CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Optimization of Intermodal Freight Train Service Schedules on Train Canals

Michael Berliner Pedersen
Teodor Gabriel Crainic

December 2007

CIRRELT-2007-51

Bureaux de Montréal :
Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince, local 2642
Québec (Québec)
Canada G1K 7P4
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Université Laval · UQÀM · HEC Montréal · École Polytechnique Montréal · Université de Montréal

Optimization of Intermodal Freight Train Service Schedules on Train Canals

Michael Berliner Pedersen[1] , Teodor Gabriel Crainic[2,*]

[1] Centre for Traffic and Transport, Technical University of Denmark, Anker Engelundsvej 1, Building 101 A, 2800 kgs., Lyngby, Denmark

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and NSERC Industrial Research Chair in Logistics Management, Université du Québec à Montréal, C.P. 8888, succursale Centre-ville, Montréal, Canada H3C 3P8

Abstract. This paper presents a mixed integer mathematical programming model to optimize schedules for intermodal trains networks. The model is based on a time-space representation of a service network but is essentially a capacitated network design model. The model assumes that the intermodal train services are routes on a network divided in train canals (aka. as train paths) as is becoming general practice in the European Union rail sector. The train canals represent time slots in which intermodal train services may be offered. Transit time and operating for intermodal rail services is an important service parameter in the competition with long-haul road transportation. Operating costs are a sum of the service operating cost and the terminal operation cost. Terminal operations play an important role in intermodal transportation. To capture this the model includes unloading, loading, transfer operation costs, and inventory costs and add capacity constraints on the number of handling operations and inventory. To model transit time the model introduces a value of time cost that multiplied with the transit time gives a measure of the value of time for commodities. This eventually leaves the delivery time of demand as an output of the model. To capture the trade-off between operational cost and the value of time cost the two are added and minimized. The model is applied to a generated network instance and nine different scenarios are solved using Xpress-MP's MIP-solver. Although the use of this 'brute force' approach with resulting mediocre solutions post-analysis of the results show that the model works in accordance to what is expected.

Keywords. Scheduled service network design, intermodal transportation, shuttle networks.

Acknowledgements. This work was supported by the Virtual Centre for Logistics and Freight Transport (CLG) in Denmark, the Danish Technical Research Council, and Atkins Denmark. The Natural Sciences and Engineering Research Council of Canada (NSERC) offered additional support through its Discovery Grant program.

---

* Corresponding author: Teodor-Gabriel.Crainic@cirrelt.ca

Dépôt légal – Bibliothèque nationale du Québec,
Bibliothèque nationale du Canada, 2007

1. Introduction – Intermodality and railways

It has been estimated that the freight traffic in the European Union will increase with 50% from 1998 to 2010 [White paper 2001]. Many parts of the European transportation networks are already operating close to their capacity level. It is estimated that around 10% of the European highway network is affected on a daily level by congestion. Several parts of the rail network have been classified as bottlenecks. The European Union has in its white paper stated how it intends to promote intermodal transportation with the rail sector as the predominant strategic transportation mode in European freight transportation, and has launched a series of infrastructure projects to improve network conditions. From 1985 to 1994 the UIRR (the organisation for European rail-road intermodal operators) experienced a growth from 1.5 million TEU to 3.5 million TEU and the ICF (European in-land intercontinental container transportation) experienced a growth from 0.9 million TEU to 1.1 million TEU ([UIRR 1995]). Thus intermodal transportation is growing, but several issues need to be considered in order to make intermodal transportation able to capture the eventual growth in freight transportation.

1.1. Influencing the modal shift from road to rail

The main focus of the European Union as stated in its white paper on transportation is to shift the modal split towards rail and sea transportation and reduce the share of freight of long-haul road transportation. The reason for this is to reduce road congestion and promote environmentally friendlier modes of transportation as rail and sea transportation. The statistics presented in this and the following sub-sections are taken from [White paper 2001] unless otherwise stated.

The CO_2 emission of the transportation sector accounts for 28% of the total European Union. Thus significant overall CO_2 emission reductions can be achieved by be reducing the emission from the transportation industry. The externalities for the three main modes of freight transportation (excluding congestion costs); rail, waterborne, and road transportation measured in Euros are estimated to be 19€ per 1000 tonne-km, 17€ per 1000 tonne-km, and 88€ per 1000 tonne-km respectively. These externalities include emissions of CO_2 among other effects such as emission of other aerosols and noise, urban nuisance, and accidents. From these figures it is clear that road transportation has a significantly higher impact on the environment than rail and waterborne transportation. It is from figures like these that the idea behind attempting to shift freight from road to rail/sea originates.

Another reason for shifting the modal split in favour of rail and sea is that removing freight from the roads onto rail and water will lower the congestion on the highway system. It is estimated that the increase in road congestion will be the cause of loss in productivity amending to 1% of the Unions GDP by 2010. Hence congestion is not just a nuisance for users of the road network, but influences the competitiveness of the whole European region.

It is believed that road transportation does not pay for all of the external costs it inflicts on the environment and on congestion ([White paper 1998]). Therefore it indirectly has a competitive advantage on that account compared to rail and waterborne transportation. For this reason many European countries are introducing road taxes (tolls). Examples are

Germany, Austria and Switzerland thereby favouring other modes of transportation. Although some initiative has been taken to make fairer competition between the different modes of transportation there are still other issues to take care of before rail can compete against long haul road transportation

1.2. Status of the European rail system

Although the European Union's transport policy encourages the use of rail it only represents 8% of the total freight volume transported in Europe. Short-sea shipping has a major cut though of 41%, while road transportation has a share of 44% and inland waterways 4%. Rail freight transportation's 8% share may be put into perspective to its share in North America of 40%.

There are a number of reasons why the modal split differs between the two continents. One reason may be the geographic and demographic characteristic of the continent. The North American continent is much larger and has its population centres located at its extremities (the coasts), while Europe's are located at the centre stretching from the British midlands through the Ruhr area to the Po Valley (so-called blue banana). This means transportation distances are larger in North America than in Europe which gives rail a competitive advantage there while trucks are more efficient on the shorter transportation distances in Europe. Furthermore North America serves as a land bridge between the west coast and the east coast for intercontinental transportation from Asia to North America and Europe.

Bottlenecks and congestion also are an issue in the European rail network. 20% of the 16.000 km of rail tracks is classified as bottlenecks. Furthermore passenger rail transportation plays a dominant role in Europe which has a preventive effect on having any extensive use of freight trains. Passenger trains have traditionally been given priority over freight trains, and hence flexibility and reliability of rail freight transportation has decreased accordingly. In contradiction to this the rail networks in North America are almost fully dedicated to freight trains.

Finally the development of the Union's member states' rail systems has traditionally been marked by nationalistic protectionist tendencies. In order to prevent access to the national rail networks to foreign rail carriers very diverse network configurations have been used. This has resulted in a collective European rail network with five different electrification systems, two different gauge systems (the rail networks on the Iberian Peninsula and in Finland differing) different labour regulations, different traffic regulations, and a continental network composed of inadequately connected national networks. The improvement of interoperability between the member states' networks now presents a major challenge in achieving an integrated continental rail network.

The liberalisation of the rail industry in Europe has lead to a separation of the traditional national rail companies into infrastructure owners and operators. Rail authorities manage infrastructure and network capacity and rail operators that operate trains according to the available capacity acquired from the rail authorities. To improve the possibility of cross-border operations, collaboration between the rail authorities of the member states is being established to create an interconnected trans-European rail network for freight trains as stated in [White paper 1996], the so-called freight freeways. Projects to eliminate bottlenecks and improve interoperability have been proposed also to improve network conditions. It has already become easier to run train operations across several countries and operators do not have to procure capacity from several capacity owners due to the One-stop-shop concept adapted to the freight freeways. With the tendency continuing, the future result in having large service networks across the European continent operated by single operators or by alliances of operators, similar to those seen in the airline industry. The resulting service networks will be complex to plan and operate while having to compete or cooperate with the road transportation industry.

Significant effort has been put into research on intermodal transportation and rail transportation. We will throughout the paper include references to research relevant to the research presented here. The review is not supposed to be a complete survey of literature on intermodal transportation. Such a survey is available in [Macharis et al. 2004]. The review will focus on some of the papers included in that survey and additional ones that are relevant to the research presented in this paper.

The contribution of this paper is to present a mathematical programming model that can plan the service networks of intermodal train operators. The model introduces terminal operations and a cost of transit time for freight in order to base the service selection on these impacting factors. The remainder of this paper is divided as such. Section 2 describes intermodal train operations and with pointers to relevant previous work related to train operations. In section 3 we describe the modelling assumptions while section 4 presents the mathematical programming model based on those assumptions. Section 5 presents a generated network instance with different data scenarios and computational results obtained by using the MIP-solver from the Xpress-MP optimizing package. The paper is concluded in section 7 with pointers to future avenues of research.

2. Intermodal transportation by train

The extension of rail infrastructure is limited compared to road infrastructure. Therefore only few pairs of customers can be served solely using train transportation and it is generally impossible to provide door-to-door deliveries using rail. Drayage moves by trucks are required to move loads from their origin to a rail terminal where it is transhipped to another rail terminal before being delivered at their destination by truck. This type of intermodal transportation, where road and train transportation are combined into a transportation chain is seen as the way forward to increase the modal split for railways in Europe ([White paper 2001]). To enable a seamless transfer between truck and train, freight is containerized as opposed to conventional trains where freight is loaded and unloaded in commodity specific wagons.

Apart from the environmental and congestion preventive effects discussed in the previous section, consolidating freight on trains achieves economies of scale and thus reduces transportation costs. The trade-off however is that the flexibility of moving containers independently by truck is lost. This loss of flexibility combined with the experienced low reliability and long transit times of rail transportation seem to be major reasons why customers choose road transportation over rail. The introduction of intermodal shuttle trains in Europe is an initiative to try and improve the reliability and transit time by rail by operating freight trains like passenger trains on a tight schedule and minimizing the number of terminal operations performed.

It is arguable whether political interest and incentive regulations such as road taxes will be enough to achieve the European Union's ambition of increasing rail's share of the modal split from the actual 8% to 15% by 2010. It must be expected that intermodal operators need to decrease operational cost and increase service levels to take on the competition against road haulage.

2.1. Analysing the intermodal trip chain

To understand how to improve operations of intermodal transportation we first analyze the events of an intermodal trip chain and compare it to long-haul trucking. Figure 1 shows an example of four containers with different customer origins and destinations that use the same intermodal train services. The four containers are transported from the four customers to rail terminal A by drayage moves. There they are transferred to a rail service going to terminal B. At terminal B they are transferred to another rail service going to terminal C. Finally at terminal C they are transferred to the trucks for the final drayage move to their end destinations. The total transit time for each of the intermodal trip chains above is a sum of:

- The transportation time of the initial drayage move to terminal A.
- Unloading time from the trucks and storing the containers in terminal A's storage place.
- Connection delay until the train arrives and is ready to be loaded.

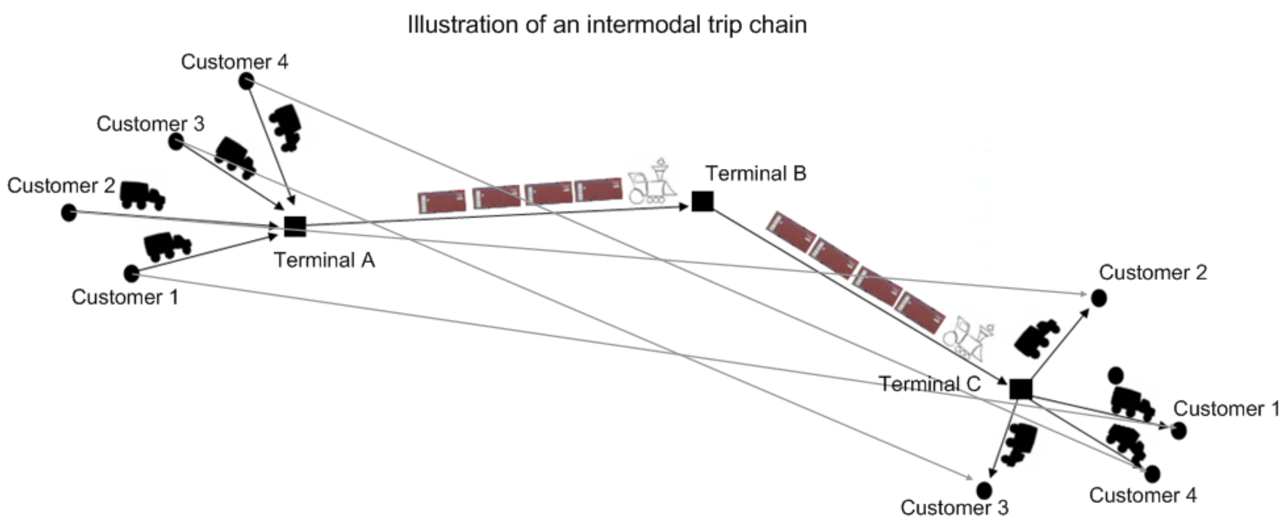

Figure 1

- Time to pluck the containers from the storage place and loading onto the train.
- Transportation time from terminal A to terminal B.
- Time to unload the containers and storing them at terminal B's storage place.
- Connection delay until the second train arrives and is ready to be loaded
- Time to pluck the containers from the storage place and loading onto the second train.
- Transportation time from terminal B to terminal C.
- Time to unload the containers and store them in terminal C's storage place.
- Connection delay for the trucks for the final drayage move to arrive.
- Time to pluck and load the containers onto the trucks.
- Transportation time of the final drayage move to the customers.

The sum of the time of the operations above has to be competitive with the direct long-haul truck moves driving directly between the points of origin and destination. The number of terminal operations and their cost are also considerable compared to having no terminal operations between the origin point and the destination point using road-haulage. In order to make intermodal trip chains competitive both the transit time and the operational cost need to be considered.

There are several possible intermodal services to offer (see [Ballis et al. 2004]), but the focus of this paper is on the intermodal shuttle trains. These are common operating practice in road-rail intermodal transportation. The characteristic of running intermodal shuttle trains is to have a fixed train make-up. That means that the train always runs with the same number of flat-bed wagons (onto which containers are loaded). The advantage of running a train with a fixed make-up means the composition of the train does not need to be changed at terminals thereby reducing handling cost and reducing the turn-around time. That enables a more efficient use of vehicle equipment. Contributions from [Turnquist et al. 1982] and [He et al 2003] propose operations research based approaches for rail yard operations. Both contributions illustrate the complexity of rail yard operations and the resulting cost and time consumption proving the reason behind running intermodal trains with a fixed train make-up.

The more efficient use of vehicle equipment in intermodal services is however subject to running trains at their capacity. There is a financial risk of providing fixed make-up trains with resulting fixed capacity if they end up running semi-empty. For intermodal operators the break-even point lies above a 90% utilization of capacity. This also poses a problem on determining the adequate train make-up. A lot of effort is put into making sure trains are operated at full capacity. Part of the process lies in marketing and sales by making sure freight demand in the service network is well balanced. In [Yan et al. 1995] an operations research based pricing method is proposed to determine optimal pricing of capacity. We will not include pricing issues in our model though, but will include a limited number of possible train make-up options.

Terminal operations are necessary to tranship containers between the different transportation modes and services. From the explanation of the events that can occur in an intermodal trip and as mentioned in [Ferreria et al. 1994] it can be deduced that terminal operations play an important role. The following events and operations may take place in a terminal:

- Arrival by train or truck, inspection
- Unloading and storage in yard
- Transhipment to other vehicle
- Stay on train for further continuation
- Loading onto train from storage

[Bostel et al. 1998] presents a model and solution method to solve the transhipment problem between two trains. The model aims to minimize the container moves between trains. [Rizzoli et al. 2002] present a simulation tool for the entire terminal process including storage operations. From the time the container arrives at the terminal by truck or train until it departs again the container does not cover any physical distance. That is why there is focus on developing technology to speed up and reduce handling cost of terminal ([Trip et al. 2002]). The unloading, transferring, and loading of containers onto trains can be done by gantry crane or by mobile crane depending on the available infrastructure at the terminal. There are only a limited number of tracks at a terminal. Thus only a limited number of trains may be present simultaneously at the terminal. The handling machinery can only perform a certain number of operations meaning that only a limited number of handling operations may occur in a given time period to avoid congestion and resulting delays. The resources available to perform terminal operations are thus limited and should to be considered when designing intermodal train services.

2.2. Characteristics of freight using intermodal train networks

The appropriate intermodal train services to offer depend largely on the demand of commodities and on the competition with road haulage. It is assumed that bulk commodities, perishable commodities, and hazardous materials are not transported by intermodal train. Due to the low value per weight unit, bulk commodities such as iron ore, coal or lumber are transported by conventional train if not by ship and are generally not accessible to intermodal transportation or long-haul road transportation. Highly perishable goods such as some dairy products are generally not transported over longer distances because of the transit time. This segment of commodities requires the flexibility and direct transportation offered by road haulage over shorter distances. Hazardous materials such as chemicals or nuclear waste have to obey certain rules and regulations and are therefore rarely transported with other commodity segments.

It is generally perceived that intermodal transportation currently is only applicable in Europe for distances over 400 km. For distances shorter than 400 km the time required to perform an intermodal trip chain makes it non-competitive compared to road transportation. For longer distance however, the transportation cost becomes a more important factor, and, assuming the train leg(s) are reliable and run at higher speeds than trucks, the additional transit time endured by terminal operations is less significant. [Nierat 1997] presents a model based on spatial theory to determine the market area of intermodal transportation. The results support the hypothesis that intermodal transportation is limited to relatively long transportation distances and to customers within reasonable distance of intermodal rail terminals and also shows that it is only for a limited region around a terminal where intermodal transportation is a cost and time efficient alternative. [Trip et al. 2002] point to the fact that the competitiveness of intermodal transportation could be improved and thus make it more attractive for shorter distances by improving terminal operations and by extending the reach of intermodal transportation networks.

What is typically seen is that intermodal train operators negotiate long term contracts with large customers such as freight forwarders or big industrial clients. With these contracts customers guarantee a certain amount of loads whilst obtaining a lower transportation price. The contracts provide stable demand for the operators and lower their financial risk. In turn the remaining capacity can be sold on the spot market. To reduce the uncertainty further Swedish operator Green Cargo for example requires customers to book in advance in order for them to be able to plan operations efficiently.

In order to compete with road transportation and attract customers we assume that intermodal operators set a transit time between origin/destination points as a strategic goal. Promising transit time sends a signal of reliability and attracts customers but also requires service standards to match the promise. A high service level achieved by running frequent trains can reduce the total transit time; however the operational cost will increase accordingly. Unless the higher service levels attract "new" demand to the system there is a significant risk of running at low utilization of capacity and thus loose the competitive edge on the operational cost. The art is to design intermodal train services that offer competitive transit times while maintaining low operational cost.

2.3. Designing intermodal network subject to infrastructure divided in train canals

Several contributions investigate modal choice and intermodal network design in a region. Such analysis can be found in [Bookbinder et al. 1998] where intermodal routing options between Canada and Mexico under NAFTA are investigated. The results of the investigation give an indication of the modal choices between pairs of 5 Canadian and 3 Mexican cities using several American cities as transhipment points. Similar analysis can be found using the STAN software package which has been applied to the São Fransisco river corridor in Brazil ([Crainic et. al. 1990]). Whether operations research methods are used or not an initial strategic analysis of a region provides an operator with a decision support which can be used to determine its network coverage area.

Given a strategic network of areas and customers to serve, the problem becomes one of choosing how often to run services. Previously, a widely adopted policy for running conventional freight trains was a "go-when-full" policy. This meant that freight trains were not scheduled and moved from their origin to their destination terminals when capacity on the rail network was available. This policy is still adopted in North-America whereas in Europe the policy is inefficient because of the large amount of passenger trains taking up the rail network capacity. The higher priority of passenger trains result in freight trains being side tracked leading to excessive transit times. For bulk commodities where low transportation costs are very important and transit time almost negligible this is not a problem. However, for time sensitive freight the long transit times are unacceptable, and it is one of the reasons for the low share of the modal split in favour of rail transportation. Contributions from [Crainic et al. 1984], [Marin et al. 1996], and [Keaton 1989] all present service network design problems with train frequencies as outputs. The competition with passenger trains for capacity on the infrastructure in the old EU15 (EU prior to expansion in 2004) means capacity is not readily available when needed. Therefore these frequency-based approaches are not appropriate for designing a service network for an European intermodal train operator.

The situation in Europe where rail business is separated into rail authorities and operators means that an intermodal train operator is not the proprietor of the infrastructure and not the sole operator using it. To overcome the issues of having several operators rail authorities have adopted a planning procedure dividing the infrastructure into so-called train paths ([Link 2005a]). The term train path is used in several operations research contributions as an actual movement of a train. However, train paths here only represent a routing possibility, which is why we will refer to them as train canals instead. Train canals are time dependent paths on the rail network. They can be compared to a time-slot or time-window within which a train must operate on the rail infrastructure. This means that there is a departure time and an arrival time associated to each of the terminals visited along the path. The division of the infrastructure into predetermined train canals prevents conflicts of trains on the network and leaves it up to the operators to acquire the train canals they need to assume their operations. Passenger trains still have priority on acquiring train canals, and passenger train operators are often involved in the process of determining train canals. However, the European national rail authorities have started to cooperate on constructing a dedicated transcontinental network of train canals for freight trains ([White paper 1996]). Although a full transcontinental network of train canals for freight is not yet implemented the main corridors have adopted the concept ([Link 2005b]) and it can be assumed that this will soon be the case for most of the European rail network. In this paper we assume that all train routing and scheduling is done according to predetermined train canals.

2.4. Service network design on train canals for intermodal trains

Several contributions can be found on train routing and scheduling and on applied service network design for train operations. A survey is presented in [Cordeau et al 1998]. [Huntley et al. 1995] and [Gorman 1998a] present service network design models with schedules for CSX transportation and Santa Fe Railways respectively. [Yano et al. 2001] present a dynamic modelling approach to schedule departures of freight and trains to and from a single terminal [Newman et al 2000] present a train routing model which includes schedules. However, freight demand is modelled to originate and is destined to rail terminals, thus drayage moves are not considered. [Nozick et al 1997] present a linear MIP-model for planning intermodal freight routing. The configuration of the train schedules is given

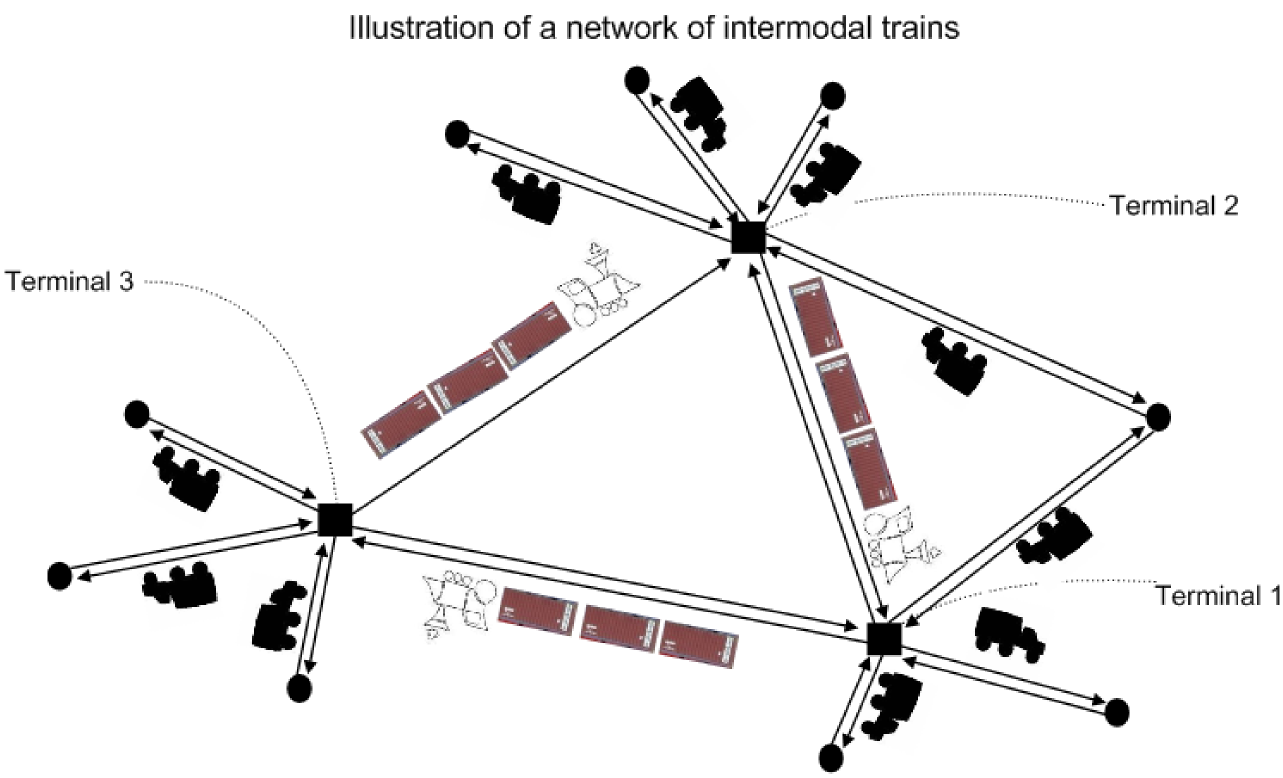

Figure 2

though. [Gorman 1998b] presents a linear MIP-model for train scheduling with limited terminal operations and [Hagani 1989] proposes a linear MIP-model for scheduling trains that is similar to the one we will propose here. The MIP-models determine the optimal scheduling on a space-time representation of a network for two types of trains including the train make-ups and empty wagon repositioning problem. The models assume that trains can run within every time-period and therefore do not account for the fact that rail network capacity may be occupied by passenger trains.

The model we present includes the limitation that trains can only be routed on available train canals. Given a set of available train canals the decision on when to provide intermodal train services indirectly becomes one of selecting appropriate train canals. The notion of being constrained by routing possibilities has made us believe that schedule synchronisation is an important issue to provide fast and reliable services on a network composed of train canals. Higher frequencies of trains in a network will reduce the overall expected transit time of freight but also require higher operational costs. By synchronizing arrivals and departures of trains in terminals it is possible to transfer containers directly from train to train. This will remove connection delays and thereby reduce the transit time without having to increase frequencies and incur higher operational cost.

It is relatively easy to synchronise the two train services shown in figure 1 such that connection delays are removed. All that is required is for the two trains to be present at terminal B simultaneously and thus achieve transfer synchronisation. Transfer synchronization becomes much more complex when a network structure of intermodal trains is considered. Figure 2 shows a simple network of intermodal trains. Synchronizing train service connections at one terminal may lead to deteriorating connections at other terminals. Capturing this network wide interaction is difficult if not impossible to do by manual planning. A quantitative method is needed to design a service network that proposes good synchronizations to reduce to transit time while keeping operational cost low at a global network level. However, synchronization becomes a complicated issue when only a limited number of routing possibilities are available meaning special attention has to be paid to train routing and terminal operations in order to minimize transit times for freight.

2.5. Problem description

The general problem is to determine an intermodal train service network for an intermodal train operator. The following items are included in the model:

- Train scheduling on train canals, including train make-up alternatives.
- Terminal operations and train synchronisation to minimize transit times.

We use a linear mixed integer mathematical programming model to model the problem. The objective of the model is to capture the trade-off between operational costs and the value of time cost incurred the freight transit times and minimize the sum of the two. The model needs to capture general operational constraints such as train routing possibilities, train canal availability, train mode capacities and terminal operation capacities. The output of the model is a train routing plan based on a selection of train canals and a freight routing plan. The costs incurred by operations and value of time determine what the optimal system configuration with respect to train routing, train synchronization, and freight routing is.

3. A Service network design model for intermodal trains

This section presents the formulation of the problem we intend to solve. The section gives a detailed description of the modelling assumptions made. An overview of the assumptions and the modelling can be seen in figure 10.

3.1. Representing the underlying network structure of train canals and train routing

One of the novelties in the modelling presented here is the use of the train canals to schedule trains. As described train canals are predefined time-dependent paths on the physical infrastructure. Assuming correctly that a train must operate within the boundaries of the available train canals, the underlying physical network and its capacity needn't be considered. We simply construct a network where arcs represent train canals. Only one train can run on a train canal so the routing and scheduling of intermodal trains can be interpreted as selection of train canal arcs.

We assume that the intermodal train operator operates on the same rail infrastructure as other rail carriers. This implies that not all train canals proposed by the rail authorities are available and that there is competition for the acquisition of them. Train canals are presently at a fixed price on a "first come, first serve" basis with priority to passenger trains. We assume that a given number of train canals are available e.g. through prior acquisition or non-acquired train canals. The acquisition process of train canals is therefore not included in the model. Design of a service network on train canals subject to the acquisition of train canals e.g. through a bidding mechanism is an interesting topic but is left for future research.

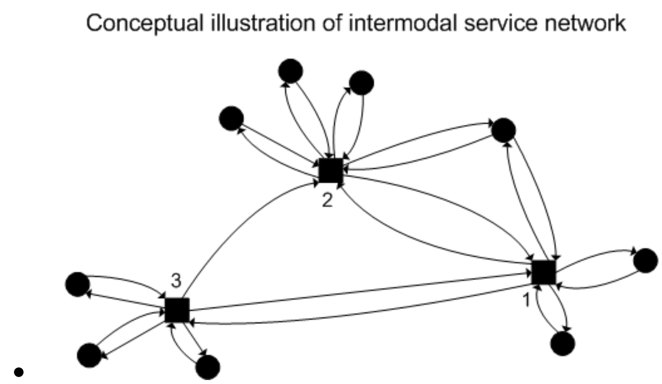

Figure 3

Figure 3 shows a conceptual view of the network in figure 2. The arcs between the terminals 1-3 (square black nodes) represent the possible train routing connections between the terminals. The arcs between the terminal nodes and the customers (round black nodes) represent the possible drayage moves between terminals and customers. To capture the time dimension of the train canals we use a space-time network representation. Figure 4 shows a space-time representation of the terminals alone from figure 3.

Each of the nodes in figure 4 represents one of the three rail terminals at a certain time interval (here days). The time periods needn't be time-wise adjacent to each other e.g. the time period could represent the opening hours of the terminal or the period of time where the majority of operations are performed. The network has a time horizon (of seven days in figure 4) and is periodic meaning it repeats itself when the time horizon is reached. It is obvious that over a long time horizon (of several years) the train schedules will change. However, passenger trains follow a repeating schedule and given their priority on the rail network intermodal trains seem to fall into that pattern too.

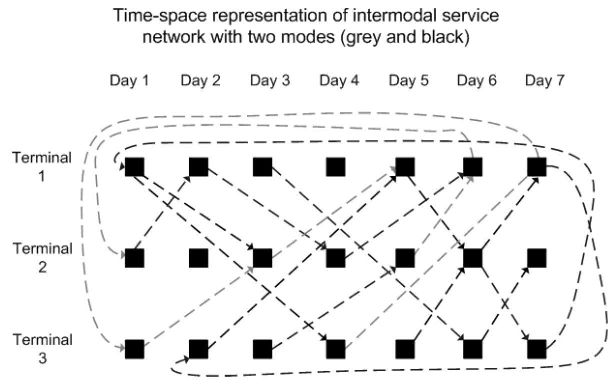

Figure 4

The time periods associated with each of the terminal nodes indicate what time periods the departure and arrival times of the train canal lie within. We simplify the representation of the arrival and departure times of by assuming that a train canal departs at the end of a node's time period and arrives at the beginning of a node's time period.

The characteristics of the train canals such as acceleration and cruising speed may set an upper bound on the number of flat-beds wagons that make up the train. To model the variability of train make-up we enumerate a number of make-up options which we shall call modes. Each mode represents a train make-up composition and has a capacity limit associated to it. If several train make-ups are possible on a train canal it can be represented by several arcs each associated to a mode. To ensure the train canal is only used once a mutual exclusion constraint may be added. We assume however that each train canal arc in the network is associated to only one mode (in figure 4 a grey and a black mode each representing different train make-ups). The capacity of the train canals arcs is inherited from the modes they are associated to.

Trains can stay at a terminal for more than one time period. This is modelled by adding train transfer arcs between succeeding terminal nodes. E.g. train transfer arcs are added for each mode between the 'terminal 1, day 1' node and the 'terminal 1, day 2' node in figure 4 etc. The arcs have been left out of figure 4 for reasons of clarity.

A fixed cost is associated to each train canal which represents the cost of train canal acquisition (if applicable) crew operating cost, maintenance and depreciation cost or leasing cost of locomotives and wagons, and fuel costs. The cost of running the train varies somewhat depending on the number of containers on the train. However, we assume that the fixed cost of is the predominant cost and thus neglect the comparatively small variable cost. The fixed cost of the transfer train arcs represents the cost of keeping the train idle at a terminal, i.e. leasing cost or deprecation and maintenance cost.

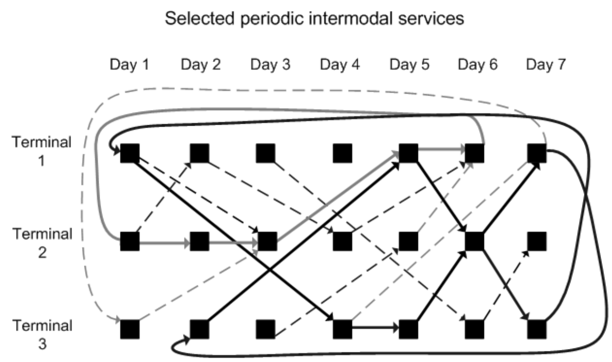

Figure 5

Figure 5 shows the train canal network from figure 3 with a selection of train canals for each mode (full lines). Notice how the nodes are balanced with respect to the number of selected train canals of each mode that enter and leave the node resulting in service cycles. The service cycles are an attribute derived from the repetitive service pattern. Since the schedule is repeated after reaching the time horizon the services must form cycles that repeat themselves. To ensure that service cycles are obeyed a constraint ensuring that trains entering a terminal in a time period must also leave it again at the end of the time period by either leaving for another terminal or to go to the terminal's next time period is added.

3.2. Representing freight demand

Because of the acquisition process and the complex operational planning required to run intermodal trains designing a service network is considered a tactical exercise performed several months in advance of actual operations. It can not be expected to have detailed knowledge of available demand at that time and it is subject to some level of uncertainty. The service network is also repetitive (e.g. weekly schedule), which means demand levels vary between repetitions. Finally, we also need to consider that demand levels are correlated with the service offered i.e. the higher the frequency of trains, the shorter expected transit times, and thus the more attracted demand.

We assume that the customers of the operator are mainly forwarding agents or large industrial customers. These customers negotiate prices with the operator to achieve lower prices while guarantying a certain amount of loads. This means we assume that the operator has some quantitative estimate of the demand potential and is able to make reasonable forecasts for customers' individual demand. We furthermore assume that the discrepancies in freight demand between the periodic repetitions are handled by the sales division of the operator. Given reasonable demand forecast, the potential to even out imbalances in demand and computational complexity of handling stochastic demand, we thus assume to have deterministic demand.

An important notion to respect when determining demand data is the correlation between the level of service offered and the level of demand. We assume that customers base their transportation choice on price and total transit time. This means that if the carrier provides frequent services resulting in shorter transit times while maintaining a constant price more customers will be attracted to use the services of the operator. Hence the level of demand is subject to the chosen train routings and schedules. However, it is reasonable to assume that the demand forecasts made are based on the existing service network and the attracted demand is minimal on a short term basis. Thus we can assume that the forecasted demand levels correspond with the optimized service network. Long term effects on demand could be simulated by establishing different forecast scenarios if the correlation between demand and service were specified.

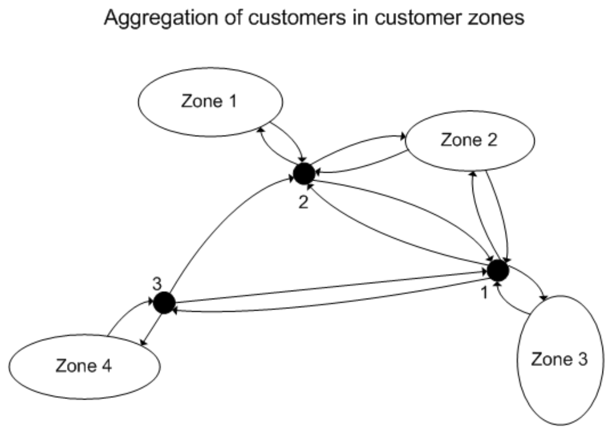

Figure 6

Although the operator is assumed to have some quantitative knowledge of the demand levels it might be on a more aggregate level. The uncertainty of data and the eventual model size makes it unreasonable to consider demand on a customer level. We believe that a realistic approach for real applications is to first aggregate customers into clusters and thus have demand based on customer zones. The demand is then forecasted as general demand potentials between each of the zones. Figure 6 represents the customers from figure 3 grouped in four customer zones. Aggregating customers into customer zones first of all reduces the size of the eventual model and also requires less detailed demand data. If the operator e.g. runs continental trains between rail hubs, customer zone data need only be forecasted on regional level.

The demand is represented as multiple commodities each with an origin customer zone and a destination customer zone. Furthermore a time of availability is associated to each commodity but we assume that commodities can be delivered to the destination customer zone at any time and thus no delivery time is associated. In the time-space representation of the network the customer zones are represented by a set of nodes that each represents a point in time (see figure 8). In the time-space representation we have multi-commodity flow problem with the peculiarity that freight has an origin node corresponding to the origin customer zone and time of availability and a set of potential destination nodes representing the destination customer zone as opposed to a single destination node.

Having no delivery time associated to the commodities means that the transit times can become very long if only operational cost are considered. Obviously that is not representative of the ambition of providing fast transit times for freight on intermodal trains. To capture the trade-off between transit times and operational cost we introduce a value of time representing the cost of the perceived transit time from a customer's point of view. The value of time increases with the total transit time which is further accentuated if freight is delivered after a promised or expected delivery time. The value of time as a function of the delivery time could look like the exponential curve shown in figure 7. As seen in the figure the value of time rises steadily until the point of promised delivery time after which it rises faster.

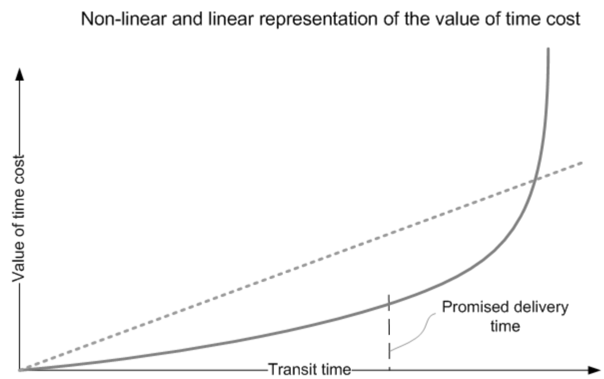

Figure 7

Freight has different values of time depending on the actual physical commodity and on the individual customers. Working on an aggregate level we assume that all commodities have the same value of time and leave further investigation on value of time for different freight commodities and heterogeneous customers to future research. Furthermore, representing the value of time by a non-linear function will add significant complexity to the model. Considering the aggregation and assumption of homogeneity of commodities it seems reasonable to neglect the non-linearity and assume a linear function for the value of time as shown in figure 7. Although the linear approximation does not give a correct representation of the cost of transit times it captures the essential difference of short transit times having low cost and low transit times having high cost. The approximation is somewhat correct for transit times that don't exceed the promised delivery time by a big margin. In figure 7 the linear approximation is an upper bound to the transit time and only diverges significantly from the "real" value of time cost after the intersection point.

It is possible to calculate a total measure of the performance of the network by adding the sum of the transit time cost for each commodity, the operating costs of running trains, and terminal operation costs. The higher the value of time the more important it will be in the measure. By adjusting the value of time the importance of fast transit versus operational cost can be controlled and thus determine the delivery time of a commodity at the customer zon***e.***

3.3. Representing drayage moves

When grouping customers into customer zones and obtaining aggregated commodities it is important to consider the distances and costs of possible connections from the customer zones to the intermodal rail terminals. As can be seen in figure 6 customer zones can be connected to one or more terminals. These connections represent possible drayage moves by truck from the customer zones to the rail terminals. It is assumed that all customers grouped in a customer zone can reach the rail terminals they are connected to.

Figure 8 shows an expansion of the network from figure 5 where customer zone nodes have been added and non-used train canals have been omitted. The (dotted line) arcs between the customer zone nodes and the terminal nodes represent the drayage moves. The drayage arcs follow the same definition as the train canal arcs in that they are assumed to arrive at the beginning of a terminal node's time period and leave at the end of it.

To perform a drayage move it is assumed that one truck is needed for each container. We make an approximation by assuming that the transportation distances and cost are equal for all customers in a customer zone and that the transportation cost is commodity indifferent. Each drayage arc can therefore be associated with a unit transportation cost**.**

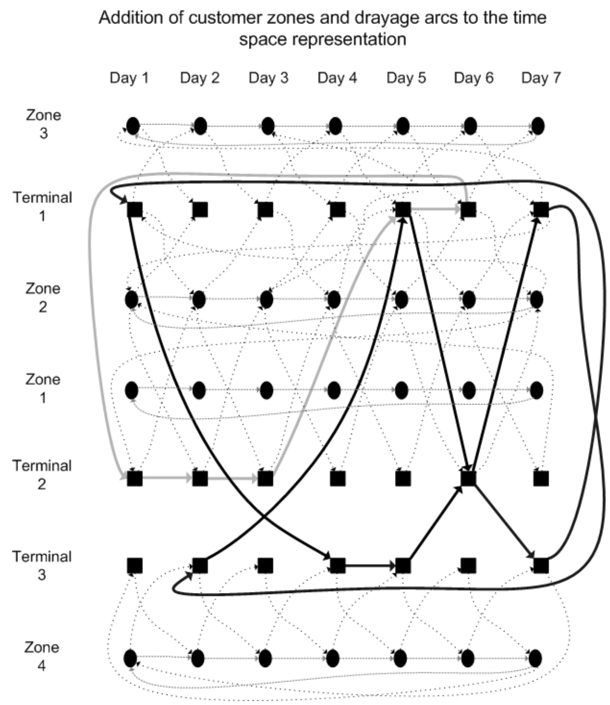

Figure 8

Given that intermodal operators are trying to achieve as seamless transportation chains as possible we assume that drayage moves are possible to perform whenever needed. That is why every terminal node is connected to a customer zone node and vice versa. The assumption is reasonable if intermodal operators plan drayage moves according to the train schedule. By assuming drayage moves are readily available at any given time the drayage arcs become un-capacitated arcs.

We assume that a commodity may be picked up any time after its time of availability. Holding arcs between the customer zone nodes are added (dashed grey lines in figure 8) to represent the possibility of commodities remaining at the customer after the time of availability until they are picked up eventually. The holding arcs have no operational cost or capacity associated to them.

3.4. Representing terminal events and operations

Making an adequate representation of terminal operations is a non-trivial task. They can not be neglected as they play an important role in the intermodal transportation chain. However, making a too detailed representation of them in a large-scale network will make the model computationally intractable. Each terminal node has a number of train canals connected to and from it representing possible train services of different modes to and from the terminal within the given time period. Furthermore a number of train transfer arcs for each mode connect from the terminal node's predecessor and to the terminal node's successor. Each mode is assumed to have a fixed train make-up, thus no operations are performed on them. However, only a limited number of trains may be present at a terminal at a given time, e.g. corresponding to the number of tracks at the terminal. Furthermore all trains that enter the terminal node must also leave it again to respect the conservation of trains.

Figure 9 shows the terminal node for terminal 1 on day 5 from figure 4. We assume that there are two available tracks at the terminal, and thus two train transfer arcs are added for each mode (grey and black horizontal full-line arrows). The remaining three train canals (inclined full line arrows) represent the train canals from figure 4 arriving from 'terminal 3, day 2' using the "black" mode, from 'terminal 2, day 3' using the "grey" mode, and leaving for 'terminal 2, day 6' using the "black" mode respectively. Using the representation in figure 9 the connection time for freight using the train canal arcs and/or train transfer arcs is not correctly represented. Both types of arcs assume that the departure time corresponds to the end of terminal nodes' time period and arrive at the beginning of them (the shaded areas in figure 9). To represent the connection time at the terminal the difference between the end time and the start time of a terminal node is added to all departing train canal arcs and departing train transfer arcs. The same procedure is done the departing drayage arcs.

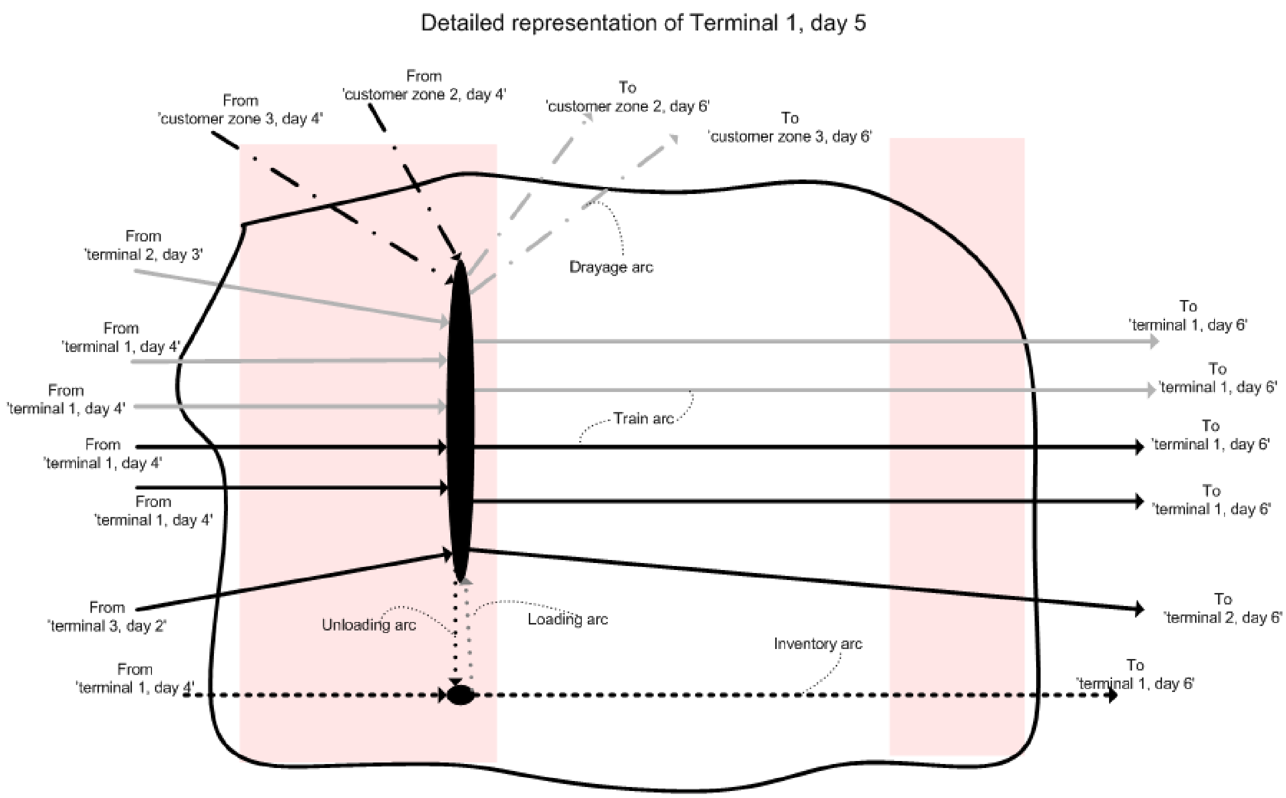

Figure 9

Commodities can arrive to and depart from a terminal either by truck (using a drayage arc) or by train (using the selected train canals). Arriving to a terminal by train freight can either stay on-board the train, be transferred to another vehicle (truck or train), or be unloaded and put in the terminals storage place for later pick-up. Freight leaving a terminal on a train could have been plucked from the terminals storage space and loaded onto the train, transferred from a truck or another train, or simply stayed put on the train itself. The same possibilities are possible for freight arriving at a terminal except for truck/train and train/truck transfer being the only meaningful vehicle transfers.

An inventory arc (short dashed black line in figure 9) is used to represent the inventory level in the terminals storage. An inventory arc of a terminal node is connected to the succeeding terminal node to represent the transition of inventory from one time period to another. To follow the same definitions for the train canal arcs and train transfer arcs inventory arcs “depart” at the beginning of a terminal nodes time period and “arrives” at the beginning of its successors time period. The “transit” time of an inventory arc is equal to the time difference between the start times of the two terminal nodes’ time periods.

As shown in figure 9 a loading arc (dotted grey line) and an unloading (dotted black line) arc is added to each node to capture the flow to and from inventory. The unloading and loading operations are assumed to occur during the entire time period. However, since the duration is captured by the departing train canal arcs, departing train transfer arcs, departing drayage arcs, and the inventory arc, the loading and unloading operations are represented as if they happen instantaneously at the beginning of the terminal node’s time period. Their “transit” times are therefore zero.

In the terminal representation shown in figure 9 it is not possible to distinguish between how the arriving and departing services are connected. Similarly it is not possible to distinguish between freight staying on a train and the vehicle transfers as there are no arcs representing the different possibilities. This means that a general assumption needs to be made on all freight arriving at a terminal. We thus assume that all freight arriving at a terminal on a service and leaving the terminal again on another service is transferred. The representation in figure 9 is thus an inaccurate representation of the possible terminal operations since staying on a train is not represented. However, a representation that could distinguish between the above mentioned operations would be much too detailed and computationally intractable. The assumption made here is an upper bound on the operational cost in a terminal and gives a better representation than assuming that no transfer operations are performed.

To capture the cost of vehicle transfer operations we add a variable vehicle transfer cost to each departing train canal arc, train transfer arc, and departing drayage arc. By doing this we implicitly assume that rail/rail and truck/rail transfers have the same cost. The cost of loading and unloading freight is different though as we assume some additional cost is incurred by moving a container to/from the storage place and storing/plucking it. An inventory transfer cost is added to the loading and unloading arcs. However, since all freight loaded continues on a departing train canal arc, train transfer arc, or drayage arc, the vehicle transfer cost is incurred. The vehicle transfer cost is thus subtracted from the inventory transfer cost on the loading arcs. Finally, a unit inventory cost is associated to the inventory arcs representing the storage cost of a container

Modelling assumptions			
Demand	deterministic	Terminals	divided into time prereiods with a start time and an end time associated to them
	constant		train arcs are connected to terminal time periods according to departure and arrival times
	aggregated figures for customer zones (clusters of customers)		train transfers connect sequential time periods
	has an origin customer zone and a time of availability		capacity limits on the number of trains in a terminal within a time period
	has a destination customer zone but no fixed delivery time		all trains that enter a terminal at a time period (on a train canal or a train transfer) must leave it again
	linear value of time cost		commodities can be unloaded into or loaded from inventory
Train canals	represens a train routing possibility between two terminals		unit cost for loading and unloading operations
	has an departure time from a terminal and an arrival time to a terminal		commodities can transfer between train canals, train transfers, and drayage arcs
	limited number		commodities staying on trains are considered as being transferred
	one train may use a train canal		unit transfer cost per container
	fixed cost of use (acquisition and operating a train on it)		capacity limit on handling operations (transfer+loading+unloading)
	has a mode associated to it giving the capacity limt		capacity limit on inventory flow
Train transfers	represents trains staying at a terminal, transferring from one period to another	Network modelling	multicommodity demand representation
	connects sequential time periods of a terminal		a commodity has an origin ndode and multiple possible destination nodes (all belonging to the same customer zone)
	limited number determined a priori		customer zones discretized in customer zone nodes
	fixed cost representing depreciation or leasing cost of having an idle train at the terminal		terminals discretized in terminal nodes
	has a mode associated to it giving the capacity limt		train canal and train transfers represented by design arcs with a binary variable associated to them
Modes	represent a given train make-up and thus a capacity limit		design arcs connect two terminal nodes
	one mode is associated to each train canal		drayage moves represents by non-capacitated arcs with no binary variable assocaited to them
Drayage moves	connects customer zones an terminal		holding arcs added between customer zone nodes to allow commodities to stay in the customer zone after the time of availability
	unit transportation cost (one truck per container unit)		invetory arcs between terminal node model the inventory level at terminals
	unlimited capacity		cyclic network representing a repetitive network after the time horizon is reached

Figure 10

The number of possible operations in a time period at a terminal is limited. Because the same resources are used for vehicle transfers, unloading, and loading, a capacity constraint can not be added to a single arc. Previously we assumed that all freight leaving on train canal arcs, train transfer arcs, or drayage arcs went through a transfer process. Additionally freight may be loaded or unloaded into inventory. Assuming that all operations take similar time (i.e. the bottleneck is at the handling operations at the trains and not on the storage transfer transportation) and remembering that loaded freight is captured by the departing vehicle arcs, the sum of all terminal operations may be written as sum of the flow on departing train canal arcs, train transfer arcs, drayage arcs, and on unloading arcs. This sum is thus restricted by the maximum handling operations capacity. The storage places also have limited space. Thus a capacity limit is imposed on the flow on the inventory arcs.

4. Mathematical formulation of the model

In this section we present the mathematical optimization model formulation based on the modelling assumptions presented in section 3. The model is an arc flow based MIP model.

4.1. Sets

In this sub-section we present the sets used in the model. The use of the different sets in the model constraints is illustrated in figures 10 to 13 and an overview of them can be seen in appendix A. We define a set of customer zones:

$$\mathcal{Z} : \textit{set of customer zones}$$

For each customer zone $z \in \mathcal{Z}$ there is a set of customer zone nodes, and we define a union of all customer zone nodes:

$$\mathcal{N}_z : \textit{set of customer zone nodes for customer zone } z$$
$$\bigcup_{z \in \mathcal{Z}} \mathcal{N}_z : \textit{union of all customer zone nodes}$$

The customer zone node sets are indexed by $k,l \in \bigcup_{z \in \mathcal{Z}} \mathcal{N}_z$. To represent the sequence of customer zone nodes, a customer zone node k's preceding neighbour and succeeding neighbour is denoted:

$$n_z^-(k) \in \mathcal{N}_z, k \in \mathcal{N}_z : \textit{preceding customer zone node for customer zone node } k$$
$$n_z^+(k) \in \mathcal{N}_z, k \in \mathcal{N}_z : \textit{succeeding customer zone node for customer zone node } k$$

As for customer zones, we determine a set of terminals, $\mathcal{S} : \textit{set of terminals}$. For each of the terminals in $s \in \mathcal{S}$ there is a set of terminal nodes, and we define a union of all terminal nodes:

$$\mathcal{N}_s : \textit{set of terminal nodes for terminal } s$$
$$\bigcup_{s \in \mathcal{S}} \mathcal{N}_s : \textit{union of all terminal nodes}$$

The terminal node sets are indexed by $h,i,j \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_s$. To represent the sequence of terminal nodes, terminal node i's preceding and succeeding neighbour is denoted:

$$n_s^-(i) \in \mathcal{N}_s, i \in \mathcal{N}_s : \textit{preceding terminal node for temrinal node } i$$
$$n_s^+(i) \in \mathcal{N}_s, i \in \mathcal{N}_s : \textit{succeeding terminal node for terminal node } i$$

We define a set off commodities that represents the demand of freight in the network.

$\mathcal{P} : \textit{set of commodities}$ **. Each commodity** $p \in \mathcal{P}$ **has an origin customer zone node and a destination customer zone associated to it:**

$$o(p) \in \bigcup_{z \in \mathcal{Z}} \mathcal{N}_z : \quad \textit{origin customer zone node of commodity } p$$

$$d(p) \in \mathcal{Z} : \quad \textit{destination customer zone of commodity } p$$

The time associated to the origin customer zone node k , t_k^{γ} , represents the time of availability of commodity p. The different train make-up strategies are represented by a set of modes: $\mathcal{M}$ *: set of modes representing a train make-up*

The only element that distinguishes train canal arcs and train transfer arcs of the same mode is that train transfer arcs connect two terminal nodes belonging to the same terminal. Thus we treat train canal arcs and train transfer arcs similarly in the mathematical formulation and denote them train arcs. For each mode $m \in \mathcal{M}$ there is a set of train arcs and each train arc connects terminal node i and terminal node j:

$\mathcal{L}_m$ *: Set of train canal arc for mode m*

$$\mathcal{L}_m : \{(i,j) \in \mathcal{L}_m \mid (i,j) \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_s\}$$

Given the train canal arc sets we for each terminal node i define the inward neighbours from terminal s, $\mathcal{N}^-_{s,m}(i)$, and the outward neighbours to terminal s, $\mathcal{N}^+_{s,m}(i)$:

$$\mathcal{N}^-_{s,m}(i) : \{ j \in \mathcal{N}^-_{s,m}(i) \mid (i,j) \in \mathcal{L}_m \}$$

inward neighbours of node i from terminal s

$$\mathcal{N}^+_{s,m}(i) : \{ j \in \mathcal{N}^+_{s,m}(i) \mid (i,j) \in \mathcal{L}_m \}$$

outward neighbours of node i from terminal s

The union of all inward terminal neighbour nodes and the union of all outward terminal nodes are also defined:

$$\bigcup_{s \in \mathcal{S}} \mathcal{N}^-_{s,m}(i) : \textit{union of all inwards terminal nodes to terminal node } i$$

$$\bigcup_{s \in \mathcal{S}} \mathcal{N}^+_{s,m}(i) : \textit{union of all outwards terminal nodes from terminal node } i \; ,$$

We define two sets of drayage arcs; a set of drayage arcs going from customer zone nodes to terminal nodes and one vice versa:

$\mathfrak{D}^{-} : \{(k,i) \in \mathfrak{D}^{-} \mid k \in \bigcup_{z \in \mathcal{Z}} \mathfrak{N}_z, i \in \bigcup_{s \in \mathcal{S}} \mathfrak{N}_s\}$

set of drayage arcs going from customer zone nodes to terminal nodes

$\mathfrak{D}^{+} : \{(i,k) \in \mathfrak{D}^{+} \mid i \in \bigcup_{s \in \mathcal{S}} \mathfrak{N}_s, k \in \bigcup_{z \in \mathcal{Z}} \mathfrak{N}_z\}$

set of drayage arcs going from terminal nodes to customer zone nodes

Given the drayage arcs sets we define $\mathfrak{N}_z^{-}(i)$ as the inward customer zone node neighbours from customer zone z of terminal node i, $\mathfrak{N}_z^{+}(i)$ as the outward customer zone node neighbours to customer zone z of terminal node i, $\mathfrak{N}_s^{+}(k)$ as the outward terminal node neighbours to terminal s from customer zone node k, and $\mathfrak{N}_s^{-}(z)$ as the outward terminal nodes from terminal s for customer zone z:

$\mathfrak{N}_z^{-}(i) : \{k \in \mathfrak{N}_z^{-}(i) \mid (k,i) \in \mathfrak{D}^{-}\}$

Set of inward customer zone node neighbours from customer zone z for terminal node i

$\mathfrak{N}_z^{+}(i) : \{k \in \mathfrak{N}_z^{+}(i) \mid (i,k) \in \mathfrak{D}^{+}\}$

Set of outward customer zone node neighbours from customer zone z for terminal node i

$\mathfrak{N}_s^{+}(k) : \{i \in \mathfrak{N}_s^{+}(k) \mid (k,i) \in \mathfrak{D}^{-}\}$

Set of outward terminal node neighbours from terminal s for customer zone node k

$\mathfrak{N}_s^{-}(z) : \{i \in \mathfrak{N}_s^{+}(z) \mid (i,k) \in \mathfrak{D}^{+}, k \in \mathfrak{N}_z\}$

Set of outward terminal node neighbours from terminal s for customer zone z

For each of the four sets defined above we define the unions as:

$\bigcup_{z \in \mathcal{Z}} \mathfrak{N}_z^{-}(i)$: *Union of all inward customer zone node neighbours for terminal node i*

$\bigcup_{z \in \mathcal{Z}} \mathfrak{N}_z^{+}(i)$: *Union of all outward customer zone node neighbours for terminal node i*

$\bigcup_{s \in \mathcal{S}} \mathfrak{N}_s^{+}(k)$: *Union of all outward terminal node neighbours for customer zone node k*

$\bigcup_{s \in \mathcal{S}} \mathfrak{N}_s^{-}(z)$: *Union of all outward terminal node neighbours for customer zone z*

4.2. Variables

This sub-section presents the variables used in the model. Appendix A includes an overview of the variables used. To represent the flow of commodities we use continuous variables. Although containers cannot be shipped in fractional numbers, the demand figures used in the model are only indications of the demand potentials and cannot be seen as actual shipment orders. There is a set of variables representing the commodity flow for holding arcs, inventory arcs, unloading arcs, loading arcs, drayage arcs, and train canal arcs:

$x^{\psi}_{i,p}$: *holding amount of commodity p at customer zone node k*

$x^{\mu}_{i,p}$: *Amount commodity p unloaded at terminal node i*

$x^{\nu}_{i,p}$: *Amount of commodity p loaded at terminal node i*

$x^{\omega}_{i,p}$: *inventory of commodity p at terminal node i*

$x^{\delta^-}_{i,k,p}$: *flow of commodity p from terminal node i to customer zone node k*

$x^{\delta^+}_{i,k,p}$: *flow of commodity p from customer zone node k to terminal node i*

$x^{\lambda}_{i,j,m,p}$: *flow of commodity p on train of mode m running on train canal arc (i, j)*

To capture the use of train arcs a binary variable is associated to each train canal arc

$$y^{\lambda}_{i,j,m} = \begin{cases} 1, \text{if a train of mode } m \text{ uses train arc } (i,j) \\ 0, \text{else} \end{cases}$$

4.3. Parameters

This sub-section presents the parameters used in the mode. An overview is shown in appendix A. We start by recalling that the value of time is assumed have a linear relationship to the transit time. We thus define b to be the unit cost per time unit representing the value of time. Each commodity p has a parameter d_p representing the total amount of commodity p that needs to be shipped from its origin customer zone node $o(p)$ to its destination customer zone $d(p)$. We introduce a parameter $a_{k,p} = a^p$ if $k = o(p)$, zero else, and a parameter $a_{z,p} = d^p$ if $z = d(p)$, zero else.

$$d_p \text{ : demand of commodity } p$$

$$a_{k,p} = \begin{cases} d_p \text{ if } k = o(p) \\ 0 \text{ else} \end{cases}$$

$$a_{z,p} = \begin{cases} d_p \text{ if } z = d(p) \\ 0 \text{ else} \end{cases}$$

Each customer zone node has a time associated to it representing the time of occurrence

t_k^{γ} *: time associated to customer zone node k*

In contrast to customer zone nodes each terminal node has a start time and an end time associated to it.

t_i^{α} *: start time associated to terminal node i*

t_i^{β} *: end time associated to terminal node i*

The transit time associated to the holding arc in customer zone node k, t_k^{ψ} , is equal to the time difference between the time of customer zone node k and the time of its succeeding customer zone node's time. However, if k is the last node in the time horizon, and thus $n_z^+(k)$ is the first, the transit time is calculated as the time horizon, T, minus the time of customer zone node k time plus the time of customer zone node $n_z^+(k)$:

$$t_k^{\psi} = \begin{cases} t_{n_z^+(k)}^{\gamma} - t_k^{\gamma}, \text{if } t_{n_z^+(k)}^{\gamma} \geq t_k^{\gamma} \\ (T - t_k^{\gamma}) + t_{n_z^+(k)}^{\gamma}, \text{if } t_{n_z^+(k)}^{\gamma} < t_k^{\gamma} \end{cases}$$

transit time of holding arc from customer zone node k

We can calculate the variable cost c_k^{ψ} associated the holding arc at customer zone node k as the transit time of the holding arc times the unit time cost:

$c_k^{\psi} = b \cdot t_k^{\psi}$ *: unit cost of holding arc from customer zone node k*

The transit time associated to the inventory arc in terminal node i, t_i^{ϖ}, is equal to the time difference between the start time of terminal node i and the time of its succeeding terminal node's start time. However, if i is the last node in the time horizon, and thus $n_s^+(i)$ is the first, the transit time is calculated as the time horizon, T, minus the start time of terminal node i time plus the start time of terminal node $n_s^+(i)$:

$$t_i^{\varpi} = \begin{cases} t_{n_s^+(i)}^{\alpha} - t_i^{\alpha}, \text{if } t_{n_s^+(i)}^{\alpha} \geq t_i^{\alpha} \\ (T - t_i^{\alpha}) + t_{n_s^+(i)}^{\alpha}, \text{if } t_{n_s^+(i)}^{\alpha} < t_i^{\alpha} \end{cases}$$

transit time of inventory arc from terminal node i

Each inventory arc has storage cost, c_i^e, associated to it. The total variable inventory cost, c_i^{ϖ}, associated with the inventory arc at terminal node i is the transit time of the holding arc times the unit time cost plus the storage cost:

$c_i^{\omega} = c_i^{e} + b \cdot t_i^{\omega}$: *unit cost for invenotry arc from terminal node i*

The loading and unloading arcs of terminal node i have no transit time but each have an inventory transfer cost, c_i^u, associated to them. The unloading cost is equal the inventory transfer cost while the loading arc cost is subtracted the vehicle transfer cost, c_i^{τ}:

$c_i^{\mu} = c_i^{u}$: *unit cost for unloading arc in terminal node i*

$c_i^{\nu} = c_i^{u} - c_i^{\tau}$: *unit cost for loading arc in terminal node i*

To constrain the inventory levels each terminal node i has a capacity limit determining the maximum number of allowed containers in its storage place:

w_i : *inventory capacity limit of terminal node i*

To constrain the number of transfer, unloading, and loading operations, each terminal node i has a handling capacity limit determining the maximum number of terminal operations in the terminal nodes time period:

u_i : *handling capacity limit of terminal node i*

Finally the train number capacity determines the maximum number of allowed trains in the terminal node:

v_i : *train capacity limit of terminal node i*

Each train arc (i,j) has a fixed cost associated to it representing the cost of routing a train on it:

$f_{i,j,m}^{\lambda}$: *fixed cost of routing a train on train canal arc (i, j) of mode m*

The transit time of train canal arc (i,j), t_{ij}^{λ}, is the sum of the actual transit time (from end time of departing terminal node i to start time of arriving terminal node j) plus the time period of the departing terminal node i. However, if the end time of terminal node i is larger than the start time of terminal node j (meaning the train runs into the following schedule repetition) the transit time is calculated as the time horizon, T, minus the start time of terminal node i time plus the start time of terminal node j:

$$t_{ij}^{\lambda} = \begin{cases} t_j^{\alpha} - t_i^{\beta} + (t_i^{\beta} - t_i^{\alpha}) = t_j^{\alpha} - t_i^{\alpha}, if \ t_j^{\alpha} \geq t_i^{\alpha} \\ (T - t_i^{\alpha}) + t_j^{\alpha}, if \ t_j^{\alpha} < t_i^{\alpha} \end{cases}$$

transit time of train canal arc (i, j)

We also need to incorporate the vehicle transfer cost incurred in terminal nodes in the train arcs. Since each arc is a departing train arc the transfer cost for the origin terminal node i, c_i^{τ}, is added to each train arc (i,j). The variable unit cost associated to each train arc is obtained by adding the vehicle transfer cost and the unit time cost multiplied by the transit time:

$$c_{ij}^{\lambda} = c_i^{\tau} + b \cdot t_{ij}^{\lambda} \ : \ unit\ cost\ \ for\ train\ canal\ arc\ (i,j)$$

The flow on train arcs is limited by the make-up defined by the arcs mode. A capacity limit is associated to each mode representing the maximum number of containers transported on a train canal using mode m:

$$q_m^{\lambda} \ : \ capacity\ of\ \ mode\ m$$

We assume that drayage arcs have the same transportation cost whether they connect terminal node i to customer zone node k or vice versa, $c_{ki}^{d+} = c_{ik}^{d-}$.The transit times differ though. Drayage arc connecting customer zone node k to terminal node i have a transit time $t_{ki}^{\delta+}$ equal to the difference between the terminal node's start time and the customer zone node's time:

$$t_{ki}^{\delta+} = \begin{cases} t_i^{\alpha} - t_k^{\gamma} \ if\ t_i^{\alpha} \geq t_k^{\gamma} \\ t_i^{\alpha} + (T - t_k^{\gamma})\ if\ t_i^{\alpha} < t_k^{\gamma} \end{cases}$$

The variable cost of using the drayage arc between customer zone node k and terminal node i is the sum of the transportation cost and the transit time multiplied by the unit time cost:

$$c_{ki}^{\delta+} = c_{ki}^{d+} + b \cdot t_{ki}^{\delta+} \ : \ unit\ cost\ of\ \ using\ drayage\ arcs \\ from\ customer\ zone\ node\ k\ to\ terminal\ node\ i$$

Drayage arc connecting customer zone node k to terminal node i have a transit time $t_{ki}^{\delta-}$ equal to the difference between the terminal node's end time and the customer zone node's time plus the time duration of the terminal node:

$$t_{ik}^{\delta-} = \begin{cases} t_k^{\gamma} - t_i^{\beta} + (t_i^{\beta} - t_i^{\alpha}) = t_k^{\gamma} - t_i^{\alpha} \ if\ t_k^{\gamma} > t_i^{\alpha} \\ t_k^{\gamma} + (T - t_i^{\alpha})\ if\ t_k^{\gamma} < t_i^{\alpha} \end{cases}$$

The vehicle transfer cost of terminal node i, c_i^{τ}, is added to the drayage arcs departing from it. The variable cost of using the drayage arc connecting terminal node i and customer zone node k is the sum of the transportation cost, the terminal vehicle transfer cost, and the transit time multiplied by the unit value of time cost:

$$c_{ik}^{\delta-} = c_{ik}^{d-} + c_i^{\tau} + b \cdot t_{ik}^{\delta-} \ : \ unit\ cost\ of\ \ using\ drayage \\ arcs\ \ from\ terminal\ node\ i\ to\ customer\ zone\ node\ k$$

4.4. Constraints

The objective is to minimize operational cost and transit time. Thus the objective function has to minimize the sum off the transportation costs, terminal operation costs, and value of time cost. The variable transportation cost, the terminal operation costs, and the value of time costs have been aggregated into unit costs associated to the flow on the respective arcs. The objective function can be written as a minimization of the sun of the fixed train canal costs multiplied by the train canal selection variable plus the sum of the unit container transit costs times the flow on the corresponding arcs:

$$\begin{aligned}\min z = &\sum_{(i,j)\in\mathcal{L}_m}\sum_{m\in\mathcal{M}} f^{\lambda}_{i,j,m} y^{\lambda}_{i,j,m} + \\ &\sum_{(i,j)\in\mathcal{L}_m}\sum_{m\in\mathcal{M}}\sum_{p\in\mathcal{P}} c^{\lambda}_{i,j} x^{\lambda}_{i,j,p,m} + \sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\omega}_{i} x^{\omega}_{i,p} + \\ &\sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\mu}_{i} x^{\mu}_{i,p} + \sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\nu}_{i} x^{\nu}_{i,p} + \\ &\sum_{(k,i)\in\mathcal{D}^+}\sum_{p\in\mathcal{P}} c^{\delta+}_{k,i} x^{\delta+}_{k,i,p} + \sum_{(i,k)\in\mathcal{D}^-}\sum_{p\in\mathcal{P}} c^{\delta-}_{i,k} x^{\delta-}_{i,k,p} + \\ &\sum_{k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z}\sum_{p\in\mathcal{P}} c^{\psi}_{k} x^{\psi}_{k,p}\end{aligned} \qquad (1)$$

We introduce a number of flow balance constraints to respect the conservation of flow in the network. Commodity p originates from a customer zone node and may continue on the holding arc or on any of the departing drayage arcs. The flow balance constraint for commodities leaving customer zone nodes becomes:

$$x^{\psi}_{n_z^-(k),p} + a_{k,p} = x^{\psi}_{k,p} + \sum_{i\in\bigcup_{s\in S}\mathcal{N}_s^+(k)} x^{\delta+}_{k,i,p} \quad \forall k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z, \forall p\in\mathcal{P} \quad (2)$$

The constraint is illustrated in figure 11.

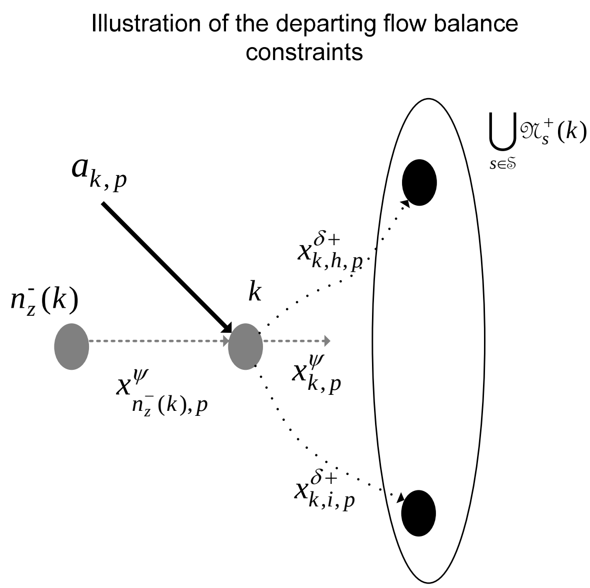

Figure 11

Since commodities have a destination customer zone, and no destination customer zone node, the flow balance constraint for commodities arriving at their destination customer zone must make sure that the sum of flow to all customer zone nodes associated to the destination customer zone must be equal to the demand of the commodity. The flow balance constraint for commodities arriving at a customer zone becomes***:***

$$\sum_{i \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_s^-(z)} \sum_{k \in \mathcal{N}_z} x_{i,k,p}^{\delta-} = a_{z,p} \quad \forall z \in \mathcal{Z}, \forall p \in \mathcal{P} \quad (3)$$

The constraint is illustrated in figure 12.

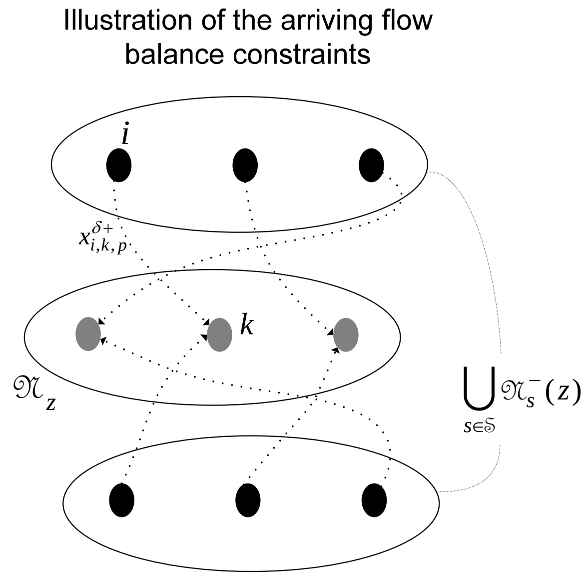

Figure 12

A commodity p may arrive at a terminal on a drayage arcs from customer zone nodes or on train arcs. Along with the commodities loaded from inventory onto other train canal arcs or drayage arcs and may also be stored back in inventory. For each terminal node i the sum of the flow of commodities arriving on arriving train arc, arriving drayage arcs, and the loading arc must be equal the sum of the flow leaving on departing train arcs, departing drayage arcs, and the unloading arc:

$$\begin{aligned}
&\sum_{m \in \mathcal{M}} \sum_{h \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_{s,m}^+(i)} x_{h,i,p,m}^{\lambda} + \\
&\sum_{k \in \bigcup_{z \in \mathcal{Z}} \mathcal{N}_z^+(i)} x_{k,i,p}^{\delta+} + x_{i,p}^{\mu} = \\
&\sum_{l \in \bigcup_{z \in \mathcal{Z}} \mathcal{N}_z^-(i)} x_{i,l,p}^{\delta-} + x_{i,p}^{\nu} + \qquad \forall i \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_s, \forall p \in \mathcal{P} \quad (4) \\
&\sum_{m \in \mathcal{M}} \sum_{j \in \bigcup_{s \in \mathcal{S}} \mathcal{N}_{s,m}^-(i)} x_{i,j,p,m}^{\lambda}
\end{aligned}$$

The constraint is illustrated in figure 13.

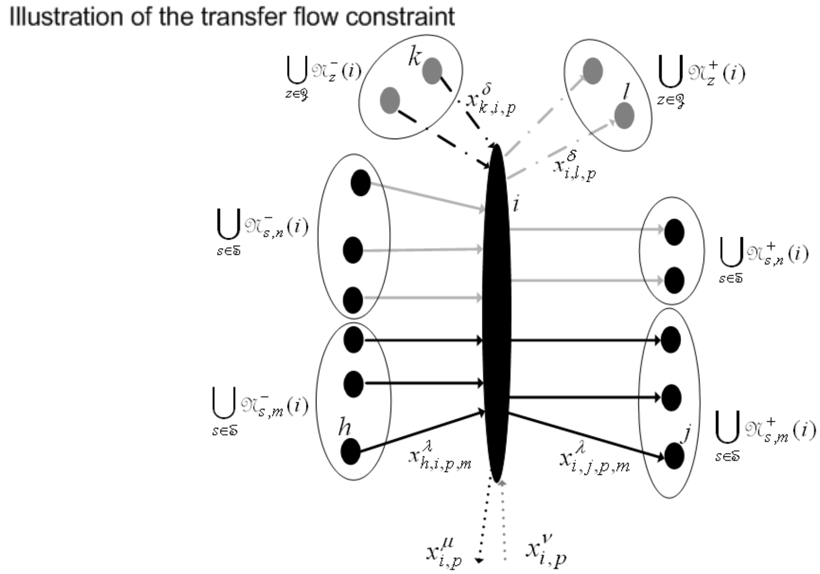

Figure 13

A flow balance constraint must also be added to ensure that there is conservation of flow in the terminals storage place. For each terminal node i the sum of the inventory flow from the predecessor (inventory level before the start of the time period) plus what is unloaded must be equal the sum of what is loaded plus what is left in inventory:

$$x^{\omega}_{n_s^+(i),i,p} + x^{\mu}_{i,p} = x^{\omega}_{i,p} + x^{\nu}_{i,p} \quad \forall i \in \bigcup_{s\in\mathcal{S}} \mathcal{N}_s, \forall p \in \mathcal{P} \quad (5)$$

The constraint is illustrated in figure 14. Notice that there are no initial and terminal conditions on the storage level due to the cyclic, repetitive schedule.

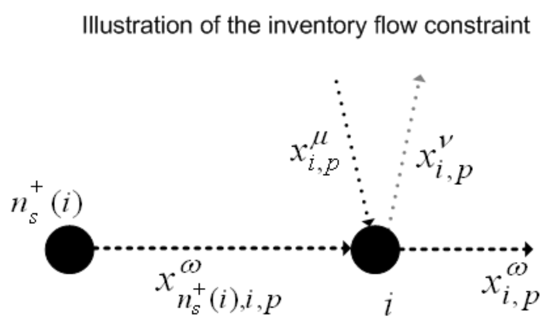

Figure 14

The capacity of terminal handlings is determined as the sum of all handling operations. All freight not remaining in inventory is handled, i.e. freight from train arcs, drayage arcs, and loading and unloading arcs. Given the balance of flow given from equation (4) the sum of all transfer operations may either be written as the sum of flow on all arriving arcs or departing arcs. We have chosen the flow on arriving arcs:

$$x^{v}_{i,p} + \sum_{m\in\mathcal{M}}\sum_{p\in\mathcal{P}}\sum_{j\in\bigcup_{s\in S}\mathcal{N}^{-}_{s,m}(i)} x^{\lambda}_{i,j,p,m} + \sum_{p\in\mathcal{P}}\sum_{k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}^{-}_{z}(i)} x^{\delta}_{i,k,p} \le u_i \qquad \forall i \in \bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (6)$$

Similarly the storage place in a terminal node has limited capacity. Thus the inventory level must be lower than the terminal node storage capacity:

$$\sum_{p\in\mathcal{P}} x^{\omega}_{ip} \le w_i \qquad \forall i \in \bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (7)$$

Finally we ensure the maximum number of trains that can be accommodated at terminal node i by setting the sum of all selected arriving train canal arcs to be less than or equal to the train number capacity

$$\sum_{m\in\mathcal{M}}\sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^{-}_{s,m}(i)} y^{\lambda}_{i,j,m} \le v_i \qquad \forall i \in \bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (8)$$

The number of trains that enter a terminal node must be equal to the number of trains leaving the terminal again. The number of trains entering and leaving a terminal node is equal to the number of selected train canal arcs. Thus the number of selected train arcs entering a terminal node must equal the number leaving it:

$$\sum_{h\in\mathcal{N}^{+}_{s,m}(i)} y^{\lambda}_{h,i,m} = \sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^{-}_{s,m}(i)} y^{\lambda}_{i,j,m} \qquad \forall i \in \bigcup_{s\in\mathcal{S}}\mathcal{N}_s, \forall m \in \mathcal{M} \qquad (9)$$

The constraint is illustrated by the selection in figure 15.

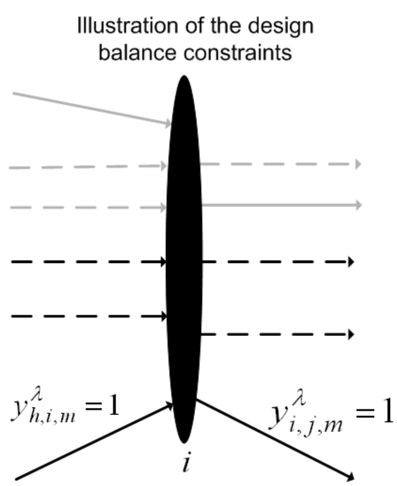

Figure 15

The maximum number of commodities that can flow on a train canal arc is constrained by the capacity of the arc's mode. By multiplying the design variable with the capacity and setting it larger or equal to the sum of the flow we ensure that the flow is less than or equal the mode capacity if the train arc is selected and zero else.

$$\sum_{p \in \mathcal{P}} x_{i,j,p,m}^{\lambda} \le q_m y_{i,j,m}^{\lambda} \quad \forall (i,j) \in \bigcup_{m \in \mathcal{M}} \mathcal{L}_m \quad (10)$$

The nine constraints (2)-(9) above plus the objective function (1) make up a MIP-model that designs an intermodal train schedule based on train canals considering terminal operations. An overview of the model can be found in appendix B. MIP-models are hard to solve for large instances. Even though the majority of the rail network capacity is reserved for passenger trains from a model point of view the remaining train canals still imply a large-scale instance for the model. Furthermore the multiple-destination choice for commodities provides looser bounds than for the traditional formulation with fixed delivery time.

5. Model implementation and result analysis

To get an indication of the complexity of the model we perform a series of computational experiments. These are conducted by using Xpress-MP's MIP solver to solve a generated test network. Model behaviour will be examined by simulating different scenarios with respect to value of time and commodity amounts.

5.1. Constructing a test data set

To test the model a test network was constructed. The network is generated in a size which may resemble a realistically sized intermodal train service network. The network consists of 25 rail terminals and 15 customer zones. The 25 rail terminals are connected by a physical rail infrastructure and the customer zones are connected to one or more of the rail terminals. The geographic lay-out of the network is illustrated in figure 16. The connections between the terminals show which modes are applied to which connections. There are three modes with capacities of 50, 100, and 200 respectively. The planning horizon used in the instance is a 7 day period and each of the customer zones and terminals are represented by 7 nodes each representing a day of the week. There are thus 105 customer zone nodes and 175 terminal nodes. The total number of nodes is 280. All terminal nodes representing one terminal are assumed to have the same capacity of trains, transfer operations, and inventory. All instance attributes are illustrated in the summary in figure 17.

The number of available train canals between terminals for each of the three modes is 253, 112, and 62 respectively. In addition 203, 112, and 63 train transfer arcs have been added for the three modes respectively in the terminals the modes connect. The total number of binary decision variables totals 805. In addition 434 drayage arcs have been added between the customer zone nodes and the terminal nodes. Finally 90 commodities were generated randomly each with one of the 105 customer zone nodes as their origin and one of the 15 customer zones as their destination. To examine the models behaviour 9 different scenarios have been created by varying the commodity amount values and the value of time each on three different levels. The origins and destinations of commodities are the same for all scenarios, but the amounts vary with a factor 1, 1.5, and 2 (low, medium, and high) respectiely. The value of time is set to 0, 5, and 10 (low, medium, and high) respectively to simulate scenarios where the value of time has no importance, has medium importance compared the operating cost, and has high importance compared to the operating costs respectively.

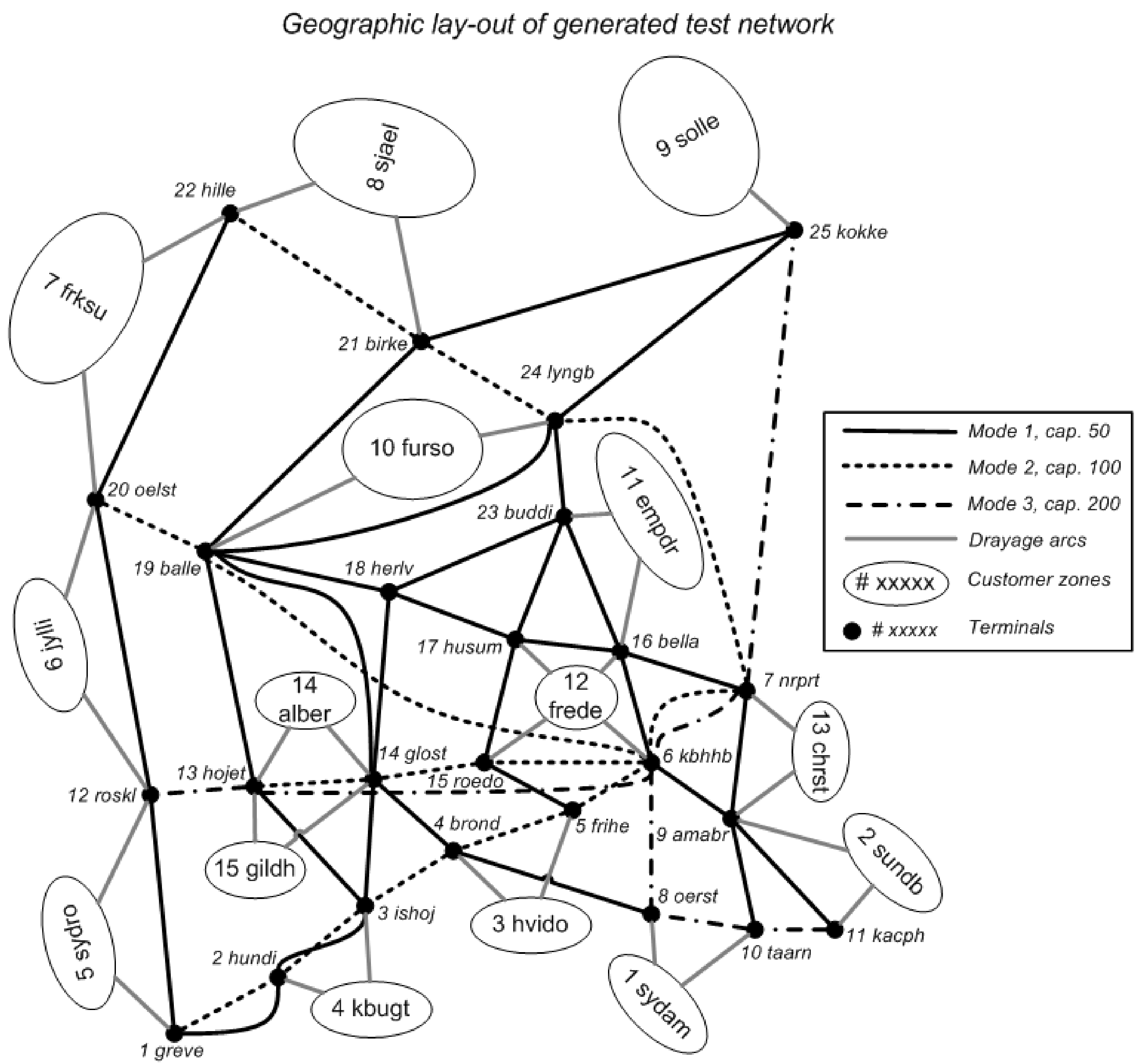
Geographic lay-out of generated test network
Mode 1, cap. 50
Mode 2, cap. 100
Mode 3, cap. 200
Drayage arcs
xxxxx
Customer zones
xxxxx
Terminals
9 solle
8 sjael
7 frksu
22 hille
25 kokke
21 birke
24 lyngb
10 furso
11 empdr
20 oelst
23 buddi
19 balle
18 herlv
6 jylli
17 husum
16 bella
12
frede
7 nrprt
14
alber
13 chrst
14 glost
6 kbhhb
13 hojet
15 roedo
12 roskl
4 brond
5 frihe
9 amabr
2 sundb
15 gildh
3 ishoj
3 hvido
8 oerst
5 sydro
2 hundi
10 taarn
11 kacph
4 kbugt
1 sydam
1 greve

Figure 16

Nodes						
Total nodes: $\mathcal{N}_z \cup \mathcal{N}_s$	280					
Terminals: $\mathcal{S}$	25					
Terminal nodes: $\mathcal{N}_s$	175					
	terminal attributes					
	Vehicle transfer cost: c_i^{τ}	Load/unloading cost: c_i^{u}	Storage cost: c_i^{e}	handling capacity: u_i	Storage capacity: w_i	train capacity: v_i
ranges	65 - 85	80 - 100	10 - 30	70 - 105	100	2 - 4
Arcs						
# of train arcs	805					
	mode 1		mode 2		mode 3	
train canal arcs $y^{\lambda}_{i,j,m}$	253		112		62	
train transfer arcs $y^{\lambda}_{i,j,m}$	203		112		63	
	Train canal cost: $f^{\lambda}_{i,j,m}$	Train transfer costs: $f^{\lambda}_{i,j,m}$	Train canal cost: $f^{\lambda}_{i,j,m}$	Train transfer costs: $f^{\lambda}_{i,j,m}$	Train canal cost: $f^{\lambda}_{i,j,m}$	Train transfer costs: $f^{\lambda}_{i,j,m}$
ranges	9585 - 13960	6000	19875 - 26780	13000	38570 - 50060	27000
# of drayage arcs: $x^{\delta+}_{k,i,p}$, $x^{\delta-}_{i,k,p}$	434					
	Unit transportation cost: c^{d+}_{ki}, c^{d-}_{ik}					
range	78 - 150					
Commodties						
# of commodities: $\mathcal{P}$	90					
	low	medium	high			
commodity amount range: d_p	3 - 6	5 - 9	6 - 12			
total amount: $\sum_{p \in \mathcal{P}} d_p$	399	618	798			
Time horizon: T	168 hours					
	low	medium	high			
value of time: b	0	5	10			
Model attributes						
binary variables	805					
continuous varaibles	168210					
constraints	44155					

Figure 17

Solving the problem using Xpress-MP

Considering the problem size depicted in figure 17 it is clear that the problem is not trivially solved. However, we use Xpress-MP's standard MIP solver to solve the problem in order to get an impression of the models complexity. To achieve feasible solutions 90 hours of CPU time (324.000 seconds) was allocated to solve each of the 9 scenarios on a using a PC with an Intel Pentium 4, 2.26Ghz processor. As a preliminary exercise some of the options available for the MIP-solver was tried out to see if any improvements in solution and solution time could be identified. However, no noteworthy effects were identified and the problem was solved using the MIP-solver's standard settings. The computational results achieved are illustrated in the table in figure 18.

Computational results on 9 scenarios									
	Scenario attributes								
flow	low			medium			high		
value of time	low	medium	high	low	medium	high	low	medium	high
Time limit	324.000 sec.								
Number of feasible solutions	2	1	2	1	1	2	2	1	3
Best feasible solution time (sec.)	285,154	184,935	139,875	222,597	158,820	258,948	234,605	167,581	313,004
Best feasible solution value	1,068,860	1,746,850	2,482,350	1,444,330	2,238,650	3,206,130	1,746,030	2,736,030	3,950,940
Lower bound at time limit	617,764	1,136,761	1,620,757	905,356	1,678,198	2,405,664	1,136,908	2,113,203	3,033,931
Lower bound/best feasible solution gap	42.20%	34.93%	34.71%	37.32%	25.04%	24.97%	34.89%	22.76%	23.21%
LP solution time (sec.)	250	113	75	275	129	79	326	127	85
LP solution	455,293	838,196	1,200,815	704,272	1,296,965	1,858,225	910,586	1,676,393	2,401,627
LP/Best feasible solution gap	57.40%	52.02%	51.63%	51.24%	42.06%	42.04%	47.85%	38.73%	39.21%
Cut generation time	25,227	9,849	8,845	12,796	9,225	6,495	13,302	8,441	6,034
Solution post cut generation	617,046	1,131,995	1,618,252	895,433	1,669,894	2,401,414	1,121,190	2,103,043	3,017,793
Cut/Best feasible solution gap	42.27%	35.20%	34.81%	38.00%	25.41%	25.10%	35.79%	23.14%	23.62%

Figure 18

The computational results show as expected that the problem is difficult to solve. Within the large time limit of 90 hours only 1-3 feasible solutions were found for the scenarios. The model is fundamentally an extension of a network design model. These are generally hard to solve because of the poor lower bounds provided from the LP-relaxation. The model here is no exception where the gap between the LP relaxation and the best found feasible solutions is between 38% and 58%. The cuts generated give reasonable improvements, 18% to 27%, on the LP-relaxation but the gaps to the best feasible solution still lies between 23% and 43%. Furthermore the cut generation times lie between 2 and 7 hours of computational time. The gaps between the lower bound achieved at the time limit of the branch-and-bound process and the best feasible solutions lie between 22% and 43% effectively only decreasing the gap by up to 1.5%. The large gaps between the bounds and the tardiness of the best found feasible solutions means that it is unreasonable to expect the best found feasible solutions to be near optimal.

Although the solutions found are not optimal or near-optimal analysis of the obtained results may still indicate whether the model works appropriately. To make the analysis of the model we have calculated a number of key-performance indicators. These are presented in the table in figure 19. The columns of the table show the performance indicators for each of the nine scenarios. The first row shows the operational cost for the best feasible solution for each of the nine scenarios. One tendency that can be seen is that the operational costs increase the higher the amount of commodities is. The other tendency that can be seen is that the operational cost increase the higher the value of time is. When the value of time increases its impact on the objective value becomes larger meaning the model chooses to include more services (resulting in higher operational costs) in order to provide faster transit times for the commodities. The second row shows the increase in operational costs from low value of time to medium value of time and from medium value of time to high value of time. Note that the operational cost actually decreases when the value of time is increased from low to medium for the medium commodity amount scenarios. An explanation to this may be that the structure of the medium/medium scenario by coincidence allows the branch-and-bound solution to find a relatively better feasible solution.

Key performance indicators for the 9 scenarios									
flow	low			medium			high		
value of time	low	medium	high	low	medium	high	low	medium	high
operating cost	1.068.860	1.144.690	1.402.110	1.444.330	1.389.170	1.728.810	1.746.030	1.758.810	2.051.580
operating cost increase (low time → med. time, med. time → high time)		7,09%	22,49%		-3,82%	24,45%		0,73%	16,65%
number of train services	35	44	64	52	61	64	60	56	72
train transfers	11	26	29	28	17	27	24	29	38
total service capacity	2850	2600	3900	3200	3650	5000	4150	4600	5300
total flow on services	1228	1193	1227	1915	1790	1754	2326	2109	2235
total service capacity utilization	43,09%	45,88%	31,46%	59,83%	49,04%	35,07%	56,05%	45,84%	42,17%
terminal nodes with hadling operations/at max capacity	80/1	86/3	90/0	89/10	76/5	87/6	90/19	87/12	99/12
Total number of handling operations	2315	2344	2201	3607	3202	2973	4223	3796	3950
Total inventory flow	1541	1054	784	2279	1520	1100	2463	1601	1167
Total transit time	126.474	112.698	100.290	193.215	157.404	135.360	217.500	179.832	173.700
transit time decrease (low time → med. time, med. time → high time)		10,89%	11,01%		18,53%	14,00%		17,32%	3,41%

Figure 19

The third and fourth row show the number of chosen train canal arcs which is equivalent to the number of services offered and the number of chosen train transfers within the same terminal's terminal nodes. As expected the number of offered services increases with the value of time and the amount of commodities in order to provide more capacity and shorter transit times. Again one exception stands out when going from low value of time to medium value of time for the high commodity amount scenarios. A further analysis shows that the more of the high capacity mode train arcs are chosen in the high/medium than in the high/low scenario. Thus effectively the available capacity (row 6) and operational costs increase, but fewer services are offered. The seventh row show the overall capacity utilization of the services offered. The utilization decreases with the value of time again indicating that more services are offered to give shorter transit times the higher the value of time is. This tendency can also be seen by looking at rows eight and nine that show the total number of handling operations and total inventory flow in terminals. The higher the value of time the fewer transfer operations are performed and the less inventory flow is seen. The higher number of services means that more direct services can be offered to the commodities resulting in fewer transfers and also means that commodities have to wait for shorter time in inventory for a service departing from a terminal. Finally row ten shows the total transit time for each of the scenarios where it is clear to see that it decreases the higher the value of time cost is. By calculating and analysing the key performance indicators shown in figure 19 it is reasonable to conclude that the model captures the trade-off between value of time and operating cost. The question is now how an operator decides on the value of time. As the value of time increases the operational costs increase and the overall transit time decreases. It is then up to the operator to decide on the trade-off between increased operational costs and increased service to the customers.

6. Future research and conclusion

In this paper a mathematical model for intermodal train scheduling was presented. The main properties of the model are that it incorporates terminal operations in terminals by setting an upper limit on handling operations and inventory and by leaving the delivery time of commodities as an output of the model. The delivery times of commodities are determined by a linear cost representing the value of time. The scheduling of trains is done on train canals that are predetermined time dependent path on the rail infrastructure as is the case in the rail sector in Europe. To model the time dimension of schedules a space-time representation were the nodes represent locations within a time period and arcs represent movements in time and space between two different locations or simply a movement in time within the same location is used.

The model was applied to a generated network using nine different scenarios with varying commodity flow and value of time costs. The model was solved using Xpress-MP's MIP-solver with results, as expected, of mediocre quality. Before any practical implementation can be considered it is necessary to design better solution methods other than the "brute force" approach used here. There are several approaches one may envision to achieve better solution methods. For one, better lower bounds may help the branch-and bound process by reducing the gap between feasible solutions and the lower bound. Second, decomposition methods such as constraint generating approaches may also be an interesting approach. Finally heuristic methods can prove to be a fast and efficient way of finding feasible solutions. By combining the different approaches the gap between feasible solution and the lower bound may be reduced in order to find and prove near optimal solutions.

However, although the model has the basic structure of a network design problem (flow balance constraints and binding capacity constraints) there are a set of constraints that have not been handled before in network design models. These are the balancing constraints stating that the number of vehicles entering a node must also leave the node again. These constraints resemble vehicle routing constraints expect for there not being a depot. The constraints effectively bind the binary decision variables together as opposed to traditional network design models where opening or closing an arc can be done independently of the other design arc. We believe the next step is to research how to handle these new set of constraints for network design models before proceeding to applying such models to real instances**.**

7. Acknowledgments

This work was supported by the Virtual Centre for Logistics and Freight Transport (CLG) in Denmark, the Danish Technical Research Council, and Atkins Denmark. The Natural Sciences and Engineering Research Council of Canada (NSERC) offered additional support through its Discovery Grant program

8. References

[Ballis et al. 2004] Ballis, A. and Golias, J. (2004). Towards the improvement of a combined transport chain performance. *European journal of operational research* 152:420-436.

[Bookbinder et al. 1998] Bookbinder, J.H. and Fox, N.S. (1998). Intermodal routing of Canada-Mexico shipments under NAFTA. *Transportation Research* 34E(4):289-303.

[Bostel et al. 1998] Bostel, N. and Dejax, P. (1998). Models and algorithms for container allocation problems on trains in a rapid transshipment shunting yard. *Transportation science* 32(4):370-379.

[Cordeau et al 1998] Cordeau, J.-F., Toth, P. and Vigo, D. (1998) A survey of optimization models for train routing and scheduling. *Transportation science* 32(4):380-404.

[Crainic et al. 1994] Crainic, T., Ferland, J.-A., and Rousseau, J.-M. (1984). A tactical planning model for rail freight transportation. *Transportation science* 18(2):165-183.

[Crainic et al. 1990] Crainic, T.G., Florian, M., and Leal, J.-E. (1990). A model for the strategic planning of national freight transportation by rail. *Transportation science* 24(1):1-24.

[Ferreira et al. 1994] Ferreira, L. and Sigut, J. (1994). Intermodal road/rail container transfer facilities: Performance and planning issues. *Proceedings of the 16th ARRB conference*, part 7.

[Gorman 1998a] Gorman, M.F. (1998). Santa Fe railway uses an operating plan model to improve its service design. *Interfaces* 28:1-12.

[Gorman 1998b] Gorman, M.F. (1998). An application of genetic and tabu searches to the freight railroad operating plan problem. *Annals of operations research* 78:51-69.

[Hagani 1989] Hagani, A.E. (1989). Formulation and solution of a combined train routing and makeup, and empty car distribution model. *Transportation Research* 23B(6):433-452.

[He et al. 2003] He, S., Song, R., and Chaudhry, S.S. (2003). An integrated dispatching model for rail yards operations. *Computers & operations research* 30:939-966.

[Huntley et al. 1995] Huntley, C.L., Brown, D.E., Sappington, D.E. (1995). Freight routing and scheduling at CSX transportation. *Interfaces* 25:58-71.

[Keaton 1989] Keaton, M.H. (1989). Designing optimal railroad operating plans: lagrangian relaxation and heuristic approaches. *Transportation Research* 23B(6):415-431.

[Links 2005a] Link to descriptions of train paths (train canals).http://www.db.de/site/bahn/en/business/track__infrastructure/international__trainpaths/international__trainpaths.html

[Links 2005b] Link to a map of the corridors adapted to using train paths (train canals). http://www.railneteurope.com/cont/corridor.aspx

[Macharis et al. 2004] Macharis, C. and Bontekoning, Y.M. (2004). Opportunities for OR in intermodal freight transport research: A review. *European journal of operational research* 153:400-416.

[Marin et al. 1996] Marin, A. and Salmeron, J. (1996). Tactical design of rail freight networks. Part 1: Exact and heuristic methods. *European journal of operational research* 90:26-44.

[Nierat 1997] Nierat, P. (1997) Market area of rail-truck terminals: Pertinence of the spatial theory. *Transportation Research* 31A(2):109-127.

[Newman et al. 2000] Newman, A.M. and Yano, C.A. (2000). Scheduling direct and indirect trains and containers in an intermodal setting. *Transportation science* 34(3):256-270.

[Nozick et al. 1997] Nozick, L.K. and Morlok, E.K. (1997). A model for medium-term operations planning in intermodal rail-truck service. *Transportation Research*, Part A 31(2):91-107.

[Rizzoli et al. 2002] Rizzoli, A.E., Fornara, N., and Gambardella, L.M. (2002). A simulation tool for combined rail/road transport in intermodal terminals. *Mathematics and computers in simulation* 59:57-71.

[Trip et al. 2002] Trip, J.J. and Bontekoning, Y. (2002). Integration of small freight flows in the intermodal transport system. *Journal of transport geography* 10:221-229.

[Turnquist et al. 1982] Turnquist, M.A. and Daskin, M.S. (1982). Queuing models of classification and connection delay in railways. *Transportation science* 16(2):207-230.

[UIRR 1995] Focus on combined transport (1995). Published by U.I.R.R. s.c., International Union of rail-road transport companies. Av. du Port 100 bte 3., B-1000 Brussels.

[White paper 1996] European Union – White Paper (1996). A strategy for revitalising the Community's railways.

[White Paper 2001] European Union - White Paper (2001). European transport policy for 2010: Time to decide. Office for official publications of the European Communities, Luxembourg.

[Yan et al. 1995] Yan, S., Bernstein, D., and Sheffi, Y. (1995). Intermodal pricing using network flow techniques. *Transportation research* 29B(3):171-180.

[Yano et al. 2001] Yano, C.A. and Newman, A.M. (2001). Scheduling trains and containers with due dates and dynamic arrivals. *Transportation science* 35(2):181-191

Appendix A – Sets, parameters, and variables used in the model

Overview of sets :

$\mathcal{Z}$	*set of customer zones*	$\mathcal{S}$	*set of terminals*
$\mathcal{N}_z$	*set of customer zone nodes for customer zone z*	$\mathcal{N}_s$	*set of terminal nodes for terminal s*
$\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z$	*all customer zone nodes*	$\bigcup_{s\in\mathcal{S}}\mathcal{N}_s$	*all terminal nodes*
$n_z^-(k)$	*preceding neighbour customer zone node*	$n_s^-(k)$	*preceding neighbour terminal node*
$n_z^+(k)$	*succeeding neighbour customer zone node*	$n_s^+(k)$	*succeeding neighbour terminal node*
$\mathcal{M}$	*set of modes*	$\mathcal{D}^+$	*set of drayage arcs going from customer zone nodes to terminal nodes*
$\mathcal{L}_m$	*set of train canal arcs for mode m*	$\mathcal{D}^-$	*set of drayage arcs going from terminal nodes to customer zone nodes*
$\mathcal{N}_{s,m}^-(i)$	*inward terminal nodes for terminal node i of train canal arcs using mode m*	$\mathcal{N}_z^-(i)$	*inward customer zone nodes for terminal node i reached by drayage arc*
$\mathcal{N}_{s,m}^+(i)$	*outward terminal nodes for terminal node i of train canal arcs using mode m*	$\mathcal{N}_z^+(i)$	*outward customer zone nodes for terminal node i reached by drayage arc*
$\mathcal{P}$	*set of commoditie s*	$\mathcal{N}_s^+(k)$	*outward terminal nodes for customer zone node k reached by drayage arc*

Overview of parameters :

b *:unit value of time cost* T *:time horizon*

d_p *:demand of commodity p*

$$a_{k,p}=\begin{cases} d_p \text{ if } k=o(p)\\ 0 \text{ else}\end{cases}$$ *demand of commdity p originating from customer zone node k*

$$a_{z,p}=\begin{cases} d_p \text{ if } z=d(p)\\ 0 \text{ else}\end{cases}$$ *demand of commodity p destined to customer zone z*

t_i^α *:start time of terminal node i* t_i^β *:end time of terminal node i* t_k^γ *:time of customer zone node k*

$$t_k^\psi=\begin{cases} t^\gamma_{n_z^+(k)}-t_k^\gamma, \text{if } t^\gamma_{n_z^+(k)}\ge t_k^\gamma\\ (T-t_k^\gamma)+t^\gamma_{n_z^+(k)}, \text{if } t^\gamma_{n_z^+(k)}< t_k^\gamma\end{cases}$$ *transit time of holding arc from customer zone node k*

$$t_i^\omega=\begin{cases} t^\alpha_{n_s^+(i)}-t_i^\alpha, \text{if } t^\alpha_{n_s^+(i)}\ge t_i^\alpha\\ (T-t_i^\alpha)+t^\alpha_{n_s^+(i)}, \text{if } t^\alpha_{n_s^+(i)}< t_i^\alpha\end{cases}$$ *transit time of inventory arc from terminal node i*

$$t_{ij}^{\lambda} = \begin{cases} t_j^{\alpha} - t_i^{\beta} + (t_i^{\beta} - t_i^{\alpha}) = t_j^{\alpha} - t_i^{\alpha}, if\ t_j^{\alpha} \geq t_i^{\alpha} \\ (T - t_i^{\alpha}) + t_j^{\alpha}, if\ t_j^{\alpha} < t_i^{\alpha} \end{cases}$$ *transit time of train canal arc (i, j)*

$$t_{ki}^{\delta+} = \begin{cases} t_i^{\alpha} - t_k^{\gamma}\ if\ t_i^{\alpha} \geq t_k^{\gamma} \\ t_i^{\alpha} + (T - t_k^{\gamma})\ if\ t_i^{\alpha} < t_k^{\gamma} \end{cases}$$ *transit time for drayage arc (k,i) going from customer zone node k to terminal node k*

$$t_{ik}^{\delta-} = \begin{cases} t_k^{\gamma} - t_i^{\beta} + (t_i^{\beta} - t_i^{\alpha}) = t_k^{\gamma} - t_i^{\alpha}\ if\ t_k^{\gamma} \geq t_i^{\alpha} \\ t_k^{\gamma} + (T - t_i^{\alpha})\ if\ t_k^{\gamma} < t_i^{\alpha} \end{cases}$$ *transit time for drayage arc (i,k) going from terminal node k to customer zone node k*

$c_k^{\psi} = b \cdot t_k^{\psi}$: *unit cost of holding arc from customer zone node k*

$c_i^{\omega} = c_i^{e} + b \cdot t_i^{\omega}$: *unit invenotry cost arc from terminal node i*

c_i^{e} :*unit storage cost for terminal node i*

c_i^{u} :*loading/unloading cost in terminal node i*

c_i^{τ} :*unit vehicle transfer cost in terminal node i*

$c_i^{\mu} = c_i^{u}$: *unit cost for unloading arc in terminal node i*

$c_i^{\nu} = c_i^{u} - c_i^{\tau}$: *unit cost for loading arc in terminal node i*

$f_{i,j,m}^{\lambda}$: *fixed cost of routing a train on train canalarc (i, j) of mode m*

$c_{ij}^{\lambda} = c_i^{\tau} + b \cdot t_{ij}^{\lambda}$:*unit cost for train canal arc (i, j)*

c_{ki}^{d+} :*unit transportaion cost from node k to i*

c_{ik}^{d-} :*unit transportaion cost from node i to k*

$c_{ki}^{\delta+} = c_{ki}^{d+} + b \cdot t_{ki}^{\delta+}$: *unit cost of using drayage arcs from customer zone node k to terminal node i*

$c_{ik}^{\delta-} = c_{ik}^{d-} + c_i^{\tau} + b \cdot t_{ik}^{\delta-}$: *unit cost of using drayage arcs from terminal node i to customer zone node k*

w_i : *inventory capacity limit of terminal node i*

u_i : *handling capacity limit of terminal node i*

v_i : *train capacity limit of terminal node i*

q_m^{λ} : *capacity of mode m*

Appendix B – Mathematical model

$$\min\ z = \sum_{(i,j)\in\mathcal{L}_m}\sum_{m\in\mathcal{M}} f^{\lambda}_{i,j,m}y^{\lambda}_{i,j,m} + \sum_{(i,j)\in\mathcal{L}_m}\sum_{m\in\mathcal{M}}\sum_{p\in\mathcal{P}} c^{\lambda}_{i,j}x^{\lambda}_{i,j,p,m} +$$

$$\sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\omega}_i x^{\omega}_{i,p} + \sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\mu}_i x^{\mu}_{i,p} + \sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s}\sum_{p\in\mathcal{P}} c^{\nu}_i x^{\nu}_{i,p} + \qquad (1)\ \textit{obejctive function}$$

$$\sum_{(k,i)\in\mathcal{D}^+}\sum_{p\in\mathcal{P}} c^{\delta+}_{k,i}x^{\delta+}_{k,i,p} + \sum_{(i,k)\in\mathcal{D}^-}\sum_{p\in\mathcal{P}} c^{\delta-}_{i,k}x^{\delta-}_{i,k,p} + \sum_{k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z}\sum_{p\in\mathcal{P}} c^{\psi}_k x^{\psi}_{k,p}$$

$$x^{\psi}_{n^-_z(k),p} + a_{k,p} = x^{\psi}_{k,p} + \sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^+_s(k)} x^{\delta+}_{k,i,p} \qquad \forall k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z, \forall p\in\mathcal{P} \qquad (2)\ \textit{departing from customer zone node flow balance constraints}$$

$$\sum_{i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^-_s(z)}\sum_{k\in\mathcal{N}_z} x^{\delta-}_{i,k,p} = a_{z,p} \qquad \forall z\in\mathcal{Z}, \forall p\in\mathcal{P} \qquad (3)\ \textit{arriving to customer zone flow balance constraints}$$

$$\sum_{m\in\mathcal{M}}\sum_{h\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^+_{s,m}(i)} x^{\lambda}_{h,i,p,m} + \sum_{k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}^+_z(i)} x^{\delta+}_{k,i,p} + x^{\mu}_{i,p} =$$

$$\sum_{l\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}^-_z(i)} x^{\delta-}_{i,l,p} + x^{\nu}_{i,p} + \sum_{m\in\mathcal{M}}\sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^-_{s,m}(i)} x^{\lambda}_{i,j,p,m} \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s, \forall p\in\mathcal{P} \qquad (4)\ \textit{terminal node flow balance constraints}$$

$$x^{\omega}_{n^+_s(i),i,p} + x^{\mu}_{i,p} = x^{\omega}_{i,p} + x^{\nu}_{i,p} \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s, \forall p\in\mathcal{P} \qquad (5)\ \textit{inventory flow balance constraints}$$

$$x^{\nu}_{i,p} + \sum_{m\in\mathcal{M}}\sum_{p\in\mathcal{P}}\sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^-_{s,m}(i)} x^{\lambda}_{i,j,p,m} + \sum_{p\in\mathcal{P}}\sum_{k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}^-_z(i)} x^{\delta}_{i,k,p} \le u_i \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (6)\ \textit{terminal handling capacity constraints}$$

$$\sum_{p\in\mathcal{P}} x^{\omega}_{ip} \le w_i \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (7)\ \textit{inventory capacity constraints}$$

$$\sum_{m\in\mathcal{M}}\sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^-_{s,m}(i)} y^{\lambda}_{i,j,m} \le v_i \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s \qquad (8)\ \textit{train number capacity constraints}$$

$$\sum_{h\in\mathcal{N}^+_{s,m}(i)} y^{\lambda}_{h,i,m} = \sum_{j\in\bigcup_{s\in\mathcal{S}}\mathcal{N}^-_{s,m}(i)} y^{\lambda}_{i,j,m} \qquad \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s, \forall m\in\mathcal{M} \qquad (9)\ \textit{train flow balance constraints}$$

$$\sum_{p\in\mathcal{P}} x^{\lambda}_{i,j,p,m} \le q_m y^{\lambda}_{i,j,m} \qquad \forall (i,j)\in\bigcup_{m\in\mathcal{M}}\mathcal{L}_m \qquad (10)\ \textit{train capacity constraints}$$

$$y^{\lambda}_{i,j,m}\in\{0,1\} \qquad \forall (i,j)\in\bigcup_{m\in\mathcal{M}}\mathcal{L}_m, \forall m\in\mathcal{M} \qquad (11)\ \textit{binary constraints}$$

$$\left.\begin{array}{l} x^{\psi}_{i,p} \\ x^{\mu}_{i,p} \\ x^{\nu}_{i,p} \\ x^{\omega}_{i,p} \\ x^{\delta-}_{i,k,p} \\ x^{\delta+}_{i,k,p} \\ x^{\lambda}_{i,j,m,p} \end{array}\right\} \ge 0 \qquad \begin{array}{l} \forall (i,j)\in\bigcup_{m\in\mathcal{M}}\mathcal{L}_m, \forall m\in\mathcal{M}, \\ \forall i\in\bigcup_{s\in\mathcal{S}}\mathcal{N}_s, k\in\bigcup_{z\in\mathcal{Z}}\mathcal{N}_z, \\ \forall p\in\mathcal{P} \end{array} \qquad (12)\ \textit{non-negativity contraints}$$