Substitution-based Equipment Balancing in Service Networks with Multiple Equipment Types

Yu Yang[1], Yassine Ridouane, Natashia Boland, Alan Erera, Martin Savelsbergh

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Dr NW, Atlanta, GA 30332

Abstract

Package express companies routinely operate multiple equipment types in their service networks. Nevertheless, the weekly schedule of movements used to transport packages through the network leads to changes in equipment inventory at the facility level, which will hinder their normal operations if it is not carefully taken care of. Traditional rebalancing seeks to move the equipment empty in the most cost efficient way. In contrast, we investigate a substitution-based equipment balancing which has been rarely considered. More specifically, we try to reduce this change, i.e., the equipment imbalance associated with the schedule of movements, by substituting the equipment types initially assigned to movements. Consequently, substantial unnecessary empty repositioning needed to restore balance can be avoided at the end. We conduct complexity analysis of the underlying optimization problems, i.e., finding the minimum imbalance of the network and minimizing the number of substitutions required to achieve the minimum network imbalance. In addition, we develop integer programming (IP) models and propose an efficient two-phase decomposition heuristic for solution. Furthermore, we perform a computational study using real-world instances to analyze the performance of the IP solution approach and assess the benefits of substitution-based equipment balancing.

Keywords: logistics, equipment balancing, complexity, decomposition

1. Introduction

Package express companies, such as FedEx and United Parcel Service, use a large and heterogeneous pool of trailers and containers in their service (linehaul) networks. A major challenge in the planning process is to ensure that the right equipment is available at the right location at the right time. This is difficult to achieve, in part, because the flow of packages between facilities in the network is not balanced. As a consequence, the companies are forced to move equipment empty, i.e., reposition equipment, which is expensive.

To reduce the complexity of the planning process, a large package express carrier typically uses a phased

Email addresses: yu.yang@ise.ufl.edu (Yu Yang), yridouane@gatech.edu (Yassine Ridouane), natashia.boland@isye.gatech.edu (Natashia Boland), alan.erera@isye.gatech.edu (Alan Erera), martin.savelsbergh@isye.gatech.edu (Martin Savelsbergh)

approach. In an initial flow planning phase, a forecast of daily origin-destination demand is used to determine origin-destination paths for packages that meet service commitments and that create consolidation opportunities (consolidation is the primary mechanism a package express carrier employs to reduce operational costs). In a load planning phase, the package flows are converted into loads, i.e., timed movements of mobile equipment (trailers or containers) loaded with shipments between terminals in the network. Some empty equipment movement decisions may also be determined in this phase. In a scheduling phase, driver dispatch schedules are created to actually move the loaded and empty equipment in the load plan; note that a load might be moved directly from its origin to destination with a single driver, or it may alternately move through one or more relay points by potentially multiple drivers. A planned set of driver schedules typically covers a period of a week and must satisfy many requirements, e.g., hours-of-service regulations and union contract rules.

In the current planning process of most major carriers, load planning may lead to *imbalances* in equipment availability over time. One common approach to measuring equipment imbalance is to consider differences in the inventory of different equipment types at a terminal at the start of the week and at the end of the week. A surplus or deficit (positive or negative change) in the inventory level of a specific equipment type at a specific terminal is called imbalance, and the total imbalance of a plan can be assessed by summing the absolute values of the imbalances for all equipment types at all terminals. When customer demand and planned operations are similar week after week, it may be desirable for plans to have as little total imbalance as possible to avoid a surplus or a shortage of equipment over time. Thus, a final phase in the planning process described above seeks to reduce the equipment imbalance associated with a load plan; this phase is the topic of this paper. More specifically, we try to decrease the total imbalance of a plan by substituting the equipment types assigned to planned loads. Equipment substitution complements the empty repositioning of equipment, but it is only possible if companies operate multiple, exchangeable equipment types. When compared with repositioning empty equipment, equipment substitution may be preferred since this may not require any additional operating cost.

In ground transportation networks, package express carriers employ operate different types of equipment, both trailers and containers, which are grouped into categories based on their size: *shorts* (trailers with a length of 28 feet), *longs* (trailers with a length ranging from 40 to 48 feet), and *extra longs* (trailers with a length of 53 feet). Equipment types can be combined into composite types. For example, a common composite type is a combination of two shorts pulled by a single truck tractor. Other composite types are three shorts or a long combined with a short. In general, the equipment assigned to a load can be substituted by a larger type as long as the origin and destination facility of the load can accommodate the new type. In some situations, an equipment type assigned to a load can be substituted by a smaller equipment type, but only if the capacity of the smaller equipment type is sufficient to accommodate the original load quantity.

When carriers develop operational load plans and schedules, they may have equipment imbalances. It should be noted that simpler trucking operations often do not face equipment imbalance problems; for example, a carrier that only uses a single trailer type for all loads and operates driver schedules that are cycles can avoid any imbalance by simply ensuring that a driver is always moving a loaded or empty trailer. More complex package carrier operations lead to imbalance when the number of loads and empties departing from a facility with a specific equipment type is different from the number of loads and empties arriving at the facility with that equipment type. As described earlier, we define the imbalance of a facility to be the sum of the absolute imbalances (surplus or deficit) across all of the equipment types at the facility, and the total system imbalance as the sum of the facility imbalances in the network. The primary goal of substitution-based equipment balancing is to minimize the total imbalance of a plan. A secondary goal is to achieve the minimum total imbalance with as few equipment substitutions as possible.

The main contributions of our research are:

- We introduce a new substitution-based equipment balancing problem for transportation carriers that operate multiple trailer and container equipment types and develop an effective integer programming (IP) based approach for its solution;
- We determine the complexity of the underlying optimization problems, i.e., finding the minimum imbalance and minimizing the number of substitutions required to achieve the minimum imbalance;
- We present a simple, but effective decomposition heuristic that yields high-quality solutions in a short amount of time; and
- We conduct a computational study, using real-world instances, to assess the efficacy of our solution approaches and to analyze the benefits of substitution-based equipment balancing.

The remainder of the paper is organized as follows. In Section 2, we review related literature. In Section 3, we introduce notation used throughout the rest of the paper and present IP formulations. Section 4 is devoted to detailed complexity analysis of the two optimization problems under several simplified settings. Section 5 discusses the computational study performed on real-world instances. Lastly, we provide some concluding remarks in Section 6.

2. Literature Review

The existing literature on equipment management in the trucking industry focuses on the design of empty repositioning strategies to balance equipment (also referred to as "empty vehicle allocation" or "redistribution"), e.g., Du and Hall (1997), Jansen et al. (2004), Song (2005, 2007), Erera et al. (2009), and Long et al. (2012). We are not aware of any literature on approaches based on equipment substitution to deal with equipment imbalance in the trucking industry. Dejax and Crainic (1987) present an overview of empty fleet management issues and strategies and introduce a taxonomy of empty flow problems related to the distribution and scheduling of empty movements. The strategies for empty equipment redistribution can be classified into two groups: (i) decentralized models where one facility operates and controls its own fleet and seeks to optimize its own performance metrics, and (ii) centralized models where an entire service network is considered and decisions are made that seek to optimize a set of global performance metrics. An example in the first group is Du and Hall (1997), who proposes a decentralized stock control policy approach for both fleet sizing and empty repositioning restricted to a given center-terminal system (such system is comprised of one center and a group of terminals connected to it). More examples of this type can be found in Song (2005, 2007). An example in the second group is Jansen et al. (2004). It describes the operational planning system POP that performs transportation planning of on average 4000 container-orders a day on trains and trucks in Germany. The objective is to minimize total cost of transportation, which is the sum of the cost of all tours. Demand uncertainty, i.e., uncertainty of anticipated future freight flows, greatly affects empty repositioning and therefore robust or stochastic optimization models are frequently used for these problems. Erera et al. (2009), for example, propose a robust recovery optimization framework that can be applied to empty repositioning problems. Long et al. (2012) introduce a two-stage stochastic programming model for an environment with uncertain demand for and supply of containers, in which the objective is to minimize the costs of repositioning containers empty. They propose a sample average approximation method using a progressive hedging heuristic.

Equipment management in other industries has also been extensively studied. For a maritime logistics setting, Li et al. (2007) consider the allocation of empty containers from supply ports to demand ports and heuristic methods are developed to obtain solutions. Boile et al. (2008) propose a mixed integer programming (MIP) formulation to minimize the cost of repositioning empty maritime shipping containers in a region by finding the optimal locations for inland depots. Chang et al. (2008) formulate an empty container substitution problem to minimize the cost of transporting empty containers. In their model, substitutions are allowed between container types based on their intended use, dimensions, and ownership. The problem is divided into dependent and independent parts and a branch-and-bound (B&B) method is applied to the dependent part. When uncertainty is involved, various approaches have been proposed to obtain the optimal policy under different settings. For example, Song and Zhang (2010) propose a fluid flow model for a single port with stochastic demands modeled by a two-state Markov chain. Zhang et al. (2014) consider multiple ports over multi-periods with stochastic demands and lost sales. Xie et al. (2017) investigate the dependence of the optimal policy on the initial inventories for a two-depot system and provide a game theory perspective. Legros et al. (2019) study the potential for consignees to manage an inventory of empty containers at their location which may be directly reused by consignors in the neighborhood when needed. Lee and Moon (2020) study the empty container repositioning problem considering foldable containers under uncertainty and propose a robust formulation that is computationally tractable. Lu et al. (2020) consider the joint decisions on pricing and empty container repositioning and develop a stochastic dynamic programming model.

It is worth noting that there are significant differences between the trucking industry and ocean transportation. Firstly, in ocean transportation, the freight flows are much more unbalanced (Theofanis and Boile, 2009), i.e., the flows are mostly one-way, thus it highly relies on empty repositioning. However, in the trucking industry, where the flows are more balanced, there is much more room for substitution to reduce imbalance instead of resorting to empty repositioning. That does not mean imbalance can be restored solely by substitution and, as mentioned, we may still need to add empty loads to fully restore balance subsequently. Secondly, the number of equipment types, i.e., containers, is relatively small in the ocean scenario. More specifically, in Chang et al. (2008), only three types of dry cargo containers with standard dimensions are considered. However, in the trucking industry, a company like UPS operates more than ten equipment types which are considered in our numerical study. As suggested by our complexity analysis and numerical experiments, the number of equipment types plays a key role in defining the problem difficulty. Thus, our problem tends to be much more difficult to solve. Lastly, substitution may invalidate an initially feasible way of packing the containers. In contrast, all trailers are compatible with the equipment (containers) considered. As a result, substitution will not cause extra costs and is more practical in the trucking industry. For other industries such as the emerging bike-sharing industry, there may not be multiple equipment types involved and thus substitution is not an option there.

In the emerging bike-sharing industry, the rebalancing is usually conducted by using trucks to carry bikes from where there is surplus to where there is deficit. The problem can thus be viewed as a special one-commodity pickup-and-delivery capacitated vehicle routing problem. Dell'Amico et al. (2014) propose several MIP formulations and branch-and-cut algorithms to obtain solutions. An effective solution method based on a clustering strategy is developed in Lv et al. (2020) to solve the same problem. Pal and Zhang (2017) study an innovative bike sharing model known as free-floating bike sharing, and propose an efficient heuristic that is tested on the one-commodity pickup and delivery traveling salesman problem. Schuijbroek et al. (2017) simultaneously determine service level requirements at each bike sharing station, and design (near-)optimal vehicle routes to rebalance the inventory. A new cluster-first route-second heuristic is proposed and shown to outperform a pure MIP formulation and a constraint programming approach. A Markov decision process approach is applied in Legros (2019) to decide at any point of time which station should be prioritized and which number of bikes should be added or removed at each station. Haider et al. (2018) try to rebalance the inventory through price incentives/disincentives which partially or fully obviates the need for a manual repositioning operation. The high-level idea shares some similarity with our work, i.e., reducing imbalance before it actually happens and thus partially getting around the need for costly repositioning efforts. The difference is that they resort to an iterative price adjustment scheme that could incur some revenue loss although the overall operational cost is reduced. In contrast, our substitution approach does not have this side effect since, as mentioned, substitutions can be done at no cost.

Lastly, in the ride-sharing industry, Ma et al. (2019) consider a ride-sharing strategy where a queueing-theoretic vehicle dispatch and idle vehicle relocation algorithms are customized for the problem. An online algorithm is developed and evaluated on both a synthetic instance and a case study of Long Island commuters to New York City. He et al. (2020) propose a distributionally robust optimization (DRO) approach that can incorporate demand temporal dependence to dynamically match the vehicle supply and travel demand for a free-float vehicle sharing system. Pouls et al. (2020) develop a forecast-driven repositioning algorithm, the core part of which is a novel MIP model. The approach is evaluated through extensive simulation studies on real-world datasets from Hamburg, New York City, and Manhattan. Deep reinforcement learning methods are developed for ride-sharing dispatching and repositioning in Holler et al. (2019) and Qin et al. (2019).

3. Notation and Formulations

We first introduce the notation used throughout the paper, which is also summarized in Table 1 for the sake of readability. A network N is represented as a directed graph, $N = (V, A)$, with each vertex representing a facility and each arc representing a load. A *load* represents a movement of equipment that is scheduled to dispatch during the planning horizon and deliver a quantity of packages for an origin-destination pair. It is also characterized by the initial equipment type assigned to the load, and this type is used to compute the initial imbalance. Let $\mathbb{Z}_{\geq 0}$ be the set of non-negative integers, $n = |V|$ be the number of vertices, and $m = |A|$ be the number of arcs. Let $\mathcal{E}$ be the set of basic equipment types and $\mathcal{C}$ be the set of equipment type configurations used operationally, which can be a basic type or a composite type formed by combining basic types. For $c \in \mathcal{C}$, $f_{ce} \in \mathbb{Z}_{\geq 0}$ indicates how many units of basic equipment type $e \in \mathcal{E}$ are used in the equipment type c. For each $v \in V$, let $\mathcal{C}_v \subseteq \mathcal{C}$ be the set of allowable equipment types at vertex v, and $\delta_v^+ = \{(v, u) \in A : u \in V\}$ and $\delta_v^- = \{(u, v) \in A : u \in V\}$ be the sets of outgoing and incoming arcs at v, respectively. Let $\sigma_v^+ := |\delta_v^+|$ and $\sigma_v^- := |\delta_v^-|$. An equipment assignment (or assignment for short) is a function $\mathscr{A} : A \to \mathcal{C}$ that assigns an equipment type to each arc. The initial assignment is denoted by $\mathscr{A}_0$. Let $\mathcal{C}_a \subseteq \mathcal{C}$ be the set of equipment types that can be assigned to arc $a \in A$, $A_{vc}^+ := \{a \in \delta_v^+ : c \in \mathcal{C}_a\}$, and $A_{vc}^- := \{a \in \delta_v^- : c \in \mathcal{C}_a\}$. For a given network N, the imbalance of a given assignment $\mathscr{A}$ is calculated by $I(\mathscr{A}) := \sum_{c \in \mathcal{C}} f_{ce} \cdot \left| \sum_{a \in A_{vc}^+} h_c\big(\mathscr{A}(a)\big) - \sum_{a \in A_{vc}^-} h_c\big(\mathscr{A}(a)\big) \right|$, where $h_c : \mathcal{C} \to \{0, 1\}$ is the indicator function that $h_c(x) = 1$ if $x = c$ and $h_c(x) = 0$ if $x \neq c$. Suppose I^* is the minimum imbalance of network N, then we say $\mathscr{A}$ is optimal for N when $I(\mathscr{A}) = I^*$.

Our goal is to find an optimal $\mathscr{A}^*$ that is closest to $\mathscr{A}_0$, i.e., $\mathscr{A}^* \in \operatorname{argmin}_{I(\mathscr{A}) = I^*} \|\mathscr{A} - \mathscr{A}_0\|$, where $\|\mathscr{A} - \mathscr{A}_0\| = |\{a \in A : \mathscr{A}(a) \neq \mathscr{A}_0(a)\}|$. We use a two-stage hierarchical optimization approach, where we compute I^* for network N in Stage 1, and find the desired $\mathscr{A}^*$ in Stage 2.

Notation	Meaning
$\mathbb{Z}_{\geq 0}$	set of non-negative integers
V	set of vertices in the network
A	set of arcs in the network
n	number of vertices
m	number of arcs
$\mathcal{E}$	set of basic equipment types
$\mathcal{C}$	set of equipment type configurations used
f_{ce}	units of basic equipment type $e \in \mathcal{E}$ used in type c
$\mathcal{C}_v$	set of allowable equipment types at vertex v
δ_v^+ (δ_v^-)	set of outgoing (incoming) arcs at v
σ_v^+ (σ_v^-)	number of outgoing (incoming) arcs at v
$\mathscr{A} : A \to \mathcal{C}$	function that assigns an equipment type to each arc
$\mathscr{A}_0$	initial assignment function
$\mathcal{C}_a$	set of equipment types that can be assigned to arc a
A_{vc}^+ (A_{vc}^-)	set of outgoing (incoming) arcs that can be assigned type c
I^*	minimum imbalance of the network
$I(\mathscr{A})$	imbalance of the assignment $\mathscr{A}$

Table 1: Notation used throughout the paper.

3.1. Finding the Minimum Imbalance

We now define the decision variables. For $a \in A$ and $c \in \mathcal{C}_a$, let

$$y_{ac} = \begin{cases} 1, & \text{if equipment } c \text{ is used on arc } a, \\ 0, & \text{otherwise.} \end{cases}$$

For $v \in V$, $e \in \mathcal{C}_v \cap \mathcal{E}$, let $r_{ve} \in \mathbb{Z}_{\geq 0}$ be the imbalance for basic equipment type e at vertex v. The following model minimizes the total imbalance I^*:

$$I^* = \min \quad \sum_{v \in V} \sum_{e \in \mathcal{C}_v \cap \mathcal{E}} r_{ve} \tag{3.1}$$

$$\text{s.t.} \quad \sum_{c \in \mathcal{C}} f_{ce} \left(\sum_{a \in A_{vc}^+} y_{ac} - \sum_{a \in A_{vc}^-} y_{ac} \right) \leq r_{ve}, \quad v \in V,\ e \in \mathcal{C}_v \cap \mathcal{E}, \tag{3.2}$$

$$\sum_{c \in \mathcal{C}} f_{ce} \left(\sum_{a \in A_{vc}^-} y_{ac} - \sum_{a \in A_{vc}^+} y_{ac} \right) \leq r_{ve}, \quad v \in V,\ e \in \mathcal{C}_v \cap \mathcal{E}, \tag{3.3}$$

$$\sum_{c \in \mathcal{C}_a} y_{ac} = 1, \quad a \in A, \tag{3.4}$$

$$y_{ac} \in \{0, 1\}, \quad a \in A,\ c \in \mathcal{C}_a. \tag{3.5}$$

Constraints (3.2) and (3.3) ensure that r_{ve} is set to the net surplus or deficit of equipment type e at vertex v, while Constraint (3.4) guarantees that exactly one equipment type is assigned to each arc (i.e., the assigned equipment type remains the same or is replaced by exactly one other equipment type). Note that the minimum imbalance induced by a plan depends on both the loads and the initial assignment, $\mathscr{A}_0$, since $\mathcal{C}_a$ depends on $\mathscr{A}_0$.

3.2. Minimizing the Number of Changes

In Stage 2, we minimize the number of changes Ω and adding Constraint (3.6) ensures that the obtained equipment assignment is an optimal assignment:

$$
\begin{aligned}
\Omega = \ & \min \sum_{a \in A} (1 - y_{a\mathscr{A}_0(a)}) \\
& \text{s.t.} \\
& \qquad \sum_{v \in V} \sum_{e \in \mathcal{C}_v \cap \mathcal{E}} r_{ve} \leq I^*, \qquad (3.6) \\
& \qquad (3.2), (3.3), (3.4), (3.5).
\end{aligned}
$$

Note that $\Omega < m$, and, thus, the two optimization models can be combined into a single optimization model as follows:

$$
\begin{aligned}
\min \quad & m \left(\sum_{v \in V} \sum_{e \in \mathcal{C}_v \cap \mathcal{E}} r_{ve} \right) + \sum_{a \in A} (1 - y_{a\mathscr{A}_0(a)}) \\
\text{s.t.} \quad & (3.2), (3.3), (3.4), (3.5).
\end{aligned}
$$

The optimization model will first try to minimize the total imbalance, computed as $\sum_{v \in V} \sum_{e \in \mathcal{C}_v \cap \mathcal{E}} r_{ve}$ – since a decrease of one unit will result in a decrease of m in the objective value, which is larger than the second term $\sum_{a \in A}(1 - y_{a\mathscr{A}_0(a)})$. After achieving the minimum imbalance, the model seeks to minimize the number of changes required and thus minimizes the overall objective value. However, for real-life instances, m can be as large as hundreds of thousands, which makes it more difficult to solve than the two-stage hierarchical optimization model.

4. Complexity of Equipment Substitution Problems

In this section, we analyze the computational complexity of some simplified settings of the equipment substitution problem. For simplicity, we only consider basic equipment types, i.e., $\mathcal{C} = \mathcal{E}$ and an assignment $\mathscr{A}$ can be viewed as a function $A \to \mathcal{E}$. We consider two variants: one where there is *full interchangeability*, i.e., the equipment assigned to all arcs in the network can be changed, and the other where there is *partial interchangeability*, i.e., the equipment assigned to a given subset of arcs cannot be changed. Furthermore, we consider settings with two and three equipment types, which are sufficient for us to establish the complexity results for even more general cases.

The findings, to be discussed next, are summarized in Table 2.

	full interchangeability		partial interchangeability	
	2 equipment types	3 equipment types	2 equipment types	3 equipment types
finding the minimum imbalance	P	P	P	NP-hard
minimizing number of changes	P	NP-hard	P	NP-hard

Table 2: Complexity of equipment substitution problems in different simplified settings.

For networks with full interchangeability, the following claim gives a necessary and sufficient condition for an assignment function $\mathscr{A}$ to be optimal. Let $\sigma_{ve}^+ = |\{a \in \delta_v^+ : \mathscr{A}(a) = e\}|$ and $\sigma_{ve}^- = |\{a \in \delta_v^- : \mathscr{A}(a) = e\}|$.

Claim 1. *Given a network N with full interchangeability and an assignment $\mathscr{A}$, then $I(\mathscr{A}) = I^*$ if and only if for all $v \in V$ and for all $e \in \mathcal{E}$, $(\sigma_{ve}^+ - \sigma_{ve}^-)(\sigma_v^+ - \sigma_v^-) \geq 0$ if $\sigma_v^+ - \sigma_v^- \neq 0$ and $\sigma_{ve}^+ - \sigma_{ve}^- = 0$ if $\sigma_v^+ - \sigma_v^- = 0$.*

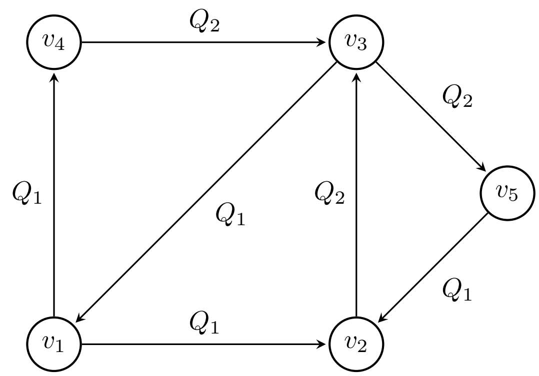
Proof. For each node v, the minimum imbalance is at least $|\sigma_v^+ - \sigma_v^-|$, which implies $\sum_{v \in V} |\sigma_v^+ - \sigma_v^-|$ is a lower bound of I^*. It is easy to see that this bound can be achieved by $\mathscr{A}(a) := e$, $\forall a \in A$ and a fixed $e \in \mathcal{E}$. By definition, for all $v \in V$, $\sum_{e \in \mathcal{E}} \sigma_{ve}^+ = \sigma_v^+$ and $\sum_{e \in \mathcal{E}} \sigma_{ve}^- = \sigma_v^-$, which implies $\sum_{e \in \mathcal{E}}(\sigma_{ve}^+ - \sigma_{ve}^-) = \sigma_v^+ - \sigma_v^-$. Now $I(\mathscr{A}) = I^*$ if and only if for all $v \in V$, $\sum_{e \in \mathcal{E}} |\sigma_{ve}^+ - \sigma_{ve}^-| = |\sigma_v^+ - \sigma_v^-|$. This, in turn, is true if and only if for all $e \in \mathcal{E}$, we have that $(\sigma_{ve}^+ - \sigma_{ve}^-)(\sigma_v^+ - \sigma_v^-) \geq 0$ when $\sigma_v^+ - \sigma_v^- \neq 0$, and that $\sigma_{ve}^+ - \sigma_{ve}^- = 0$ when $\sigma_v^+ - \sigma_v^- = 0$. ■

4.1. Two-Equipment Type Networks With Full Interchangeability

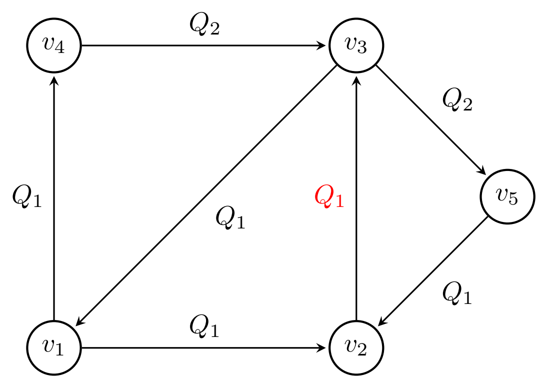
Given a network N with two equipment types, i.e., $\mathcal{E} = \{Q_1, Q_2\}$, and full interchangeability, it is easy to see that both $\mathscr{A}(a) = Q_1$ for all $a \in A$ and $\mathscr{A}(a) = Q_2$ for all $a \in A$ achieve minimum imbalance. However, it is not obvious how to find an assignment that minimizes the number of changes required to achieve minimum imbalance. We start by giving an example that demonstrates that the natural greedy heuristic, which identifies in each iteration an arc for which swapping the equipment type reduces the imbalance the most and then performs that swap, does not necessarily yield the desired $\mathscr{A}^*$. Note that it can be easily modified to provide a counterexample for two-equipment type networks with partial interchangeability. If the equipment on (v_1, v_2) is not allowed to be changed, the arguments still hold.

4.1.1. A Counterexample for the Greedy Heuristic

In Iteration 1, the greedy heuristic may change the equipment type on (v_2, v_3) from Q_2 to Q_1 as it reduces the imbalance by 4 (the maximum possible).

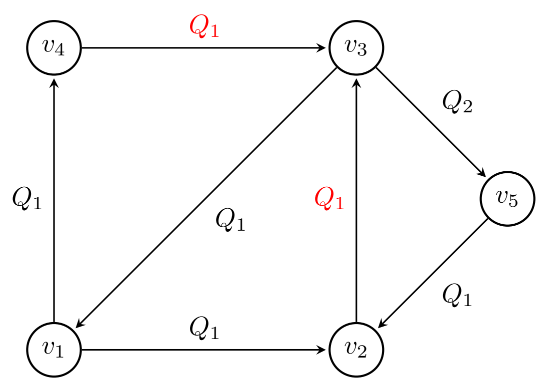


Imbalance	Q_1	Q_2
v_1	1	0
v_2	2	1
v_3	1	1
v_4	1	1
v_5	1	1
Total	10	



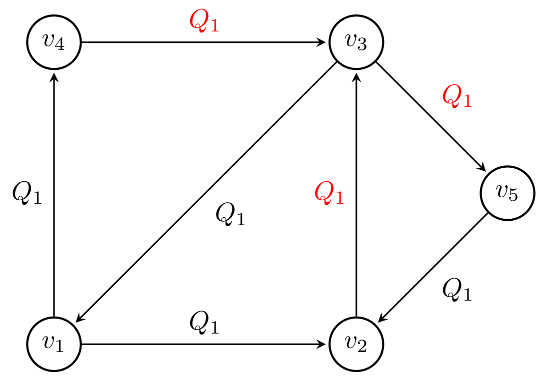
Imbalance	Q_1	Q_2
v_1	1	0
v_2	1	0
v_3	0	0
v_4	1	1
v_5	1	1
Total	6	

In Iteration 2, the greedy heuristic may change the equipment type on (v_4, v_3) from Q_2 to Q_1 even though it does not reduce the imbalance.



Imbalance	Q_1	Q_2
v_1	1	0
v_2	1	0
v_3	1	1
v_4	0	0
v_5	1	1
Total	6	

Finally, in Iteration 3, the greedy heuristic will change the equipment type on (v_3, v_5) from Q_2 to Q_1, which reduces the imbalance by 4. The minimum imbalance is achieved in three iterations.



Imbalance	Q_1	Q_2
v_1	1	0
v_2	1	0
v_3	0	0
v_4	0	0
v_5	0	0
Total	2	

However, it is easy to see that the minimum imbalance can be reduced to 2 by only 2 equipment type changes: change the equipment type on (v_5, v_2) from Q_1 to Q_2 and on (v_4, v_3) from Q_2 to Q_1.

To show that an optimal assignment closest to the initial assignment can be found in polynomial time, we formulate the problem as an integer program and show that the coefficient matrix is totally unimodular (TU). Because the TU property guarantees that solving the linear programming (LP) relaxation will yield an optimal integer solution and LP has been proved to be polynomially solvable by L. G. Khachiyan in 1979.

4.1.2. IP Formulation and Polynomial Solvability

Given that there are only two equipment types, a simpler formulation for the Stage 2 problem can be constructed. Let $A_1 := \{a \in A : \mathscr{A}_0(a) = Q_1\}$ and $A_2 := \{a \in A : \mathscr{A}_0(a) = Q_2\}$, and for all $a \in A$ let

$$x_a = \begin{cases} 1 & \text{if equipment type } Q_1 \text{ is assigned to arc } a, \\ 0 & \text{otherwise.} \end{cases}$$

Minimizing the number of changes required to achieve the minimum imbalance can be modeled as follows

$$\min \quad \sum_{a \in A_1} (1 - x_a) + \sum_{a \in A_2} x_a$$

$$\text{s.t.} \quad 0 \leq \sum_{a \in \delta_v^+} x_a - \sum_{a \in \delta_v^-} x_a \leq \sigma_v^+ - \sigma_v^-, \quad v \in V, \sigma_v^+ > \sigma_v^-, \tag{4.1}$$

$$\sigma_v^+ - \sigma_v^- \leq \sum_{a \in \delta_v^+} x_a - \sum_{a \in \delta_v^-} x_a \leq 0, \quad v \in V, \sigma_v^+ < \sigma_v^-, \tag{4.2}$$

$$\sum_{a \in \delta_v^+} x_a - \sum_{a \in \delta_v^-} x_a = 0, \quad v \in V, \sigma_v^+ = \sigma_v^-, \tag{4.3}$$

$$x_a \in \{0, 1\}, \quad a \in A.$$

Validity of the formulation follows from the fact that Constraints (4.1), (4.2), and (4.3) guarantee that the condition in Claim 1 is satisfied. Actually, (4.1) implies that when $\sigma_v^+ - \sigma_v^- > 0$, we have that $\sigma_{vQ_1}^+ - \sigma_{vQ_1}^- =$

$\sum_{a\in\delta_v^+} x_a - \sum_{a\in\delta_v^-} x_a \geq 0$ and $\sigma_{vQ_2}^+ - \sigma_{vQ_2}^- = \sigma_v^+ - \sigma_v^- - \left(\sum_{a\in\delta_v^+} x_a - \sum_{a\in\delta_v^-} x_a\right) \geq 0$. Similarly, (4.2) ensures that when $\sigma_v^+ - \sigma_v^- < 0$, both $\sigma_{vQ_1}^+ - \sigma_{vQ_1}^- \leq 0$ and $\sigma_{vQ_2}^+ - \sigma_{vQ_2}^- \leq 0$ hold. When $\sigma_v^+ - \sigma_v^- = 0$, (4.3) ensures that $\sigma_{vQ_1}^+ - \sigma_{vQ_1}^- = \sigma_{vQ_2}^+ - \sigma_{vQ_2}^- = 0$. In addition, the objective function forces the solution to be closest to the initial assignment $\mathscr{A}_0$.

The above IP can be written in the following form

$$\begin{aligned} \min \quad & w^T x \\ \text{s.t.} \quad & l \leq Lx \leq u, \\ & x \in \{0,1\}^m, \end{aligned}$$

where L is the node-arc incidence matrix of the network. Because a node-arc incidence matrix of a network is TU (Wolsey and Nemhauser, 1999) and $l, u \in \mathbb{Z}^n$, the LP relaxation will return an integral solution. Since LPs can be solved in polynomial time, we have shown its polynomial solvability.

4.2. Two-Equipment Type Networks With Partial Interchangeability

As before, let $A_1 := \{a \in A : \mathscr{A}_0(a) = Q_1\}$ and $A_2 := \{a \in A : \mathscr{A}_0(a) = Q_2\}$. In this setting, we assume that the equipment types on the two given sets $X_1 \subseteq A_1$ and $X_2 \subseteq A_2$ cannot be changed, i.e., the equipment type Q_1 assigned to arcs in X_1 cannot be changed to Q_2, and the equipment Q_2 assigned to arcs in X_2 cannot be changed to Q_1.

4.2.1. Finding the Minimum Imbalance

Determining the minimum imbalance I^*, when certain equipment type assignments cannot be changed, is no longer trivial, but can be accomplished using the following IP:

$$\min \quad \sum_{v\in V}(r_{v1} + r_{v2})$$

$$\text{s.t.} \quad \sum_{a\in\delta_v^+\setminus(X_1\cup X_2)} x_a + |\delta_v^+ \cap X_1| - \sum_{a\in\delta_v^-\setminus(X_1\cup X_2)} x_a - |\delta_v^- \cap X_1| \leq r_{v1}, \quad v \in V, \tag{4.4}$$

$$\sum_{a\in\delta_v^+\setminus(X_1\cup X_2)} (1 - x_a) + |\delta_v^+ \cap X_2| - \sum_{a\in\delta_v^-\setminus(X_1\cup X_2)} (1 - x_a) - |\delta_v^- \cap X_2| \leq r_{v2}, \quad v \in V, \tag{4.5}$$

$$\sum_{a\in\delta_v^-\setminus(X_1\cup X_2)} x_a + |\delta_v^- \cap X_1| - \sum_{a\in\delta_v^+\setminus(X_1\cup X_2)} x_a - |\delta_v^+ \cap X_1| \leq r_{v1}, \quad v \in V, \tag{4.6}$$

$$\sum_{a\in\delta_v^-\setminus(X_1\cup X_2)} (1 - x_a) - |\delta_v^- \cap X_2| - \sum_{a\in\delta_v^+\setminus(X_1\cup X_2)} (1 - x_a) - |\delta_v^+ \cap X_2| \leq r_{v2}, \quad v \in V, \tag{4.7}$$

$$x_a \in \{0,1\}, \quad a \in A \setminus (X_1 \cup X_2), \tag{4.8}$$

$$r_{v1}, r_{v2} \in \mathbb{R}, \quad v \in V, \tag{4.9}$$

where

$$x_a = \begin{cases} 1 & \text{if equipment type } Q_1 \text{ is assigned to arc } a, \\ 0 & \text{otherwise,} \end{cases}$$

and $r_{v1}, r_{v2} \in \mathbb{Z}_{\geq 0}$ compute the imbalance for equipment type Q_1, Q_2 at vertex v, respectively. Constraint (4.4) requires that r_{v1} is no smaller than the net incoming Q_1 at vertex v, computed by the total incoming minus total outgoing Q_1 at vertex v. In addition, constraint (4.6) requires that r_{v1} is greater than or equal to the net outgoing Q_1 at vertex v. Minimizing r_{v1} will thus force it to equal the imbalance of Q_1 at vertex v. Similarly, (4.5) and (4.7), combined with minimizing r_{v2} in the objective function, ensure that r_{v2} equals the imbalance of Q_2 at vertex v.

The LP relaxation of the above IP can be written in the following form:

$$\begin{aligned} \min \quad & (0, w) \begin{pmatrix} x \\ r \end{pmatrix} \\ \text{s.t.} \quad & \begin{pmatrix} B & -I \\ -B & -I \end{pmatrix} \begin{pmatrix} x \\ r \end{pmatrix} \leq \begin{pmatrix} d \\ -d \end{pmatrix}, \\ & 0 \leq x_a \leq 1, \quad a \in A \setminus (X_1 \cup X_2), \\ & r \in \mathbb{R}^{2n}, \end{aligned}$$

where $B = \begin{pmatrix} \hat{L} \\ -\hat{L} \end{pmatrix}$, and $\hat{L} = (l_a)_{a \in A \setminus (X_1 \cup X_2)}$ is a submatrix of the node-arc incidence matrix L.

Claim 2. *Polyhedron* $P := \left\{ (x, r) : \begin{pmatrix} B & -I \\ -B & -I \end{pmatrix} \begin{pmatrix} x \\ r \end{pmatrix} \leq \begin{pmatrix} d \\ -d \end{pmatrix}, \quad 0 \leq x \leq 1, r \in \mathbb{R} \right\}$ *has integral extreme points.*

Proof. B has $2n$ rows, and l_a is the column of L that corresponds to arc a. Since L is TU, so are $\hat{L}$ and B. Suppose (x^*, r^*) is an extreme point of P and $\mathcal{I} \subseteq [4n]$ is the index set of active constraints at (x^*, r^*). Note that if there exists a $j \in [2n]$ such that both j and $j + 2n$ are in $\mathcal{I}$, then $b_j x^* - u_j r^* = d_j$ and $-b_j x^* - u_j r^* = -d_j$, where b_j is the j-th row of B and u_j is the j-th unit vector. Then $u_j r^* = 0$ and only one of the constraints is necessary. Let $\mathcal{I}_1 = \{j : j \in \mathcal{I} \cap [2n], j+2n \notin \mathcal{I}\}$, $\mathcal{I}_2 = \{j : j \in \mathcal{I} \cap ([4n] \setminus [2n]), j - 2n \notin \mathcal{I}\}$, $\mathcal{I}_3 = \{j : j \in \mathcal{I} \cap [2n], j + 2n \in \mathcal{I}\}$, $H_1 = \{j : r^*_j = 0\}$, $H_2 = [2n] \setminus H_1$, $r^1 = (r^*_j)_{j \in H_1}$, $r^2 = (r^*_j)_{j \in H_2}$. Then (x^*, r^2) is an extreme point of

$$\hat{P} = \left\{ (x, y) : \begin{pmatrix} \widehat{B} & -I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \hat{d}, \quad 0 \leq x \leq 1, y \in \mathbb{R} \right\},$$

where $\widehat{B} = \begin{pmatrix} (b_j)_{j \in \mathcal{I}_1 \cup \mathcal{I}_3} \\ -(b_j)_{j \in \mathcal{I}_2} \end{pmatrix}$ is a submatrix of B, $\hat{d} = \begin{pmatrix} (d_j)_{j \in \mathcal{I}_1 \cup \mathcal{I}_3} \\ -(d_j)_{j \in \mathcal{I}_2} \end{pmatrix} \in \mathbb{Z}^{|\mathcal{I}_1|+|\mathcal{I}_2|+|\mathcal{I}_3|}$. Since B is TU, so are $\widehat{B}$ and $\begin{pmatrix} \widehat{B} & -I \end{pmatrix}$. Therefore, every extreme point of $\hat{P}$ is integral and (x^*, r^2) is integral. Consequently, (x^*, r^*) is integral (since $r^1 = 0$). ■

4.2.2. Minimizing the Number of Changes

The integrality property derived in the previous section relies on the structure of the coefficient matrix. Therefore, if we modify the objective, but keep the coefficient matrix the same, the solution to the LP relaxation will still be integral. As mentioned in Section 3, we can combine the optimization models of the two stages into a single optimization model by changing the objective. In the case of two equipment types this results in

$$\begin{aligned} \min \quad & \sum_{v \in V} m(r_{v,1} + r_{v,2}) + \sum_{a \in A_1 \setminus X_1} (1 - x_a) + \sum_{a \in A_2 \setminus X_2} x_a \\ \text{s.t.} \quad & (4.4), (4.5), (4.6), (4.7), (4.8), (4.9). \end{aligned}$$

Observing that $\sum_{a \in A_1 \setminus X_1} (1 - x_a) + \sum_{a \in A_2 \setminus X_2} x_a \leq |A_1 \setminus X_1| + |A_2 \setminus X_2| = |A \setminus (X_1 \cup X_2)| < m$, the optimal solution is guaranteed to achieve the minimum imbalance while minimizing the number of changes made to the original assignment.

4.3. Two-Equipment Type Models: Additional Results

A more general question, for a given $K \in \mathbb{Z}_{\geq 0}$ and a given $\mathscr{A}_0$, is whether an $\mathscr{A}'$ can be found such that $I(\mathscr{A}') \leq K$ and $\|\mathscr{A} - \mathscr{A}'\|$ is minimum in a network with two equipment types and either full or partial interchangeability. Next, we will show that to solve this problem, it suffices to solve K integer minimum cost circulation problems in an auxiliary network $G' = (V', E')$, which has size polynomial in the size of the original network G, and is constructed as follows.

We add three vertices $\{s, t, s'\}$ to V, and thus $V' = V \cup \{s, t, s'\}$. Let $E' = E \cup E^+ \cup E^- \cup \{s's, ts'\}$, where $E^+ = \{sv : \sigma_v^+ > \sigma_v^-, v \in V\}$ and $E^+ = \{vt : \sigma_v^- > \sigma_v^+, v \in V\}$. Define the capacity function $u : E' \to \mathbb{Z}_{\geq 0}$ and cost function $c : E' \to \mathbb{Z}$ as follows

$$u(e) = \begin{cases} \sigma_v^+ - \sigma_v^-, & \text{if } e \in E^+, \\ \sigma_v^- - \sigma_v^+, & \text{if } e \in E^-, \\ 1, & \text{if } e \in E, \\ k, & \text{if } e = s's \text{ or } ts', \end{cases} \qquad c(e) = \begin{cases} -1, & \text{if } e \in A_1, \\ 1, & \text{if } e \in A_2, \\ 0, & \text{if } e \in E^+ \cup E^- \cup \{s's, ts'\}, \end{cases}$$

where $k \in \mathbb{Z}_{\geq 0}, k \leq K$. It is easy to see that there exists an integer circulation in G'. Suppose we have a minimum cost integer circulation $f_k : E' \to \mathbb{Z}_{\geq 0}$ of G' and define the assignment function $\mathscr{A}^k$ as

$$\mathscr{A}^k(e) = \begin{cases} Q_1, & \text{if } e \in E, f_k(e) = 1, \\ Q_2, & \text{if } e \in E, f_k(e) = 0. \end{cases}$$

By construction, the imbalance of equipment Q_1 for $\mathscr{A}^k$ is less than or equal to k. Thus if we solve this problem for each $k \in \{0, 1, \ldots, K\}$, then we are able to find an $\mathscr{A}'$ such that $I(\mathscr{A}') \leq K$ and thus $T := \{0 \leq k \leq K : I(\mathscr{A}^k) \leq K\}$ is not empty. Let $k^* = \operatorname{argmin}_{k \in T} \|\mathscr{A}^k - \mathscr{A}^0\|$. Since for each k, f_k achieves the minimum cost, then $\mathscr{A}^{k^*}$ is closest to $\mathscr{A}_0$ with $I(\mathscr{A}^{k^*}) \leq K$ over all assignment functions. Due to the integrality of capacities, we can find the minimum cost integer circulation in strongly polynomial-time (Tardos, 1985). Given that $K \leq 2m$, finding the desired $\mathscr{A}' := \mathscr{A}^{k^*}$ can be done in polynomial time.

In case of partial interchangeability, the cost function has to be modified slightly:

$$c(e) = \begin{cases} -1, & \text{if } e \in A_1 \setminus X_1, \\ 1, & \text{if } e \in A_2 \setminus X_2, \\ -m, & \text{if } e \in X_1, \\ m, & \text{if } e \in X_2, \\ 0, & \text{if } e \in E^+ \cup E^- \cup \{s's, ts'\}, \end{cases}$$

where a large cost m or $-m$ is assigned to the arcs on which the equipment cannot be changed. If for some $0 \leq k \leq K$, the cost of f_k, $c(f_k) \geq m$, then the equipment on some of the fixed arcs has been changed, and thus the corresponding $\mathscr{A}^k$ is not feasible. Let $\hat{T} = \{0 \leq k \leq K : I(\mathscr{A}^k) \leq K, \ c(f_k) < m\}$. If $\hat{T} = \emptyset$, then the desired assignment function $\mathscr{A}'$ does not exist, otherwise, let $\hat{k}^* = \operatorname{argmin}_{k \in \hat{T}} \|\mathscr{A}^k - \mathscr{A}_0\|$, and $\mathscr{A}' := \mathscr{A}^{\hat{k}^*}$ is the desired assignment function.

4.4. Three-Equipment Type Networks

We call a network N balanced if for any vertex $v \in V$, $\sigma_v^+ = \sigma_v^-$. The following proposition characterizes the difficulty of finding an assignment function $\mathscr{A}$ such that $I(\mathscr{A}) = 0$ in a balanced three-equipment network with partial interchangeability.

Proposition 1. *The problem of deciding whether there exists an assignment $\mathscr{A}$ such that $I(\mathscr{A}) = 0$ for a balanced network N with three equipment types and a set $S \subseteq A$ of arcs on which the equipment type cannot be changed is NP-complete.*

Proof. Transformation from Simple D2CIF, which is known to be NP-complete (Even et al., 1975).

Simple D2CIF: Given a directed graph $D = (V, A)$, with two source vertices $s_1, s_2 \in V$, two sink vertices

$t_1, t_2 \in V$, each arc $a \in A$ having capacity one, and two non-negative integers R_1 and R_2. Do there exist two flow functions f_1 and f_2, both $A \to \mathbb{N}$, such that f_1 sends R_1 units of flow from s_1 to t_1 and f_2 sends R_2 units of flow from s_2 to t_2?

Without loss of generality, we assume $\sigma^-_{s_1} = \sigma^-_{s_2} = \sigma^+_{t_1} = \sigma^+_{t_2} = 0$ for D. We construct $D' = (V', A')$ as follows. Add a vertex s to V. For each $v \in V$, if $\sigma^+_v > \sigma^-_v$, add $\sigma^+_v - \sigma^-_v$ parallel arcs from s to v and if $\sigma^+_v < \sigma^-_v$, add $\sigma^-_v - \sigma^+_v$ parallel arcs from v to s. For $v \in V$, let A_{vs} and A_{sv} denote the set of parallel arcs from v to s, and from s to v, respectively.

Let the three equipment types be Q_1, Q_2 and Q_3. Let $\hat{A}_{ss_1} \subseteq A_{ss_1}$, $\hat{A}_{ss_2} \subseteq A_{ss_2}$, $\hat{A}_{t_1s} \subseteq A_{t_1s}$, and $\hat{A}_{t_2s} \subseteq A_{t_2s}$ with $|\hat{A}_{ss_1}| = |\hat{A}_{t_1s}| = R_1$, $|\hat{A}_{ss_2}| = |\hat{A}_{t_2s}| = R_2$. In the initial equipment assignment, let the equipment assigned to arcs in $\hat{A}_{ss_1}$ and $\hat{A}_{t_1s}$ be Q_1, in $\hat{A}_{ss_2}$ and $\hat{A}_{t_2s}$ be Q_2, and in $A_{ss_1} \setminus \hat{A}_{ss_1}$, $A_{ss_2} \setminus \hat{A}_{ss_2}$, $A_{t_1s} \setminus \hat{A}_{t_1s}$, $A_{t_2s} \setminus \hat{A}_{t_2s}$, and in A_{sv}, and A_{vs} for $v \in V \setminus \{s_1, s_2, t_1, t_2\}$ be Q_3. Furthermore, let these sets be such that the equipment type cannot be changed. For convenience, let Y_1 and Y_2 denote the sets of arcs with equipment type Q_1 and Q_2, respectively, for which the equipment type cannot be changed.

For a YES-instance of Simple D2CIF, we have the required flow functions f_1 and f_2. We may assume f_1 does not send any flow from s_2 to t_2 and f_2 does not send any flow from s_1 to t_1 (otherwise, we can simply delete those flows from f_1 and f_2). Zero imbalance can be achieved by $\mathscr{A} : A' \to \{Q_1, Q_2, Q_3\}$ defined as follows:

$$\mathscr{A}(a') = \begin{cases} Q_1, & a' \in \{a \in A : f_1(a) = 1\} \cup Y_1, \\ Q_2, & a' \in \{a \in A : f_2(a) = 1\} \cup Y_2, \\ Q_3, & \text{otherwise.} \end{cases}$$

For equipment type Q_1, since f_1 is a flow, by conservation of flow for $v \in V \setminus \{s_1, s_2, t_1, t_2\}$, Q_1 is balanced. By the definition of $\mathscr{A}$, it is also balanced for $\{s_1, s_2, t_1, t_2, s\}$. Similarly, equipment type Q_2 is balanced. Since for all $v \in V'$, $\sigma^+(v) = \sigma^-(v)$, equipment type Q_3 is also balanced, and, thus, the imbalance is zero.

Now suppose there exists an assignment $\mathscr{A}$ for D' that achieves zero imbalance, then the corresponding instance of Simple D2CIF is YES-instance, because the two flow functions f_1 and f_2 can be constructed as follows:

$$f_1(a) = \begin{cases} 1, & a \in \{a \in A : \mathscr{A}(a) = Q_1\}, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$f_2(a) = \begin{cases} 1, & a \in \{a \in A : \mathscr{A}(a) = Q_2\}, \\ 0, & \text{otherwise.} \end{cases}$$

Since $M(D', \mathscr{A}) = 0$, for all $v \in V \setminus \{s_1, s_2, t_1, t_2\}$, the imbalance for Q_1 and Q_2 is zero, and, thus, there is conservation of flow. For s_1, since on R_1 arcs from s to s_1 the equipment type is Q_1 and on the remaining arcs from s to s_1 the equipment type is Q_3, and the equipment type on these arcs cannot be changed, we have $\sum_{a\in A\,:\,a\ni s_1} f_1(a) = R_1$. Similarly, we have $\sum_{a\in A\,:\,a\ni t_1} f_1(a) = R_1$, $\sum_{a\in A\,:\,a\ni s_2} f_2(a) = R_2$, and $\sum_{a\in A\,:\,a\ni t_2} f_2(a) = R_2$. Therefore, f_1 and f_2 are the desired flow functions. ■

Proposition 1 implies that in the case of partial interchangeability, both the problem of finding the minimum imbalance and the problem of finding the minimum number of changes required to reach minimum imbalance are NP-hard. Next, we show that the problem of finding the minimum number of changes required to reach minimum imbalance is also NP-hard for the case of full interchangeability.

Proposition 2. *The problem of deciding whether there exists an assignment $\mathscr{A}$ such that $I(\mathscr{A}) = I^*$ and $\|\mathscr{A} - \mathscr{A}_0\| \leq K$, for a given $K \in \mathbb{Z}_{\geq 0}$ and a network N with three equipment types and full interchangeability is NP-complete.*

Proof. For convenience, the decision problem in Proposition 1 is referred to as 3EPI (3-equipment with partial interchangeability). We prove the proposition by providing a transformation from 3EPI.

Given a balanced network $D = (V, A)$, three disjoint subsets Y_1, Y_2, and Y_3 of A such that the equipment types on arcs in Y_1, Y_2, and Y_3, are Q_1, Q_2, and Q_3, respectively, and cannot be changed, and $K \in \mathbb{Z}_{\geq 0}$, we construct a directed graph $D' = (V', A')$ as follows. For all arcs $a = (u, v) \in Y_i$, $i = 1, 2, 3$, replace a by a directed path P_a of length $m = |A|$ from u to v by adding $m-1$ intermediate vertices and $m-1$ arcs. Note that $|V'| \leq mn$ and $|A'| \leq m^2$ and the size of D' is polynomial in the size of D. Let the initial assignment for $a' \in A'$ be

$$\mathscr{A}_0(a') := \begin{cases} Q_1, & \text{if } a' \in P_a \text{ for } a \in Y_1, \\ Q_2, & \text{if } a' \in P_a \text{ for } a \in Y_2, \\ Q_3, & \text{otherwise.} \end{cases}$$

By construction, D' is also a balanced network. Let $\hat{I}^*$ be the minimum imbalance of D' and $\hat{I}(\mathscr{A})$ be the imbalance of an assignment of D'. Observe that in any assignment $\mathscr{A}'$ with $\hat{I}(\mathscr{A}) = 0$ and $\|\mathscr{A}' - \mathscr{A}_0\| \leq m-1$, the equipment type on arcs $a' \in p_a$ for $a \in Y_1 \cup Y_2 \cup Y_3$ cannot have changed. Therefore, if we can find an $\mathscr{A}$ with $I(\mathscr{A}) = 0$ and $\|\mathscr{A}' - \mathscr{A}_0\| \leq m-1$, then the instance of 3EPI is a YES-instance, verified by assignment

$$\mathscr{A}(a) := \begin{cases} Q_1, & \text{if } \mathscr{A}'(a) = Q_1, \\ Q_2, & \text{if } \mathscr{A}'(a) = Q_2, \\ Q_3, & \text{otherwise.} \end{cases}$$

If, on the other hand, for all $\mathscr{A}$ with $\hat{I}(\mathscr{A}) = 0$, we have $\|\mathscr{A}' - \mathscr{A}_0\| \geq m$, then the instance of 3EPI is a NO-instance. ■

5. Computational Study

Our computational study uses real-world instances derived from weekly load plans and weekly driver schedules for a package express network.

5.1. Equipment and Substitution Matrices

There are three categories of equipment: *Shorts* (**S**), *Longs* (**L**), and *Extra Longs* (**XL**). Category **S** contains three equipment types: W, WW, and SC, category **L** contains nine equipment types: Z, ZZ, S, Y, YY, O, OO, TMF, and TMB, and category **XL** contains two equipment types: ZZZ and LC. To support strategic and tactical analysis, it is beneficial to be able to accommodate different sets of substitution rules. This is accomplished by the introduction of an equipment substitution matrix (ESM). An ESM is 0-1 matrix, where a 1 in the i-th row and j-th column means that the equipment type corresponding to the i-th row can be substituted with the equipment type corresponding to the j-th column, and a 0 means that this substitution is not allowed. To determine the set of allowable equipment types $\mathcal{C}_a$ for a load a, we consider the equipment substitution matrix, whether the facilities at the origin and destination of the load have equipment type restrictions, and the size of the load (a smaller equipment type is allowed only if the load fits). The most restrictive equipment substitution matrix, ESM1, is shown in Table 3. It does not consider composite equipment types and it does not allow substituting a short equipment type with a larger equipment type (i.e., and equipment type in **L** or **XL**). The latter requirement tries to avoid the use of lightly utilized equipment since the size of equipment in category **S** is much smaller that the size of equipment in categories **L** and **XL**. Note that equipment types TMF and TMB are only allowed to be swapped with each other.

	W	WW	SC	Z	ZZ	S	Y	YY	O	OO	TMF	TMB	ZZZ	LC
W	1	1	1	0	0	0	0	0	0	0	0	0	0	0
WW	1	1	1	0	0	0	0	0	0	0	0	0	0	0
SC	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Z	1	1	1	1	1	1	1	1	1	1	0	0	1	1
ZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1
S	1	1	1	1	1	1	1	1	1	1	0	0	1	1
Y	1	1	1	1	1	1	1	1	1	1	0	0	1	1
YY	1	1	1	1	1	1	1	1	1	1	0	0	1	1
O	1	1	1	1	1	1	1	1	1	1	0	0	1	1
OO	1	1	1	1	1	1	1	1	1	1	0	0	1	1
TMF	0	0	0	0	0	0	0	0	0	0	1	1	0	0
TMB	0	0	0	0	0	0	0	0	0	0	1	1	0	0
ZZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1
LC	1	1	1	1	1	1	1	1	1	1	0	0	1	1

Table 3: ESM1.

Equipment substitution matrix ESM2 introduces composite equipment type 2WW, which consists of two shorts, specifically two pieces of equipment type WW, with capacity close to the capacity of the equipment types in category **XL**. As shown in Table 4, all equipment types, except TMF and TMB, can be substituted with equipment type 2WW, which gives much more flexibility, but, as a result, also makes the optimization models more difficult to solve. As none of the loads in the system initially have a composite equipment type, there is no need to include a row for equipment type 2WW in the equipment substitution matrix.

	W	WW	SC	Z	ZZ	S	Y	YY	O	OO	TMF	TMB	ZZZ	LC	2WW
W	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1
WW	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1
SC	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1
Z	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
ZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
S	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
Y	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
YY	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
O	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
OO	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
TMF	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
TMB	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
ZZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
LC	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1

Table 4: ESM2.

Equipment substitution matrix ESM3 introduces bobtails (BT). When a load is labeled as a bobtail, it means that a tractor moves without pulling any trailer. Assigning an equipment type to a bobtail, i.e., having the tractor pull one or more (empty) trailers, can be an effective way to reduce the imbalance. Furthermore, we allow short equipment types to be substituted with large or extra large equipment types.

	W	WW	SC	Z	ZZ	S	Y	YY	O	OO	TMF	TMB	ZZZ	LC	2WW	BT
W	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
WW	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
SC	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
Z	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
ZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
S	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
Y	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
YY	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
O	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
OO	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
TMF	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
TMB	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
ZZZ	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
LC	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0
BT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 5: ESM3.

5.2. Instances

We use ten instances in our computational experiments derived from weekly load plans and driver schedules. Detailed information about the instances can be found in Table 6. Each weekly load plan contains all the loads that are scheduled to be dispatched during the week. The loads are of three types: (a) loaded, in which case an equipment type is assigned and a volume is specified (as a percentage of trailer capacity), (b) empty, in which case an equipment type is assigned, but there is no volume specified, and (c) bobtail, in which case no equipment type is specified. The instances have about 3350 facilities and about 300 thousand loads, with about 40% of these being loaded, 30% being empty, and 30% being bobtails. The initial imbalance in the instances is around 4,000. The load plans have been modified slightly, to ensure that balance can be restored. (In the appendix, we present necessary and sufficient conditions to guarantee that balance can restored.)

Instance	# Loads	# Facilities	# Equipment Types	Initial Imbalance	Empty Loads (%)	Bobtails (%)
I1	302,344	3,254	15	4,098	29.40	28.14
I2	301,509	3,982	15	3,982	29.33	28.05
I3	301,270	3,270	15	4,006	29.19	28.08
I4	300,963	3,273	15	4,106	29.09	28.10
I5	300,298	3,285	15	4,126	29.03	27.95
I6	300,013	3,275	15	3,948	28.93	27.85
I7	299,519	3,281	15	4,078	28.80	27.79
I8	299,519	3,280	15	3,870	28.83	27.63
I9	299,107	3,291	15	3,786	28.89	27.38
I10	299,415	3,285	15	3,566	28.89	27.38

Table 6: Information on the instances used in the computational experiments.

All the models are coded in C++. Gurobi 8.1 with default settings is used for solving the MIPs. All experiments were run in a 20-core machine with Intel(R) Xeon(R) 2.30GHz processors and 256GB of RAM. The optimality tolerance is set to 0.005 for Stage 1 and 0.1 for Stage 2. No time limit was enforced.

5.3. A Simple Decomposition Heuristic

When the composite equipment type 2WW is not allowed (ESM1), the coefficient matrix of the optimization problems presented in Section 3.1 and 3.2 will be a (0,1)-matrix, and the optimization problems can be solved relatively easily in practice (even though the problem is **NP-hard**). However, when the composite equipment type 2WW is allowed, the coefficient matrices will no longer be (0,1)-matrices, as the composite equipment type consists of two basic equipment types. As a consequence, solving the optimization problems is much more difficult and requires much more computing time. Therefore, we have developed a simple, but computationally effective, two-phase decomposition heuristic. In the first phase, we only allow equipment to be substituted with short or composite equipment types (**S** and **C**), which makes the optimization problem more tractable. In the second phase, the substitutions identified in the first phase are fixed and we seek to further reduce the imbalance by equipment substitutions among the **S**, **L**, and **XL** categories. Computational experiments show that the heuristic produces high-quality solutions in a short amount of time.

We believe the idea of the decomposition heuristic can be more general than what is present above. More specifically, when dealing with an intractable decision problem, we may decompose the problem by first focusing on part of the decisions, which should reduce the problem to a tractable size. A rule of thumb in defining the restricted problem is to decompose the sources of difficulty. Subsequently, the decisions obtained in the first phase are fixed and the options that have been disabled are enabled in the second phase, which produces another tractable problem. The result of solving these two phases can yield a near-optimal solution to the original problem of interest. This idea can be easily generalized to decomposition heuristics involving more than two phases.

5.4. Analysis

In the first set of the experiments, we focus on finding the minimum imbalance with a minimum number of substitutions. Furthermore, we evaluate the performance of the decomposition heuristic. In the second set of experiments, we assess the benefits of complementing empty repositioning with equipment substitution when restoring balance.

5.4.1. Equipment Substitution

Let I_0 and $\hat{I}$ be the initial imbalance and minimum imbalance, respectively, and let the imbalance reduction be $\Delta I = (I_0 - \hat{I})/I_0$. Let N_s be the number of substitutions in the Stage 1 solution, let $\hat{N}_s$ be the minimum number of substitutions required to reach minimum imbalance, and let the substitution reduction be $\Delta N_s := (N_s - \hat{N}_s)/N_s$.

Table 7 shows the optimization results using ESM1. We observe that, on average, the imbalance can be

Instance	Stage 1				Stage 2			
	I_0	$\hat{I}$	$\Delta I(\%)$	Time (s)	N_s	$\hat{N}_s$	$\Delta N_s(\%)$	Time (s)
I1	4,098	3,600	12.15	18	50,441	275	99.45	139
I2	3,982	3,474	12.76	16	50,858	268	99.47	176
I3	4,006	3,476	13.23	19	50,032	295	99.41	255
I4	4,106	3,542	13.74	20	50,094	289	99.42	783
I5	4,126	3,564	13.62	21	49,554	316	99.36	110
I6	3,948	3,426	13.22	16	49,857	311	99.38	221
I7	4,078	3,600	11.72	17	48,765	288	99.41	154
I8	3,870	3,498	9.61	17	49,346	222	99.55	60
I9	3,786	3,450	8.87	16	50,809	192	99.62	81
I10	3,566	3,344	6.23	15	49,298	165	99.67	202

Table 7: Optimizarion results using equipment substitution matrix ESM1.

reduced by about 11% and that this requires, on average, fewer than 300 substitutions. We also observe that the Stage 2 optimization takes, on average, more than 10 times as long as the Stage 1 optimization.

Tables 8 and 9 show the optimization and heuristic results using ESM2, respectively. ESM2 allows, on top of ESM1, the possibility to substitute a load with composite equipment type 2WW, which provides more flexibility.

Instance	Stage 1				Stage 2			
	I_0	$\hat{I}$	$\Delta I(\%)$	Time (s)	N_s	$\hat{N}_s$	$\Delta N_s(\%)$	Time (s)
I1	4,098	2,010	50.95	1,391	111,817	2,851	97.45	8,623
I2	3,982	1,894	52.44	4,257	120,030	2,817	97.65	50,338
I3	4,006	1,835	54.19	1,656	114,860	2,939	97.44	8,937
I4	4,106	1,913	53.41	822	114,541	2,928	97.44	43,061
I5	4,126	1,968	52.30	1,669	115,582	2,885	97.50	18,931
I6	3,948	1,916	51.47	2,410	114,831	2,735	97.62	11,581
I7	4,078	2,147	47.35	2,161	115,228	2,811	97.56	6,413
I8	3,870	2,075	46.38	2,017	116,725	2,933	97.49	18,702
I9	3,786	2,036	46.22	2,158	112,962	2,813	97.51	9,489
I10	3,566	1,882	47.22	3,817	112,454	2,982	97.35	15,713

Table 8: Optimization results using equipment substitution matrix ESM2.

Instance	Stage 1				Stage 2			
	I_0	$\hat{I}$	$\Delta I(\%)$	Time (s)	N_s	$\hat{N}_s$	$\Delta N_s(\%)$	Time (s)
I1	4,098	2,092	48.95	42	54,763	2,791	94.90	152
I2	3,982	1,968	50.58	44	55,247	2,845	94.85	150
I3	4,006	1,941	51.55	47	55,982	2,846	94.92	277
I4	4,106	1,971	52.00	44	55,597	2,937	94.72	229
I5	4,126	2,034	50.70	43	55,478	2,887	94.80	191
I6	3,948	1,977	49.92	44	56,446	2,813	95.02	361
I7	4,078	2,217	45.64	43	55,514	2,758	95.03	242
I8	3,870	2,150	44.44	47	55,770	2,866	94.86	292
I9	3,786	2,080	45.06	47	56,572	2,863	94.94	265
I10	3,566	1,948	45.37	51	56,181	2,859	94.91	187

Table 9: Heuristic results using equipment substitution matrix ESM2.

We observe that, on average, the imbalance can be reduced by about 50% and that this requires, on average, fewer than 3000 substitutions. We also observe that the decomposition heuristic performs well, achieving, on average, an imbalance reduction of about 48% and also requiring, on average, fewer than 3000 substitutions to achieve this. However, we see a dramatic reduction in computing time, i.e., on average by more than 98%.

Tables 10 and 11 show the optimization and heuristic results, respectively, using equipment substitution matrix ESM3. ESM3 allows, on top of ESM2, the substitution of a bobtail movement to any equipment type.

We observe that, on average, the imbalance can be reduced by about 67% and that this requires, on average, about 3500 substitutions. We also observe that the decomposition heuristic continues to perform well in terms of imbalance reduction, achieving, on average, a reduction of about 62%. However, its behavior varies in terms of number of substitutions required. In three of the ten instances, the number of substitutions required is around 14,000. On the other hand, optimization starts to become computationally prohibitive

Instance	Stage 1				Stage 2			
	I_0	$\hat{I}$	$\Delta I(\%)$	Time (s)	N_s	$\hat{N}_s$	$\Delta N_s(\%)$	Time (s)
I1	4,098	1,306	68.13	32,710	196,070	3,394	98.27	97,721
I2	3,982	1,228	69.16	38,124	161,337	3,519	97.82	159,458
I3	4,006	1,225	69.42	29,418	170,337	3,422	97.99	49,144
I4	4,106	1,213	70.46	16,373	170,302	3,485	97.95	76,702
I5	4,126	1,262	69.41	18,206	176,428	3,457	98.04	100,794
I6	3,948	1,176	70.21	21,099	166,709	3,581	97.85	39,000
I7	4,078	1,439	64.71	23,640	163,384	3,456	97.88	144,742
I8	3,870	1,387	64.16	18,394	170,292	3,500	97.94	162,128
I9	3,786	1,321	65.11	40,144	163,418	3,768	97.69	85,082
I10	3,566	1,215	65.93	36,830	188,755	3,877	97.95	187,588

Table 10: Optimization results using equipment substitution matrix ESM3.

Instance	Stage 1				Stage 2			
	I_0	$\hat{I}$	$\Delta I(\%)$	Time (s)	N_s	$\hat{N}_s$	$\Delta N_s(\%)$	Time (s)
I1	4,098	1,552	62.13	590	58,964	3,423	94.19	1,101
I2	3,982	1,474	62.98	567	58,187	3,342	94.26	908
I3	4006	1451	63.78	672	55323	3179	94.25	884
I4	4,106	1,477	64.03	639	57,882	3,303	94.29	971
I5	4,126	1,514	63.31	1,696	191,098	14,334	92.50	168
I6	3,948	1,437	63.60	717	58,261	3,367	94.22	960
I7	4,078	1,673	58.97	1,336	191,357	13,893	92.74	131
I8	3,870	1,645	57.49	914	57,730	3,119	94.60	1,609
I9	3,786	1,575	58.40	831	59,217	3,308	94.41	727
I10	3,566	1,418	60.24	1,595	189,423	14,467	92.36	130

Table 11: Heuristic results using equipment substitution matrix ESM3.

with some instances requiring more than 60 hours of computing (Stage 1 plus Stage 2). Most instances require less than 30 minutes of computing time using the decomposition heuristic.

For ease of comparison, we present a few critical statistics for the different equipment substitution matrices in Table 12. These statistics clearly show the benefits derived from allowing more flexible substitution rules and that the performance of the decomposition heuristic is good, but deteriorates slightly when the flexibility increases.

We next explore in detail for a single instance the relationship between the imbalance reduction and the number of substitutions required to achieve that reduction. More specifically, Figure 1 shows this relationship for Instance 1 (using ESM3). We see that the closer we get to the maximum possible imbalance reduction, the larger the number of substitutions required. For instance, to reduce the initial imbalance by 50%, about 1,250 substitutions are required. However, to reduce the imbalance from 50% to 70%, about 2,300 additional substitutions are required.

Instance	ESM1		ESM2				ESM3			
	Exact		Exact		Heuristic		Exact		Heuristic	
	$\Delta I(\%)$	$\hat{N}_s$	$\Delta I(\%)$	$\hat{N}_s$	$\Delta I(\%)$	$\hat{N}_s$	$\Delta I(\%)$	$\hat{N}_s$	$\Delta I(\%)$	$\hat{N}_s$
I1	12.15	275	50.95	2,851	48.95	2,791	68.13	3,394	62.13	3,423
I2	12.76	268	52.44	2,817	50.58	2,845	69.16	3,519	62.98	3,342
I3	13.23	295	54.19	2,939	51.55	2,846	69.42	3,422	63.78	3,179
I4	13.74	289	53.41	2,928	52.00	2,937	70.46	3,485	64.03	3,303
I5	13.62	316	52.30	2,885	50.70	2,887	69.41	3,457	63.31	14,334
I6	13.22	311	51.47	2,735	49.92	2,813	70.21	3,581	63.60	3,367
I7	11.72	288	47.35	2,811	45.64	2,758	64.71	3,456	58.97	13,893
I8	9.61	222	46.38	2,933	44.44	2,866	64.16	3,500	57.49	3,119
I9	8.87	192	46.22	2,813	45.06	2,863	65.11	3,768	58.40	3,308
I10	6.23	165	47.22	2,982	45.37	2,859	65.93	3,877	60.24	14,467

Table 12: A few critical statistics for substitution matrices ESM1, ESM2 and ESM3.

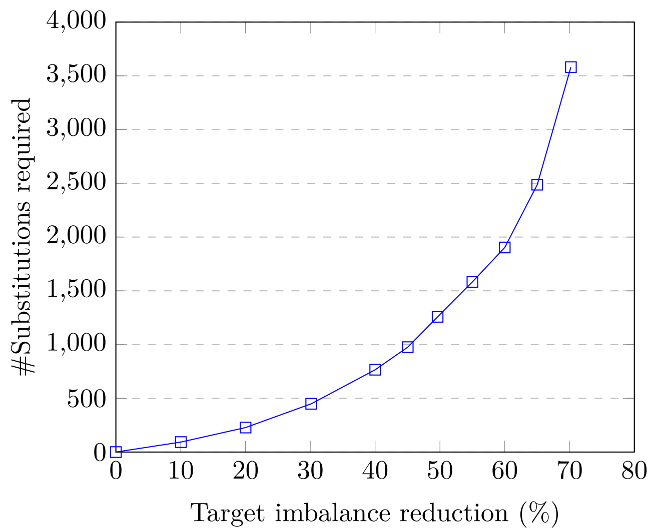


Figure 1: Relationship between the imbalance reduction and the number of substitutions required for Instance 1.

5.4.2. Restoring Balance

Next, we assess the benefits of complementing empty repositioning with equipment substitution. To do so, we compare two scenarios: (a) restoring balance of a given load plan by adding empty loads, and (b) minimizing the imbalance of a given load plan by substituting equipment and then adding empty loads. In both scenarios, when adding empty loads, we minimize the number of empty load miles added (this is a minimum cost flow problem). When minimizing the imbalance of a given load plan by substituting equipment, we use equipment substitution matrix ESM3. In Table 13, we report for both scenarios the number of empty loads added and the number of empty load miles added as well as the reduction in empty load miles when equipment substitutions are performed before empty loads are added.

We observe that the number of empty loads that need to be added to restore balance reduces significantly

Instance	Empty repositioning		Equipment substitutions + Empty repositioning		Miles reduction (%)
	# Empties	# Miles	# Empties	# Miles	
I1	14,082	52,523	6,930	30,559	41.82
I2	14,012	51,460	6,799	32,185	37.46
I3	13,548	48,996	6,237	29,675	39.43
I4	14,154	50,382	6,552	29,713	41.02
I5	14,189	48,715	7,073	32,617	33.05
I6	13,392	44,613	6,493	29,125	34.72
I7	14,730	52,757	7,647	34,013	35.53
I8	14,477	62,710	7,905	45,398	27.61
I9	14,172	50,478	7,593	33,509	33.62
I10	12,956	45,024	6,960	28,603	36.47

Table 13: Restoring balance with and without equipment substitutions.

when we first reduce the imbalance by equipment substitutions, and, more importantly, that the number of empty load miles added reduces significantly, on average by about 36%.

6. Final Remarks

In this paper, we approach this challenge of achieving equipment balance, i.e., seeking to have the same equipment at a facility at the end of the week as at the start of the week, via a novel substitution-based method. We carefully analyze the computational complexity for several simplified settings and draw the conclusion that the problem is in general **NP-hard** to solve when there are no less than three types of equipment in the network of interest. Furthermore, we develop a MIP model and propose an efficient heuristic for solving instances with real data. Our numerical experiments demonstrate that the heuristic works well in practice. Nevertheless, what happens during the week and seasonal changes in package flows are ignored since our current focus is only on the condition at the end of the week.

A natural future research direction, therefore, is inventory-aware equipment management, in which time is modeled explicitly, e.g., days for planning periods of one or more weeks, and weeks for planning periods of one or more quarters. We are currently pursuing this research direction.

Acknowledgment

We are grateful for the support of the Transportation Analytics and Operations Research team at UPS. We want to thank them for the many informative discussions and help with questions regarding equipment planning in transportation networks.

References

Boile, M., Theofanis, S., Baveja, A., Mittal, N., 2008. Regional repositioning of empty containers: Case for inland depots. Transportation Research Record 2066, 31–40.

Chang, H., Jula, H., Chassiakos, A., Ioannou, P., 2008. A heuristic solution for the empty container substitution problem. Transportation Research Part E: Logistics and Transportation Review 44, 203–216.

Dejax, P.J., Crainic, T.G., 1987. Survey paper—a review of empty flows and fleet management models in freight transportation. Transportation science 21, 227–248.

Dell'Amico, M., Hadjicostantinou, E., Iori, M., Novellani, S., 2014. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. Omega 45, 7–19.

Du, Y., Hall, R., 1997. Fleet sizing and empty equipment redistribution for center-terminal transportation networks. Management Science 43, 145–157.

Erera, A.L., Morales, J.C., Savelsbergh, M., 2009. Robust optimization for empty repositioning problems. Operations Research 57, 468–483.

Even, S., Itai, A., Shamir, A., 1975. On the complexity of time table and multi-commodity flow problems, in: 16th Annual Symposium on Foundations of Computer Science (sfcs 1975), IEEE. pp. 184–193.

Haider, Z., Nikolaev, A., Kang, J.E., Kwon, C., 2018. Inventory rebalancing through pricing in public bike sharing systems. European Journal of Operational Research 270, 103–117.

He, L., Hu, Z., Zhang, M., 2020. Robust repositioning for vehicle sharing. Manufacturing & Service Operations Management 22, 241–256.

Hoffman, A.J., 2003. Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, in: Selected Papers Of Alan J Hoffman: With Commentary. World Scientific, pp. 244–258.

Holler, J., Vuorio, R., Qin, Z., Tang, X., Jiao, Y., Jin, T., Singh, S., Wang, C., Ye, J., 2019. Deep reinforcement learning for multi-driver vehicle dispatching and repositioning problem, in: 2019 IEEE International Conference on Data Mining (ICDM), IEEE. pp. 1090–1095.

Jansen, B., Swinkels, P.C., Teeuwen, G.J., de Fluiter, B.v.A., Fleuren, H.A., 2004. Operational planning of a large-scale multi-modal transportation system. European Journal of Operational Research 156, 41–53.

Lee, S., Moon, I., 2020. Robust empty container repositioning considering foldable containers. European Journal of Operational Research 280, 909–925.

Legros, B., 2019. Dynamic repositioning strategy in a bike-sharing system; how to prioritize and how to rebalance a bike station. European Journal of Operational Research 272, 740–753.

Legros, B., Bouchery, Y., Fransoo, J., 2019. A time-based policy for empty container management by consignees. Production and Operations Management 28, 1503–1527.

Li, J.A., Leung, S.C., Wu, Y., Liu, K., 2007. Allocation of empty containers between multi-ports. European Journal of Operational Research 182, 400–412.

Long, Y., Lee, L.H., Chew, E.P., 2012. The sample average approximation method for empty container repositioning with uncertainties. European Journal of Operational Research 222, 65–75.

Lu, T., Lee, C.Y., Lee, L.H., 2020. Coordinating pricing and empty container repositioning in two-depot shipping systems. Transportation Science 54, 1697–1713.

Lv, C., Zhang, C., Lian, K., Ren, Y., Meng, L., 2020. A hybrid algorithm for the static bike-sharing repositioning problem based on an effective clustering strategy. Transportation Research Part B: Methodological 140, 1–21.

Ma, T.Y., Rasulkhani, S., Chow, J.Y., Klein, S., 2019. A dynamic ridesharing dispatch and idle vehicle repositioning strategy with integrated transit transfers. Transportation Research Part E: Logistics and Transportation Review 128, 417–442.

Pal, A., Zhang, Y., 2017. Free-floating bike sharing: Solving real-life large-scale static rebalancing problems. Transportation Research Part C: Emerging Technologies 80, 92–116.

Pouls, M., Meyer, A., Ahuja, N., 2020. Idle vehicle repositioning for dynamic ride-sharing, in: International Conference on Computational Logistics, Springer. pp. 507–521.

Qin, Z.T., Tang, X., Jiao, Y., Zhang, F., Wang, C., Li, Q.T., 2019. Deep reinforcement learning for ride-sharing dispatching and repositioning, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence, pp. 6566–6568.

Schuijbroek, J., Hampshire, R.C., Van Hoeve, W.J., 2017. Inventory rebalancing and vehicle routing in bike sharing systems. European Journal of Operational Research 257, 992–1004.

Song, D.P., 2005. Optimal threshold control of empty vehicle redistribution in two depot service systems. IEEE Transactions on Automatic Control 50, 87–90.

Song, D.P., 2007. Characterizing optimal empty container reposition policy in periodic-review shuttle service systems. Journal of the Operational Research Society 58, 122–133.

Song, D.P., Zhang, Q., 2010. A fluid flow model for empty container repositioning policy with a single port and stochastic demand. SIAM Journal on Control and Optimization 48, 3623–3642.

Tardos, É., 1985. A strongly polynomial minimum cost circulation algorithm. Combinatorica 5, 247–255.

Theofanis, S., Boile, M., 2009. Empty marine container logistics: facts, issues and management strategies. GeoJournal 74, 51.

Wolsey, L.A., Nemhauser, G.L., 1999. Integer and combinatorial optimization. volume 55. John Wiley & Sons.

Xie, Y., Liang, X., Ma, L., Yan, H., 2017. Empty container management and coordination in intermodal transport. European Journal of Operational Research 257, 223–232.

Zhang, B., Ng, C.T., Cheng, T., 2014. Multi-period empty container repositioning with stochastic demand and lost sales. Journal of the Operational Research Society 65, 302–319.

Appendix. Restoring Balance

As mentioned in the introduction, companies restore balance by introducing empty loads that send equipment from facilities with an excess of equipment to facilities with a deficit of equipment. This incurs extra costs as it may involve creating new driver schedules and additional transportation costs. A natural question to ask is whether it is always possible to restore balance by introducing empty loads. Next, we give a necessary and sufficient condition on the service network that guarantees that balance can be restored by introducing empty loads.

We say that a load plan with a single equipment type can be balanced if it is possible to reduce the equipment imbalance to zero by adding empty loads. We first provide an example of a service network that cannot be balanced. Consider network $N = (V, A_0)$ with $V = \{v_1, v_2, v_3, v_4\}$ and $A_0 = \{v_1 \to v_2, v_2 \to v_1, v_2 \to v_3, v_3 \to v_4, v_4 \to v_3\}$ as shown in Figure 2.

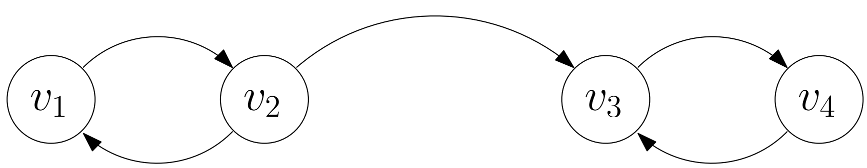
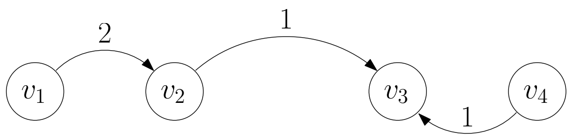
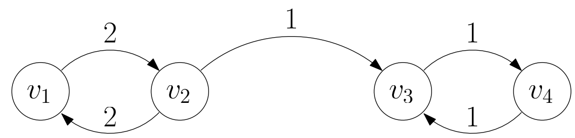


Figure 2: Example of a network with 4 nodes and 5 arcs.

Let load plan $\mathcal{L}$ have two loads on arc $v_1 \to v_2$, one load on arc $v_2 \to v_3$, and one load on arc $v_4 \to v_3$. Hence, the initial imbalance is 6 (2 at node v_1, 1 at node v_2, 2 at node v_3, and 1 at node v_4). The imbalance can be reduced to 2 by adding two empty loads on the arc $v_2 \to v_1$ and one empty load on arc $v_3 \to v_4$ (see Figure 3a). It is not possible to reduce the imbalance to zero because there is no path from v_3 to v_2.



(a) Initial load plan $\mathcal{L}$ with imbalance 6



(b) Load plan adjusted with new empty loads (blue arcs) with final imbalance 2

Figure 3: Example where empty repositioning does not yield zero imbalance.

We see that the fact that there is no path from v_3 to v_2 using arcs in A_0 makes it impossible to create a "cycle of loads" to reduce the imbalance. Based on this observation, we give a necessary and sufficient condition that guarantees that a given load plan (with a single equipment type) can be balanced.

Claim 3. *Given a network $N = (V, A_0)$ and a load plan $\mathcal{L}$, then $\mathcal{L}$ can be balanced if and only if every load $\ell \in \mathcal{L}$ belongs to a directed cycle in A_0.*

Proof. Let $|\mathcal{L}(a)|$ be the number of loads on arc $a \in A_0$ in the plan $\mathcal{L}$. The problem of adding empty loads to reach zero imbalance can be viewed as one to find a circulation f such that for any arc $a \in A_0$, the flow satisfies $|\mathcal{L}(a)| \leq f(a)$ and $f(a) \in \mathbb{Z}_+$.

Sufficiency: For any subset $U \subseteq V$, if $\delta^{\text{in}}(U) \neq \emptyset$, then $\delta^{\text{out}}(U) \neq \emptyset$ since each arc belongs to a cycle. Let $d(a) = |\mathcal{L}(a)| \in \mathbb{Z}_+$ and $c(a) = M \in \mathbb{Z}_+$ large enough, $\forall a \in A_0$, then $d(\delta^{\text{in}}(U)) \leq c(\delta^{\text{out}}(U))$. According to Hoffman's circulation theorem (Hoffman, 2003), such circulation f exists.

Necessity: If such a circulation exists, then it can be decomposed into a set of cycles. ■