Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.
Simultaneous Planning of Container and Vehicle-Routes Using Model Predictive Control*

Rie B. Larsen, Bilge Atasoy and Rudy R. Negenborn

Abstract—When containers are transported on a-modal bookings, the transport supplier can decide which combination of trucks, trains, ships, etc. to use. This gives the flexibility to transport suppliers to route the containers in accordance with the current state of the synchromodal transport network. At the same time, it enables the transport providers to route their vehicles in real time based on the current need for transportation. The interdependency of the routes of containers and of vehicles has not yet been discussed explicitly in the synchromodal literature. The aim of this paper is thus to illustrate the effect of planning the routes of containers and trucks as one integrated problem. This is addressed with a model predictive control planning method. Simulation experiments of a synchronmodal hinterland network are used to illustrate the method’s potential.

I. INTRODUCTION

The longer it takes from the moment a plan is made until it is implemented, the larger is the risk that something unexpected will happen. In container transport this unexpected event could be extreme weather delaying a barge, or an extra control check by customs delaying a container. In traditional container transport, such events are handled manually, hence making direct truck transport the easiest mode to use. Truck transport is however often the least environmentally friendly and the most man-hour consuming mode of transport. From a societal and economical perspective it is therefore desirable to use other modes of transport such as rail and water instead. Multi-modal, intermodal and synchronmodal transport, as well as the physical internet, supply chain logistics, etc. are all concepts that enable such a shift away from simplistic solutions and towards overall efficient solutions.

The shift towards an overall efficient approach creates new challenges on both the strategic, network design level, on the tactic, flow scheduling level and on the operational, specific movements level. For synchronmodal transport it can be argued that the time-horizon of decisions taken on the tactical level becomes closer to the time-horizon of decisions on the operational level [1], when the flows and services can be re-planned based on online information. A key enabler for this change is the concept of a-modal bookings where the service of transport is bought instead of a specific transport order, and allows the supplier to change this decision during the execution of the transport.

In the literature, it is already well established that the freedom to take last minute decisions is important for achieving the positive effects of synchronmodality, together with the ability to make efficient plans. Smart planning, disruption handling, dynamic switching, and demand aggregation are in [2] identified to be the four categories of necessary action to obtain synchronmodality. Real-time switching and integrated planning are also in the literature review [3] found to be among the 8 most important properties of synchronmodality.

To utilize the ability to make last minute decisions, several methods for routing containers from origin to destination in synchronmodal networks have been proposed. A framework to find the $k$ shortest paths for routing containers through a synchronmodal transport network where barges and trains are departing according to a schedule is presented in [4]. This framework does not reconsider decisions on future actions automatically, but the ability to do so when disruptions occur is discussed. In [5], last minute decisions are used to route commodity flows online over a network with scheduled barge and train services, assuming truck capacity is infinite and instantly available. An optimization-based receding horizon approach is used, which is shown to outperform a greedy approach for different prediction accuracy levels. In [6] a similar problem is treated by learning a preferred policy with Approximate Dynamic Programming. In [7], a very comprehensive overview of the Operations Research planning models used in multi-., inter-, and synchro-modal transport can be found.

The research on dynamic transport planning tackles the problem of container transport from the vehicle owner’s perspective. Here the planning approaches try to accommodate future events by, e.g., optimizing over different scenarios as in [8]. In [9] an approximate dynamic programming method that incorporates probabilistic knowledge of future events is discussed. Another approach is to assume an accurate plan of the future demand is available when, e.g., the sailing schedules and truck flows are planned, knowing that if the realized future demand exceeds the planned capacity, an expensive alternative can be used ad hoc, see, e.g., [10] and [11]. A general overview of the dynamic vehicle routing problem literature can be found in [12]. Most papers in this category do not relate themselves to inter- or synchro-modal transport. Some do however accommodate transshipments in their models, e.g., [13] and [14], and cover thus some of the challenges of intermodal transport planning.

It is however seldom discussed that synchronmodal, a-
modal, bookings give the transport supplier the opportunity to reroute and reschedule containers and vehicles simultaneously. The authors of [15] state that “the flexibility in transportation routes may be used in conjunction with the operational fleet deployment problem. This creates new and more complex optimisation challenges”, but the statement is not explored further. A planning model that besides routing containers also decides if a specific service is operated or not is presented in [8]. The services are however not routed, which, e.g., for a scenario with more import than export will lead to overcapacity of empty vehicles on the import side.

To address the challenge of integrating container and truck planning in synchromodal transport networks, this paper proposes the use of model predictive control (MPC). For a multi-commodity, synchromodal transport network considering multi-type trucks as well as scheduled trains and barges, the method is used to re-optimize the integrated plan online. This paper is organized as follows. In Section II the transport network model is introduced. Section III presents the control algorithm used for the simultaneous, real-time planning. This control method is compared in Section IV to a real-time container routing method, which is presented in the same section. Finally in Section V the conclusions and directions of future research are discussed.

II. MODEL DESCRIPTION

The transport network is modelled as a continuous state, discrete time, state-space commodity flow model of a hinterland network. Vehicles and containers are modelled on separate networks that are coupled by the constraint that containers can only flow on a directed arc if there is at least the same number of trucks flowing on the same arc. If one of the nodes is a train or barge node (a scheduled service), no trucks are required. The main features of the model are:

- Demand is modelled as containers available to the network and needed from the network. Unsatisfied demand is penalized. The demand is regarded fully known over the planning horizon.
- Commodity flows are considered to be continuous variables. This simplifies the model and can capture the desired level of accuracy, see [16].
- Unscheduled vehicles, with trucks as example, are also modelled as continuous variable flow. This again allows for balance between model complexity and accuracy.
- Each scheduled service is modelled separately. Two trains serving the same route are modelled as two nodes.
- A limited number of containers can be trucked to and from a node and a limited number can be loaded to and from the scheduled services at any given time.
- Terminal operating hours, truck drivers resting hours, etc., are not considered.

The model supports multi-commodity flows for both import and export. The demand profiles at the destinations are created based on time windows for each single container, but as commodity flows are considered, one container of a certain commodity can replace another. This assumption is also used in, e.g., [5]. Trucks are modelled in the same fashion as containers, allowing to distinguish different kinds of vehicles. Each truck network includes a free parking node that represents the trucks that are not being used in the network but are available to the network.

It is assumed that when a travel is started, it is also fulfilled. In other words, that no decisions can be taken when a truck or container is on an arc. The scheduled services (e.g., barge and train) are modelled as nodes with time dependent arc capacities that corresponds to the timetable of the respective connection. When the scheduled service is at a terminal, it has a predetermined time slot to unload and hereafter a predetermined time slot to load before it departs according to schedule.

The mathematical description of the transport network is kept general, while the specifications of the network used as example can be found in Section IV and in Figure 1. The state \( x_i \) of each node \( i \in \mathcal{N} \) in the system at every time step \( k \) is given by:

\[
    x_i(k) = \begin{bmatrix}
        x_i^c(k) \\
        u_i^v(k) \\
        v_i^v(k)
    \end{bmatrix},
\]

where \( x_i^c(k) \in \mathbb{R}_{\geq 0}^{n_c} \) is the quantity of the \( n_c \) different commodities stacked at node \( i \) and \( x_i^v(k) \in \mathbb{R}_{\geq 0}^{n_v} \) is the quantity of the \( n_v \) different vehicle types parked at node \( i \). The vector \( u_i^v(k) \) is the amount of containers that are on the way to node \( i \) by truck at time step \( k \). It is necessary to keep a record of the containers that are on the way to node \( i \) but have not yet arrived, since each arc in the truck network is associated with a travel time \( \tau_{ji} \) that acts as a delay. Formally, \( u_i^v(k) = [u_{ji}(k-1) \ldots u_{ji}(k-\tau_{ji}) \ldots u_{ji}(k-1)]^T \ldots [v_{ji}(k-\tau_{ji}) \ldots v_{ji}(k)]^T, \{j \cdots j'\} = \mathcal{T}_i \) where \( u_{ji}(k) \in \mathbb{R}_{\geq 0}^{n_i} \) is the volume of containers that leave node \( j \) at time step \( k \) on the arc to node \( i \). The set \( \mathcal{T}_i \) contains all nodes with a truck connection to node \( i \). Likewise, \( v_i^v(k) = [v_{ji}(k-1) \ldots v_{ji}(k-\tau_{ji}) \ldots v_{ji}(k-1)]^T \ldots [v_{ji}(k-\tau_{ji}) \ldots v_{ji}(k)]^T, \{j \cdots j'\} = \mathcal{T}_i \) is the amount of trucks that are on the way to node \( i \) at time step \( k \). Here, \( v_{ji}(k) \in \mathbb{R}_{\geq 0}^{n_i} \) is the amount of trucks that leave node \( j \) towards \( i \) at time step \( k \).

The demand is modelled as virtual destination nodes \( d \in \mathcal{D} \) with arcs to their adjacent network nodes that...
have unlimited capacity and zero travel time. This way, containers with a given destination node can arrive at the
destination location (the adjacent network node) before it is
needed without extra cost, other than the storage cost and
capacity limitations. The unsatisfied demand (both available
and needed containers) at the virtual destination nodes is
penalized. We say that node $i$ has outgoing demand when $i$
is the origin of the commodity and that node $i$ has incoming
demand when $i$ is the destination. The virtual destination
nodes have different dynamics than the the nodes in the
network, namely

$$x^i_d(k + 1) = x^i_d(k) - u_{di}(k) - u_{id}(k) + d_i(k),$$  \hspace{1cm} (2)$$

where $x^i_d(k) \in \mathbb{R}_{\geq 0}^n$ is the amount of incoming and outgoing
demand at time step $k$. Both incoming and outgoing demand are
modelled as positive values, since the commodities are
defined based on destination. The variable $u_{id}(k) \in \mathbb{R}_{\geq 0}^n$
is the containers that were available at network node $i$ that
are used to satisfy the incoming demand at time step $k$, and likewise, $u_{di}(k) \in \mathbb{R}_{\geq 0}^n$ is the containers used
to satisfy the outgoing demand. Demand satisfaction can be
postponed (hence the integral dynamics), and the new
demands $d_i(k) \in \mathbb{R}_{\geq 0}^n$, that can be satisfied from time
step $k$, act as disturbances to the system and are thus not
controllable.

The remaining nodes in the network are described as in (1)
and have the same dynamics. For describing the dynamics
three sets are defined for each node $i$: $T_i$ as introduced earlier,
$S_i$ and $D_i$. The set $S_i$ contains all nodes to which $i$ is linked
via a time-dependent arc connection. If node $i$ is a scheduled
service, $S_i$ contains the terminals it serves, and if node $i$ is a
terminal, $S_i$ contains the scheduled services that depart from
here. Notice that if $i$ is a scheduled service $T_i = \emptyset$. Likewise
$D_i$ contains the adjacent destination node for node $i$.
This set contains maximum one element. The dynamics of $x^i_v(k)$
is

$$x^i_v(k + 1) = x^i_v(k) + \sum_{j \in T_i} (u_{jji}(k) - \tau_{jji} - u_{ijj}(k))$$

$$+ \sum_{s \in S_i} (u_{isi}(k) - u_{is}(k)) + \sum_{d \in D_i} (u_{di}(k) - u_{id}(k)), \hspace{1cm} (3)$$

where the control action $u_{is}(k) \in \mathbb{R}_{\geq 0}^n$ is the containers
moved from node $i$ over a time-dependent connection to node
$s$. If node $i$ is a barge, $u_{is}(k)$ is unloading containers at
terminal $s$. $u_{is}(k) \in \mathbb{R}_{\geq 0}^n$ is the reverse movement.

As there are no scheduled services nor demand in the truck
network the dynamics hereof is given by:

$$x^i_v(k + 1) = x^i_v(k) + \sum_{j \in T_i} (u_{jji}(k) - \tau_{jji} - u_{ijj}(k)).$$  \hspace{1cm} (4)$$

The two networks are connected by the constraint that
containers cannot be moved without a truck if they are
transported on a truck-arc.

$$1_n u_{ijj}(k) \leq 1_n v_{ijj}(k) \forall j \in T_i.$$  \hspace{1cm} (5)$$

The bold $1_a = \{1\}^a$ is a row vector of size $a$ with all ones.

The network is furthermore constrained by capacities:

$$1_n x^i_v(k) \leq c^i_v$$ \hspace{1cm} (6)

$$x^i_v(k) \leq c^i_v$$ \hspace{1cm} (7)

$$1_n \sum_{j \in T_i} (u_{jji}(k - \tau_{jji}) + u_{ijj}(k)) \leq c^i_m$$ \hspace{1cm} (8)

$$1_n u_{isi}(k) \leq c_{si}(k), \ s \in S_i$$ \hspace{1cm} (9)

$$1_n u_{ isi}(k) \leq c_{is}(k), \ s \in S_i, \hspace{1cm} (10)$$

where, at location $i$, the scalar $c^i_v$ is the maximum number of
containers that can be stored, $c^i_v \in \mathbb{R}_{\geq 0}^n$ is the maximum
number of vehicles of each kind that can be parked ((7) is
to be satisfied element wise). The scalar $c^i_m$ is the maximum
number of containers that can be moved to and from trucks
within one time step at location $i$. The schedules of the barge
and train connections are implemented by the time varying
crane speeds $c_{si}(k)$ and $c_{is}(k)$. To illustrate, assume $i$ is a
barge and $s$ is a terminal. When the barge is at the terminal
and can be unloaded $c_{si}(k) = 0$ and $c_{is}(k) \neq 0$, and when
the barge can be loaded $c_{si}(k) \neq 0$ and $c_{is}(k) = 0$, otherwise
$c_{si}(k) = 0$ and $c_{is}(k) = 0$.

### III. Proposed Control Method

To achieve an efficient execution of container transport
and truck routing that can adapt to delays online, a convex
MPC is proposed. The control variables are, for all $i \in \mathcal{N}$,
the amount of departing trucks and the containers they bring,$u_{ijj}(k), \forall j \in T_i$ and $u_{isi}(k), \forall j \in T_i$, the quantity to load
and unload for scheduled services, $u_{isi}(k), \forall j \in S_i$, and the
amount of demand to satisfy $u_{idi}(k)$ and $u_{id}(k), d \in D_i$.

It is assumed that the controller has an accurate model
for the dynamics of the transport system, including time-
invariant travel times and access to accurate information of
the state of the global system every $\Delta T$ minute. Furthermore,
a prediction of the future demand is assumed available to the
controller. At each time $t = i \Delta T$, $i \in \mathbb{N}$ the controller gets
up to date information and uses it to find the sequence of
decisions that will minimize a cost function over a prediction
horizon $T_p$. Only the decisions that require an action at this
time step $t = i \Delta T$ are implemented, and when $t = (i + 1) \Delta T$, the process starts over.

To evaluate what the best sequence of decisions is, the
MPC controller solves optimization problem (11), where the
measured state (1) for node $i$ at time $t$ is denoted by $\tilde{x}_i(t)$.
The decision vector $U$ contains all inputs $u_{ijj}(k)$, $v_{ijj}(k)$,
$u_{idi}(k)$ and $u_{id}(k)$ for all $i \in \mathcal{N}$ and $k \in [0, T_p - 1]$. The
time-invariant weight $M^p_i$ is the cost of storing a container
at node $i$, while $M^v_i$ is the cost of parking a truck. $M^s_{ij}$ is
the cost of a truck journey from $i$ to $j$ and $M^c_{ij}$ is the cost
associated with moving a container from a stack to a truck
or vice versa. Moving a container to or from a scheduled
service has the cost $M^s_i$, which is only paid at the terminals,
transport by scheduled service is paid per container per time
step as the container storage cost $M^c_i$. The cost of unfulfilled
demand is a quadratic term scaled by $M^s_i$, which lets small
delays be significantly cheaper than large delays.
Typically, MPC ensures recursive feasibility of the optimization problem and stability of the controlled system by special constraints and costs at the end of the prediction horizon [17]. The synchronomodal transport system described in this paper is inherently marginally stable and recursively feasible, but as the actions taken within the prediction horizon will effect the state of the system in the future and thus the long-term (infinity) cost, considerations regarding the two concepts are important. The methods to address these challenges often impose conservatism that will cause underutilization of the scheduled services in the synchronomodal transport system, see e.g., [18]. A way to address the long-term cost of the MPC problems, when no formulation of the expected infinity costs and constraints exist, is to use a long prediction horizon, see, e.g., [19] or [20]. The current literature on this assume different symmetric cost functions around a reference point (here the global zeros-state) that lies in the interior of the feasible set. If the transport cost is formulated based on absolute numbers and the reference point is set to be a vector of very small positive numbers instead of a zero vector, then the assumptions hold and only the time-varying constraints prevent a calculation of the necessary length of the prediction horizon. In the proposed method the prediction horizon is chosen long enough to cover the longest travel time in the network (both for trucks and scheduled vehicles). To ensure the controller sees the consequences of its decisions, only trucks that will arrive within the prediction horizon are allowed to depart (13). A way to decrease the size of the optimization problem is described in [21], where the time periods related to each MPC time step \( k \) is increased for increasing \( k \).

IV. EXPERIMENTS

To evaluate the potential benefits of simultaneous routing of containers and trucks, simulation experiments of a hinterland transport scenario have been carried out. Experiments are first done with the planning method presented in Section III and hereafter with a benchmark planning method that considers truck capacity infinite and instantly available.

The benchmark planning method is based on the same MPC problem constraints but with cost function (15), where the travel-cost \( M^c_{ij} \) is per container instead of per truck, and the parking cost \( M^p_{i}x^p_i = 0 \). Furthermore to discourage empty movements of trucks, the handling cost \( M^h_i \) is per departing truck instead of per departing container. The simulation scenario was furthermore altered so that a sufficient number of trucks are available at all locations. In the scenario used with the proposed planning method, all trucks were initially at their respective free parking locations (node 8 and 9).
\[
\min_{U} \sum_{k=0}^{T_p} \left( \sum_{i \in N} \left( M^c_i x^c_i + \sum_{j \in T_i} (M^f_{ij} u_{ij}(k) + M^l_{ij} v_{ij}(k)) + \sum_{s \in S_i \cap N_i} (M^e_i (u_{si}(k) + u_{sl}(k))) \right) + \sum_{d \in D} (x^d_i(k))^T M^d x^d_i(k) \right)
\]

s.t. \quad (12), (13), (14)

Fig. 2. Demand profile at the three virtual destination nodes. The quantity of new demand \( d_i(k) \) is shown over time steps. Each color represents the corresponding commodity’s portion of the total demand. Outgoing demand is shown as positive and the incoming demand as negative.

The demand profiles for the virtual demand nodes 3, 5, and 10 were generated based on individual transport orders with a transportation time of minimum 40 time steps. In the scenario, significantly more containers are imported (destination 3 and 5) than exported (destination 10/ship). This scenario is chosen since it frequently occurs when transport of full and empty containers are considered as decoupled problems, as in e.g. [22]. The profile also emphasizes the difference between the proposed and the benchmark methods. It is assumed that the controller has access to an accurate demand prediction for the prediction horizon. The profiles can be seen in Figure 2. The simulations start empty with no containers at any stack or vehicle.

The simulation of 600 iterations was performed in Matlab, solved in Mosek and compiled with Yalmip’s optimizer [23]. Including initialization the computations took 30 to 31 minutes for both the proposed and the benchmark method. Since the MPC reoptimizes every 15 minutes, this is well within the computational budget.

B. Results

In Figure 3, the total number of vehicles driving in the network is presented over time. In the left figure, the port-area vehicles are shown (yellow network in Figure 1), while the right shows the vehicles in the hinterland network (green network in Figure 1). The shaded area is the used vehicles for each time step the y-axis shows the number of vehicles that has departed a node but has not yet reached the next node. The shaded area is the total number of vehicles for the proposed method with the light part being the portion of empty vehicles. From the benchmark method only information on vehicles that transport containers exist, this is shown with a red line.

Fig. 3. Number of driving vehicles \( \sum_{i \in N} \sum_{j \in N} \sum_{l=0}^{T_v-1} v_{ij}(k-l) \), i.e., for each time step the y-axis shows the number of vehicles that has departed a node but has not yet reached the next node. The shaded area is the total number of vehicles for the proposed method with the light part being the portion of empty vehicles. From the benchmark method only information on vehicles that transport containers exist, this is shown with a red line.

In Figure 4, the occupation at the barge and the train. The colors indicate the four different commodities. Top: proposed method. Bottom: benchmark method.

Fig. 4. The occupation at the barge and the train. The colors indicate the four different commodities. Top: proposed method. Bottom: benchmark method.

empty trucks, and thus only the full vehicles are shown. In the port area, it is clear that the periodic arrival of ships with demand to satisfy creates peaks in the number of driving trucks. These peaks are higher for the benchmark method, where the maximum number of port vehicles used at a time step was 31.2 while it was 11.7 for the proposed method. The total number of port vehicles driving with a container for a time-step was 1279 for the benchmark method and 1156 for the proposed method. One could argue that since the benchmark method has no parking fee \( M^p_i = 0 \cdot 1_{n_i} \), the difference between using a truck and parking a truck is larger in the benchmark method. To ensure the results are not due to this effect, a second experiment was made with the proposed
method, but with $M_{ij}^t = \tau_{ij} [24]$. The results obtained with this method follow the same trends as the results for the proposed method, but shows larger peaks of up to 21.9 port vehicles per time step. This is still significantly less than the benchmark method, and it can thus be concluded that the proposed method provides a more even usage of trucks than the benchmark method.

In the hinterland network, more vehicles are used with the proposed method than the benchmark method. This is mainly because the import is larger than the barge capacity and significantly larger than the export. In the lower row in Figure 4, it can be seen that the benchmark method fills the barge and train whenever it is time-wise reasonable. After $t = 137\Delta T$ the barge arrives at node 4 after the ship has departed, so due to the high penalty on unsatisfied demand, no export is transported by barge. In the top row, the corresponding figures for the proposed method are seen. The largest difference between the results is the transport of export by train. In the proposed method, the trucks have to drive from the hinterland to the port area in order to transport the import that exceeds the barge and train capacity. Since the difference between driving an empty and a full vehicle is the loading fee $M_i^t = 3 \cdot 1_{\text{unit}}$, this is a cheaper alternative than the spot price on the barge and train. This results in 29910 hinterland vehicles driving a time step with a container with the proposed method. For the benchmark method it is 22331 vehicles driving a time step. However, the maximum number of vehicles driving on the hinterland network during one time step is very similar with 80 for the proposed method and 81 for the benchmark method.

V. CONCLUSION AND FUTURE RESEARCH

The often used assumption that trucks are instantly available at any location in the synchromodal network significantly changes what the optimal actions are. The proposed method schedules trucks and containers simultaneously and successfully smooths out peaks in the needed number of trucks. In the simulated scenario, where significantly more containers are imported than exported the proposed method favours using empty truck capacity for export instead of spot-priced barge and train capacity. This is very different from the benchmark method that does not consider the cost of driving empty. Both methods are based on MPC, which has only been used for container transportation planning in a few instances, and never for planning containers and trucks simultaneously. Future research should compare results where the schedules of barge and train are not fixed, but are decided by the central control. It is furthermore important to investigate how the existing knowledge on MPC can improve the efficiency of synchromodal transport under stochastic scenarios and thus with imperfect predictions.

REFERENCES