

Optimization of Collection and Consolidation Operations in Cross-Border Multi-modal Distribution Networks

Gulcin Ermis¹ and Francesco Alesiani¹ and Konstantinos Gkiotsalitis²

Abstract—This study introduces a model to solve a dynamic network optimization model on a heterogeneous graph. We use this model to optimize the collection and consolidation operations on a cross-country multi-modal distribution network. The model's dynamic objects are trucks, trailers, orders, unvisited collection and customs check points. Information about dynamic objects is extracted from a real-time database. The model's static objects include objects that are known in advance, such as warehouses. The constraints of the problem include due dates, vehicle capacity, availability of vehicles, and precedence constraints of visiting locations. We propose a mixed-integer programming model and provide a solution using mathheuristics. We decompose the master MIP model into subproblems that can be solved to optimality with LP solvers. We also reduce the graph complexity by variable fixing due to optimized subproblems or by bounding the maximum number of paths to be selected due to the solutions of priority-based bin packing algorithms. Finally, we convert the resulting problem into a bipartite matching problem by expanding the graph nodes which can then be solved in polynomial time. We implement our solution method on real-time data retrieved from the tracking system of a third-party logistics company. Experiments show that our solution method significantly outperforms other heuristics in terms of solution quality which is measured with respect to lateness, empty kilometers traveled, travel times, number of required/used vehicles, load factors, and ratio of served orders.

Keywords: mathheuristics; multi-modal; consolidation; cross-border; logistics.

I. INTRODUCTION

We study the consolidation and collection operations in cross-border multi-modal transportation networks. We provide a centralized solution for a highly complex system that includes: (a) ad-hoc demands with due-dates, (b) a heterogeneous and dynamic structure of vehicle fleet that is simultaneously reused both in import and export directions, (c) precedence relations in collection routes due to the need to visit customs check locations after collecting goods and before reaching to warehouses or to pick up vehicle components such as trailers, (d) rail and sea port schedules, (e) value, volume and weight of products to be packed, and (f) packing operations at the consolidation hubs. This system also allows to share vehicles among different consolidation

¹NEC Laboratories Europe GmbH, Kurfürsten-Anlage 36, 69115 Heidelberg, Germany

²University of Twente, Dept. of Civil Engineering, 7522LW Enschede, The Netherlands

*Supported by the European Union's Horizon 2020 - EU.3.4. SOCIETAL CHALLENGES - Smart, Green And Integrated Transport programme under grant agreement number 769141. This publication reflects only the authors' view. The European Union is not responsible for any use that may be made of the information it contains.

centers when there is a need at one but redundancy at another one. We also consider the capacities of vehicles, traffic conditions that cause variability in travel times, time windows due to the working times of warehouses or the deadlines of deliveries, and the availability of vehicles at a certain time point. We provide a mixed-integer programming model for the centralized problem and solve it by decomposition to several linear programming models in such a way that their optimized solutions help to variable fixing in the master problem or provide input for another sub-problem. This decomposition approach reduces the complexity considerably.

In the composition phase, we first transform the hybrid graph of the master problem into a bi-partite matching network. We use the solutions from the decomposition phase as inputs and we replace the nodes with non-unit values by many nodes with unit values. We then solve the bi-partite matching problem to optimality and employ heuristics to match the outputs with the ones of a priority-based bin packing solution considering some constraints to provide a final solution. The final solution is a matching of truck, trailer and orders inside the packs before the long haul delivery starts. In conclusion, by using LP structures in the solution we provide a very short computational time including also the pre-processing of highly complex real-time data. This allows using a dynamic automated replanning approach instantly in a system of ever-changing resources, conditions and information. Besides the computational complexity improvements, our numerical experiments prove that many different key performance indicators (KPIs) are considerably improved compared to the use of baseline methods, such as classic heuristics.

II. RELATED STUDIES

Multi-modal logistics networks play an important role in reducing CO₂ emissions. Thus, rail-road-sea combinations are usually preferred over road transportation. One of the studies that take CO₂ emission costs into consideration in designing a logistics network with optimal mode combinations is the work of [1]. Zhang et al. [1] consider time windows for deliveries. In our case, time windows constraints occur because of the due dates of orders and strict time schedules at rail or sea ports. Another study on the design of a multi-modal network considering the CO₂ emissions is the work of [2]. This study showed the superiority of a multi-modal network in automobile transportation that encourages the usage of Roll-on/Roll-off (Ro-Ro) ships by solving a mixed-integer programming model resulting in reduced CO₂ emissions. Zhou et al. [3] solved the capacitated multi-

modal network flow problem that minimizes operational costs and CO₂ emissions. Although the capacitated multi-modal network flow problem is NP-Hard, they showed the efficiency in practical scales.

Rather than mode selection, our study helps to increase the use of the multi-modal route taking into consideration critical time windows caused by the strict schedules at critical hubs, like rail and sea ports. From this perspective, the work of [4] presents a closer approach to our work since their goal was to improve consolidation operations to reduce the disruptions of flow of goods on a preferred multi-modal route. Bowen et al. [4] provided a genetic algorithm to solve the NP-Hard transit consolidation problem in multi-modal logistics. Their aim was to improve the operations of unloading, loading, sorting and packing small batches of cargo with multiple origins and multiple destinations at the consolidation center. At the broader level, the problem analyzed in [4] was highly related to our problem where we have multiple sources and destinations of partially loaded vehicles which need to unload the cargo at a consolidation center to be further consolidated and distributed. While they focused on a local approach related to the operations at the consolidation center, we provide a broader level solution by solving a hybrid model including also the collection of goods or empty trailers, visiting customs check points when necessary before reaching consolidation centers, packing and matching free trucks, trailers, and packed goods at the consolidation center, and considering the due dates of order dispatches.

In cross border chains, using multi-modal transportation and consolidation hubs is crucial. Lin et al. [5] studied the impact of cross-docking hubs in Europe and inter-modal rail-based distribution in China–Europe container trades. Tyan et al. [6] evaluated the freight consolidation policies in global third party logistics. Fontaine et al. [7] incorporated inbound and outbound demands, the combination of traditional road-based carriers with modes of mass transport, consolidation centers, and the size of heterogeneous vehicles into a scheduled service network design formulation that was solved with Benders decomposition algorithm. Third party logistics companies play an important role in carriage of goods using intermodal transportation networks, handling consolidation or warehousing operations and supporting customs checks in cross-border logistics. Belieres et al. [8] proposed a matheuristic for logistics service network design problem encountered by third-party logistics providers. Serrano et al. [9] focused on distribution and operation planning in cross-docking platforms for overseas delivery consolidation in automotive industry.

Another common issue that needs to be dealt with in complex transportation networks is the uncertainty and disruptions. In the problem we analyze, ad-hoc demands occur any time without a predicted pattern due to the online orders and the large variety of commodities carried by the logistics provider. In addition, there can be disruptions in transportation flows that might stem from traffic conditions, cancellations or breakdowns in transportation means, waiting times

for customs checks, etc. We handle this problem by providing a fast dynamic re-planning algorithm that continuously tracks the position and status of vehicles, the demand changes and every other traceable input. The algorithm processes the ever-changing information and provides a solution according to the changing conditions.

In past studies, Gao et al. [10] provided a bi-objective stochastic multi modal transportation planning model to deal with uncertainty and multi-commodity re-balancing in humanitarian logistics. Mafiroh et al. [11] showed the advantage of multi-modal relief distribution as a response to a disaster where they evaluated the mode combinations, airplane–truck, sea vessels–truck, and airplane–helicopter by analyzing the data from an earthquake. Haghani et al. [12] provided heuristics to solve a complex humanitarian logistics problem which was defined as a multi-commodity, multi-modal network flow problem with time windows. Sheu et al. [13] proposed a multi-objective mixed-integer programming model to design a demand-driven centralized emergency supply network that integrates three sub-networks (shelter network, medical network, and distribution network) in order to respond to large-scale natural disasters. Zhang et al. [14] solved multiple-resource and multiple-depot emergency response problems considering secondary disasters using heuristics. Garrido et al. [15] provided a stochastic approach to optimize inventory levels for emergency supplies and availability of vehicles in order to deliver enough supplies to satisfy demands with a given probability in flood emergency. Moreno et al. [16] proposed stochastic mixed-integer programming models to solve multi-period facility location and transportation problems with re-use of vehicles in multi-commodity and multi-modal relief chains. Most of these studies in humanitarian logistics attempt to increase the level of readiness before a disaster, thus they employ stochastic or robust optimization methods. In contrast, our study contributes to the state-of-the-art by handling the sudden changes by fast reoptimization algorithms.

III. PROBLEM SPECIFICATION

We deal with a real time network optimization problem. We consider a cross-country multi-modal distribution network. In this multi-modal logistics network, the goods are carried via different transportation modes, such as block trains, Ro-Ro ships, speedy vehicles, trucks and trailers. The main hubs are the locations where the goods are handled and transferred between different modes. These hubs are: 1) rail ports, 2) sea ports, and 3) central warehouses. The partially loaded trucks oftentimes visit warehouses to consolidate their loads. Thus, these warehouses are critical hubs where collection ends and delivery starts. Once the orders are stored at the warehouses, packed into trailers and assigned to trucks, they can head to rail or sea ports or continue via road until reaching their final destination.

The problem is to optimize the collection and packing operations in this multi-modal logistics network. The input is a real-time dataset gathered from the tracking system of a third party logistics company. The output is an optimized

global planning solution, such as which trucks will pick up which trailers, which companies will be visited to collect goods, which are the selected paths to reach the warehouses, which truck-trailer pair will accommodate each packed group of orders, which orders are packed in which groups, which are the arrival times of the truck-trailer pairs to load the assigned groups of orders.

The design of a solver requires a fast dynamic re-planning approach since the status and positions of the moving trucks, the availability of vehicles at specific locations at required time windows, the traffic conditions or the demands change constantly. Every time the tracking system is queried, there occurs a new instance of this complex problem with different resources and constraints. In the distribution network, warehouses are the main hubs which are visited when the trucks are unloading goods collected from different regions before the re-packing of orders and the actual distribution starts. Before the latest delivery time is reached, reshuffled orders have to be packed into free trucks with attached trailers at these warehouses. However, only a subset of the owned vehicles can arrive at these warehouses at the required times since, at these times, they might perform other operations. This creates a waiting time for consolidated orders to be picked up by a free vehicle with an attached trailer. Only after a good match of the arrival time of a usable vehicle and the ready time of a re-shuffled pack match, considering also the due-dates of orders packed within different containers, a loaded truck can leave the warehouse for delivery to customers. Delivery can be either by road directly to customers, or to a block train station that ends up at a sea port after which loaded trucks are transferred to a Ro-Ro ship to reach their final destinations.

A summary of the features of the objects of the distribution network is provided below.

1) *Warehouses:* They are static objects and they serve as central hubs for trucks carrying orders in import and export directions. The locations, opening hours and working days of these objects are known in advance and are constant.

2) *Fleet characteristics:* Figure 1 represents fleet characteristics.

Set 1 – Free trucks with no attached trailers: These trucks are available at different geo-locations at the time of tracking.

Set 2 – Free trailers: These trailers are not attached to any truck and they are available at different geo-locations at the time of tracking. Note that a truck in set 1 has to visit the geo-location of a trailer in set 2 and pick it up before visiting any warehouse.

Set 3 – Trucks in use: These trucks have attached trailers but, at the time of tracking, they are not available since they are still collecting orders and have unvisited pick-up and customs check locations. The geo-locations of the trucks at the time of tracking are known. The final destination of a truck in this group might be a warehouse, a company, or a customs check point. Only after visiting this final destination, the truck becomes free and can be sent to a/another warehouse from its final destination. Note that a truck in set 3 has to complete its tour on unvisited pick-up

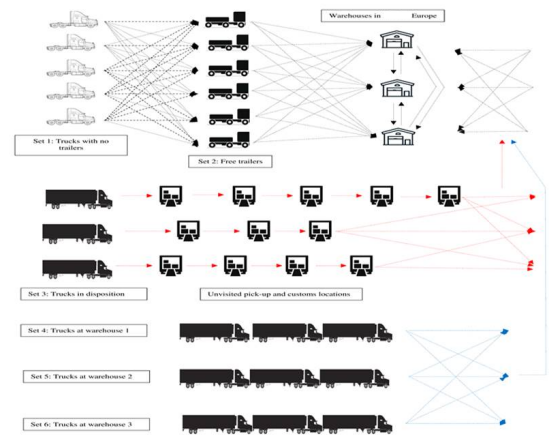


Fig. 1. Graph generation

and customs check locations and reach its final destination before being sent to warehouse.

Set 4 – Free trucks with attached trailers: At each one of the warehouses, there are available trucks with attached trailers at the time of tracking. The trucks in set 4 can be used directly at their own warehouses or redirected to other warehouses.

3) *Customs:* Customs checks play an important role in cross-border transportation networks. In our case, each truck collecting goods has to visit a customs check location before heading to its final destination that belongs to another customs union.

4) *Orders:* A real time retrieval of data from the tracking system gives a detailed information on the orders being carried around in import and export directions, waiting at depots, sea or rail ports of the cross-border multi modal distribution network. To solve the logistics problem, we extract the set of orders which have entered into warehouses by the time of tracking. The extracted orders are the ones which were collected from logistic companies before the time of tracking and they are still waiting at the warehouses by the time of tracking in order to be packed into trailers and get assigned a delivery plan to rail ports and sea ports.

IV. DESIGN CONSIDERATIONS

A comprehensive pre-processing algorithm is applied at each data retrieval from the tracking system to extract the real time instance of a dynamic graph. Using this processed information, we design a graph mostly having dynamic vertices whose positions or features change over time and which leads us to define an instance of a special transportation problem. In this graph, the current locations of moving trucks, different locations of free trailers at the time of tracking, unvisited pick up locations or customs check points of vehicles at the time of tracking are dynamic vertices which will not be at the same status after a time has passed. The status of a vehicle, such as whether it has an attached trailer or not, whether it is occupied or free, where it is or what is its task

and destination at the moment, and which is its preceding route before reaching to its final destination, is extracted by complex queries. Similar is the case for whether a trailer is free or not and where it is located, can that be matched with an available truck, ad-hoc orders, their values, amounts, due-dates, from which warehouse they have to be dispatched, have they already been delivered or still waiting, and their current status i.e., loaded, being transported on the collection or delivery route, entered a specific warehouse. A summary of specific features are given below.

A. Node generation

We generate a graph where each node represents a geo-location of an object. Any node of this graph corresponds to one of the following: (i) The real time GPS position of a free or non-free truck, (ii) Geo-location of a free trailer, (iii) Geo-location of an unvisited collection point (a company) for a non-free truck, (iv) Geo-location of an unvisited customs location (customs) for a non-free truck, (v) Geo-location of a final destination (a company or a warehouse) for a non-free truck, (vi) Geo-location of a central warehouse.

The nodes defined in (vi) are *static* since the central hubs of the logistics companies do not change over time. The nodes defined in (i) to (v) are *dynamic*. This is due to the moving structures of vehicles, i.e., GPS positions of trucks, the location where a free trailer exists, and the continuous change over time of the unvisited truck locations. Another factor affecting this is the dynamic planning. Customs locations or companies to be visited by trucks are planned differently at different times of the day or in different days based on ad-hoc demands and the fleet approaching the region at the time of planning. Figure 1 denoted a simple representation of the graph and fleet structure.

B. Graph features

The origins of the graph are the nodes representing the current GPS positions of non-free and free trucks with attached trailers and free trucks with no attached trailers. The sink nodes are the central warehouses. Precedence relations exist between the nodes representing the locations of free trucks with no attached trailers and free trailers since a truck has to pick up a trailer first before visiting a central warehouse. Similarly, the paths including the unvisited companies, customs locations and destinations of non-free trucks are connected by directed links. Thus, the graph includes both arcs and edges between the nodes.

The travel times between the nodes are obtained only when required by using an API that gets the geo-locations of any two nodes as input and returns the result by either (1) requesting HTTP responses from a website tracking the traffic and shortest routes in real time or (2) calling an an approximation function based on *haversine* distance.

A *source* is a free empty truck with an attached trailer. To have a usable source, a free truck has to pick up a trailer first or a non-free truck has to finish its collection from several companies, visit the customs check and reach to a warehouse or another destination. From that point sources

can be allocated to required areas. In our terminology, a flow is a vehicle with or without a trailer, free or non-free, that is following a path from one node to another. Demands are the numbers of free vehicles with trailers required at warehouses within specific time windows.

V. MATHEMATICAL MODEL

We provide a two-staged approach in modeling the entire system involving (1) the classical bin-packing model adapted to our case where time dependency of orders are considered, and (2) a mixed-integer programming (MIP) model that takes the solution of the bin-packing model for each warehouse as input. The latter one incorporates the features of entire system.

A. Bin packing

We first solve a bin packing problem to find the minimum required number of required truck/trailer combinations at each warehouse $w \in W$, where W is the set of warehouses.

Let $O^{|T|w}$ be the set of orders that have been delivered to warehouse, $w \in W$, within a time period $|T|$ and waiting to be consolidated and packed for delivery regarding their due-dates and values. Let p_j be the payweight of order $j \in O^{|T|w}$ and Q be the capacity of a trailer in terms of payweights. We denote the number of vehicles required at warehouse, w , within a time period, $|T|$, by $B^{|T|w}$. We aim to minimize this value for each warehouse, $w \in W$, where W is the set of warehouses. The mathematical model given in (1) to (5) is a classic bin-packing model [17] that is adjusted in line with the explained structure of this problem. In this model, y_i is the binary decision variable which is equal to 1 when trailer i is used, and x_{ij} is the binary decision variable which is equal to 1 when order j is assigned to trailer i . We minimize $B^{|T|w}$ for each warehouse, $w \in W$, by repeatedly solving the following mathematical model.

$$\min B^{|T|w} = \sum_{i=1}^{|O^{|T|w}|} y_i \quad (1)$$

subject to:

$$\sum_{j=1}^{|O^{|T|w}|} p_j x_{ij} \leq Q y_i \quad \forall i \in \{1, \dots, |O^{|T|w}|\} \quad (2)$$

$$\sum_{i=1}^{|O^{|T|w}|} x_{ij} = 1 \quad \forall j \in \{1, \dots, |O^{|T|w}|\} \quad (3)$$

$$x_{ij} \in \{0, 1\} \quad \forall i, j \in \{1, \dots, |O^{|T|w}|\} \quad (4)$$

$$y_i \in \{0, 1\} \quad \forall i \in \{1, \dots, |O^{|T|w}|\} \quad (5)$$

B. MIP model for network optimization

The network model is formulated as a mixed-integer programming model. Let A be the set of nodes, which denote the geo-locations of trucks in Set 1 (free trucks that have no attached trailers), and B be the set of nodes, which denote the geo-locations of trailers in Set 2 (free trailers). There exist a precedence relation between the nodes in set A and set B which requires that a node in set A has to be visited

before a node in set B since a trailer has to be picked up by a truck. Thus, the links between the nodes of A and B are directed from A to B and the reverse links do not exist on the input graph. The travel time from node $i \in A$ to node $j \in B$ is denoted as t_{ij} .

Let C be the set of nodes that represent the geolocations of trucks in Set 4 (free trucks that have attached empty trailers and are located at different warehouses). Set C contains unique node representations for each truck even when they share a common location. That is, a separate node is added to graph for the geolocation of each truck even though the geolocation is the same as the geolocation of another truck. This case is valid for all types of vehicles, however it is most likely to be observed for Set 4.

We define a set D of nodes, which represent the current geo-locations of trucks traveling in their collection routes (trucks in Set 3), and a set E of nodes, which correspond to their final destinations. Between each truck's location $i \in D$ and the associated destination location $j \in E$, there exists a predetermined path including the nodes representing the locations of companies and the customs locations. To avoid showing every node and link between the current location i and destination location j of this path on the graph, we use a single directed link, (i, j) , which denotes the entire path with a travel time value of t_{ij}^d , $i \in D$, $j \in E$, $i \prec j$. Therefore, every node in set E has exactly one predecessor node in set D and every node in set D has exactly one successor in set E , such that $[(i \prec j) \rightarrow (v \not\prec j \wedge i \not\prec z) \mid \forall (i \in D \wedge v \in D \setminus \{i\}), \forall (j \in E \wedge z \in E \setminus \{j\})]$ and $[(j \not\prec i) \mid \forall (i \in D \wedge j \in E)]$. We obtain the travel time value, $t_{ij} = t_{ij}^d$, of the entire path by traversing the path by a preprocessing algorithm using a *haversine* function-based approximation method. The *haversine*-based method considers the curve of the earth in computing distance and travel time. Alternatively, we can receive for each link of the path iterative responses from web using an API to have a more accurate estimate.

Finally, we define the set F , to represent the locations of the warehouses, which are the sink nodes of the graph. $\forall i \in B \cup E \cup C$ and $\forall k \in F$, travel time, t_{ik} , is the travel time between the geo-locations represented by nodes i and k .

Each warehouse location, $k \in F$, has to be visited by a number of trucks, b_k , within a time period T . Each $k \in F$ is a vertex corresponding to the geo-location of a warehouse $w \in W$ that is shown as $k \leftarrow geoloc^w$ and b_k is equal to $B^{|T|w}$ which is the minimum required number of trucks to pack the orders at warehouse $w \in W$ within a time period T and obtained by solving the bin packing problem defined in Eqs.1-5.

We denote the current time by R^C and tracking time for truck location by r_i^t , for $i \in A \cup E$. x_{ij} are the binary decision variables. The objective is to minimize total travel time with respect to precedence and other constraints. Lateness, number of vehicles, and load factors are also improved by solving the explained models at different stages.

The mathematical program is formalized in Eqs.6-23. The

objective function in Eq.6 minimizes the total travel time of the selected paths and edges. Eq.7 is a constraint related to the number of usable vehicles that has to reach to each warehouse. The constraint in Eq.8 ensures that for each node in set B , the number of incoming arcs from the nodes in set A is equal to number of outgoing arcs to set F . Constraints in Eqs. 9 and 10 enforce the number of outgoing arcs to set F from a node in set B and number of incoming arcs to a node in set B from set A be at most 1. Due to the constraint in Eq.11, if any path from a node in set D to node in set E is selected, then there has to be one outgoing link from the associated node in set E to anyone of the nodes in set F .

Eq.12 bounds the number of vehicles going from each node in set E to the nodes in set F to be at most 1. Eqs.13 and 14 are *big-M* constraints that restrict the time a vehicle has to reach to a warehouse. With respect to Eq.15, one vehicle at a warehouse can visit at most one of the other warehouses and, due to Eq.16, if there is such a visit, it should be within the time window T . Eqs.17-22 ensure that some of the links in the graph are directional by setting the binary decision variables to 0 for forbidden flows. This also denotes the precedence constraints among the nodes. Eq.23 are the binary decision variables which represent the flows on directed or undirected links.

$$\min \sum_{i \in D, j \in E} t_{ij}^d x_{ij} + \sum_{i \in A, j \in B} t_{ij} x_{ij} + \sum_{i \in B \cup E \cup C, k \in F} t_{ik} x_{ik} \quad (6)$$

subject to:

$$\sum_{i \in B \cup C \cup E} x_{ik} \geq b_k \quad \forall k \in F \quad (7)$$

$$\sum_{i \in A} x_{ij} = \sum_{k \in F} x_{jk} \quad \forall j \in B \quad (8)$$

$$\sum_{k \in F} x_{jk} \leq 1 \quad \forall j \in B \quad (9)$$

$$\sum_{i \in A} x_{ij} \leq 1 \quad \forall i \in A \quad (10)$$

$$x_{ij} = \sum_{k \in F} x_{jk} \quad \forall i \in D, \forall j \in E, i \prec j \quad (11)$$

$$\sum_{k \in F} x_{jk} \leq 1 \quad \forall j \in E \quad (12)$$

$$(r_i^t + t_{ij})x_{ij} + t_{jk}x_{jk} \leq R^C + T + (1 - x_{ij})M + (1 - x_{jk})M \quad \forall i \in A, \forall j \in B, \forall k \in F \quad (13)$$

$$(r_i^t + t_{ij}^d)x_{ij} + t_{jk}x_{jk} \leq R^C + T + (1 - x_{ij})M + (1 - x_{jk})M \quad \forall i \in D, \forall j \in E, \forall k \in F, i \prec j \quad (14)$$

$$\sum_{k \in F} x_{ik} \leq 1 \quad \forall i \in C \quad (15)$$

$$t_{ik}x_{ik} \leq T + (1 - x_{ik})M \quad \forall i \in C, \forall k \in F \quad (16)$$

$$x_{ik} = 0 \quad \forall i \in A, \forall k \in F \quad (17)$$

$$x_{ji} = 0 \quad \forall j \in B, \forall i \in A \quad (18)$$

$$x_{ki} = 0 \quad \forall k \in F, \forall i \in A \cup B \cup C \cup D \cup E \quad (19)$$

$$x_{ij} = 0 \quad \forall i \in D, \forall j \in E, i < j \quad (20)$$

$$x_{ji} = 0 \quad \forall j \in E, \forall i \in D \quad (21)$$

$$x_{ik} = 0 \quad \forall i \in D, \forall k \in F \quad (22)$$

$$x_{ij} \in \{0, 1\} \quad \forall i, j \in A \cup B \cup C \cup D \cup E \cup F \quad (23)$$

VI. SOLUTION METHOD

We first call the pre-processing algorithm described in Section IV to extract the nodes and graph features, orders that have entered into warehouses by the time of tracking and their due-dates, fleet characteristics regarding their status, location, movement type, availability at the time of tracking, customs check locations, loading and unloading locations.

Second, we solve the time dependent bin-packing model given in Section V-A for each warehouse, using a modified version of an approximation algorithm, *first fit decreasing* [18], to minimize the required number of vehicles at each depot. Instead of using a pure first fit decreasing approach we use priority weights to consider also the due-dates of orders. Complexity of the procedure is $O(|W|n^2)$, where $|W|$ is the number of warehouses and n is the number of orders. The outputs are used as inputs for the MIP model.

To solve the MIP model, we relax some of the constraints by setting them as penalty terms in the objective function. The reason is that the problem instances are able to create very different scenarios, thus in some cases there might not be a feasible solution of the model due to the limited resources. In such cases, we are interested in reaching our goal as much as possible by introducing the relaxed constraints into the objective function. These relaxations occur for the constraints in Eqs. 6, 13, 14 since sending enough number of vehicles to warehouses or catching the required time windows might not be possible at all times. After that, we perform the following sequential steps:

- 1) we solve a truck-trailer matching problem between sets A and B with an LP Solver to send the trucks to the closest trailers.
- 2) we then fix the decision variables related to selected links by 1. We obtain an estimated time of travel on the routes between sets D and E and reduce them into single links. Estimation procedure requires traversing a pre-determined path between D and E , thus has a complexity of $O(n)$. We are able to repeat this procedure both by calling a time function derived from haversine distance or an API that retrieves the actual travel times based on traffic conditions. Later, we replace each node with a non-unit value by many nodes with unit values and generate the related links. Then, we solve a bipartite matching problem on the expanded network using an LP Solver.
- 3) we find which truck-trailer groups are visiting which warehouses by compressing the expanded nodes again.
- 4) finally, with a heuristic we try to match the truck-trailer pairs with the most appropriate packs of orders considering the due-dates of orders within the packs

and the travel time of truck-trailer pairs that will reach to warehouses using the selected paths.

This method is highly flexible since it eliminates the risk of infeasibility by relaxing strict constraints. It is also successful in approaching targets and it is highly efficient in terms of computational times due to the use of LP models. All procedures we use are polynomially solvable.

VII. COMPUTATIONAL RESULTS.

Our solution method in Section VI (alg^p) includes a complex pre-processing algorithm to extract from a vast amount of information heterogeneous graphs representing the entire system. We repeat the experiments for the cases where travel distances and times between the nodes of the graph are either computed by haversine function based formula or retrieved by an API that takes the geo-locations as input and returns the actual distances and travel times based on traffic or other external conditions. In both cases computations are not performed beforehand but a travel time or distance function is called only when there is a need while the algorithm is being performed. Dynamic nodes of the graph and their locations are not always the same and depend on the time of the tracking. One example to this case is that some nodes denote the current locations of the trucks that are on the move and will have traveled more next time we retrieve information. Considering this and also the traffic conditions, which affect travel times, every time a different instance and scenario occurs.

We performed the experiments on a 2.8 GHz PC with Intel Core i5 processor and 16 GB of RAM. CPU time was c.a. 180 seconds for all experiments when travel times are computed by haversine distance based formula. With the API, CPU time is longer since there occurs a communication delay for each http request and response to retrieve an actual travel time considering real time environmental conditions. Thus, we can run the suggested solver every 3-5 minutes in the former case and every 10-15 minutes in the latter case. There was no significant difference between the solution qualities of the two approaches. The performance of our solution method (alg^p) against intuitive heuristics (alg^b) is presented in Table I. Table I shows significant improvement in all of the key performance indicators compared to simple intuitive heuristics (alg^b). These KPIs are:

- $KPI_1 = ((\sum_{j \in CU_t} p_j)/Q)/|UT|$ is the load factor over actually used trailers,
- $KPI_2 = ((\sum_{j \in CR_{tw}} p_j)/Q)/|RT|$ is the load factor over required number of trailers,
- $KPI_3 = |RT| = \sum_{w \in W} |rt_w| = \sum_{i \in S_w} y_{iw}$ is the Number of required vehicles,
- $KPI_4 = (\sum_{w \in W} \max[|rt_w| - |ut_w|, 0])/|rt_w|/|W|$ is the Ratio of unserved packs of orders,
- $KPI_5 = (\sum_{(v,g) \in SVG} \max[t_{vg} - \min_{o \in OTW_{go}}, 0])/ \min(|V|, |G|)$ is the average lateness for served packs,
- $KPI_6 = \sum_{i \in A, j \in B} x_{ij} t_{ij} + \sum_{i \in B, j \in F} x_{ij} t_{ij} + \sum_{i \in C, j \in F} x_{ij} t_{ij}$ is the total empty travel time of vehicles,

- $KPI_7 = \sum_{i \in A, j \in B} x_{ij} t_{ij} + \sum_{i \in B, j \in F} x_{ij} t_{ij} + \sum_{i \in C, j \in F} x_{ij} t_{ij} + \sum_{i \in E, j \in F} x_{ij} t_{ij} + \sum_{i \in D, j \in E} x_{ij} t_{ij}^d$
- $KPI_8 = \sum_{i \in A, j \in B} x_{ij} dist_{ij} + \sum_{i \in B, j \in F} x_{ij} dist_{ij} + \sum_{i \in C, j \in F} x_{ij} dist_{ij}$ is the Empty kilometers of vehicles.

where Q is the trailer capacity, CU_t be the set of order ids inside an actually used trailer $t \in UT$, CR_t be the set of order ids inside a required trailer $t \in RT$, UT be the set of indices of trailers that are actually used to serve the packed orders and $|UT|$ be the cardinality of UT , RT be the set of indices of required trailers to serve all packed orders and $|RT|$ be the cardinality of RT , and $\sum_{j \in CU_t} p_j$ be the sum of payweights of orders that are assigned to trailer t .

TABLE I
PERFORMANCE

KPIs	units	alg ^b baseline	alg ^p Section VI	% improvement	CPU time (sec)
KPI_1	%	85.1	96.7	13.6	180
KPI_2	%	79.8	98.2	18.4	180
KPI_3	trailers	120	99	17.5	180
KPI_4	%	58.2	19.3	66.8	180
KPI_5	minutes	29	0	100	180
KPI_6	minutes	17594	6300	64.1	180
KPI_7	minutes	61264	32007	47.8	180
KPI_8	km	3909.8	1400	64.1	180

VIII. CONCLUSION

In this study we proposed an optimization approach for the collection and consolidation operations in cross-border multi-modal distribution networks. We proposed a mixed-integer programming model and provided a solution using matheuristics. We then implemented our solution method on real-time data retrieved from the tracking system of a third-party logistics company. Experiments show that our solution method significantly outperforms other heuristics in terms of solution quality which is measured with respect to lateness, empty kilometers traveled, travel times, number of required/used vehicles, load factors, and ratio of served orders.

Results show that proposed solution method is highly efficient in computational terms and improves many important KPIs by 13% to 67% when compared against an intuitive heuristic. Thus, our solution approach can be easily used in practice for fast global re-planning in other inter-modal logistics networks. In future research, we are planning to perform additional experiments using more baseline heuristic methods and different networks.

REFERENCES

[1] D. Zhang, R. He, S. Li, and Z. Wang, "A multimodal logistics service network design with time windows and environmental concerns," *PLOS ONE*, vol. 12, no. 9, pp. 1–19, 09 2017. [Online]. Available: <https://doi.org/10.1371/journal.pone.0185001>

[2] B. Dong, M. Christiansen, K. Fagerholt, and S. Chandra, "Design of a sustainable maritime multi-modal distribution network – case study from automotive logistics," *Transportation Research Part E: Logistics and Transportation Review*, vol. 143, p. 102086, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1366554520307365>

[3] M. Zhou, Y. Duan, W. Yang, Y. Pan, and M. Zhou, "Capacitated multi-modal network flow models for minimizing total operational cost and co2e emission," *Computers and Industrial Engineering*, vol. 126, pp. 361–377, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0360835218304650>

[4] L. Bowen, Y. Bin, Z. Xiaolin, and L. Jun, "Operational optimization of transit consolidation in multimodal transport," *Computers and Industrial Engineering*, vol. 129, pp. 454–464, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0360835219300750>

[5] N. Lin, H. M. Hjelle, and R. Bergqvist, "The impact of an upstream buyer consolidation and downstream intermodal rail-based solution on logistics cost in the china-europe container trades," *Case Studies on Transport Policy*, vol. 8, no. 3, pp. 1073–1086, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2213624X19300884>

[6] J. C. Tyan, F.-K. Wang, and T. C. Du, "An evaluation of freight consolidation policies in global third party logistics," *Omega*, vol. 31, no. 1, pp. 55–62, 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0305048302000944>

[7] P. Fontaine, T. G. Crainic, O. Jabali, and W. Rei, "Scheduled service network design with resource management for two-tier multimodal city logistics," *European Journal of Operational Research*, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221721000874>

[8] S. Belieres, M. Hewitt, N. Jozefowicz, and F. Semet, "A time-expanded network reduction matheuristic for the logistics service network design problem," *Transportation Research Part E: Logistics and Transportation Review*, vol. 147, p. 102203, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1366554520308450>

[9] C. Serrano, X. Delorme, and A. Dolgui, "Cross-dock distribution and operation planning for overseas delivery consolidation: A case study in the automotive industry," *CIRP Journal of Manufacturing Science and Technology*, vol. 33, pp. 71–81, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1755581721000274>

[10] X. Gao, X. Jin, P. Zheng, and C. Cui, "Multi-modal transportation planning for multi-commodity rebalancing under uncertainty in humanitarian logistics," *Advanced Engineering Informatics*, vol. 47, p. 101223, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1474034620301920>

[11] M. F. Maghfiroh and S. Hanaoka, "Multi-modal relief distribution model for disaster response operations," *Progress in Disaster Science*, vol. 6, p. 100095, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2590061720300326>

[12] A. Haghani and S.-C. Oh, "Formulation and solution of a multi-commodity, multi-modal network flow model for disaster relief operations," *Transportation Research Part A: Policy and Practice*, vol. 30, no. 3, pp. 231–250, 1996. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0965856495000208>

[13] J.-B. Sheu and C. Pan, "A method for designing centralized emergency supply network to respond to large-scale natural disasters," *Transportation Research Part B: Methodological*, vol. 67, pp. 284–305, 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0191261514000927>

[14] J.-H. Zhang, J. Li, and Z.-P. Liu, "Multiple-resource and multiple-depot emergency response problem considering secondary disasters," *Expert Systems with Applications*, vol. 39, no. 12, pp. 11066–11071, 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417412004861>

[15] R. A. Garrido, P. Lamas, and F. J. Pino, "A stochastic programming approach for floods emergency logistics," *Transportation Research Part E: Logistics and Transportation Review*, vol. 75, pp. 18–31, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1366554514002087>

[16] A. Moreno, D. Alem, and D. Ferreira, "Heuristic approaches for the multiperiod location-transportation problem with reuse of vehicles in emergency logistics," *Computers and Operations Research*, vol. 69, pp. 79–96, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0305054815002828>

[17] S. Martello and P. Toth, "Bin-packing problem," in *Knapsack problems: algorithms and computer implementations*. Chichester, UK: John Wiley and Sons, ISBN 0471924202, 1990, ch. 8, pp. 221–245.

[18] D. Johnson, "Near-optimal bin-packing algorithms," 1973.