Scheduling Movements in the Network of an Express Service Provider

Ilse Louwerse   Jos Mijnarends   Ineke Meuffels   Dennis Huisman
Hein Fleuren

Econometric Institute Report EI2012-08

May 11, 2012

Abstract

Express service providers manage shipments from senders to receivers under strict service level agreements. Such shipments are usually not sufficient to justify a single transportation, so it is preferred to maximize consolidation of these shipments to reduce cost. The consolidation is organized via depots and hubs: depots are local sorting centers that take care of the collection and delivery of the parcels at the customers, and hubs are used to consolidate the transportation between the depots. A single transportation between two locations, carried out by a certain vehicle at a specific time, is defined as a movement. In this paper, we address the problem of scheduling all movements in an express network at minimum cost. Our approach allows to impose restrictions on the number of arriving/departing movements at the hubs so that sufficient handling capacity is ensured. As the movement scheduling problem is complex, it is divided into two parts: one part concerns the movements between depots and hubs; the other part considers the movements between the hubs. We use a column generation approach and a local search algorithm to solve these two subproblems, respectively. Computational experiments show that by using this approach the total transportation costs are decreased.

Key words Express service provider · Movement rescheduling · Integer programming · Column generation · Local search

1 Introduction and Motivation

An express service provider offers time-guaranteed delivery of parcels to its customers. This concerns collection of the parcels at the customers, transportation via the network of the express service provider, and delivery of the parcels at the receivers. Local sorting centers called *depots* take care of the collection and delivery of the parcels to the senders and receivers. It is not cost efficient to transport all parcels directly from their origin depot to their destination depot, as usually the number of depots is large and the amount of parcels that needs to be transported between each pair of depots is relatively small.

1

Therefore, *hubs* are used to consolidate the transport between depots. At the hubs, parcels are unloaded from the arriving vehicles, sorted, and loaded on the departing vehicles.

Express service providers offer multiple *service types* to their customers. The service types differ in the date and time at which parcels need to be at their destination depot. We define a *flow* as all parcels between a pair of depots with the same service type. The origin cut-off time denotes the latest time a flow should be available at its origin depot, while the destination cut-off time denotes the latest time a flow has to arrive at its destination depot. For each flow a *route*, describing the sequence of locations visited when transporting the flow from its origin depot to its destination depot, has to be determined. The cut-off times determine which routes are possible for each flow.

Several vehicle types are used to transport the parcels along the routes. The vehicle types differ in speed, capacity and cost per kilometer. The preferred vehicle type depends on the number of parcels that needs to be transported and the time available to transport the parcels. In an express network, a number of vehicles move between the different types of locations according to a given schedule. One such transportation between two locations done by a certain vehicle type on a given time is called a *movement*. The challenge of scheduling all the movements in an express network at minimum cost is referred to as the *movement scheduling problem*.

The movement scheduling problem considered in this paper originates from one of the largest express providers in the world, as it was expected that its current scheduling process could be improvement. The current process is not supported by any modeling tools and movement schedules are improved locally, based on practical experiences. This research aims to develop a model that solves the movement scheduling problem.

For practical reasons, it might be desired to restrict the changes in the number of arrivals and departures at the hub locations. To ensure that the new movement scheme can be implemented on short term, the current hub capacities should be sufficient to handle the new movement scheme. In this research, we present additional constraints to take this into account. However, our goal is to find the optimal movement scheme, which can be obtained by relaxing these restrictions.

The solution approach that is presented in this paper divides the movement scheduling problem into two parts, such that for each part a well-known mathematical formulation can be developed. The first part concerns scheduling the movements between depots and hubs; we present a composite variable formulation and a column generation algorithm used to solve it. The second part concerns the problem of scheduling the movements between hubs, which is formulated as a network loading problem and solved with a local search algorithm.

The remainder of this paper is organized as follows. In Section 2 we discuss literature related to our problem, and we introduce the overall solution approach that we use. The mathematical formulations are given in Section 3. Section 4 describes the solution approach in more detail. In Section 5 the results of several experiments based on data of an express service provider are presented. The conclusions and some recommendations for further research are mentioned in Section 6.

2   Literature Review and Solution Framework

The movement scheduling problem has similarities to the service network design problem, which concerns the decision how to move the parcels from their origins to their destinations and how to route the vehicles. The movement scheduling problem can be seen as an extension of the service network design problem on three levels. First of all, the time component has to be integrated in the model. To accomplish this, the planning horizon is discretized and the physical network is replicated for each period. Such dynamic networks are presented in Farvolden and Powell (1994) and Pedersen et al. (2009). The second feature that has to be included is the general empty balancing of vehicles, also called repositioning. We refer to Dejax and Crainic (1987) for an overview of models for empty vehicle repositioning. The third feature concerns the design variables, which are the variables concerning the allocation of capacity to the arcs. In most network design models, these are binary variables. However, in our model several vehicles can be assigned to an arc, resulting in integer variables. Crainic (2000) describes a model with integer variables modeling discrete choice design decisions.

To our knowledge, there is no literature presenting problem structures corresponding to our problem. However, several papers deal with problems that have similarities with the problem described in this paper. Such problems are presented in Armacost et al. (2002), Armacost et al. (2004) and Barnhart and Schneur (1996). Two simplifications are made in these papers. First, the presented models concern an one hub structure. Second, sort start and end times at the hub are introduced as the latest time parcels can arrive at and leave from the hub, respectively. As a result of these sort start and end times two separate problems can be solved: one to determine the depot-hub routes and one to determine the hub-depot routes. Usually, the networks of the express service provider that motivate our research concern more than one hub without a distinct hub sort as vehicles arrive and departe within the same time window, so that the presented models are insufficient.

The research presented in this paper concerns multi-hub express network in which we determine the routings of the flows, the vehicles to be used and the schedules of the resulting movements. As we want to construct routes that go via several hubs, it is not realistic to use hub windows. Besides, we want to take into account a heterogeneous vehicle fleet, different service types, exact timing of movements and parcels, and routes via several depots and hubs. These additional aspects make it difficult to solve the movement rescheduling problem at once. Therefore, we divide the problem in several subproblems. For each subproblem a well-known mathematical formulation can be developed.

2.1   Subproblems

Recall that for each flow a route has to be determined and that the cut-off times determine the routes that are possible for each flow. Two types of routes can be distinguished: *direct routes* and *non-direct routes*. In case the cut-off times are very tight and any possible route via at least one hub is infeasible, a direct route between the origin and

destination depot of the flow is used. Large flows between depots can also be a reason to use a direct route. In any other case, the parcels can be consolidated and routed via one or several hubs.

Non-direct routes visit one or more intermediate locations, of which at least one is a hub. Non-direct routes are divided in routes from depots to hubs (DH-routes), routes between two hubs (HH-routes), and routes from hubs to depots (HD-routes). The problem of finding the non-direct routes is divided into two parts: the first part concerns movements from depots to hubs and vice versa (DH/HD-problem) and the second part concerns the movements between the hubs (HH-problem). As flows are routed via sequences of DH-, HH- and HD-routes, the problems are connected to each other.

A DH-route starts at a depot, possibly visits other depots to load parcels, and ends at a hub. During the pickup process, parcels can only be loaded and not unloaded, as sorting of parcels occurs at the hubs. Similarly, HD-routes start at a hub and visit one or more depots where parcels are unloaded. During the delivery process, parcels can only be unloaded. If a DH-route or a HD-route visits several depots, some processing time has to be taken into account at these intermediate depots to load or unload parcels. Once a DH-route ends and the parcels arrive at the hub, the parcels are unloaded, sorted, and loaded on a vehicle that moves either to another hub or to the destination depot. A composite variable approach as described in Armacost et al. (2002) is used to formulate the DH/HD-problem as a set partitioning problem. This formulation has the flexibility to handle practical constraints that cannot be easily incorporated in traditional formulations. We solve DH/HD-problem with a column generation algorithm. Barnhart et al. (1998) show that column generation can be successfully applied to set partitioning problems with a large number of variables.

HH-routes are used to route the parcels between the hubs. A HH-route visits at least two hubs. The first and last hubs on the HH-route are the hub at which the DH-route ends and the hub at which the HD-route starts, respectively. At each hub along a HH-route it is possible to both unload and load parcels. The HH-problem is formulated as a network loading problem. Network loading problems (NLPs) model the design of capacitated networks with no variable flow cost, and with facilities of fixed capacity available to carry flow. For an overview of models and algorithms of network design problems we refer to Magnanti and Wong (1984) and Magnanti and Mirchandani (1993). Despite the importance of the NLP in a variety of settings (freight transportation, telecommunications industry), the available research on it is limited. The primary focus in the literature has been on the uncapacitated variant. The algorithms and heuristics developed for the uncapacitated problem, as for instance presented in Holmberg and Hellstrand (1998), do usually not work very well for the capacitated problem. The gap between the integer solution and the linear programming solution is small for the uncapacitated variant, but not for the capacitated variant as shown by Leung et al. (1990). We will use a local search algorithm to solve the NLP.

To transport the parcels via a non-direct route, sequences of DH-, HH- and HD-routes have to be constructed. Two routes can be connected if the end location of the
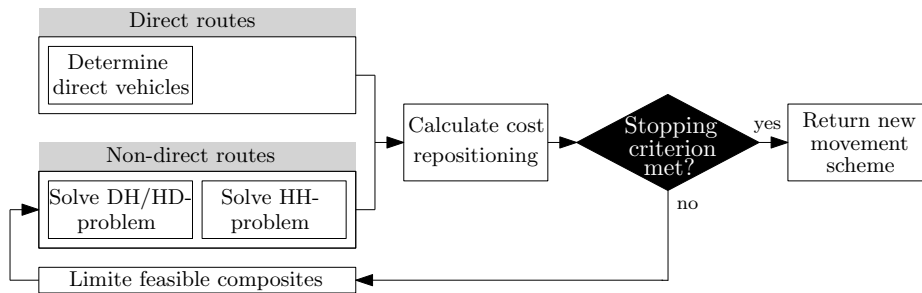
Figure 1: Overall algorithm

first route and the start location of the second route are the same, and the end time of the first route is before the start time of the second route.

Figure 1 shows the overall algorithm used to solve the movement rescheduling problem. Consecutively, the DH/HD-problem and the HH-problem are solved. Below, the column generation and local search algorithms used to solve these problems will be discussed.

Because the movement schedule is carried out on a daily basis, the number of incoming and outgoing vehicles at each location during a day has to be equal. If this is not the case, some vehicles need to be repositioned which results in additional cost. Once the DH/HD-problem and the HH-problem are solved, the repositioning cost has to be determined. The repositioning problem can be formulated as a flow problem and the optimal solution can be obtained by solving the corresponding linear programming relaxation.

The subproblems are solved sequentially which means that the solution of one subproblem depends on the solutions of one or several other subproblems. Therefore, there has to be some interaction between the subproblems and an iterative procedure is defined. A stopping criteria is defined and as long as the stopping criteria is not met, a new movement scheme is created. We try to improve the overall solution by adding additional constraints to the DH/HD-problem such that some of the composites used in the current solution are not feasible in the next solution. We decide which constraints to add to the DH/HD-problem based on the solution of the HH-problem. It is relatively expensive to transport the flows that are on a hub movement of a small vehicle type or a hub movement with a low utilization as the distances between the hubs are relatively large. Therefore, we try to decrease the total transportation cost by restricting these flows to be transported via only one hub.

3 Mathematical Formulations

This section describes the assumptions of the movement scheduling problem and presents the mathematical formulations of the DH/HD-problem and the HH-problem.

5

3.1 Assumptions

First of all, it is assumed that all locations are given and that the function of a location, i.e. whether a location is a depot or a hub, cannot be changed. Second, services are given and as express providers offer guaranteed services to their customers we assume that these services are feasible with respect to the cut-off times. The third assumption is that the number of vehicles available of each type is infinite. If there would be a limited number of available vehicles of each type, additional restrictions can be added to the model and our solution approach would still be applicable. Fourth, it is assumed that the costs of the different vehicle types are such that it is always cheaper to use one vehicle of a larger type than two vehicles of a smaller type. Moreover, the capacity of the hubs can be finite, defined by additional restrictions that impose a maximum number of movements that can arrive at and depart from a hub.

For all parcels that are not routed directly, a non-direct route, consisting of a DH-route, a HH-route and a HD-route, needs to be selected. Because of modeling reasons, it is assumed that all parcels in a flow have to be transported via the same DH- and HD-route. However, in general this assumption is not restrictive as the total flows at the depots are usually low so that only a few vehicles are used. For the flows between the hubs, the assumptions are less strict as the total flows between the hubs are larger and therefore flows between hubs are often consolidated. Hence, it is assumed that the parcels in a flow can use different HH-routes.

To limit the possible number of DH- and HD-routes, there are some restrictions on the depots and hubs that can be on the same route. The maximum number of depots on a DH/HD-route is two. Hence, each DH- and HD-route consists of one or two depots and one hub. Besides, only depots and hubs that are located near each other can be on the same route. These restrictions can be validated as in general the timing in express networks is tight, so that it is unlikely that the DH- and HD-routes can visit more than two depots or that they can go via locations that are not located near each other. For the HH-routes, we do not restrict the number of hubs the parcels can visit in a route.

3.2 DH/HD-problem

The DH/HD-problem is to assign each flow to a DH- and a HD-route such that all parcels arrive at their destination depot in time, feasible combinations of DH- and HD routes are chosen, and the transportation cost is minimized. The problem can be formulated using the composite variable approach as described in Armacost et al. (2002). The composites represent a combination of vehicles on a route with sufficient capacity to transport a feasible combination of flows. Compared to general network design problems in which both flow variables and design variables have to be modeled explicitly, the use of composite variables results in a dramatic decrease in problem size as the flow variables are removed as explicit decision variables and embedded in the design variables. Using composite variables, the problem can be modeled as a set partitioning model. Armacost et al. (2002) show that such a model formulation is stronger than the conventional network design model formulation by showing that the integer programming formula-

tions are equivalent and that the composite variable formulation has a stronger linear programming relaxation.

An overview of the notation of the model is given below.

Sets

$L = D \cup H$	set of locations, index i
$H \subset L$	set of hub locations, index h
$D \subset L$	set of depot locations, index j, k
V	set of vehicle types, index v
S	set of services, index s
P	set of DH-composites, index p
Q	set of HD-composites, index q

Parameters

r_h	maximum number of DH- and HD-movements at hub h
c_p	cost of DH-composite p
d_q	cost of HD-composite q
m_{phv}	number of vehicles v arriving at hub location h on DH-composite p
n_{qhv}	number of vehicles v departing from hub location h on HD-composite q

$$v_{pjis} = \begin{cases} 1 & \text{if DH-composite } p \text{ transports flow } s \text{ from depot } j \text{ to location } i; \\ 0 & \text{otherwise.} \end{cases}$$

$$w_{qijs} = \begin{cases} 1 & \text{if HD-composite } q \text{ transports flow } s \text{ from location } i \text{ to depot } j; \\ 0 & \text{otherwise.} \end{cases}$$

$$e_{pq} = \begin{cases} 1 & \text{if DH-composite } p \text{ can be used in combination with HD-composite } q; \\ 0 & \text{otherwise.} \end{cases}$$

Variables

$$x_p = \begin{cases} 1 & \text{if DH-composite } p \text{ is used;} \\ 0 & \text{otherwise.} \end{cases}$$

$$y_q = \begin{cases} 1 & \text{if HD-composite } q \text{ is used;} \\ 0 & \text{otherwise.} \end{cases}$$

Parameters e_{pq} are used to ensure that feasible combinations of DH- and HD-composites are chosen. DH-composite p can be used before HD-composite q if there is enough time available between the end time of composite p and the start time of composite q to transport the parcels between the end hub of composite p and the start hub of composite q, and to process the parcels at the hub(s).

The DH/HD-problem can be formulated as follows.

$$\min \sum_{p \in P} c_p x_p + \sum_{q \in Q} d_q y_q \qquad (1)$$

$$\sum_{p \in P} v_{pjis} x_p = 1 \qquad \forall i \in L, j \in D, i \neq j, s \in S \qquad (2)$$

$$\sum_{q \in Q} w_{qijs} y_q = 1 \qquad \forall i \in L, j \in D, i \neq j, s \in S \qquad (3)$$

$$\sum_{v \in V} \left(\sum_{p \in P} m_{phv} x_p + \sum_{q \in Q} n_{qhv} y_q \right) \leq r_h \qquad \forall h \in H \qquad (4)$$

$$\sum_{p \in P} \sum_{q \in Q} v_{pjks} w_{qkjs} e_{pq} x_p y_q = 1 \qquad \forall j, k \in D, j \neq k, s \in S \qquad (5)$$

$$x_p, y_q \in \{0,1\} \qquad \forall p \in P, q \in Q \qquad (6)$$

The objective function (1) minimizes the cost of transporting the parcels between the depots and hubs. Constraints (2) and (3) specify that one DH-composite and one HD-composite have to be selected for each flow. A maximum on the number of DH- and HD-movements at each hub location is described by constraints (4). Constraints (5) ensure that for each flow that has to be transported between two depots a feasible combination of a DH- and a HD-composite is selected. Quadratic constraints (5) can be linearized by introducing binary variables z_{pq}, defined for each possible combination of a DH- and a HD-composite. Using these variables, constraints (5) can be replaced by the following set of constraints (7)-(9).

$$\sum_{p \in P} \sum_{q \in Q} v_{pjks} w_{qkjs} e_{pq} z_{pq} = 1 \qquad \forall j, k \in D, j \neq k, s \in S \qquad (7)$$

$$z_{pq} \leq x_p \qquad \forall p \in P, q \in Q \qquad (8)$$

$$z_{pq} \leq y_q \qquad \forall p \in P, q \in Q \qquad (9)$$

3.3 HH-problem

The hub problem is to transport the flows between the hubs, at minimum cost and given the DH- and HD-route of the flow. We model the hub problem as a network loading problem. The time dimension is explicitly taken into account by using a time-expanded directed graph. Each node in the graph represents a hub in a given time period. Directed arcs are added to link nodes. An arc connecting two nodes representing the same hub represents freight waiting at the hub to be loaded onto an outbound vehicle. An arc connecting two nodes representing different hubs represents a transportation service between two hubs. The set of arcs is divided in several subsets, one for each vehicle type.

Sets

N	set of nodes, index i,j
$N^h \subset N$	set of nodes representing hub h
A	set of arcs, index (i,j)
F	set of flows, index f

Parameters

n_h	maximum number of vehicles at hub h
k^f	total number of parcels of flow f
c_{ij}	cost of moving one vehicle over arc (i,j)
q_{ij}	capacity of one vehicle on arc (i,j)

Variables

x_{ij}^f	number of parcels of flow f on arc (i,j)
y_{ij}	number of vehicles on arc (i,j)

The origin and destination hub of flow f are represented by $O(f)$ and $D(f)$, respectively. Below, the arc-based formulation of the hub problem is presented.

$$\min \sum_{(i,j)\in A} c_{ij} y_{ij} \tag{10}$$

$$\sum_{j\in N} x_{ij}^f - \sum_{j\in N} x_{ji}^f = \begin{cases} k^f & \text{if } i = O(f); \\ -k^f & \text{if } i = D(f); \\ 0 & \text{otherwise.} \end{cases} \quad \forall f \in F \tag{11}$$

$$\sum_{f\in F} x_{ij}^f \leq q_{ij} y_{ij} \quad \forall (i,j) \in A \tag{12}$$

$$\sum_{i\in N^h} \left(\sum_{j\in N\setminus N^h} y_{ij} + \sum_{j\in N\setminus N^h} y_{ji} \right) \leq n_h \quad \forall h \in H \tag{13}$$

$$x_{ij}^f \in \mathbb{R}_+ \quad \forall f \in F, (i,j) \in A \tag{14}$$

$$y_{ij} \in \mathbb{Z}_+ \quad \forall (i,j) \in A \tag{15}$$

The objective function (10) minimizes the total transportation cost. Constraints (11) are the flow conservation constraints, and constraints (12) specify that the flow on an arc cannot exceed the capacity of the vehicles on the arc. A maximum number of vehicles that can arrive at and depart from hub h is described by constraints (13). The integrality of the variables x_{ij}^f is ensured by the integrality of the vehicle capacities.

4 Solution Methods

In this section we discuss the column generation algorithm used to solve the DH/HD-problem and the local search algorithm used to solve the HH-problem. We refer to Lübbecke and Desrosiers (2005) and Aarts and Lenstra (2003) for a general discussion of topics related to column generation and local search.

4.1 Column Generation Algorithm

The algorithm starts with creating a set of initial columns for the DH/HD-problem. We use the following heuristic to do this. Each flow goes either via the hub closest to the origin depot and the hub closest to the destination depot, via one of these hubs or via the fastest hub (i.e. the hub that results in the earliest arrival time at the destination depot).

We apply column generation to the linear programming relaxation of the DH/HD-problem (1)-(6), without constraints (5). Let π_{jis}^*, μ_{ijs}^* and σ_h^* be the optimal values of the dual variables associated with constraints (2), (3) and (4), respectively. The reduced cost for non-basic composites p and d are:

$$\theta_p = c_p - \sum_{i \in L} \sum_{j \in D, i \neq j} \sum_{s \in S} v_{pjis} \pi_{jis}^* - \sum_{v \in V} \sum_{h \in H} m_{phv} \sigma_h^* \tag{16}$$

$$\lambda_q = d_q - \sum_{i \in L} \sum_{j \in D, i \neq j} \sum_{s \in S} w_{qijs} \mu_{ijs}^* - \sum_{v \in V} \sum_{h \in H} n_{qhv} \sigma_h^* \tag{17}$$

The pricing problems of the DH- and HD-problems can be solved separately. The problems are very similar and we only present the pricing problem of the HD-composites in the section. Each flow is represented by nodes in the graph. These nodes are called flow nodes. Besides, a source node and several sink nodes are added. The source node does not represent a physical location, but is added to ensure a start location for each shortest path. The sink nodes represent the hubs. The source is connected to all flow nodes and the flow nodes are connected to two sinks (representing the two nearest hubs). Besides, each flow node is connected to each other flow node that can be in the same composite. This means that each flow node is only connected to nodes representing the same origin depot and a different destination location and/or service type, or nodes representing one of its two nearest depots.

Figure 2 shows a representation of such a directed acyclic graph for the DH-problem for a network with two depot locations (j and k), one service type (s), and one hub (h). The graph can be seen as consisting of two similar sub graphs: An upper graph and a lower graph. In the upper graph, there is a node from the source to each flow node. Besides, each flow node is connected to the other flow node representing the same origin depot and to the hub. In the lower graph, each flow node is connected to the other flow node representing the same origin depot and to the hub. Besides, the flow nodes in the upper graph are connected to the flow nodes in the lower graph representing the other depot.
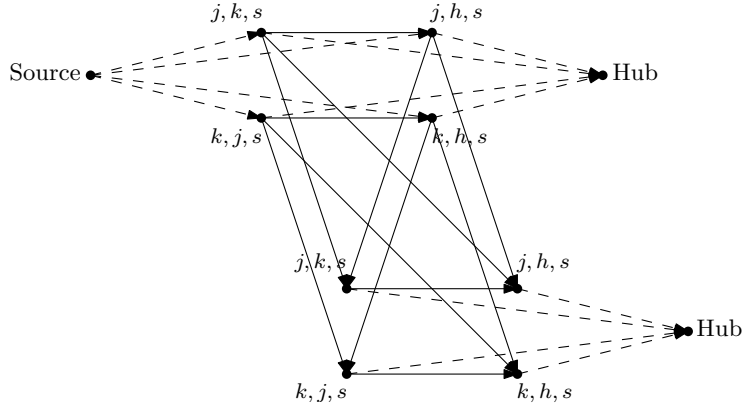
Figure 2: Directed acyclic graph

Paths in the graph go from the source via one or several flow nodes to a hub node. Each path represents a feasible composite. An approximation of the reduced cost of a composite is obtained as the path cost, if the arc weights are given as follows:

- An arc from the source to a node n representing flow from depot j to location i and service s is labeled with cost $-\pi_{jis}^*$ plus the minimum cost of transporting flow f_{jis}.

- An arc from node n to node n' representing flow from depot j' to location i' and service s' is labeled with cost $-\pi_{j'i's'}^*$ plus the minimum cost of transporting flow $f_{j'i's'}$.

- An arc from node n to a sink node m representing hub h is labeled with cost $-\sigma_h^*$.

Equation (16) shows that the reduced cost of a DH-composite consists of the cost of transporting the total flow on the composite, minus the dual variables of the flows on the composite and the dual variable of the hub at which the composite ends. As we do not know beforehand which flows will be on the composite, we cannot take the exact cost of the composite into account in the shortest path problem. To take the cost to some extent into account, we add the minimum cost of transporting the flows to the arc weights. This minimum cost is defined as the cost of the largest vehicle divided by the capacity of the largest vehicle, and multiplied by the distance between the origin depot of the flow and the nearest hub of this depot.

Including an arc in a composite means including the node at which the arc terminates in the composite. If a path only visits flow nodes in the upper sub graph, the composite contains one depot. If a path visits flow nodes in both the upper and the lower sub graphs, the composite contains two depots. A shortest path problem is solved to find paths with negative reduced cost.

To solve the shortest path problem, we use the single-source shortest path algorithm for directed acyclic graphs as described in Cormen et al. (2001). The algorithm produces

a shortest path from the source node to every other node in the graph. Note that we are only interested in the shortest paths from the source to the hubs and that the algorithm produces several columns with negative reduced cost. In each iteration, several columns representing a shortest paths between the source and a hub can be added if they have negative reduced cost.

4.2 Local Search Algorithm

Given a solution of the DH/HD-problem, the HH-problem is solved. The origin and destination hub, as well as the arrival time at the origin hub and the departure time from the destination hub of each flow are given by the solution of the DH/HD-problem. We solve the HH-problem with a local search algorithm.

In the initial solution, all parcels go directly from their origin hub h^o to their destination hub h^d. The time parcel p arrives at its origin hub is denoted as α_p, and the latest time it has to leave its origin hub to arrive at its destination hub on time is denoted as β_p. A vehicle is scheduled if there are enough parcels available at the origin hub to fill a largest vehicle, or at the time some parcels have to leave the origin hub to arrive at the destination hub in time. If there are enough parcels to fill a largest vehicle, the parcels with the earliest departure times from the destination hub are assigned to this vehicle first. Parcels that do not necessarily have to leave the origin hub are scheduled on the movement as long as there is capacity left. The vehicle type chosen is the smallest vehicle type such that the capacity of the vehicle exceeds the number of parcels that needs to be scheduled. The steps used to determine the movements of the initial solution between each pair of origin hub and destination hub are summarized below.

1. Denote the set P as all parcels p between h^o and h^d.

2. Set $t_0 = \min_{p \in P}\{\alpha_p\}$ and $t_1 = \max_{p \in P}\{\beta_p\}$.

3. Set the current time $t = t_0$ and the total flow at h^0 waiting to be transported $f = 0$.

4. While $t \leq t_1$ do:

   (a) Add all parcels p for which $t = \alpha_p$ to f.

   (b) Check whether f is large enough to fill one or more largest vehicles. Schedule the movement(s) and subtract the parcels assigned to the vehicle(s) from f.

   (c) If there are parcels p in f for which $t = \beta_p$, schedule a movement and subtract the parcels from f.

   (d) Set $t = t + 1$.

Solutions in the neighborhood of the current solution are obtained by performing small modifications to the current solution. We define small modifications as rerouting parcels that go from hub h^o to hub h^d in the current solution to going from hub h^o via hub h^i to hub h^d in the new solution. Once parcels are rerouted, some movements between

the hub pairs $h^o - h^d$, $h^o - h^i$ and $h^i - h^d$ have to be rescheduled. In the neighborhood search, we do not consider parcels that are scheduled on movements consisting of the largest vehicle with no capacity left. The reason is that it is always most cost efficient to use a largest vehicle and the shortest route if there are enough parcels.

The movements that are candidates to be rescheduled are the movements of the smaller vehicle types, and the movements with a low utility (defined as the ratio of the flow on the movement and the capacity of the movement). To start each neighborhood search, we order the movements within each group of vehicle type by increasing utility, starting with the group of movements with the smallest vehicle type. Denote this movement set as M. We start with the first movement $m \in M$ and reschedule this movement via hubs h^i for which the driving time between hubs h^o and h^i, the processing time at hub h^i and the driving time between hubs h^i and h^d does not exceed the available time between the arrival of the parcels at hub h^o and the departure from hub h^d. It would be computationally too intensive to consider each movement $m \in M$ and each intermediate hub h^i. Therefore, we start with the first movement $m \in M$ and consider for this movement all possible intermediate hubs h^i. As long as no movement scheme with lower cost is found, the algorithm continues with the next movement $m \in M$, again considering all possible intermediate hubs. We denote with z the current solution, with z^* the best solution found so far, and with $c(z)$ the cost of solution z. The steps of the local search algorithm are summarized below.

1. Set $z^* = z$.

2. While $c(z) \leq c(z^*)$ do:

   (a) Get the first movement $m \in M$.
   (b) For each possible intermediate hub $h^i \in H$, $h^i \notin \{h^o, h^d\}$ do:
       i. Reroute the parcels scheduled on m via h^i.
       ii. Calculate the cost of the new movement scheme z.
       iii. Set $z^* = z$ if $c(z) < c(z^*)$.
   (c) Remove m from M.

We have found a (local) optimal solution if, at the end of the algorithm, the set M is empty and $c(z) \leq c(z^*)$.

5 Computational Results

We use three networks to test the solution approach presented in Section 2. In this section we first describe the test networks and the parameter settings. Next, numerical results are presented, which show that our solution approach can be used to decrease total transportation cost. We used CPLEX 12.1 on a 3.0 GHz processor with 3 GB of memory to solve the LP and MIP models in the column generation algorithm, and the repositioning model.

5.1 Networks and Parameter Settings

The solution approach is applied to three data instances. The first data instance originates from the Global Optimisation Game (GO-Game). This game is designed to show the challenges faced by an express service provider and is based on a simplified network. For more information on this game, we refer to Meuffels et al. (2010). The other two test networks are modified data instances of an express service provider. Table 1 shows the main characteristics of the networks. The processing times, and the number of services and vehicle types is the same for all networks. The number of locations, parcels and flows differ. While the GO Network is mainly introduced for illustrative purposes, Network 1 and Network 2 are representable networks of express service providers.

Two service types are defined: premium and normal. Table 2 shows that the number of parcels with premium service is small compared to the number of parcels with normal service. The origin cut-off times of the parcels of normal service are earlier than the cut-off times of the parcels of premium service. The opposite holds for the destination cut-off times.

In the column generation algorithm we specify a maximum number of columns that can be generated per iteration. Besides, we limit the number of columns in the restricted

Table 1: Characteristics of the networks

Characteristics	GO Network	Network 1	Network 2
Locations	10	44	60
Hubs	2	4	10
Processing time hubs (in minutes)	60	60	60
Processing time depots (in minutes)	15	15	15
Services	2	2	2
Vehicle types	2	2	2
Number of parcels	100,000	800,000	1,100,000
Number of flows	180	1,088	1,245
Maximum distance	537	909	1,215
Average distance	254	369	539

Table 2: Characteristics of the service types

		GO Network	Network 1	Network 2
Normal	% parcels	85	87.5	90
	Origin cut-off time	day 1, 20:00	day 1, 18:00	day 1, 18:00
	Destination cut-off time	day 2, 7:00	day 2, 11:00	day 2, 12:00
Premium	% parcels	15	12.5	10
	Origin cut-off time	day 1, 20:00	day 1, 20:00	day 1, 20:00
	Destination cut-off time	day 2, 6:00	day 2, 8:00	day 2, 10:00

master problem by deleting in each iteration the columns of which the reduced cost are positive and above a certain threshold. These columns are kept in a separate pool. At the beginning of each iteration we check whether some columns in this pool have negative reduced cost and should be added again to the restricted master problem. We run experiments with different values for the threshold for removing columns from the network and for the maximum number of columns added per iteration. Based on these experiments, we set the maximum number of columns added in each iteration to 25 for all three networks. For the GO Network the threshold for removing columns is set to 250, for Network 1 and Network 2 this threshold is set to 500. For the GO Network and Network 1 we obtain the best results using these settings. For Network 2 we are not able to run experiments with higher values for these thresholds due to memory issues. We expect that increasing the values of the thresholds results in better solutions.

5.2 Numerical Results

As our goal is to obtain the optimal movement scheme, we relax the constraints that limit the number of movements at each hub. The cost of the original solution and the cost of the best solutions found by our solution method are shown in Tables 3-5. The original solution represents the cost of the movement schemes similar to the current movement schemes of the express service provider. In the current movement schemes, most parcels are transported via the hub closest to the origin depot and the hub closest to the destination depot, and routes via several depots are not possible. The initial solution is the movement scheme without applying the column generation algorithm to find better DH- and HD-composites. The initial set of DH- and HD-composites is determined as described in Section 4.1.

The cost of the best movement schemes found with our solution method are 3.4%, 12.9% and 18.6% lower than the cost of the original solutions for respectively the GO Network, Network 1 and Network 2. Applying the column generation algorithm to the initial solutions of the GO Network and Network 1 does not result in a decrease in cost. When the column generation algorithm terminates because no additional columns with negative reduced cost can be identified, the gap between the lower bound and the solution of the DH/HD-problem was 0% and 3.2% for respectively the GO Network and Network 1. This indicates that the best solution found by our solution method is close to the optimal solution of the DH/HD-problem. Applying the column generation algorithm for Network 2, the DH/HD cost decreases by 4.8% and the total cost by 2.8%. The gap between the lower bound and the solution of the DH/HD-problem was about 50%. As the gap is large, we expect that better solutions can be obtained with higher threshold values.

For each network, the time needed to compute the initial solution is only a few milliseconds. The total running time of the algorithm for the GO Network is two seconds. For Network 1, the running times of the first and second iteration are 188 and 74 seconds, and for Network 2 these are equal to 273 and 471 seconds. Hence, the best solution of the GO Network and Network 1 are found within one second, and the best solution for Network 2 is found in less than five minutes.

15

We developed an iterative procedure with the goal to decrease the total transportation cost by rescheduling the movements with low utilization in the HH-problem. The experiments show that this procedure does result in a decrease in HH cost, but that this decrease is offset by an increase in DH/HD-cost. For the three networks, the movement scheme with the lowest total transportation cost is either the movement scheme of the initial solution or the movement scheme found in the first iteration.

Table 3: Transportation cost, GO Network

	Total	Direct	DH/HD	HH	Repos.
Original solution	87,879	5,916	67,884	5,208	8,871
Initial solution	84,893	34,473	33,663	13,671	3,086
Iteration 1	84,893	34,473	33,663	13,671	3,086
Iteration 2	85,463	34,473	35,226	12,369	3,395
Iteration 3	85,194	34,473	37,041	11,067	2,613
Iteration 4	85,289	34,473	38,133	9,765	2,918
Iteration 5	86,258	34,473	38,625	9,765	3,395

Table 4: Transportation cost, Network 1

	Total	Direct	DH/HD	HH	Repos.
Original solution	46,895	689	25,406	15,468	5,333
Initial solution	40,866	569	23,053	13,703	3,541
Iteration 1	40,866	569	23,053	13,703	3,541
Iteration 2	42,032	569	24,649	12,894	3,921

Table 5: Transportation cost, Network 2

	Total	Direct	DH/HD	HH	Repos.
Original solution	164,863	9,993	70,501	60,901	23,468
Initial solution	138,061	6,735	49,063	55,745	26,518
Iteration 1	134,190	6,735	46,708	55,745	25,002
Iteration 2	141,480	6,735	57,853	51,891	25,001

Table 6: Average transportation cost, simulated cases of Network 1

	Total	Direct	DH/HD	HH	Repos.
Initial solution	41,203	676	23,080	13,647	3,800
Iteration 1	40,978	676	22,949	13,647	3,706

For Network 1 and Network 2, ten additional cases were constructed by varying the flows of the networks. The number of parcels of a flow is randomly generated from a normal distribution with mean equal to the number of parcels of this flow in the original network and standard deviation equal to a quarter of the mean. Also for the simulated cases of Network 2 we are only able to apply the column generation algorithm with low threshold values due to memory issues. Hence, we do not show the results of these experiments. For Network 1 the results of the simulated cases, shown in Table 6, are similar to the results shown in Table 4. The minimum, average and maximum gap between the lower bound and the solution of the DH/HD-problem in the first iteration are equal to 2%, 2.6% and 3.3%, respectively. Only the first iteration is presented as also for the simulated case we are not able to improve the movement scheme by using the iterative procedure.

6   Conclusions and Recommendations

This paper proposes a solution approach to schedule movements in the network of an express service provider. Hereby, we simultaneously take into account a heterogeneous vehicle fleet, different service types, exact timing of movements and parcels, and routes via several depots and hubs. The solution approach consists of a column generation algorithm, a local search algorithm and an iterative procedure.

The solution approach is tested on three data instances, of which one is a simplified network and the other two are modified instances of networks of an express service provider. We show that using our solution approach the total transportation cost is decreased by 3.4%, 12.9% and 18.6%, respectively.

The results indicate that our solution approach is well suitable to apply to small- and middle-sized networks. Due to memory issues we are not able to apply the solution approach with the optimal parameter settings to the large network. However, the results look promising and we expect that by using another implementation of the solution approach good results can be obtained for large networks as well.

For our test instances, the best movement scheme is always found in the initial solution or in the first iteration. This means that the computation time needed to find the best solution is limited. In case longer computation times are allowed as well, the iterative procedure can be used to decrease the transportation cost in subsequent iterations of the algorithm. Considering different rules for the iterative procedure is a subject for further research.

Another topic for further research is the pricing problem of the column generation algorithm. First, we use an approximation algorithm to solve the pricing problem. To prove optimality, the pricing problem has to be solved exactly. Second, in the shortest path problem we use an approximation of the transportation cost of the composites as it is not possible to calculate the exact cost beforehand without knowing which flows are on the composite. A topic for further research is whether it is possible to reformulate the pricing problem and take the exact transportation cost into account.

References

E.H.L. Aarts and J.K. Lenstra. *Local search in combinatorial optimization*. Princeton University Press, Princeton, 2003.

A.P. Armacost, C. Barnhart, and K.A. Ware. Composite variable formulations for express shipment service network design. *Transportation Science*, 36(1):1–20, 2002.

A.P. Armacost, C. Barnhart, K.A. Ware, and A.M. Wilson. Ups optimizes its air network. *Interfaces*, 34(1):15–25, 2004.

C. Barnhart and R.R. Schneur. Air network design for express shipment service. *Operations Research*, 44(6):852–863, 1996.

C. Barnhart, E.L. Johnson, G.L. Nemhauser, M.W.P. Savelsbergh, and P.H. Vance. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 46(3):316–329, 1998.

T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. *Introduction to algorithms*. The MIT University Press, Cambridge, 2001.

T.G. Crainic. Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272–288, 2000.

P.J. Dejax and T.G. Crainic. Survey papera review of empty flows and fleet management models in freight transportation. *Transportation Science*, 21(4):227–248, 1987.

J.M. Farvolden and W.B. Powell. Subgradient methods for the service network design problem. *Transportation Science*, 28(3):256–272, 1994.

K. Holmberg and J. Hellstrand. Solving the uncapacitated network design problem by a lagrangean heuristic and branch-and-bound. *Operations Research*, 46(2):247–259, 1998.

J.M.Y. Leung, T.L. Magnanti, and V. Singhal. Routing in point-to-point delivery systems: Formulations and solution heuristics. *Transportation science*, 24(4):245–260, 1990.

M.E. Lübbecke and J. Desrosiers. Selected topics in column generation. *Operations Research*, 53(6):1007–1023, 2005.

T.L. Magnanti and P. Mirchandani. Shortest paths, single origin-destination network design, and associated polyhedra. *Networks*, 23(2):103–121, 1993.

T.L. Magnanti and R.T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18(1):1–55, 1984.

I. Meuffels, H. Fleuren, J. Poppelaars, H. Hoornenborg, and F De Rooij. The design of express networks in a nutshell - playing the global optimisation game (go-game). *OR News*, 39:6–8, 2010.

M.B. Pedersen, T.G. Crainic, and O.B.G. Madsen. Models and tabu search metaheuristics for service network design with asset-balance requirements. *Transportation Science*, 43(2):158–177, 2009.