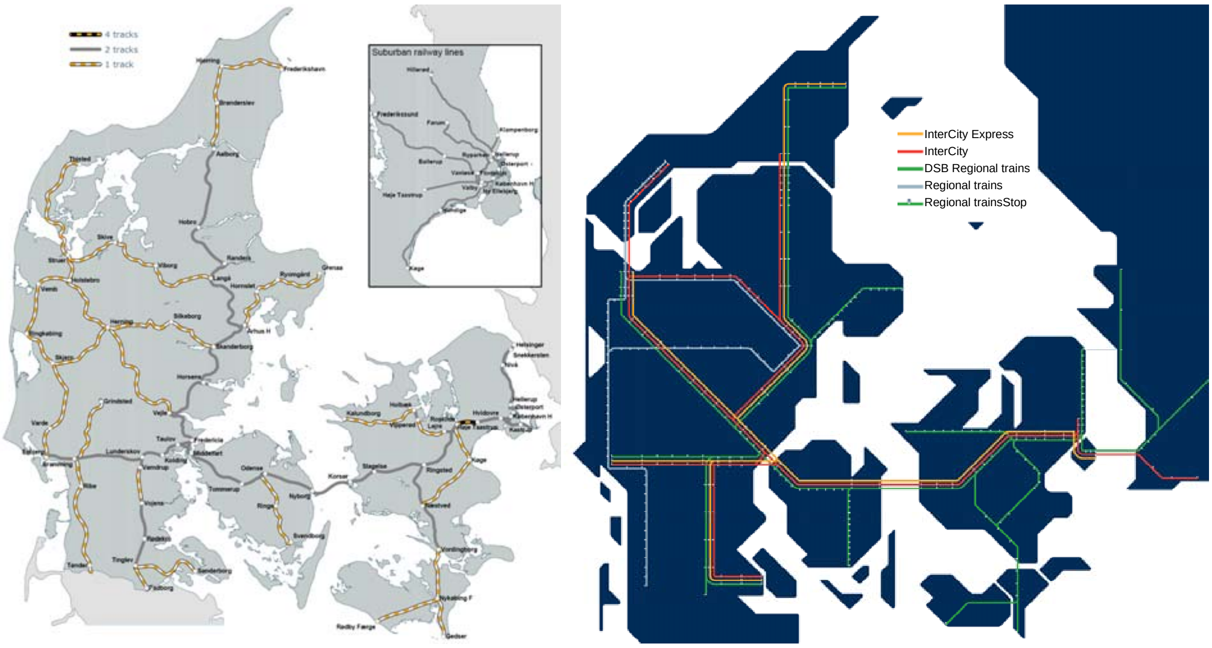


Figure 3: The Danish railway infrastructure. Based on (Rail Net Denmark 2006) and (DSB 2007)

3 Network effects for trains

Network effects for trains can be illustrated by queuing time. Queuing time is the difference in running time when comparing a single train on a line with a situation with many trains on the line. Queuing time on railway lines occurs when the traffic intensity is close to the capacity level due to e.g. mixed operation (slow and fast trains). When close to the capacity level, the operation speeds of fast trains must/will adapt to the slower trains cf. figure 4. This will increase the travel time for the trains that under free conditions could run at higher speeds (Salling, Landex 2006).

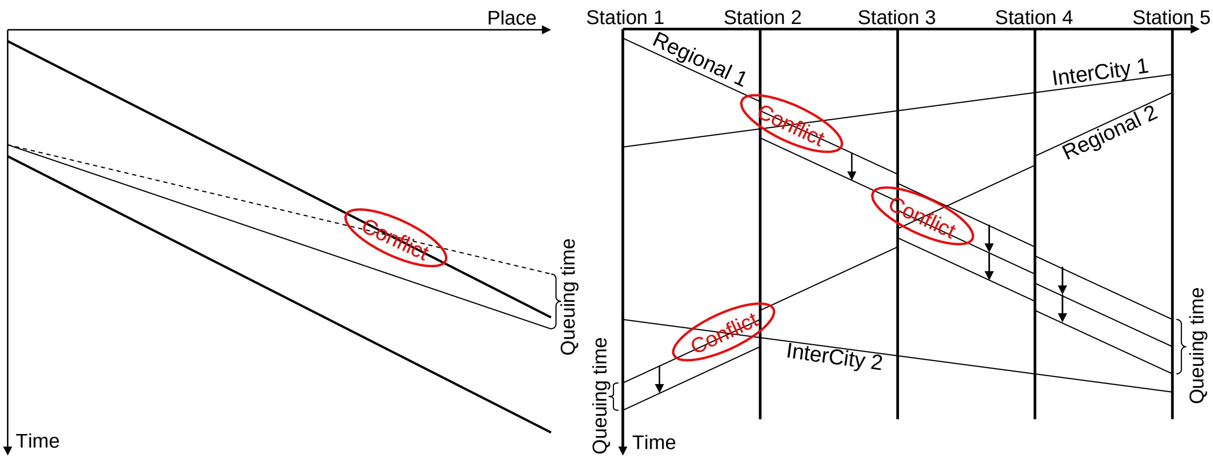


Figure 4: Extended running time (queuing time) for trains due to other trains on the railway line (double track to the right and single track on the left). Partly based on (Salling, Landex 2006)

To calculate the queuing time for trains the Danish developed SCAN model (Strategic Capacity Analysis of Network) ((Kaas 1998b) and (Kaas 1998a)) can be used (a similar function is found in the German tool UX-SIMU). SCAN is a computer tool for calculation of capacity in a railway network. In SCAN capacity is measured as average queuing time in a sample of candidate timetables for a given infrastructure alternative.

SCAN can be used in the strategic planning process where the exact infrastructure and timetable is not determined. Therefore, the system is based on a structure where it is only necessary to know the plan of operation (i.e. the number of trains within each category), the infrastructure in a simple way and the main dynamics of the rolling stock (Kaas 1998a). Based on the infrastructure and the plan of operation different timetables are simulated and the queuing time is calculated, cf. figure 5.

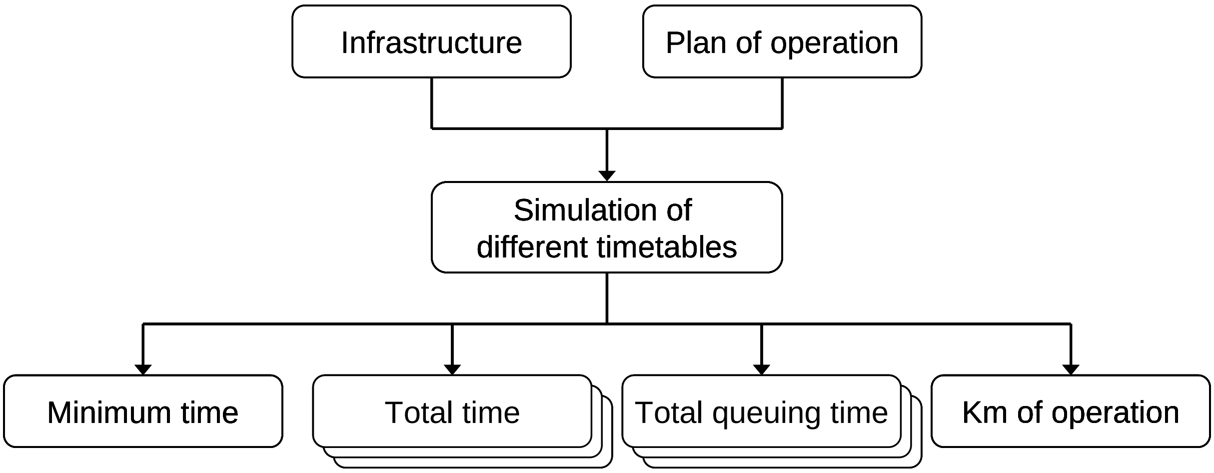


Figure 5: Calculation of queuing time

Examining a large number of different timetables based on the same plan of operation will result in different queuing times. These different queuing times can then be ordered according to the queuing time as shown on figure 6. It is then possible to see the span in queuing time and choose the timetable that has the lowest queuing time and still fulfils other potential requirements for the timetable – e.g. possible transfers between trains.

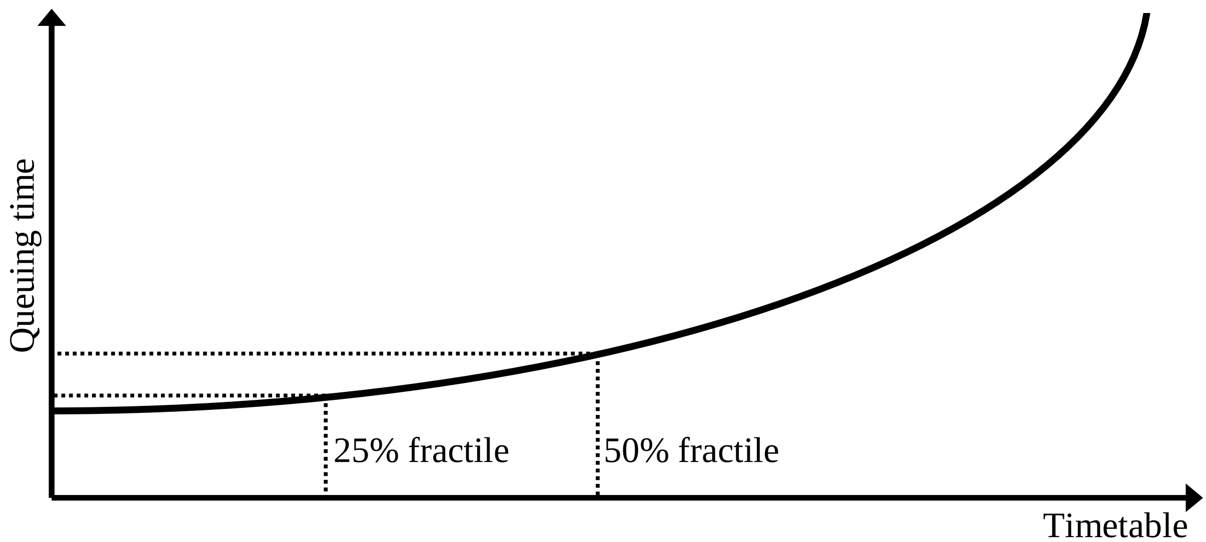
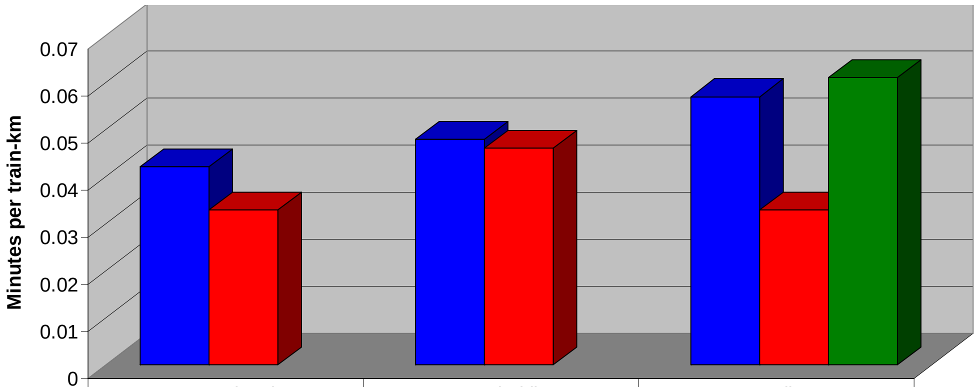


Figure 6: Sorting the timetables according the queuing time – including 25% and 50% fractiles of the timetables

Based on figure 6 the queuing time of a plan of operation (on a given infrastructure) can be evaluated. However, the final (chosen) timetable is not necessarily the timetable with the lowest queuing time as other considerations are taken in the timetabling process. In this way e.g. the 25% or 50% fractile of the timetables can be determining for the expected queuing time of the plan of operation.

Previous analyses ((Hansen 2004), (Hansen, Landex & Kaas 2006) and (Landex, Kaas & Hansen 2006)) have shown that both the size of the network and the connections between trains (correspondences) influences the network effects, cf. figure 7.



	Current situation	Extended line	New line
East Denmark	0.042	0.048	0.057
Copenhagen-Ringsted	0.033	0.046	0.033
East Denmark with fixed transfers			0.061

Figure 7: Queuing time/Network effects Copenhagen-Ringsted. Based on (Hansen 2004), (Hansen, Landex & Kaas 2006) and (Landex, Kaas & Hansen 2006)

Figure 7 illustrates network effects in terms of queuing time for the Danish railway line between Copenhagen and Ringsted for different scenarios. Two different analysis areas have been identified:

- The entire Eastern Denmark, until Lillebælt
- The infrastructure between Copenhagen and Ringsted only

It appears from figure 7 that the queuing time – or network effects – is increasing with the size of the analysis area. Also transfers between trains (correspondences) increase the queuing time.

The reason for the increase in queuing time is the higher complexity of the operation. Correspondences reduce the degree of freedom in the timetabling which result in a higher risk of queuing time. To avoid an increase in queuing time; timetable planners have to be more precise when timetabling for larger networks (and networks with correspondences) than for a railway line with no track connection to other railway lines.

4 Network effects for passengers

The network effects – or queuing time – above is described only for the trains and not the passengers. Network effects for passengers are more complex to calculate than the network effects for the trains. The higher complexity is because the passengers are affected of the networks effects of the trains. Furthermore, passengers often have more options for a journey. The waiting time at the station(s) should also be included in the calculation of network effects for the passengers. To be able to calculate the network effects for the passengers it is necessary to know the timetables – and thereby the network effects for the trains.

As described in section 3 transfers is a network effect. Transfers in public transport networks are unavoidable as it is not possible to design a network where all passengers can travel the direct way from their origin to their destination. It is not all transfers in (larger) public transport networks that will have good correspondences as improving one transfer/correspondence might worsen other correspondences.

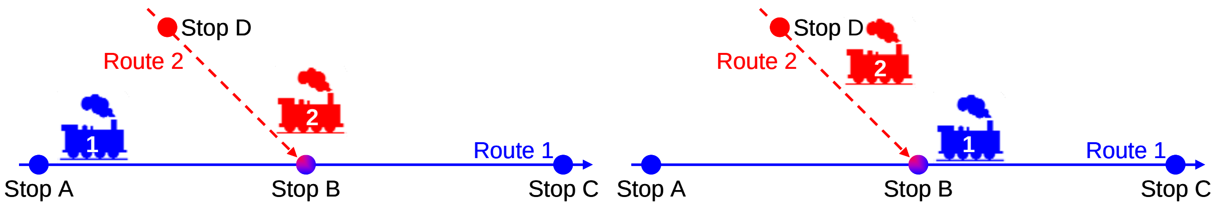


Figure 8: Correspondence between two routes (left) and no correspondance (right)

Figure 8 illustrates a simple railway network with and without correspondence at "Stop B". The travel time between "Stop D" and "Stop C" varies depending on the timetable and thereby the correspondence at "Stop B", cf. table 1. The timetables can in this case be optimized to minimize the travel time for passengers travelling from "Stop D" to "Stop C". In the optimized timetable in table 1 there is 2 minutes of transfer time even though one minute is enough. The extra transfer time in the timetable is to reduce the risk of missing the next train if the first train is delayed.

Table 1: Timetable scenarios for simple railway network (needed transfer time is 1 minute)

	Scenario 1			Scenario 2			"Optimized"		
Stop D	6	–	–	12	–	–	8	–	–
Stop A	–	8	28	–	8	28	–	8	28
Stop B	10	14	34	16	14	34	12	14	34
Stop C	–	18	38	–	18	38	–	18	38
Total time D→C	12 minutes			26 minutes			10 minutes		

The example above is straightforward to overview and optimize but for more complex networks the optimization of correspondences becomes complex. Figure 9 shows a journey with two transfers. In the beginning and in the end of the journey there are train routes with half-hour frequency but in-between there is a 5-minute-frequency train route. Examining the transfers independently there is good correspondences at both stops but the passengers in the left example in figure 9 will have no correspondence at the second transfer due to the long waiting time while there is a correspondence on the example to the right in figure 9.

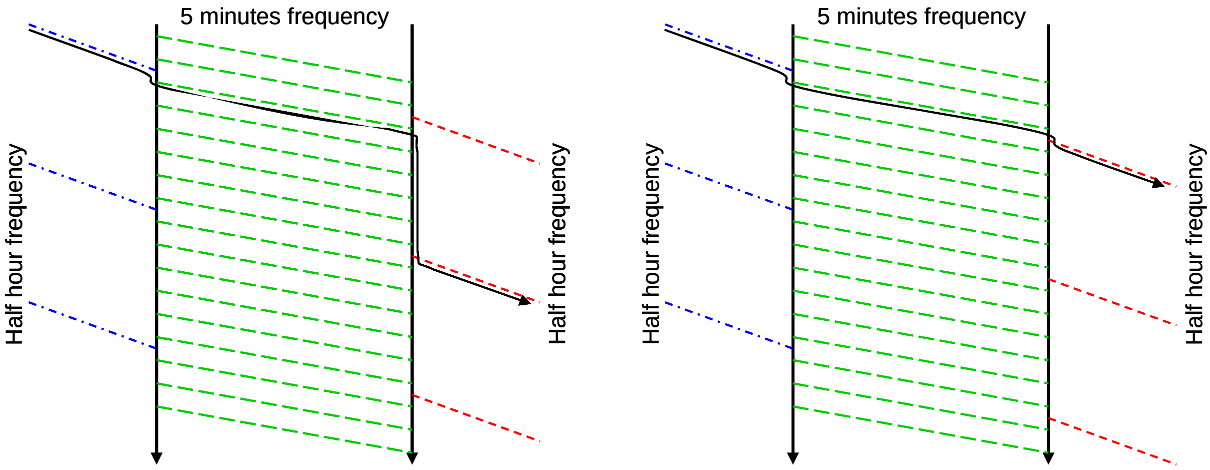


Figure 9: Journey with two transfers: lack of correspondence (left) and correspondence (right)

The network effects of the passengers can due to the dependency on the infrastructure and the timetables be estimated as the (additional) time the passengers spend in the system. This measurement for the network effects is similar to the queuing time measurement for the trains.

As the amount of time – or network effects – varies depending on the amount of lost correspondences, the network effects depends on the punctuality of the railway system. To take the punctuality of the railway system into account when calculating the network effects for the passengers it is necessary to simulate the (candidate) timetables.

4.1 Calculation of network effects for the passengers

Network effects for passengers is basically the delays of the passengers compared to the "optimal" timetable. This definition of network effects for the passengers is very similar to the queuing time for network effects for the trains. The network effects for passengers should include the trains' risk of delays in the operation too. This is because correspondences is a network effect and that passengers might loose their correspondences if the trains are delayed.

Passenger delays can be calculated in different ways, cf. table 2. The simplest way of calculating passenger delays is denoted the 0th generation models. Here the train delays are examined and eventually multiplied with the number of passengers. The 0th generation passenger delay models have the disadvantage that they do not take passengers route choice into account and that the models count passengers who (due to the delays) have reached an earlier train than planned as delayed too[1].

Table 2: Methods to calculate passenger delays (Nielsen, Landex & Frederiksen 2007)

	Train delays (0^{th} generation)	Cross section delays (0^{th} generation)	Counting train delays (0^{th} generation)	Optimal route choice model (1^{st} generation)	1½ generation model	Passenger delay model (2^{nd} generation)	Passenger delay model (3^{rd} generation)
Considerations of passenger delays	No	Partly	Partly	Partly	Partly	Yes	Yes
Complexity of the method	Very simple	Low	Low	Medium	Medium	High	High
Needs of information on passenger demand	No	Average aligning passengers	Counted passengers	OD matrix	OD matrix	OD matrix	OD matrix
Passengers may predict delays in the future (full information is assumed)	No	No	No	Yes	Yes	Partly	Can be incorporated
Passengers may arrive before time if a better con-nection emerge	No	No	No	Yes	Yes	Yes	Yes
Accuracy	Very low	Quite low	Fairly low	Low	Medium	Medium	High
Bias	Mostly under-estimate delays	Will quite often under-estimate delays	Will fairly often under-estimate delays	Large under-estimation of delays	Under-estimate delays	No sys-tematic bias	No system-atic bias

1^{st}, 1½ and 2^{nd} generation passenger delay models uses route choice models to estimate the passenger delays. However, the passengers know the delays before they occur. For 2^{nd} generation passenger models the passengers only know the probability of delays based on experience and can plan their route according to that. Using 3^{rd} generation passenger delay models the passengers do not know delays before they occur why they cannot react on the delays before they know about them.

Estimating the network effects of the passengers without delays 1^{st} generation models and above can be used. Thus all these models can calculate the time the passengers spend in the railway network. However, 3^{rd} generation models are the best if the network effects should be calculated in case of train delays or if sensitivity analyses have to be carried out. As 3^{rd} generation models require the same work effort as 1^{st}, 1½ and 2^{nd} generation models it is

[1] The paradox that passengers due to train delays are travelling earlier than planned is described in (Landex, Nielsen 2006b)

recommended that 3rd generation models are used as it is possible to extend the analyses in the future.

4.1.1 3rd generation passenger delay models

3rd generation models assume that passengers are planning their optimal desired route according to the official timetable (or by incorporating expected delays using a 2nd generation model). However, if delays occur over a certain threshold during the trip, the passengers are assumed to reconsider the route at that point in time and space along the route. If a train is completely cancelled, the passengers reconsider their choice without a threshold.

The main benefit of 3rd generation models is that it is more realistic and precise than the prior generations of passenger delay models. The disbenefit is that it is more complicated to implement, and that the calculation time is larger. This is because the route choice model has to be re-run at the point in time and space where the schedule is delayed.

The model uses the optimal paths (or paths taking expected delays into account) in the planned timetable for two purposes:

1. To compare planned travel times with the ones in the realized timetable
2. To estimate an a priori path choice strategy for the passengers.

A 1st or 2nd generation passenger delay model is therefore used to calculate the initial solution for the 3rd generation model.

A core assumption is that the paths the passengers choose in the a priori path choice strategy are stored as a sequence of lines (each with a specific run) and transfer stations. The passengers are then assumed to try to follow the same sequence of transfer stations and lines as planned, but the passengers may use different train runs for each line. The difference in passenger's time between first and the second route choice assignment equals the passenger's delays. The workflow of 3rd generation passenger delay models is shown in figure 10.

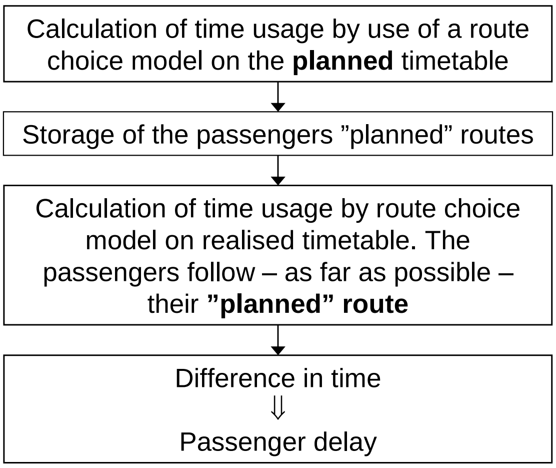


Figure 10: Workflow of 3rd generation passenger delay models

This approach calculating passenger's routes is somewhat similar to a rule-based assignment. To make this feasible, the rule-based network and diachronically graph interact by pointer structures that are built in memory as the graph is built (somewhat similar to the principles in (Nielsen, Frederiksen 2006). To ease the formulation if the model, it distinguish between whether the planned routes contain transfers or not (Nielsen, Frederiksen 2005).

4.1.2 Calculating passenger delays

As described in section 3 it is relatively straightforward to calculate the network effects for trains. However, cases with small network effects for trains do not necessarily result in small network effects for passengers – and vice versa.

Although the SCAN model calculates the network effects for trains in terms of queuing time (Kaas 1998a), it can be used as an input to calculate the network effects for passengers too. This is because the output timetables of SCAN can be used together with a route choice model to calculate the passenger's time usage in the railway network in case of no delays. In this way the network effects for the passengers can be determined as the difference between the times used in the actual analyzed timetable and the best analyzed timetable.

A problem with the modelling approach described above is that the SCAN model does not use time supplements why the model only can be used to evaluate the plan of operation. Alternatively, the North American Train Performance Calculator (TPC) (White 2007) can be used instead to generate a large amount of timetables which can be investigated. However, the TPC model is developed for North American conditions where there is no regular timetable – the trains are operated more or less improvised (White 2005). While including time supplements in the SCAN model and/or adapting the TPC model for regular timetables the methodology is well suited for strategic analyzes in the Danish/European content.

Simulation models based on future plans of operation are well suited for strategic analyzes but it is difficult to examine where the problems are most severe. Therefore, it is difficult to examine where the infrastructure should be improved and the effect of the improvement. Furthermore, the strategic analyzes do not take (risk of) delays into account why the results do not reflect the actual operation.

To reflect the actual operation and to be able to examine problems in the infrastructure, "traditional" simulation is necessary. Therefore, it is necessary to build up the infrastructure and timetables before simulating the operation in case of disturbances and then evaluate the results for both trains and passengers, cf. figure 11.

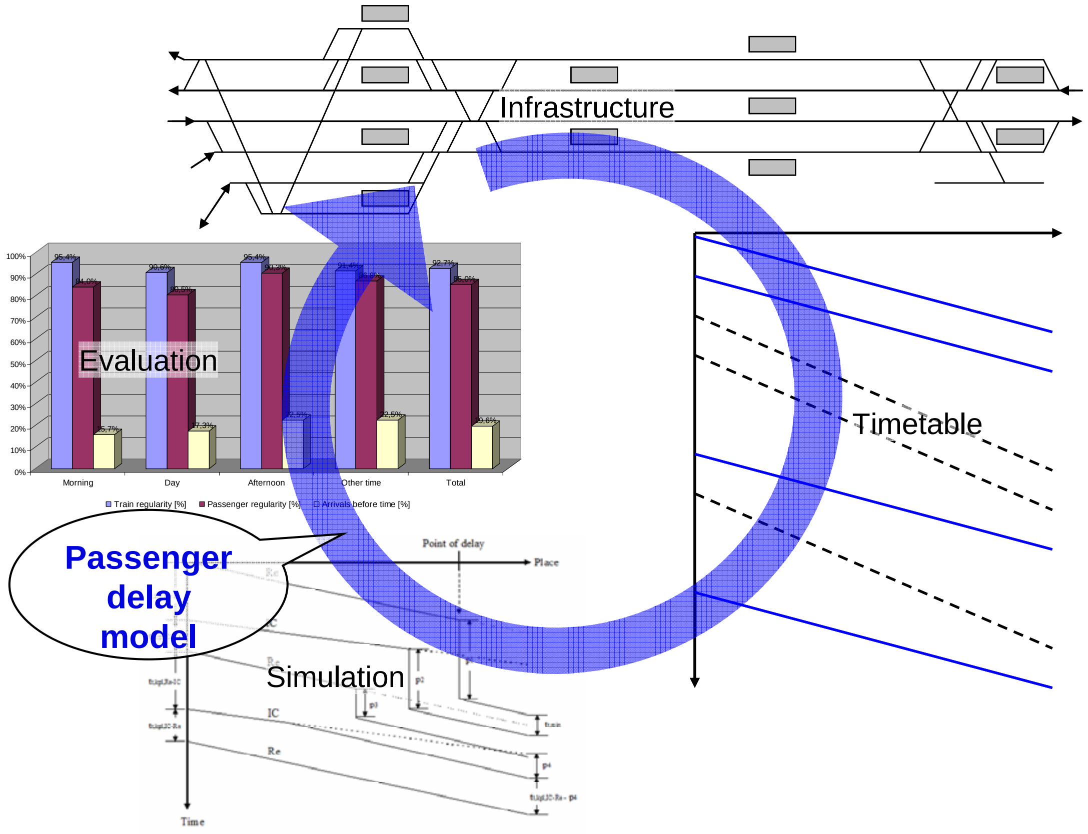


Figure 11: Traditional simulation of railway traffic with passenger delays. Based on (Landex, Nielsen 2006a)

The different infrastructures and timetables will result in different amount of time for the passengers. These different amounts of time can then be compared used to evaluate the network effects of the passengers. However, traditional simulation projects are time consuming but by combining microscopic and macroscopic models – so-called meso models – as e.g. done by Railnet Austria (Sewcyk, Radtke & Wilfinger 2007) can reduce the workload of simulations.

5 Discussion

Network effects of passengers can be used to improve the timetables for the passengers. This can be done by comparing timetables from different years and evaluate different travel relations together with the total time spend in the railway system.

It is not only possible to evaluate previous and present timetables. By examining different candidate timetables it is possible to examine the network effects of future timetables. In this way different timetable strategies can be examined – e.g. an additional overtaking. When examining the network effects of an additional overtaking it is possible to evaluate both the time gain for the passengers in the fast train and the time loss for the passengers in the train that is overtaken. This examination can either be done locally for a single railway line or for the entire system including transfers to/from other trains.

Improving the timetables not taking the risk of delays into account can result in a too optimized timetable where even small delays will result in lost correspondences for the passengers etc. To take common delays into account simulation of the timetables with a typical delay distribution can be performed. The time for the passengers – or the network effects – can then be calculated based on the simulated timetables. In this way it is possible to optimize the timetables for the passengers and make the timetables robust for network effects and future delays.

In the longer term this approach can also be used in the centralized control offices to decide if a train should wait for a delayed train to obtain the correspondence. This is because the simulation of the traffic combined with calculating the network effects – and thereby time – for the passengers can be used to evaluate the consequences of different scenarios. In this way it will be possible to improve the operation of the trains – and although a train might become more delayed the passengers will arrive more punctual.

Although a decision support system for centralized control offices based on network effects for passengers has a distant prospect, calculation of network effects for passengers can improve the operation on short term too. This is because network effects for passengers can be taken into account when planning for contingency operation. When the troubled operation then occurs and the timetable for contingency operation is taken into operation the network effects of the passengers will be taken into account implicitly.

6 Conclusion

Railway operation is often affected by network effects as a change in one part of the network can influence other parts of the network. This influence can even be far away from where the original change was made. The network effects occur because the train routes (often) are quite long and that the railway system has a high degree of interdependencies as trains cannot cross/overtake each other everywhere in the network.

Network effects can affect both trains and passengers. Network effects for trains can be measured by queuing time for the trains while the network effects for the passengers can be measured as passenger delays compared to the optimal timetable. It is more complex to calculate network effects for passengers as the network effects for the passengers depend on the network effects for the trains. Moreover, delays in the operation can enlarge the network effects for the passengers as correspondences might be lost.

This article suggests methods to calculate network effects for trains and passengers. Using these methods to calculate network effects for different candidate timetables it is to test different timetable strategies and choose the best strategy for the final timetable. In this way it is possible to improve the timetables for both the operator(s) and the passengers. In the longer term the approach can also be used in case of contingency operation. Here an evaluation of the network effects can be used to choose the dispatching strategy, which results in the fewest network effects.

7 References

DSB 2007, “Trains in Denmark”, *Timetables from DSB*, Copenhagen, Denmark.

Hansen, S. 2004, *Large Transport Infrastructure Investments and their Strategic Impacts with a Special Focus on Enterprises*, Centre for Traffic and Transport, Technical University of Denmark.

Hansen, S., Landex, A. & Kaas, A.H. 2006, "The network effects of railway investments", *Computers in Railways X*, eds. J. Allen, C.A. Brebbia, A.F. Rumsey, G. Sciutto, S. Sone & C.J. Goodman, WITpress, Prague, Czech Republic, pp. 45.

Kaas, A.H. 1998a, "Development and practical use of a capacity model for railway networks", *Structural Integrity and Passenger Safety*, ed. C.A. Brebbia, WITpress, Great Britain, pp. 73.

Kaas, A.H. 1998b, *Methods to calculate capacity of railways*, Institute for Planning, Technical University of Denmark.

Landex, A., Kaas, A.H. & Hansen, S. 2006, *Railway Operation*, Centre for Traffic and Transport, Technical University of Denmark.

Landex, A. & Nielsen, O.A. 2006a, "Modelling expected train passenger delays on large scale railway networks", *7th World Congress on Railway Research*, UIC, Montreal, Canada.

Landex, A. & Nielsen, O.A. 2006b, "Simulation of disturbances and modelling of expected train passenger delays", *Computers in Railways X*, eds. J. Allen, C.A. Brebbia, A.F. Rumsey, G. Sciutto, S. Sone & C.J. Goodman, WITpress, Prague, Czech Republic, pp. 521.

Nielsen, O.A. & Frederiksen, R.D. 2006, "Optimisation of timetable-based stochastic transit assignment models based on MSA", *Annals of Operations Research*, vol. 114, no. 1, pp. 263-285.

Nielsen, O.A. & Frederiksen, R.D. 2005, "Modelling train passenger delays", *Symposium on The Reliability of Travelling and the Robustness of Transport Systems*, ed. H.J. Zuylen, The Netherlands Research School for Transport, Infrastructure and Logistics, The Netherlands, pp. 137.

Nielsen, O.A., Landex, A. & Frederiksen, R.D. 2007, "Passenger Delay Models for Rail Networks" in *Proceedings volume of the second workshop on the Schedule-based approach in Dynamic Transit Modelling (SBDTM 2005)* Kluwer Academic.

Rail Net Denmark 2006, "Network statement", *Report from Rail Net Denmark*, Copenhagen, Denmark.

Salling, K.B. & Landex, A. 2006, "Computer based ex-ante evaluation of the planned railway line between Copenhagen and Ringsted by use of a new Decision Support System named COSIMA-DSS", *Computers in Railways X*, eds. J. Allen, C.A. Brebbia, A.F. Rumsey, G. Sciutto, S. Sone & C.J. Goodman, WITpress, Prague, Czech Republic, pp. 65.

Sewcyk, B., Radtke, A. & Wilfinger, G. 2007, "Combining Microscopic and Macroscopic Infrastructure Planning Models", *2nd International Seminar on Railway Operations Modeling and Analysis*, eds. I.A. Hansen, A. Radtke, J.P. Pachl & E. Wendler, IAROR, Germany.

White, T.A. 2007, "The development and use of dynamic traffic management simulation in North America", *2nd International Seminar on Railway Operations Research,* eds. I.A. Hansen, A. Radtke, J.P. Pachl & E. Wendler, IAROR, Germany.

White, T.A. 2005, "North American Experience with Timetable-Free Railway Operation", *1st International Seminar on Railway Operations Modelling and Analysis*, eds. I.A. Hansen, F.M. Dekking, R.M.P. Goverde, B. Heidergott & L.E. Meester, IAROR, The Netherlands.