Denne artikel er publiceret i det elektroniske tidsskrift
Artikler fra Trafikdage på Aalborg Universitet
(Proceedings from the Annual Transport Conference
at Aalborg University)
ISSN 1603-9696
www.trafikdage.dk/artikelarkiv

Joint modeling of schedule- and frequency-based services in public transport assignment models

Morten Eltved[A]*, Otto Anker Nielsen[A], Thomas Kjær Rasmussen[A], Rasmus Dyhr Frederiksen[B]
A: DTU Management Engineering, Bygningstorvet 116, 2800 Kongens Lyngby
B: Rapidis APS, Tobaksvejen 21 2. mf, 2860 Søborg
*Corresponding author: Morteneltved@hotmail.com

Abstract

Public transport networks today are getting increasingly complex with many lines and possibilities to go from origin to destination. When passengers make their route choice in a public transport network they cannot depart at the minute they want, but must wait until the first departure on the considered line. But how do passengers plan their route if they have a combination of high frequent and low frequent services under the assumption that passengers do not consult the timetable for high frequent services?

This paper describes a framework to include both schedule- and frequency-based services in a joint model. Such networks are found in most major cities and especially in the greater Copenhagen area where there is a mix of frequency-based services such as A-busses and the metro and schedule-based services as the S-train and local bus lines. Four different transfer types are identified when transferring between schedule- and frequency-based services. These include a type where the passenger transfers from a frequency-based to a schedule-based line. In this case the passenger has a probability to reach the first departure on the schedule-based line, but can in some cases also miss the first departure and must wait for the next departure.

A choice set generation method is developed using the event dominance principle to exclude alternatives which are above a certain threshold. This gives a choice set which is used in a discrete choice model (MNL). On this basis, it is possible to distribute the flow across the different alternatives. Two example cases are used to show the methodology: DTU to Copenhagen Airport and DTU to Brønshøj. The results indicate that there the framework can handle the two types off lines. It is found that the desired departure time, parameters in the utility function and the choice of threshold is crucial to find the correct choice set and distribution of flow across the alternatives. But there is also found improvement points in the choice set generation technique, but especially the discrete choice modelling should be investigated further to include that passengers can take decisions en route.

Introduction and background

This paper concerns the ideas and possible methods for creating a unified schedule- and frequency-based public transport assignment model. This paper is based on the master thesis of the corresponding author where the aim was to develop, test and discuss possible methodologies to make a joint model, which can capture the passengers' behavior in a mixed network in the best way.

In the era of public transport assignment models there has been two main ways to represent the supply in the model – frequency- or schedule-based. The supply side started as being described with frequencies, implying that the passenger would assume that a line would run with a given headway but without knowing the exact timetables (Nökel and Wekeck, 2007). This allowed the path searches to be done in a static model with no time dimension. Later the models developed into schedule-based models where each run on a line is included with the times when it passes different stops on the route. The schedule-based models allowed the modeler to better represent coordination between lines and thereby describing the passengers route choice in a more detailed way (Liu et al., 2010).

Why is a joint model interesting?

The advantage in the schedule-based models is that it can model the coordination between lines much better, and in low frequent systems this coordination is very important for the passengers' route choice. But it requires more data input and more calculation time because the model should run for several possible departure times. The frequency-based models are much simpler and requires less input and calculation time. They work well with high frequent services, where the passenger might not consult the public timetable (if available), but if there are low frequent services in the network it is very difficult to estimate the transfer times between two lines.

As mentioned in Gentile and Nökel (2016, chap. 6) there is currently no framework to include these two different ways of representing the supply. They propose to take a decision on whether the network is better described with schedules or frequencies and choose the best option. They say that no simple way of making a combined model is available, proposing that one should assume truncated waiting time distributions for schedule-based services. But this is an assumption which requires many assumptions on what the truncated distribution should be and for which services the passenger behavior should be treated as schedule-based behavior. Another proposal in Gentile and Nökel (2016, chap. 6) is to rely on a schedule-based network and incorporate passenger cost functions for frequency-based lines, while still keeping a completely schedule-based network. This is a possible solution for modeling the current system, but when changing systems, it requires the modeler to take decisions on the exact departure times of the line which might influence the transfer times to corresponding lines.

There are 3 main aims for developing a joint model, which are stated below. The aims are described further in the subsections below.

- Modelling the existing situation more realistically
- Creating more reliable forecasts
- Simplifying modelling of future scenarios

Modelling todays situation more realistically

The public transport service in the greater Copenhagen area is a mix of schedule- and frequency-based lines. There are many schedule-based lines such as local busses, S-train and regional trains, but there is also bus lines and metro lines which operate with a published frequency rather than exact departures (DSB, 2016; Metro, 2017; Movia, 2017). Such an example is shown in Figure 1 below. Here line 3A runs with a published frequency of 6-8 minutes, but it is not informed to the customer when the exact departure of the line is. Therefore, the passenger needs to include additional buffer time, if the passenger needs to reach a schedule-based line or must finish the trip before a certain time. If the passenger instead takes a trip with 6A the departure time is given, and the passenger can therefore time his arrival and find the transfer time to other schedule-based lines.

3A (mandag - fredag)

Nordhavn St.	Trianglen	Elmegade	Åboulevard	Vesterbrogade	Sønder Boulevard	Sydhavn St.	Kgs. Enghave (Valbyparken)
0.00	0.05	0.10	0.13	0.19	0.22	0.24	0.28
00	**05**	**10**	**13**	**19**	**22**	**24**	**28**
30	**35**	**40**	**43**	**49**	**52**	**54**	**58**
4.30	4.35	4.40	4.43	4.49	4.52	4.54	4.58
4.50	4.55	5.00	5.03	5.09	5.12	5.14	5.18
5.10	5.15	5.20	5.23	5.29	5.32	5.34	5.39
5.30	5.35	5.40	5.43	5.49	5.52	5.54	5.59
5.40	5.45	5.50	5.53	5.59	6.02	6.04	6.09
5.50	5.55	6.00	6.03	6.09	6.12	6.14	6.19
6.00	6.05	6.10	6.13	6.19	6.22	6.24	6.29
6.10	6.15	6.20	6.23	6.29	6.32	6.34	6.39
6.19	6.24	6.29	6.32	6.38	6.41	6.44	6.50
6.27	6.32	6.37	6.40	6.46	6.49	6.52	6.58
6 - 8 min.							
18.49	18.55	19.01	19.04	19.11	19.15	19.18	19.24
18.59	19.05	19.11	19.14	19.21	19.25	19.28	19.34
19.09	19.15	19.21	19.24	19.31	19.35	19.38	19.44
19.20	19.26	19.32	19.35	19.41	19.45	19.48	19.54
19.30	19.36	19.42	19.45	19.51	19.55	19.58	20.04
19.40	19.46	19.52	19.55	20.01	20.05	20.08	20.14

6A (mandag - fredag)

Rødovrehallen	Damhustorvet	Zoologisk Have	Hovedbanegården	Nørreport St.	Bispebjerg St.	Emdrup Torv	Buddinge St.
0.17	0.24	0.31	0.39	0.46	0.56	1.02	1.12
0.47	0.54	1.01	1.09	1.16	1.26	1.32	—
17	**24**	**31**	**39**	**46**	**56**	**02**	**12**
47	**54**	**01**	**09**	**16**	**26**	**32**	**—**
3.47	3.54	4.01	4.09	4.16	4.26	4.32	—
4.17	4.24	4.31	4.39	4.46	4.56	5.02	5.12
—	—	—	—	5.01	5.11	5.17	—
4.47	4.54	5.01	5.09	5.16	5.26	5.32	5.42
5.02	5.09	5.16	5.24	5.31	5.41	5.47	—
5.14	5.21	5.28	5.36	5.43	5.53	6.00	6.10
5.25	5.32	5.39	5.47	5.55	6.06	6.13	6.23
5.36	5.43	5.50	5.58	6.06	6.17	6.24	6.34
5.45	5.53	6.00	6.08	6.16	6.27	6.34	6.45
5.55	6.03	6.10	6.18	6.26	6.37	6.44	6.55
6.04	6.12	6.19	6.27	6.35	6.47	6.54	7.05
6.12	6.20	6.27	6.35	6.43	6.55	7.02	7.14
6.21	6.29	6.36	6.45	6.53	7.05	7.12	7.24
6.27	6.35	6.42	6.51	6.59	7.12	7.19	—
—	—	—	—	7.03	7.16	7.23	7.35
6.34	6.43	6.51	7.00	7.08	7.21	7.29	—
—	—	—	—	7.12	7.25	7.32	7.44

Figure 1 - Example of a frequency-based line (3A) and a schedule-based line (6A) (Movia, 2017)

But even when there is no published schedule-based timetable, the passenger can find the departure time using Rejseplanen (Rejseplanen, 2017). An example of a trip with two lines which are both frequency-based in the public timetables is shown in Figure 2. This trip is including bus line 8A which runs every 5-8 minutes in peak hours without specific departure times and the metro which runs every 4 minutes to the airport. But the passenger is presented with specific times for both lines, and the passenger can therefore more exactly plan his trip. However, when both lines are high frequent many passengers will not consult a travel planner before departure, and would therefore just experience the trip as a frequency-based trip. So, this could indicate a need for a model including frequency-based services, but there is not much research showing how many passengers take decisions according to actual departure times or just the frequency of a line, and this should be carried out to further back up the argument that some passengers do not consult a travel planner before departing.

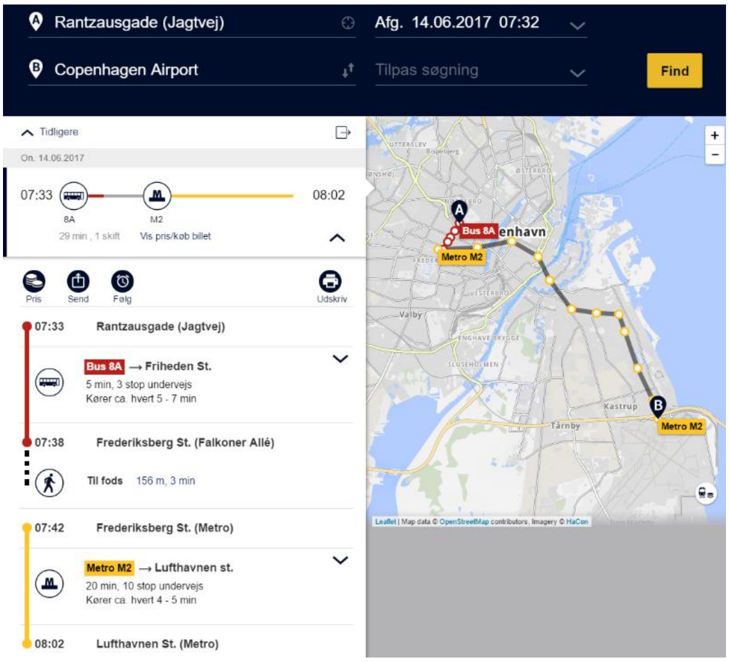

Figure 2 - Example of journey plan, where exact runs are published instead of frequencies (Rejseplanen, 2017)

Making more reliable forecasts

The second reason for considering a unified model is to make more reliable forecasts. Modeling the future is always subject to uncertainty about what the network supply will be in each year. Taking an example in the Copenhagen network, the bus line 150S which runs in the corridor from Kokkedal via DTU to Nørreport is today a schedule-based service in the published timetable. The bus runs with a headway of five minutes in the morning peak hour but with a given schedule that is available to the passengers. In the future, it might be known that 150S will also run with a headway of five minutes, but it is not known if the actual departure times will be the same. To get a more reliable forecast it is suggested to change the high frequency services to frequency-based lines and thereby not having to deal with coordination of different connecting lines.

Making it easier to model future scenarios

On top of the more reliable forecasts it could also be an advantage to have frequency-based lines in different scenarios for the future. The introduction of a new line in a network requires many assumptions on which lines the new line should coordinate with. Imagine that the bus company wants to make a forecasting scenario where some buses that corresponds to 150S have a higher frequency than today. This could result in some waiting times at transfers being longer (or shorter) because the arrival time at a given stop is changed. If the line 150S and the lines being changed were instead described as a frequency-based service, the waiting time at transfers would only change according to the service which is transferred to and would therefore be a more average consideration of the future scenario without a possible loss of good connections.

Scope of work

The study of public transport assignment models includes a large variety of challenges and many different model types. So, to clarify how this work fit into previous research and what has been the aim, the following five points describe the limitations of this work:

1. In difference to large scale models only one OD-relation is investigated in each example
2. Delays and in-vehicle crowding are not considered
3. It is assumed that all services depart in discrete minutes, which is also the case for the Danish National Transport Model (NTM)
4. There should be no use of simulation in both the choice set generation and assignment of passengers
5. The aim of this work has been to create a model with two stages: first a choice set generation phase which should feed into a discrete choice model to distribute the flow across alternatives

Methodology

This section describes the developed methodology and the main ideas behind the framework. The backbone of the model is a choice set generation phase which feeds into a discrete choice model. First different transfer types are described followed by a justification of choice of headway distribution. Then it is described how the choice set is generated and how this can be used in a discrete choice model.

There is a need to define the different transfer types to be able to incorporate all transfer types in the choice set generation phase. The four transfer types are identified between the schedule- and frequency-based services and these are described in Table 1. The schedule- to schedule-based transfer is deterministic given by the timetable because no delays are assumed. When transferring from a frequency- to a schedule-based service there is a probability to catch the first departure of a line (a candidate) and for this an associated transfer time. This is the case because the passenger does not know the exact departure time of the frequency-based service and therefore cannot be certain to catch the first candidate on a line. For the two last cases (3 and 4) a statistical distribution gives the transfer time. This also means, that if the passenger takes two frequency-based services in a row the distributions need to be additive (convolutive).

Table 1 - Overview of transfer types between schedule- and frequency-based services

Transfer to/from	*SB line*	*FB line*
SB line	The transfer time is deterministic (Case 1)	Transfer time from statistical distribution (Case 3)
FB line	Probability to catch the next service and an associated transfer time (Case 2)	Transfer time from statistical distribution (Case 4)

The convolution of two independent statistical distributions is not straightforward for many distributions. The most used statistical distribution in frequency-based models in the research is the exponential distribution (Chriqui and Robillard, 1975; Schmöcker et al., 2013). Li et. al (2015) estimated headway distributions for buses based on real life data and found that the exponential distribution did not describe the headway as good as for example gamma or Erlang distributions. The reason that these distributions give a good fit is because of the phenomenon known as bus bunching (Gentile and Noekel, 2016, chap. 2). All these distributions can in some way be convoluted, but in a model with no delays the distributions mentioned are not suitable. They all have long tails and thereby there would be a probability to wait for longer than one headway. The natural distribution to choose would be the discrete uniform distribution, because it describes regular headway and can be convoluted to a close form for up to 3 independent distributions. However, for four independent distributions there is no closed form, as the distributions approaches a normal distribution. So, for this work the binomial distribution is chosen. This distribution can be described with a maximum and a mean and multiple independent distribution with different means can be convoluted.

How to distribute flow between origin and destination

In previous studies by Friedrich et al. (2001) and Hoogendoorn-Lanser et al. (2014) they used Branch & Bound techniques to generate a choice set in a schedule-based network. They incorporated different behavioral constraints to limit the choice set and generate realistic alternatives. In Hoogendoorn-Lanser et al. (2014) they developed different criteria that should be fulfilled by a feasible route. If one criterion was not met, the route could not be included in the choice set. Obvious criteria were departure after arrival time at a node, and that routes should not include circles. After this criteria on time, space and money were used to prune out non-reasonable alternatives. The idea of using different rules is the base to the following algorithm to generate a proper choice set in a mixed network with schedule- and frequency-based services. The idea is to build upon the event dominance principle in Florian (2004, 1999) with introduction of a relaxation of the event dominance considering a threshold for including non-optimal paths.

The event dominance algorithm is, in its original implementation, used to find the shortest path in a schedule-based network. It finds only the best path in terms of cost of a route, however, with a slight modification more routes can be found if there is not a strict event dominance at each node. The original event dominance algorithm is described in Algorithm 1.

while an event at end node has not been checked **do**
Find cheapest event in heap, which has not been checked
for all possible event types going from the node considered **do**
Check if event can be inserted based on criteria on both time and cost compared to the events already at the tonode;
if can be inserted **then**
Insert event
Check if any event already existing event is dominated and remove dominated events
end
end

Algorithm 1 - Basic algorithm, where different rules can be applied

The events in the algorithm is meant as movements in the network. In this paper, these events are the following:

- Access or egress network on connectors
- Ride on line (schedule- or frequency-based)
- Transfer between stops
- Waiting at stop (handled implicitly)
- Transfer penalty (cost handled implicitly)

An earlier arriving event is denoted E_1 compared to the later arriving event E_2. It is important to note that all events are only checked against later events and at that arrives at the same stop. The new events are inserted if both of the two following rules are not violated.

$$E_1 TotalCost + (E_2 EndTime - E_1 EndTime)\beta_{\mathrm{wait}} \leq E_2 TotalCost$$

$$E_1 TotalCost + E_1 TotalCost * \beta_{slack} \leq E_2 TotalCost$$

The first rule ensures that if it is cheaper to arrive later, than arriving early to a stop and waiting, then the later arriving event is still inserted, as it can possibly make a good connection to the next service from the stop. The other criterion allows non optimal paths to be included. The slack parameter which defines the threshold is in this work set to 20%, which means that events which is 20% worse at a stop is discarded. As mentioned in the beginning of the section there are cases where it is not 100% certain to catch the first departure on a schedule-based line when coming from a frequency-based line. In these cases, both the first and second candidate must be under the threshold at the destination and if they don't, the path is not included in the final choice set.

The cost of each alternative is decided by the parameters used in the model. In this work the parameters from NTM is used, but there should naturally be new estimated model parameters when having decided on the final model type. To finally distribute the flow across the alternatives a discrete choice model is used. In this work a standard MNL model is used, but other models were discussed. First, it was discussed if a nested structure where each decision to board a vehicle would be a nest in the model. Another option was the recursive logit or recursive nested logit model where the flow is distributed from the destination and then back to the origin (Fosgerau et al., 2013; Mai et al., 2015). These recursive models seem applicable, as they would possibly be able to capture that a passenger does not reach the first candidate and therefore chooses another line instead, whereas the MNL model only include full alternatives. So, for future research these models should be considered. Also a restricted model should be given consideration, as the threshold in the choice set generation is similar to the work in Rasmussen et al. (2017).

Results

The developed framework is tested on two real life examples which include both schedule- and frequency-based services. The first example from DTU to Copenhagen Airport was simplified and used for developing the methodology, while the second example from DTU to Brønshøj is larger and was used to test different aspects of the developed methodology.

Results for example from DTU to Copenhagen Airport

The first example is a trip from DTU to Copenhagen Airport, where there are four main alternatives. The first choice is whether to go on the bus directly to Nørreport or to take the bus to Lyngby Station and from there take the S-train to Nørreport. From Nørreport there is the option of taking either the Metro or the Regional train to the airport. The driving times of the different lines and their frequencies and type of line is shown in Figure 3 below. There is both options of combining schedule- and frequency-based services, but there is also the option to only take for example frequency-based lines.

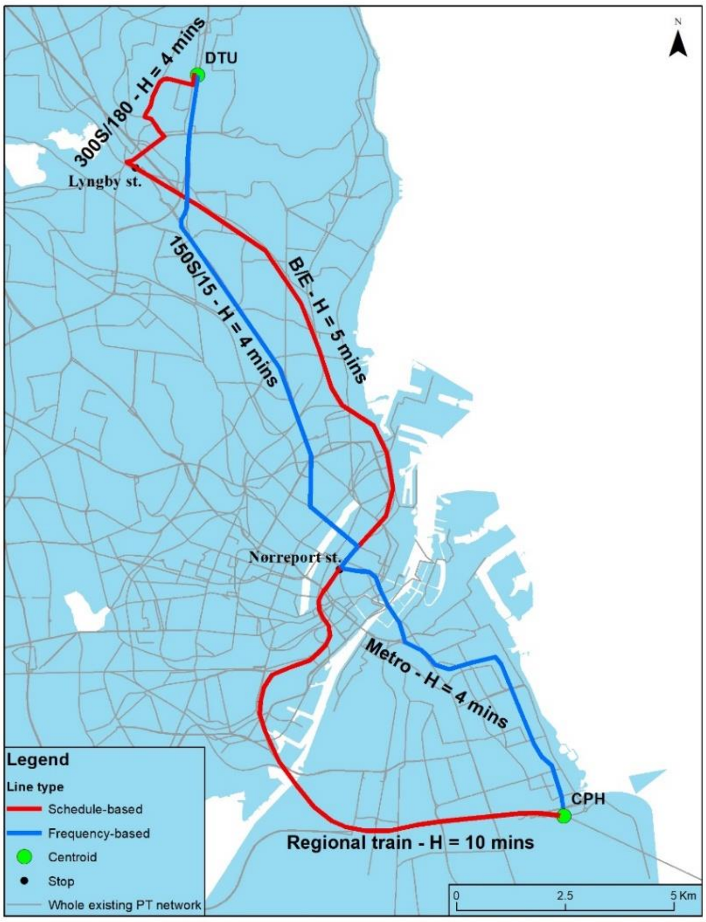

Figure 3 - Overview of example from DTU to Copenhagen Airport

Base results with all alternatives and the described methodology with departure at 7.30am

To show that the described framework can be used to get results, the difference between using all alternatives or using the developed framework is shown in Figure 4 and Figure 5. In the cost of each alternative is shown. Here it is clear that the best alternative is to take 150S/15E and then the metro. When using the developed framework, the alternative of going via Lyngby St. and taking the S-train is not included in the choice set because the cost is more than 20% higher than the least costly route. The flow change is in this case very small, since the excluded alternative is not very attractive in both travel time and travel cost.

Table 2 - Overall results of flow with all alternatives and with the developed model – Departure at 7.30am

Path sequence	*Arrival time*	*Cost*	*Probability (All alt. 7.30)*	*Probability 7.30*
150S/15E -> Metro	08:21	10.12	77%	78%
150S/15E -> Regional train	08:28	12.05	11%	11%
300S/180 -> B -> Metro	08:31	12.09	11%	11%
300S/180 -> B -> Regional	08:38	14.01	2%	-

Figure 4 - Flow with all alternatives

Figure 5 - Flow when having a threshold of 20%

Effect of changing departure time and preferences

The resulting flows are naturally depending on the preferences of the passengers and the desired departure time. In Figure 6 below, the flow for a desired departure time is shown. The main change is, that the alternative with 150S/15E and the regional train is not included, because there is only 69% probability to catch the first alternative of the regional train and if the first departure is not caught the cost exceeds the threshold. This is maybe not behaviorally realistic, as some passengers might just wait for the next departure if they don't catch the first departure. But this is a result of the strict implementation of the threshold.

Another test to see if the methodology works is to change the preferences for in vehicle time in train and S-train. These parameters are reduced by 25% and the threshold is changed to 30%. The change to a higher preference for train results in all paths being generated. However, the alternative of using 150S/15E and the regional train is only included in the choice set because of the threshold of 30%, as the cost of the second candidate is 20.4% higher than the shortest path. When comparing the flow of this example with the previous example with standard parameters and 20% slack, the biggest difference is between Nørreport and CPH where 30% of the flow now uses the regional train. There is also 14% more probability to use the S-train instead of using 150S/15E. These results show, that the algorithm can work with different parameters and that the results are depending on the threshold defined by the modeler.

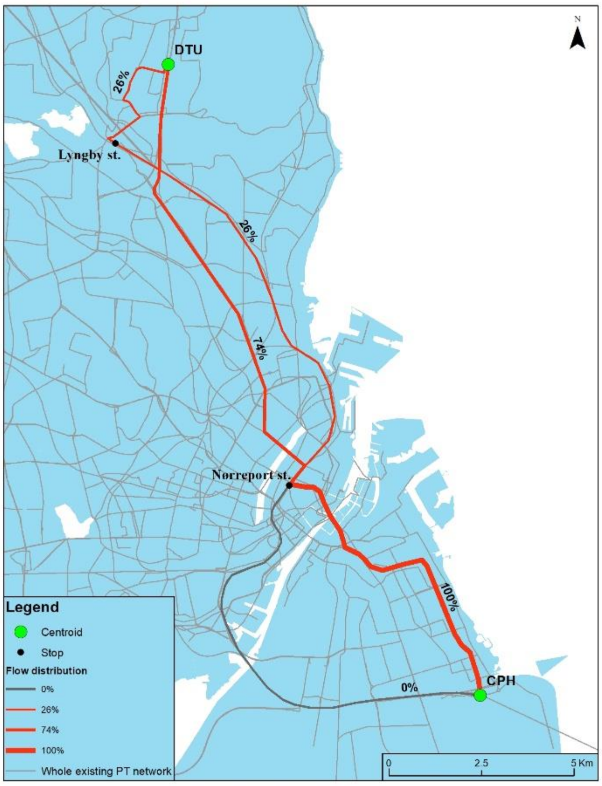

Figure 6 - Flow distribution for 7.35am

Figure 7 - Flow distribution with higher preference for train

Results for example from DTU to Brønshøj

The example from DTU to Brønshøj is not a relation where many passengers are assumed travel. But it has many good features for the use of testing the methodology:

- No obvious shortest path
- Many different lines and many equally good alternatives
- Mix of schedule- and frequency-based lines in the network
- Relatively small example, where most alternatives can be enumerated

The full overview of lines included in the example is shown in Figure 8 below. There are many options of going via for example Herlev, Ryparken or Tuborgvej, but all alternatives require at least one transfer. Some of the paths are a combination of schedule- and frequency-based services, but there are also many options where only schedule-based services are needed.

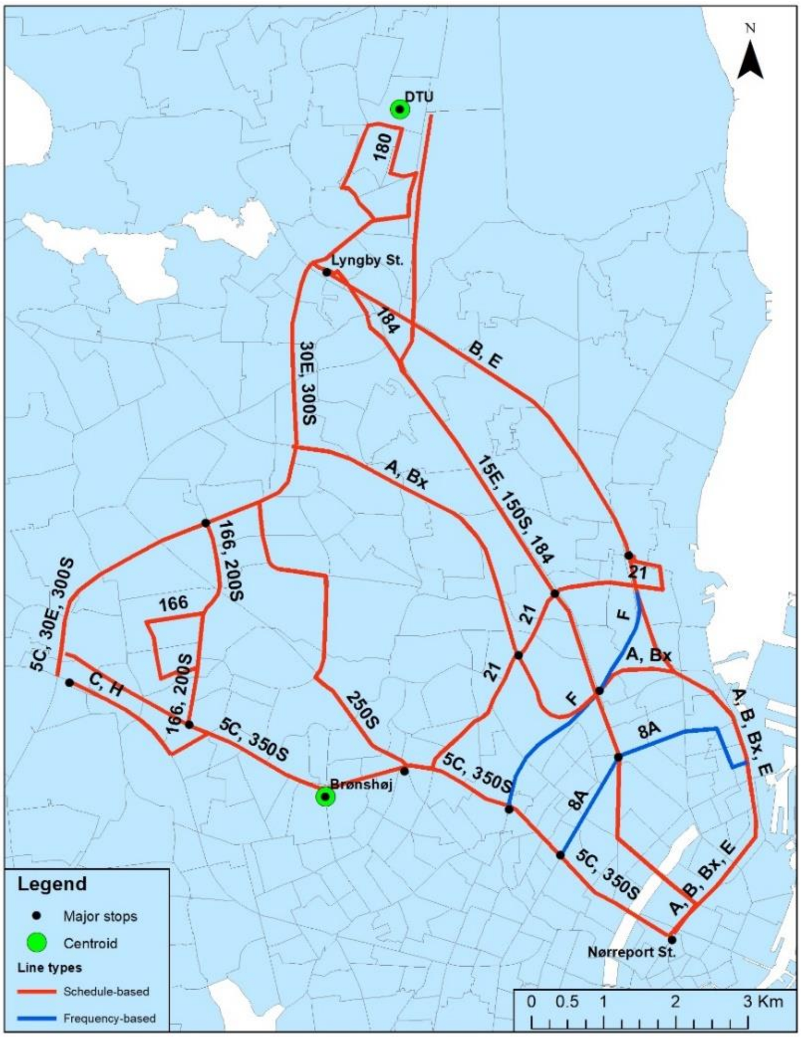

Figure 8 - Overview of lines included in the example DTU-Brønshøj

Base results

In Figure 9 the route shares for a desired departure time at 7.30am are shown. The main alternative is to take 300S to Herlev and transfer to 5C which gives an arrival time at 8:19 AM. The two lines are perfectly coordinated in Herlev, where the passenger has no waiting time when transferring. There are paths which arrive earlier (8:15), but because these paths require two transfers, they are less attractive than the path via Herlev. The choice set consists of eight different alternatives including the alternative of going to Lyngby Station and via Hellerup to Nørrebro and from then to Brønshøj which include three transfers. However, because the in-vehicle time parameter for S-train is around 20% lower than the parameter for bus, the path is still attractive. The weighted arrival time with the probability of taking each path is 8:18am and the weighted cost is 10.97. These values can be used to compare with the other examples on this network.

Almost all possible paths are included in the choice set, which proves that the algorithm works and that a slack of 20% is, in this case, enough to generate the alternatives that a passenger would usually consider. It is important to note, that no paths are 100% overlapping. This also means, that the path of taking 150S to Ryparken and going to Nørrebro Station and then taking 350S is included, while the alternative where the trip would end with 5C is not included. This is a case where, if a passenger doesn't catch the first departing 350S (transfer from the frequency-based F-line), the passenger would probably consider taking 5C instead if it departs before the next bus on line 350S. This problem is further described in the example with frequency aggregation.

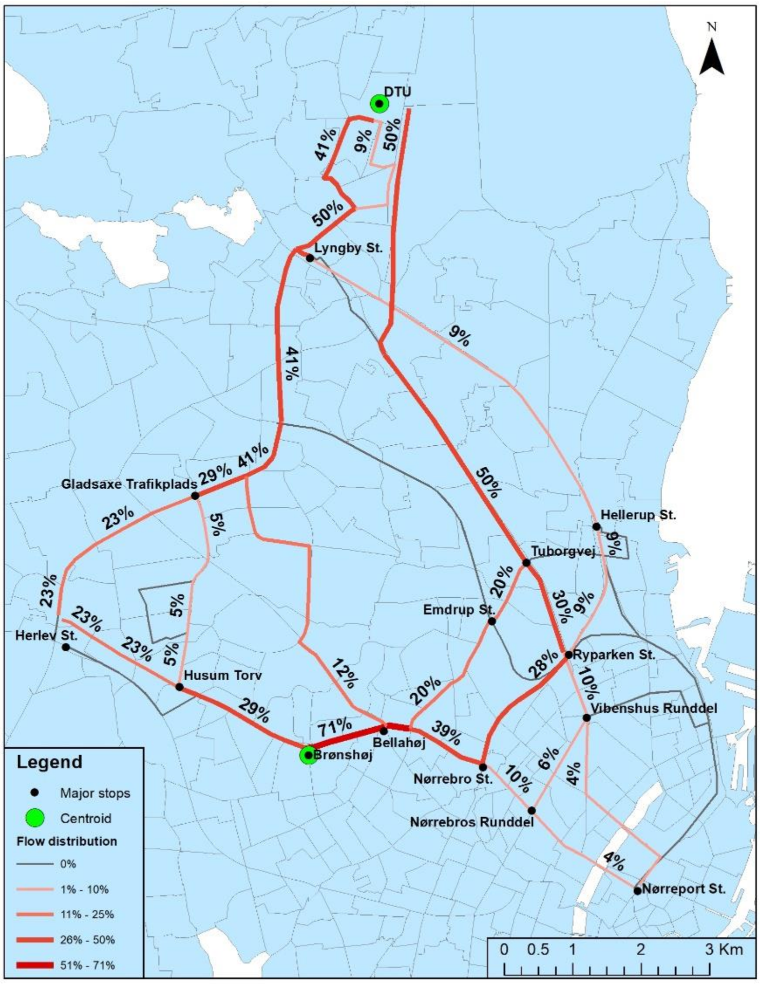

Figure 9 - Flow distribution for departure at 7.30am

Changing high frequent lines to frequency-based lines

As a test of how the framework reacts to changes in the network three lines are changed from schedule-based services to frequency-based services. These lines are 5C, 150S and 350S, which runs with respectively 4, 5 and 6 minutes headway. They are changed so that they run with the same frequency as now, but just without knowing the exact departure times. In contrast to the original network it is now possible to have paths with only frequency-based services in the choice set and the passenger do not have any option to arrive to Brønshøj using a schedule-based service.

The assignment on the new network leads to slightly different paths in terms of the line level and one new path has replaced the path including line 21. The new path is going via Herlev Station to Husum Station and includes a long transfer to Husum Torv and then finishing the trip on 5C. The reason that the path including line 21 is not included, is because the probability of catching the first candidate on line 21 from line 150S is 97% and the second candidate is not within the threshold, which means that this alternative is excluded. The shares between the paths are quite like the assignment in the existing network with the probability from the alternative with line 21 spread across the other alternatives. The arrival times are in general a few minutes later than in the existing network with a weighted arrival time at 8.20am. The costs are similarly a bit higher at 11.25 compared to the existing network which was 10.97. The higher costs are primarily the result of some well-coordinated transfers not being as good as it were with schedule-based services. The difference in route shares between the existing alternative and the new more frequency-based network can be seen in Figure 10.

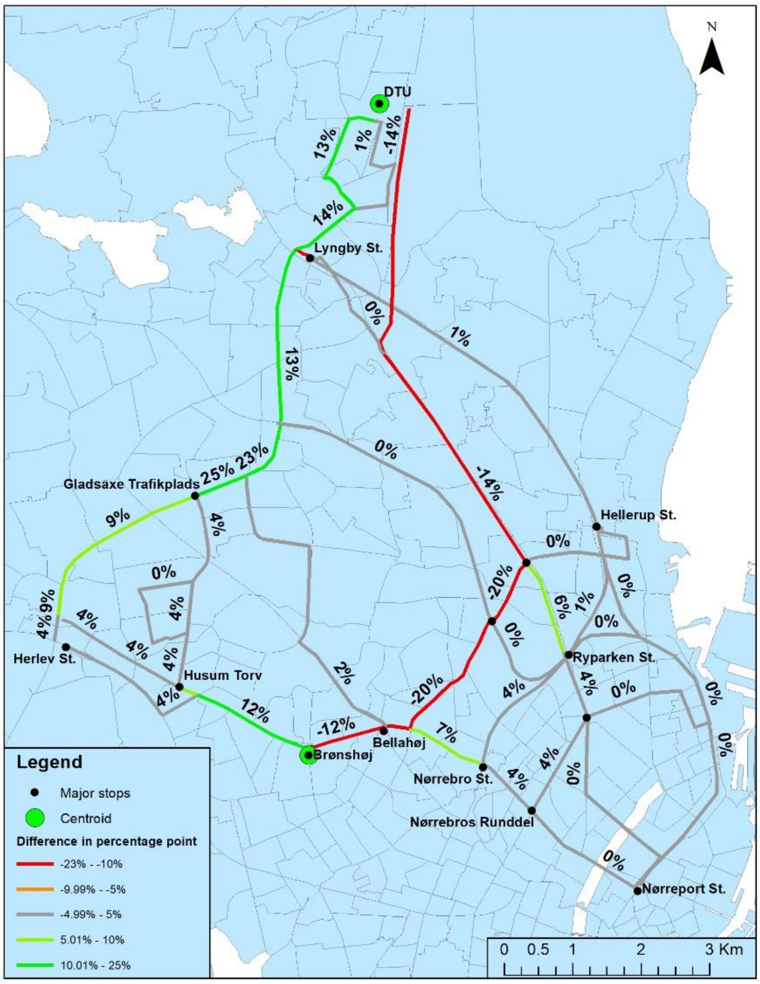

Figure 10 - Flow distribution for more frequency-based network with departure at 7.30am

Frequency aggregation of the network

As mentioned in the description of the results for the existing network there are a few corridors, where the passenger could possibly perceive two parallel lines as the same and therefore just board the first departing line. This is the case for 5C and 350S and for 150S and 15E. The aggregation is done on stop level, meaning that direct arcs between all stops on the line are created. If any of these direct arcs are similar on stop level between the two lines then they are aggregated. The specific rules are, that they are aggregated if the relative difference is less than 15% or the absolute difference in driving time is less than or equal to two minutes. The frequency of the aggregated line is then the sum of the two frequencies and the driving time is a weighted average compared to the number of departures pr. hour.

In Figure 11 the difference in flow compared to the flow in the existing network is shown. The alternative with 300S and 5C is the most attractive alternative as it is the case in the existing network. The cost for the path is the same for the two examples, but the other paths in the frequency aggregated network have higher costs and the choice probability is therefore higher in this example. Moreover, the alternative via Nørreport is not included in the choice set for the frequency aggregated network because it has a cost which is 20.5% higher than the best alternative and is therefore just outside the threshold. Overall the flows are not very different from the original flows, but the higher costs for alternatives going via the highway gives a slight change in the flow distribution. The weighted arrival time is 8.18am and the weighted cost is 11.09 and therefore closer to the cost of the existing network compared to the more frequency-based model. This result indicate that it might be a good opportunity to frequency aggregate and this is also discussed in the discussion below along with discussions of other aspects of the results in this section.

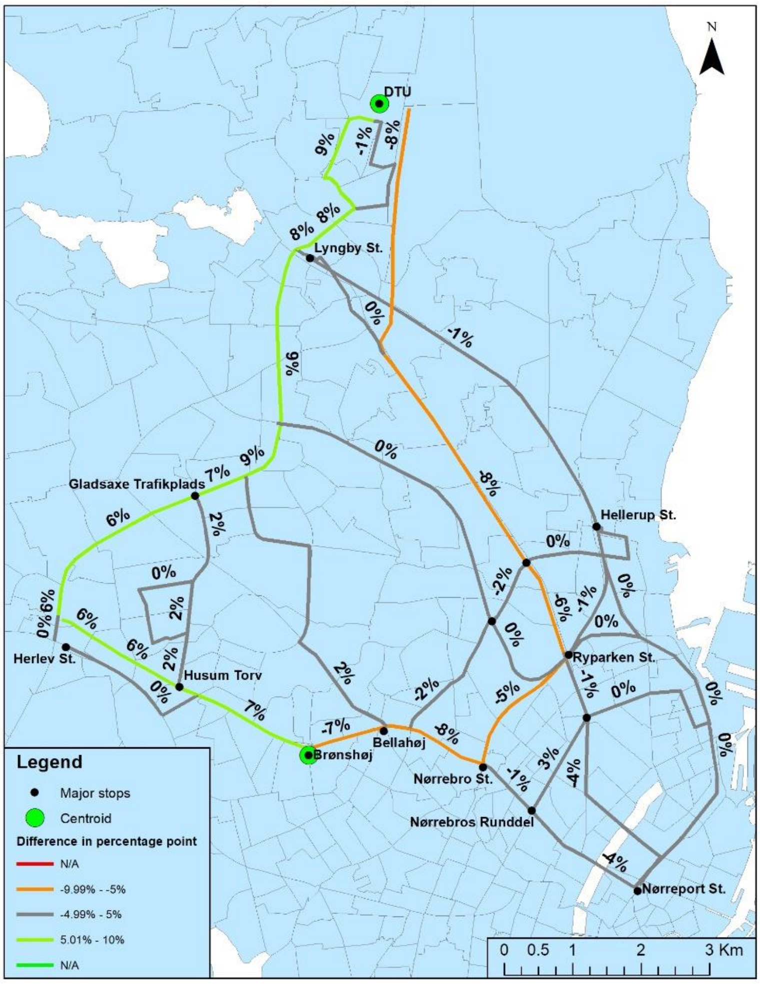

Figure 11 - Flow distribution difference for frequency aggregated network compared to existing network
Positive indicate more flow in frequency aggregated scenario

Discussion

The results in the previous section show that it is possible to generate alternatives and to find reasonable flows from origin to destination. But it also shows that the model is very sensitive to the setting of the threshold and how the not 100% certain paths are included. There should be made a big effort to come up with a more dynamic model where passengers are not assumed to have taken the final choice of their route at the origin, but they should also be able to make a new decision along the route if they do not catch the first candidate of a line. There is a clear problem in how to generate realistic alternatives and this work should be made carefully and validated through SP- or RP-studies to find out how passengers do take decisions on their route choice. But there are also many possibilities to extend the current model and this is discussed in the remainder of this discussion.

Extensions to the model

There are several natural extensions to the presented model. The assumption of the work for this paper has been that all services are regular and that in-vehicle congestion does not influence passengers' route choice. But this has been shown in other papers, see i.e. Nielsen (2000) or Li (2015), that passengers are affected by delays and in vehicle congestion. It is well known that buses in for example Copenhagen are not always running regularly, and this can influence passengers' route choice.

The inclusion of delays in the developed framework, where all paths must with 100% certainty arrive within a threshold of 20%, would lead to some paths which include transfers from frequency- to schedule-based services being less likely to be included compared to a model with regular headways. The tails of the distributions would give some probability of arriving much later compared to a model with bounded distributions (i.e. regular headways). If either a gamma or exponential distribution was used for example, assumptions could be made, that paths with only 95% certainty should be within the threshold. It would then have the disadvantage, that it is not ensured that all paths that will constitute 100% probability to catch a line at a given stop would be included. But these paths could potentially be generated in a later search after having established that the 95% path did arrive at the destination. This could however potentially lead to many extra searches in a large network and should be investigated further to see if it is a potential solution.

When including delays in the network this should naturally also affect the schedule-based lines. The investigation of empirical delay distributions would possibly tell that the delay distributions should be different for schedule- and frequency-based services. Introducing delays for schedule-based services would lead to a different transfer pattern between schedule-based services. If a schedule-based service can be delayed the transfer time between two schedule-based services will no longer be deterministic. There will instead be a probability to reach the service and an associated transfer time. If the first service is not reached then the second candidate of the line should be considered just as the case of a transfer from a frequency-based service to a schedule-based service in a regular network. An updated table for the transfer types is shown below and is only changed for schedule-based transfers, because the other types already includes a probabilistic part.

Table 3 - Updated transfer types when including delays

Transfer to/from	*SB line*	*FB line*
SB line	Probability to catch the next service and an associated transfer time	Transfer time from statistical distribution
FB line	Probability to catch the next service and an associated transfer time	Transfer time from statistical distribution

The above-mentioned inclusion of delays should also be complimented with the inclusion of in-vehicle congestion. Trafikstyrelsen (2012) made a made a brief note on the topic and described that some lines in the morning peak are in lack of seats. This part would be quite simple to incorporate in the utility specification; however, the fail-to-board probability would not be as easy to incorporate. Trafikstyrelsen mentioned the fail-to-board probability in the note and states that some passengers in the metro in the peak hours are not able to board the first departing service. To handle this, the waiting time distributions could be skewed or in some way changed to reflect that the average transfer time is higher.

Modelling low frequent lines as frequency-based lines

This paper has only dealt with high frequent frequency-based services. Another scope of a unified model could be to have low frequent schedule-based services incorporated as frequency-based services. In this way, there wouldn't be any need to adjust these lines if there were major changes in for example the intercity train timetable. The main part in this problem is, that the feeder buses to the intercity trains are usually coordinated with the arrival and departures of the trains. If a large change in the intercity timetable is made, then all connecting buses should be adjusted accordingly. However, this process is manual and requires, that the modeler adjusts each connecting bus, which then leads to connections with other buses being missed. So, it would require a large amount of work to recode all the timetables.

The situation of large changes in timetables is not something that happens every year. But if for example DSB wants to evaluate three different timetable proposals for the future, the evaluation would not be correct unless the buses were adjusted to the new timetable. The problem regarding modeling these low frequent services as frequencies is, that it is hard to know whether a bus and a train are coordinated if they are modeled as frequency-based lines. It could also be, that only every second train on a line is coordinated with a specific feeder bus, meaning that if the bus was frequency-based the passenger could have two possible options in the choice set, because the frequency-based line could be reached from both train runs. The idea of modeling low frequency systems as frequency-based services could help to model future scenarios, but it also requires much attention before it can be made consistently.

How to use the developed framework in large scale models

One of the aims for the work behind this paper was to provide recommendations for how to implement frequency-based services into the Danish National Transport Model. There is naturally a large gap from the work made for this paper and to a final implementation. But however, there are already a few things which can possibly be used in the further development of NTM.

From the examples shown in this paper it is evident that the desired departure time is important for which paths are included in the choice set. The best departure time for a given path depends on when the first service departs, and this should be taken into consideration when generating the choice set. One way to find the best departure time would be to brute-force all departure times, but this is computationally not very attractive. So, if there could be ways to analytically find the best desired departure time without testing all possibilities this could potentially create a shorter computation time. The finding of the best desired departure time also relates to the probabilities of catching the first service. If the passengers want to be sure to catch a service, then that should be reflected in the desired departure time. But some passengers might just want to minimize the total expected travel cost (utility) and both types of passengers should be reflected in the parameters used in the model.

But the problem about the best desired departure time is not the only problem to be solved. In a large network, there are several million OD-relations to solve when including launches at different time periods. Even though there is a possibility to run the proposed algorithm for choice set generation, it might be too time consuming when having millions of searches to make. One way of solving this could be to make an initial search in a completely frequency aggregated frequency-based network, where different paths and the sequence of visited stops could be enumerated either using constrained enumeration or by using the proposed algorithm with slack. These stop sequences could then be used in a time expanded network, and for each desired departure time a search for each stop sequence could be made to define what the cost of that path would be for the desired departure time. In that way, the search for paths at a given time would be very fast and provide a choice set with different alternatives which could also include frequency-based services. An example of how a framework for large scale modelling could look like is shown below.

1. Frequency aggregate all stop pairs within a timeperiod
2. Find stop sequences for all OD relations in the timeperiod given some slack
3. For each OD relation AND each stop sequence AND launch find the best path using the stop sequence, by also including probability to reach schedule-based services
4. Store all paths for each launch and check whether some paths are only relevant at some given launch time compared to hidden wait time vs. launch earlier or later
5. For the choice set for each launch distribute flow across choices using a RUM-based model
6. Calculate level of service
7. If including congestion make flow averaging and check for convergence
    a. Else stop
8. If not converged go to step 3 (or step 2 if the flow averaging made large changes)
    a. Else stop

Conclusion

This paper has provided a possible framework for a unified model including both schedule- and frequency-based services. As it was described in the introduction there are currently no consistent way of modelling a joint network. The proposed model tries to model passengers route choice without taking the explicit departure time for frequency-based services into consideration, which is the same assumption as for strictly frequency-based models.

The description of the different transfer types revealed four main transfer types. The transfer types to a frequency-based service gives average transfer times and it is therefore not possible to create as well coordinated transfers as with schedule-based services. The choice set generation was shown to be difficult to solve in a behaviorally realistic way when including frequency-based services. The assumption that passengers only choose paths which are guaranteed to be within a certain threshold describes a risk averse attitude, which is not described in any literature or estimation of any model.

The choice of a MNL model as choice model described the flow across different alternatives, but did not capture correlation and overlap between the alternatives. In the examples shown in this work a path could constitute of several candidate paths which imitates that passengers only consider simple paths using the same lines. However, the idea of using a recursive model to distribute the flow could potentially lead to a more realistic choice model, because it would imitate that passengers can reevaluate their route choice in case the first candidate of a line is not reached.

The results showed that the algorithm could generate multiple paths and that it was possible to identify the unique paths to be included in the choice set. The local slack gives the reasonable effect, that passengers do not consider detours in the start of a trip, which seems behaviorally realistic. The results also showed that a change of some lines from schedule- to frequency-based services led to a different flow, and that the assumption that passengers only consider a route if it is always below the threshold can exclude some reasonable alternatives from the choice set. The example of frequency aggregation provided reasonable results and did not change the overall flow considerably.

The discussion introduced different extensions to the model, whereof the part on delays is especially interesting, as the route choice could change depending on the different delay distributions for each service type and line. Finally, recommendations for large scale modeling were provided including an idea for a possible framework for large scale modeling. This paper does not provide a full model which can work on larger networks, and there should therefore be put a considerable amount of work into creating a more consistent model where it does not change the level of service whether a line is modelled schedule- or frequency-based.

References

Chriqui, C., Robillard, P., 1975. Common Bus Lines. Transp. Sci. 9, 115–121.

DSB, 2016. S-tog Timetable 2017.

Florian, M., 2004. Finding Shortest Time-dependent Paths in Schedule-based Transit Networks: A Label Setting Algorithm, in: Wilson, N., Nuzzolo, A. (Eds.), Schedule-Based Dynamic Transit Modeling: Theory and Applications. Kluwer Academic Publishers, pp. 43–52.

Florian, M., 1999. Deterministic Time Table Transit Assignment, in: First Asian EMME/2 Users Group Meeting. Shanghai, p. 15.

Fosgerau, M., Frejinger, E., Karlstrom, A., 2013. A link based network route choice model with unrestricted choice set. Transp. Res. Part B Methodol. 56, 70–80. doi:10.1016/j.trb.2013.07.012

Friedrich, M., Hofsaess, I., Wekeck, S., 2001. Timetable-Based Transit Assignment Using Branch and Bound Techniques. Transp. Res. Rec. 1752, 100–107.

Gentile, G., Noekel, K., 2016. Modelling Public Transport Passenger Flows in the Era of Intelligent Transport Systems. Springer. doi:10.1007/978-3-319-25082-3

Hoogendoorn-Lanser, S., Bovy, P., Van Nes, R., 2014. Application of Constrained Enumeration Approach to Multimodal Choice Set Generation. Transp. Res. Rec. J. Transp. Res. Board 50–57. doi:10.3141/2014-07

Li, Q., Chen, P., Nie, Y., 2015. Finding optimal hyperpaths in large transit networks with realistic headway distributions. Eur. J. Oper. Res. 240, 98–108. doi:10.1016/j.ejor.2014.06.046

Liu, Y., Bunker, J., Ferreira, L., Ttrv, F., 2010. Transit Users' Route-Choice Modelling in Transit Assignment: A Review. Transp. Rev. 30, 753–769. doi:10.1080/01441641003744261

Mai, T., Fosgerau, M., Frejinger, E., 2015. A nested recursive logit model for route choice analysis. Transp. Res. Part B 75, 100–112. doi:10.1016/j.trb.2015.03.015

Metro, 2017. Timetable for the Metro in Copenhagen [WWW Document]. URL http://www.m.dk/#!/om+metroen/rejseinformation/koereplan (accessed 1.16.17).

Movia, 2017. Samlet buskøreplan for Movia. København.

Nielsen, O.A., 2000. A stochastic transit assignment model considering differences in passengers utility functions. Transp. Res. Part B 34, 377–402.

Nökel, K., Wekeck, S., 2007. Choice Models in Frequency-based Transit Assignment.

Rasmussen, T.K., Nielsen, O.A., Watling, D.P., Prato, C.G., 2017. The restricted stochastic user equilibrium with threshold model: Large-scale application and parameter testing. Eur. J. Transp. Infrastruct. Res. EJTIR 17, 1–24.

Rejseplanen, 2017. Rejseplanen [WWW Document]. URL Rejseplanen.dk

Schmöcker, J.-D., Shimamoto, H., Kurauchi, F., 2013. Generation and calibration of transit hyperpaths. Transp. Res. Part C 36, 406–418. doi:10.1016/j.trc.2013.06.014

Trafikstyrelsen, 2012. Kapacitet og trængsel i den kollektive trafik i hovedstadsområdet.