A novel hybrid-link-based container routing model

Shuaian Wang

University of Wollongong, shuaian@uow.edu.au

Publication Details

A novel hybrid-link-based container routing model

Abstract
Container routing determines how to transport containers from their origins to their destinations in a liner shipping network. Container routing needs to be solved a number of times as a subproblem in tactical-level decision planning of liner shipping operations. Container routing is similar to the multi-commodity flow problem. This research proposes a novel hybrid-link-based model that nests the existing origin-link-based and destination-link-based models as special cases. Moreover, the hybrid-link-based model is at least as compact as the origin-to-destination-link-based, origin-link-based and destination-link-based models in the literature.

Keywords
routing, container, link, hybrid, model, novel

Disciplines
Engineering | Science and Technology Studies

Publication Details

This journal article is available at Research Online: http://ro.uow.edu.au/eispapers/1825
A novel hybrid-link-based container routing model

Shuaian Wang *

School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia

Abstract

Container routing determines how to transport containers from their origins to their destinations in a liner shipping network. Container routing needs to be solved a number of times as a subproblem in tactical-level decision planning of liner shipping operations. Container routing is similar to the multi-commodity flow problem. This research proposes a novel hybrid-link-based model that nests the existing origin-link-based and destination-link-based models as special cases. Moreover, the hybrid-link-based model is at least as compact as the origin-to-destination-link-based, origin-link-based and destination-link-based models in the literature.

Key Words: maritime transportation; liner shipping; container routing; multi-commodity flow; hybrid-link-based model; totally unimodular matrix

* Corresponding author
E-mail address: wangshuaian@gmail.com
1 Introduction

Liner shipping companies deploy containerships on regularly scheduled services to transport containers. Containerships in liner shipping have to sail according to the planned schedule no matter whether they are fully loaded or not. Liner shipping services are usually weekly, which means that each port of call is visited on the same day of every week. Once the weekly liner shipping services are designed, they are operated for a period of three to six months. Therefore, it is important for liner shipping companies to design efficient services as a large proportion of the total operating cost is fixed once the services are designed.

Liner shipping decision problems can be classified as strategic, tactical, and operational (Christiansen et al., 2004, 2013; Meng et al., 2013). Fleet size and mix (e.g., Meng and Wang, 2011), alliance strategy (e.g., Agarwal and Ergun, 2010) and network design (e.g., Fagerholt, 1999, 2004; Shintani et al., 2007; Imai et al., 2009; Gelareh et al., 2010; Gelareh and Pisinger, 2011; Reinhardt and Pisinger, 2012; Plum et al., 2013) are strategic-level decision problems. Network alteration (e.g., Wang and Meng, 2013), fleet deployment (e.g., Meng and Wang, 2012; Wang and Meng, 2012a), schedule design (e.g., Qi and Song, 2012), and speed optimization (e.g., Psaraftis and Kontovas, 2013) are tactical decision problems. Operational decisions include problems such as container booking/routing (e.g., Song and Dong, 2013) and ship rescheduling (e.g., Yan et al., 2009; Brouer et al., 2013b).

Container routing occurs at both the operational level and the tactical level. Container routing determines how to transport containers from their origins to their destinations in a liner shipping network (Wang et al., 2013b). Take Fig. 1 as an example. It shows a liner shipping network consisting of three ship routes with fixed port rotations. Containers from Singapore to Hong Kong can be transported either on ship route 1 or ship route 2. If there are many containers to be transported from Singapore to Jakarta, then containers from Singapore to Hong Kong should be transported on ship route 2 to reserve the capacity on ship route 1 for containers from Singapore to Jakarta. In addition to different ship routes on which
containers can be transported from origin to destination, another complicating factor is transshipment. For instance, containers from Hong Kong to Colombo can be transported on ship route 2, or they can be transported on ship route 1 to Singapore and transshipped to ship route 2 and then transported to Colombo. The choice of direct shipment on ship route 2 is preferable because the latter involves a high transshipment cost at Singapore. However, if there are many containers to be transported from Hong Kong to Xiamen or from Xiamen to Singapore, then the choice of transshipment at Singapore from ship route 1 to ship route 2 has to be adopted. Consequently, it is not an easy task to determine the optimal container routing.

![An illustrative liner shipping network](Source: Wang, 2013)

Container routing is not only significant to liner shipping companies as an independent problem at the operational level, more often than not, it serves as a subproblem in a number of tactical-level decision problems such as network alteration and fleet deployment. In the tactical-level decision problems, container routing must be solved iteratively and hence the computational efficiency of container routing models is vital. Consequently, it is important to develop models for container routing that are compact and can be easily solved.

Container routing problems are very similar to multi-commodity flow problems (MCF) studied in the field of operations research (e.g., Tomlin, 1966; Ahuja et al., 1993; Gamst et al., 2010). MCF can be solved in polynomial time. However, there are often extra conditions that have to be satisfied, making the problem NP-hard. Moreover, the sizes of many practical
applications are extremely large. Therefore, a number of specialized algorithms have been developed, most of which use a decomposition strategy that is based on duality and relaxation of coupling constraints. The main motivations for decomposition are (i) to reduce the problem to smaller sub-problems and (ii) to parallelize or distribute computations. We refer to Ouorou et al. (2000) for an overview of solution algorithms on MCF.

The container routing problem and MCF are not identical. For instance, the MCF with an upper-bounded path length is NP-hard (Gamst, et al., 2010). However, the container routing problem with an upper-bounded path travel time is still polynomially solvable. This is because unlike MCF, the liner shipping services have a weekly frequency.

There are generally two types of container routing models (or MCF models): path-based models (Brouer et al., 2011; Song and Dong, 2012; Wang and Meng, 2012b; Wang et al., 2013a) and link-based models. Path-based models need to enumerate all possible paths or generate dynamically the profitable paths for containers to be transported from origin to destination. By contrast, the number of variables in link-based models increases polynomially with the size of the liner shipping network. The advantage of path-based models is that side constraints can be easily handled.

It should be noted that in network design for MCF, it is usually to determine whether a link should be added or not. That is, the network design under MCF is to determine whether the capacity of a link is 0 (no construction cost) or a fixed positive number (with a fixed construction cost). Hence, both link-based models and path-based models are used in network design under MCF. In liner shipping network design, a link cannot be added separately, because a shipping service is a loop, where the links are connected, have the same capacity, and provide a weekly frequency. As a result, path-based model may be difficult to handle in liner shipping network design. Hence, to date we are unaware of any studies on liner shipping network design that uses path-based formulations for container routing. Recently, Plum et al. (2013) made a breakthrough in liner shipping network design by proposing an exact solution
method for the most general problem settings. Their container routing model is somewhere between link-based and path-based formulations.

The objective of this research is to develop a novel and compact link-based model, which we call hybrid-link-based model. It can be more efficiently solved by general-purpose commercial solvers than other link-based models and can be applied to many situations.

The remainder of the paper is organized as follows: Section 2 reviews existing link-based models in the literature. Section 3 proposes a novel hybrid-link-based model that requires fewer variables than existing models. Section 4 develops a linear programming model to obtain the optimal choice of origins and destinations for the hybrid-link-based model. Section 5 reports numerical experiments. Section 6 concludes.

2 Existing link-based models

Before presenting existing link-based models, we describe the container routing problem and define relevant parameters.

Consider a set $\mathcal{R}$ of ship routes, regularly serving a group of ports denoted by the set $\mathcal{P}$. Ship route $r \in \mathcal{R}$ can be expressed as:

$$
p_{r_1} \rightarrow p_{r_2} \rightarrow \cdots \rightarrow p_{r_{N_r}} \rightarrow p_{r_1}
$$

where $N_r$ is the number of ports of call and $p_{r_i}$ is the $i$th port of call, $i = 1, 2, \ldots, N_r$. Define $I_r := \{1, 2, \ldots, N_r\}$. The voyage from port of call $i$ to port of call $i + 1$ is called leg $i$ and leg $N_r$ is the voyage from port of call $N_r$ to the first port of call. In Fig. 1 three ship routes are shown: ship route 1 has three legs, ship route 2 has five legs, and ship route 3 has three legs. Each ship route has a weekly service frequency, which means that each port of call is visited on the same day every week. A string of homogeneous ships with a capacity of $V_r$ (twenty-foot equivalent units, abbreviated as TEUs) is deployed on ship route $r$ to maintain the weekly frequency. We do not consider the restriction of channels (Qu and Meng, 2012), and hence the ship capacity $V_r$ is the capacity of voyage legs.
Represent by \( W \) the set of origin-to-destination (OD) port pairs, \( W \subset \mathcal{P} \times \mathcal{P} \). The demand for OD pair \((o,d) \in W\) is deterministic (Meng and Qu, 2012; Meng et al., 2012; Wang et al., 2013c) for simplicity, and is denoted by \( q_{od} \) (TEUs/week). The penalty cost for not shipping a container in OD pair \((o,d)\) is \( g_{od} \) (USD/TEU). Containers can be transshipped at any port from origins to destinations. The load, transshipment, discharge cost (USD/TEU) at port \( p \in \mathcal{P} \) is denoted by \( \hat{c}_p \), \( \bar{c}_p \) and \( \hat{c}_p \), respectively.

The container routing problem aims to determine how many containers in each OD pair to transport and how to transport the containers to minimize the sum of container handling cost and penalty for not fulfilling the demand.

There are three types of link-based container routing models: OD-link-based, origin-link-based and destination-link-based. We elaborate on these three types of models below. It should be mentioned that to make this paper concise, factors that are not directly related to the contribution of the paper such as empty container repositioning are not included in the models.

2.1 OD-link-based model

Agarwal and Ergun (2008) applied an OD-link-based formulation in network design and Brouer et al. (2011) employed the OD-link-based model for container routing.

Before describing the OD-link-based model, we first need to define the set of origin ports, which is the set of ports that are origin of at least one OD pair:

\[
O := \{ p \in \mathcal{P} \mid \text{there exists } d \in \mathcal{P} \text{ satisfying } (p,d) \in W \} \tag{2}
\]

and the set of destination ports:

\[
D := \{ p \in \mathcal{P} \mid \text{there exists } o \in \mathcal{P} \text{ satisfying } (o,p) \in W \} \tag{3}
\]

The decision variables are as follows. \( \hat{z}_{ri}^{od} \) and \( \bar{z}_{ri}^{od} \) are the volume of containers (TEU/week) from \((o,d) \in W\) loaded and discharged at port of call \( i \) on ship route \( r \), respectively (note that when calculating \( \hat{z}_{ri}^{od} \) and \( \bar{z}_{ri}^{od} \), a transshipped container is considered as being discharged once and being loaded once); \( f_{ri}^{od} \) is the volume of containers (TEU/week) from \((o,d) \in W\) flowing on leg \( i \) on ship route \( r \) (we define \( f_{r0}^{od} := f_{rN}^{od} \)). \( y_{od} \) and \( x_{od} \) are the fulfilled and unfulfilled demand (TEU/week) for \((o,d) \in W\), respectively;
\( \hat{z}_p, \tilde{z}_p, \) and \( \bar{z}_p \) are the total volume of loaded, discharged, and transshipped containers (TEU/week) at port \( p \in P \), respectively. The OD-link-based model is:

\[
\text{[OD-model]} \quad \min_{z_{ri}, z_{ro}, \bar{z}_{ri}, \bar{z}_{ro}, y^o, x^o} \sum_{p \in P} \left( \hat{z}_p \hat{c}_p + \tilde{z}_p \tilde{c}_p + \bar{z}_p \bar{c}_p \right) + \sum_{(o,d) \in W} g^{od} x^{od}
\]

subject to:

\[
f^{od}_{r,i+1} + z^{od}_{ri} = f^{od}_{ri} + z^{od}_{ro}, \quad \forall r \in R, \forall i \in I_r, \forall (o,d) \in W (5)
\]

\[
\hat{z}_p = \begin{cases} \sum_{(p,d) \in W} y^{od}, & \forall p \in O \\ 0, & \forall p \in P \setminus O \end{cases}
\]

\[
\tilde{z}_p = \begin{cases} \sum_{(o,p) \in W} y^{op}, & \forall p \in D \\ 0, & \forall p \in P \setminus D \end{cases}
\]

\[
\bar{z}_p = \sum_{r \in R, i \in I_r, p_o = p} \sum_{(o,d) \in W} \left( z^{od}_{ri} - \hat{z}_p, \forall p \in P \right)
\]

\[
\sum_{r \in R, i \in I_r, p_o = p} \sum_{(o,d) \in W} \left( z^{od}_{ri} - \bar{z}_{ri} \right) = \begin{cases} y^{od}, & p = o \\ -y^{od}, & p = d, \forall (o,d) \in W, \forall p \in P \\ 0, & \text{otherwise} \end{cases}
\]

\[
\sum_{(o,d) \in W} f^{od}_{ri} \leq V_r, \forall r \in R, \forall i \in I_r
\]

\[
y^{od} + x^{od} = q^{od}, \forall (o,d) \in W (11)
\]

\[
z^{od}_{ri} \geq 0, \tilde{z}^{od}_{ri} \geq 0, f^{od}_{ri} \geq 0, \forall r \in R, \forall i \in I_r, \forall (o,d) \in W (12)
\]

\[
y^{od} \geq 0, x^{od} \geq 0, \forall (o,d) \in W (13)
\]

The objective function (4) minimizes the sum of container handling cost and penalty cost. Constraint (5) is container flow conservation equation. Constraints (6)–(8) define the total volume of loaded, discharged, and transshipped containers at port \( p \in P \), respectively. Constraint (9) computes the fulfilled demand. Constraint (10) imposes ship capacity constraint on each leg of each ship route. Constraint (11) defines the container shipment demand. Constraints (12) and (13) define the domains for the decision variables. Note that as aforementioned, in most cases container routing is a subproblem in tactical-level decisions. Hence, the volume of containers is modeled as a continuous number rather than an integer in
constraints (12) and (13) because the error caused by such an approximation is much smaller than the prediction error of the container shipment demand.

The number of flow variables (e.g. $f_{od}$) in the OD-link-based model has the magnitude of $|W|^r \sum_{r \in R} N_r$. The total numbers of variables and constraints are reported in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of variables</th>
<th>Number of constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD model</td>
<td>$3</td>
<td>W</td>
</tr>
<tr>
<td>Origin model</td>
<td>$3</td>
<td>O</td>
</tr>
<tr>
<td>Destination model</td>
<td>$3</td>
<td>D</td>
</tr>
<tr>
<td>Hybrid model</td>
<td>$3(</td>
<td>O</td>
</tr>
</tbody>
</table>

2.2 Origin-link-based model

The origin-link-based model is more compact than the OD-link-based model. The origin-link-based model is applied in network design by Alvarez (2009) and fleet deployment by Wang and Meng (2012a) and Wang (2013).

In the origin-link-based model, we use $\hat{z}_{r_i}$ and $\overline{z}_{r_i}$ to represent the total volume of containers with origin port $o \in O$ and any destination loaded and discharged at port of call $i$ on ship route $r$, respectively (transshipped containers are also considered) and use $f_{r_i}$ to denote the total volume of containers with origin port $o \in O$ and any destination flowing on leg $i$ of ship route $r$. The meanings of $y_{od}$, $x_{od}$, $\hat{z}_p$, $\overline{z}_p$, and $\tilde{z}_p$ are the same as the OD-link-based model. The origin-link-based model is equivalent to the OD-link-based model and is formulated as:

\[
\text{[origin-model]} \quad \min_{y_{od}, x_{od}, \hat{z}_p, \overline{z}_p, \tilde{z}_p} \sum_{(r_i)_{i \in I}} \left( \hat{z}_{r_i} \hat{z}_p + \overline{z}_{r_i} \overline{z}_p + \tilde{z}_{r_i} \tilde{z}_p \right) \quad + \sum_{(o,d) \in W} \tilde{g}_{od} x_{od} \quad (14)
\]

subject to:

\[
f_{r_{i+1}} + \hat{z}_{r_i} = f_{r_i} + \overline{z}_{r_i}, \forall r \in R, \forall i \in I_r, \forall o \in O \quad (15)
\]

\[
\hat{z}_p = \begin{cases} 
\sum_{(p,d) \in W} y_{od}, \forall p \in O \\
0, \forall p \in P \setminus O 
\end{cases} \quad (16)
\]
\[ z_p = \begin{cases} \sum_{(o,p) \in W} y_{op}, & \forall p \in D \\ 0, & \forall p \in \mathcal{P} \setminus D \end{cases} \tag{17} \]

\[ z_p = \sum_{r \in \mathcal{R}} \sum_{i \in I_r, p_{o,p} = p} \sum_{o \in \mathcal{O}} \sum_{p \in \mathcal{P}} (z_{pi}^o - \bar{z}_p), \forall p \in \mathcal{P} \tag{18} \]

\[ \sum_{r \in \mathcal{R}} \sum_{i \in I_r, p_{o,p} = p} (z_{pi}^o - \bar{z}_p) = \begin{cases} \sum_{(o,d) \in \mathcal{W}} y_{od}, & p = o \\ -y_{op}, & (o, p) \in \mathcal{W}, \forall o \in \mathcal{O}, \forall p \in \mathcal{P} \\ 0, & \text{otherwise} \end{cases} \tag{19} \]

\[ \sum_{o \in \mathcal{O}} f_{n_o}^o \leq V_r, \forall r \in \mathcal{R}, \forall i \in I_r \tag{20} \]

\[ y_{od} + x_{od} = q_{od}, \forall (o,d) \in \mathcal{W} \tag{21} \]

\[ \bar{z}_{n_o}^o \geq 0, \bar{z}_{n_o}^o \geq 0, f_{n_o}^o \geq 0, \forall r \in \mathcal{R}, \forall i \in I_r, \forall o \in \mathcal{O} \tag{22} \]

\[ y_{od} \geq 0, x_{od} \geq 0, \forall (o,d) \in \mathcal{W} \tag{23} \]

In the origin-link-based model, the number of flow variables (e.g. \( f_{n_o}^o \)) has the magnitude of \(|O| \sum_{r \in \mathcal{R}} N_r\). Evidently, \(|O| \leq |W|\). Therefore, as indicated in Table 1, the numbers of variables and constraints in the origin-link-based model are smaller than those in the OD-link-based model. Hence, the origin-link-based model is at least as compact as the OD-link-based model. We conduct a further analysis of the number of variables in the origin-link-based and OD-link-based models. We have \(|O| \leq |\mathcal{P}|\). In reality it is quite often that \(|O|\) is only slightly smaller than or even the same as \(|\mathcal{P}|\), as most ports in a liner shipping network are origin port in at least one OD pair. We also have \(|\mathcal{W}| \leq |\mathcal{P}|\), where “≤” rather than “<” because a port cannot have demand to itself (more exactly, \(|\mathcal{W}| \leq |\mathcal{P}|^2 - |\mathcal{O}|\)). The relations of \(|\mathcal{W}|\) and \(|\mathcal{P}|\) in the literature are shown in Table 2. In all the instances in the literature, we have \(|\mathcal{W}| < |\mathcal{P}|\). Hence, in practice the origin-link-based model is much more compact as it has fewer variables than the OD-link-based model.
Table 2 Relations between the number of ports and the number of OD pairs in literature

| Number of ports | Number of OD pairs | | | |
|----------------|------------------|| | |
| 46             | 652              | 0.31 |
| 12             | 22               | 0.15 |
| 19             | 38               | 0.11 |
| 39             | 369              | 0.24 |
| 45             | 722              | 0.36 |
| 111            | 4000             | 0.32 |
| 47             | 1764             | 0.80 |
| 197            | 9630             | 0.25 |

2.3 Destination-link-based model

The destination-link-based model is very similar to the origin-link-based model. The ODlink-based, origin-link-based, and destination-link-based models are all equivalent. The destination-link-based model is used in network design by Brouer et al. (2013a), Wang and Meng (2013), and Liu et al. (2014), and container routing by Bell et al. (2011, 2013).

In the destination-link-based model, we use \( \hat{d}_{ri} \) and \( \hat{z}_{ri} \) to represent the total volume of containers with destination port \( d \in D \) and any origin loaded and discharged at port of call \( i \) on ship route \( r \), respectively (transshipped containers are also considered) and use \( f_{rd} \) to denote the total volume of containers with destination port \( d \in D \) and any origin flowing on leg \( i \) of ship route \( r \). The meanings of \( y_{op} \), \( x_{op} \), \( \hat{z}_{p} \), \( \hat{z}_{p} \), and \( \hat{z}_{p} \) are the same as the OD-linkbased model. The destination-link-based model is:

\[
\text{[destination-model]} \quad \min_{f_{rd}, y_{op}, x_{op}, \hat{d}_{ri}, \hat{z}_{ri}, \hat{z}_{p}} \left( \sum_{p \in O} \left( \hat{z}_{p} \hat{x}_{op} + \hat{z}_{p} \hat{x}_{op} + \hat{z}_{p} \hat{x}_{op} \right) + \sum_{(o,p) \in W} g_{op} \right) \quad (24)
\]

subject to:

\[
f_{rd_{i+1}} + \hat{z}_{ri} = f_{rd} + \hat{z}_{ri}, \forall r \in R, \forall i \in I, \forall d \in D \quad (25)
\]

\[
\hat{z}_{p} = \begin{cases} \sum_{(o,p) \in W} y_{op}, & \forall p \in O \\ 0, & \forall p \in P \setminus O \end{cases} \quad (26)
\]

\[
\hat{z}_{p} = \begin{cases} \sum_{(o,p) \in W} y_{op}, & \forall p \in D \\ 0, & \forall p \in P \setminus D \end{cases} \quad (27)
\]

\[
\hat{z}_{p} = \sum_{i \in I} \sum_{r \in R \setminus r_{i} \setminus i \in D} \hat{d}_{ri} - \hat{z}_{p}, \forall p \in P \quad (28)
\]
\[
\sum_{r \in R} \sum_{i \in I^r : p_i = p} \left( z_{ri}^d - z_n^d \right) = \begin{cases} 
- \sum_{(o,d) \in W} y_{o,d}^d, & p = d \\
0, & \text{otherwise}
\end{cases}
\]

\[
\sum_{d \in D} f_{n}^d \leq V_r, \forall r \in R, \forall i \in I_r
\]

\[
y_{o,d}^d + x_{o,d}^d = q_{o,d}^d, \forall (o,d) \in W
\]

\[
z_{o,n}^d \geq 0, z_{n}^d \geq 0, f_{n}^d \geq 0, \forall r \in R, \forall i \in I_r, \forall d \in D
\]

\[
y_{o,d}^d \geq 0, x_{o,d}^d \geq 0, \forall (o,d) \in W
\]

In the destination-link-based model, the number of flow variables (e.g. \( f_{n}^d \)) has the magnitude of \( |D| \sum_{r \in R} N_r \), which is not greater than \( |P| \sum_{r \in R} N_r \). Similar to the origin-link-based model, the numbers of variables and constraints in the destination-link-based model are also not greater than those in the OD-link-based model, as indicated by Table 1.

3 A hybrid-link-based model

The origin-link-based and destination-link-based models may not have the smallest number of decision variables. For example, suppose that there are 7 ports \( p_1, p_2, \ldots, p_7 \) and the set of OD pairs is shown by arrows in Fig. 2a. The number of origins \( |O| = 5 \) and the number of destinations \( |D| = 6 \). That is, the number of \( f_{o}^a \) (or \( f_{n}^d \)) is \( 5 \sum_{r \in R} N_r \) (or \( 6 \sum_{r \in R} N_r \)). However, if we use origin-link-based formulation for origin port \( p_4 \) (and the OD pairs \( (p_1, p_2), (p_4, p_5), (p_6, p_7) \)), and destination-link-based formulation for destination port \( p_1 \) (and the OD pairs \( (p_2, p_4) \) and \( (p_3, p_7) \)) and \( p_7 \) (and the OD pairs \( (p_5, p_7) \) and \( (p_6, p_7) \)), then the number of flow variables is only \( 3 \sum_{r \in R} N_r \). Mathematically, we let \( \bar{O} \) be the set of origin ports and \( \bar{D} \) be the set of destination ports for the hybrid-link-based model, \( \bar{O} \subseteq O, \bar{D} \subseteq D \). In the example of Fig. 2a, we have

\[
\bar{O} = \{ p_4 \}
\]

\[
\bar{D} = \{ p_1, p_7 \}
\]
We further define $\mathcal{O}^d$ as the set of ports $p \in \mathcal{P}$ where the OD pair $(p, d) \in \mathcal{W}$ is assigned to destination port $d \in \mathcal{D}$, and $\mathcal{O}^o$ as the set of ports $p \in \mathcal{P}$ where the OD pair $(o, p) \in \mathcal{W}$ is assigned to origin port $o \in \mathcal{O}$. In the example of Fig. 2a, we have

$$\mathcal{D}^p = \{p_2, p_3, p_5, p_6\}$$
(36)

$$\mathcal{O}^p = \{p_2, p_3\}$$
(37)

$$\mathcal{O}^p = \{p_5, p_6\}$$
(38)

<table>
<thead>
<tr>
<th>origin</th>
<th>destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>$p_1$</td>
</tr>
<tr>
<td>$p_2$</td>
<td>$p_2$</td>
</tr>
<tr>
<td>$p_3$</td>
<td>$p_3$</td>
</tr>
<tr>
<td>$p_4$</td>
<td>$p_4$</td>
</tr>
<tr>
<td>$p_5$</td>
<td>$p_5$</td>
</tr>
<tr>
<td>$p_6$</td>
<td>$p_6$</td>
</tr>
<tr>
<td>$p_7$</td>
<td>$p_7$</td>
</tr>
</tbody>
</table>

(a)

<table>
<thead>
<tr>
<th>origin</th>
<th>destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>$p_1$</td>
</tr>
<tr>
<td>$p_2$</td>
<td>$p_2$</td>
</tr>
<tr>
<td>$p_3$</td>
<td>$p_3$</td>
</tr>
<tr>
<td>$p_4$</td>
<td>$p_4$</td>
</tr>
<tr>
<td>$p_5$</td>
<td>$p_5$</td>
</tr>
<tr>
<td>$p_6$</td>
<td>$p_6$</td>
</tr>
<tr>
<td>$p_7$</td>
<td>$p_7$</td>
</tr>
</tbody>
</table>

(b)

Fig. 2 Motivation of the hybrid-link-based model

Note that in the example of Fig. 2b, the OD pair $(p_4, p_1)$ can either be assigned to origin port $p_4$ or destination port $p_1$ but not both. If it is assigned to origin port $p_4$, we have

$$\mathcal{D}^p = \{p_1, p_2, p_3, p_5, p_6\}$$
(39)

$$\mathcal{O}^p = \{p_2, p_3\}$$
(40)

$$\mathcal{O}^p = \{p_5, p_6\}$$
(41)

If it is assigned to destination port $p_1$, we have

$$\mathcal{D}^p = \{p_2, p_3, p_5, p_6\}$$
(42)

$$\mathcal{O}^p = \{p_2, p_3, p_4\}$$
(43)
Suppose that we already know the sets $\bar{O}$ and $\bar{D}$, and have determined the sets $\bar{O}^d$, $d \in \bar{D}$ and $\bar{D}^o$, $o \in \bar{O}$. In the hybrid-link-based model, we use $\bar{z}^o_i$ and $\bar{z}^o_i$ to represent the total volume of containers with origin port $o \in \bar{O}$ and any destination loaded and discharged at port of call $i$ on ship route $r$, respectively (transshipped containers are also considered) and use $f_{ri}^o$ to denote the total volume of containers with origin port $o \in \bar{O}$ and any destination flowing on leg $i$ of ship route $r$. We use $\bar{z}^d_i$ and $\bar{z}^d_i$ to represent the total volume of containers with destination port $d \in \bar{D}$ and any origin loaded and discharged at port of call $i$ on ship route $r$, respectively (transshipped containers are also considered) and use $f_{ri}^d$ to denote the total volume of containers with destination port $d \in \bar{D}$ and any origin flowing on leg $i$ of ship route $r$. The meanings of $x^{od}$, $\bar{x}^{od}$, $\bar{z}_p^o$, $\bar{z}_p^d$, and $\bar{z}_p$ are the same as the OD-link-based model. The hybrid-link-based model is:

\[
\min_{\bar{z}_p^o, \bar{z}_p^d, f_{ri}^o, f_{ri}^d, y^{od}, y^{dp}} \sum_{o \in O} \left( \bar{z}_p^o \bar{c}_p + \bar{z}_p^d \bar{c}_p + \bar{z}_p \bar{c}_p \right) + \sum_{(o,d) \in W} g^{od} \bar{z}^{od}
\]

subject to:

\[
f_{ri}^o + \bar{z}^o_i = f_{ri}^o + \bar{z}^o_i, \forall r \in R, \forall i \in I_r, \forall o \in \bar{O}
\]

\[
f_{ri}^d + \bar{z}^d_i = f_{ri}^d + \bar{z}^d_i, \forall r \in R, \forall i \in I_r, \forall d \in \bar{D}
\]

\[
\bar{z}_p^o = \begin{cases} 
\sum_{(o,d) \in W} y^{od}, & \forall p \in O \\
0, & \forall p \in P \setminus O 
\end{cases}
\]

\[
\bar{z}_p^d = \begin{cases} 
\sum_{(o,p) \in W} y^{dp}, & \forall p \in D \\
0, & \forall p \in P \setminus D
\end{cases}
\]

\[
\bar{z}_p = \sum_{r \in R, i \in I_r, p_o = p} \left( \sum_{o \in O} \bar{z}_p^o + \sum_{d \in D} \bar{z}_p^d \right) - \bar{z}_p, \forall p \in P
\]

\[
\sum_{r \in R, i \in I_r, p_o = p} \left( \bar{z}_p^o - \bar{z}_p^o \right) = \begin{cases} 
\sum_{d \in D} y^{od}, & p = o \\
-y^{op}, & p \in \bar{D}^o, \forall o \in \bar{O}, \forall p \in P \\
0, & \text{otherwise}
\end{cases}
\]
In the hybrid-link-based model, the number of flow variables (e.g. $f_{ni}$ and $f_{di}$) has the magnitude of $(|\bar{O}|+|\bar{D}|) \sum_{r\in R} N_r$. We can always find a definition of sets $\bar{O}$ and $\bar{D}$ such that $|\bar{O}|+|\bar{D}| \leq \min \{|O|, |D|\}$ (for example, if $|O| \leq |D|$ , we can define $\bar{O} := O$ and $\bar{D} := \emptyset$). In other words, the hybrid-link-based model is at least as compact as the origin-link-based model and the destination-link-based model, provided that the sets $\bar{O}$ and $\bar{D}$ are appropriately defined. The numbers of variables and constraints of the hybrid-link-based model are shown in Table 1.

4 Optimal choice of origins and destinations for the hybrid-link-based model

To find the sets $\bar{O}$ and $\bar{D}$ such that $|\bar{O}|+|\bar{D}|$ is minimized, we need the following variables. Let $\theta_p \in \{0,1\}$ be a binary variable which equals 1 if and only if port $p \in P$ is selected as an element in set $\bar{O}$ and $\pi_p \in \{0,1\}$ be a binary variable which equals 1 if and only if port $p \in P$ is selected as an element in set $\bar{D}$. The sets $\bar{O}$ and $\bar{D}$ must satisfy that for each OD pair $(o,d) \in W$ at least one of the following two conditions holds: (i) $o \in \bar{O}$ and (ii) $d \in \bar{D}$. The model for optimizing the choice of origins and destinations for the hybrid-link-based formulation is:

\[ \text{[Choice-model 1]} \min_{\theta_p, \pi_p} \sum_{p \in \bar{P}} \theta_p + \sum_{p \in \bar{P}} \pi_p \]
subject to:

\[ 0_o + \pi_d \geq 1, \forall (o,d) \in W \]
\[ \theta_p \in \{0,1\}, \forall p \in \mathcal{P} \]  
\[ \pi_p \in \{0,1\}, \forall p \in \mathcal{P} \]  

The above model is very similar to the set covering problem. The set covering problem can be described as follows: given a set of elements (e.g., \( U = \{1,2,3,4,5\} \)) which is called the universe, and a set \( S \) of sets whose union equals the universe (e.g., \( S = \{\{1,2\},\{1,3,5\},\{2,4,5\},\{3,4\},\{1,5\}\} \)), find the smallest subset of \( S \) the union of which contains all elements of the universe (here the smallest subset of \( S \) is \( \{\{1,3,5\},\{2,4,5\}\} \), the cardinality of which is 2). In the above parameter optimization problem, the universe is the set of all OD pairs \( \mathcal{W} \), and the set \( S \) has 21\( \mathcal{P} \) elements and is:

\[ S = \bigcup_{p \in \mathcal{P}} \left\{ (o,d) \in \mathcal{W} \mid o = p \right\} \cup \left\{ (o,d) \in \mathcal{W} \mid d = p \right\} \]  

(62)

Minimizing \( \sum_{p \in \mathcal{P}} \theta_p + \sum_{p \in \mathcal{P}} \pi_p \) is equivalent to finding the smallest subset of \( S \) the union of which contains all OD pairs. Since the set covering problem is NP-complete, one might conjecture that the integer linear programming model (58) is also difficult to solve. In fact, the number of binary decision variables in model (58) is \( 2^{2|\mathcal{P}|} \). If the network has 200 ports, then the number of all combinations of decision variables is roughly \( 2.6 \times 10^{120} \).

By carefully examining the problem structure, we find that the parameter optimization problem is a special case of the set covering problem, but not vice versa. In particular, in the parameter optimization problem, each element in the universe is an element of exactly two elements (which are sets) in set \( S \). That is, the element of the universe \( (o^*,d^*) \in \mathcal{W} \) is an element of the two sets of \( \{ (o,d) \in \mathcal{W} \mid o = o^* \} \) and \( \{ (o,d) \in \mathcal{W} \mid d = d^* \} \), which are elements of set \( S \). As a result of this property, the parameter optimization problem is easy to solve, as shown below.

Represent by \( \mathbb{Z} \) the set of integers. It is easy to prove that the model (58) is equivalent to the following model by removing the upper bound “1” on \( \theta_p \) and \( \pi_p \):

[Choice-model 2]

\[ \min_{\theta_p, \pi_p} \sum_{p \in \mathcal{P}} \theta_p + \sum_{p \in \mathcal{P}} \pi_p \]  

subject to:

\[ \theta_o + \pi_d \geq 1, \forall (o,d) \in \mathcal{W} \]  

(63)

(64)
\[ \theta_p \geq 0, \forall p \in P \]  
\[ \pi_p \geq 0, \forall p \in P \]  
\[ \theta_p \in \mathbb{Z}, \pi_p \in \mathbb{Z}, \forall p \in P \]  

**Proposition 1**: The coefficient matrix of the linear programming relaxation of model (63) excluding the lower bound constraints (65) and (66), which is the coefficient matrix of constraint (64), is totally unimodular.

**Proof**: Let \( A \) be the coefficient matrix of constraint (64). \( A^T \) is a \( 2|P| \times |W| \) matrix. Each element of \( A^T \) is either 0 or 1 and each column contains two 1’s. If we divide \( A^T \) into two matrices: the top \(|P|\) rows constitute one matrix and the bottom \(|P|\) rows constitute the other matrix, then each matrix has exactly one element of 1 in each column. Consequently, \( A^T \) is totally unimodular and thereby \( A \) is totally unimodular. □

Since the right-hand side coefficients of constraints (64)–(66) are all integers, all the extreme point optimal solutions to model (63) are integers. Hence, the integrality constraint in Eq. (67) can be dropped. In other words, model (63) can be easily solved as a linear programming problem.

Hence, we can conclude that little additional effort is needed for identifying the optimal choice of origins and destinations in the hybrid-link-based model. This is because on one side, the optimal choice of origins and destinations is a linear programming problem; on the other side, in contrast to container routing which needs to be solved a number of times as a subproblem in tactical-level decision planning of liner shipping operations, the optimal choice of origins and destinations needs to be made only once.

5 **Numerical experiments**

We have demonstrated that in terms of the numbers of decision variables and constraints, the hybrid-link-based model is as least as good as the origin-link-based model and destination-link-based model, which are at least as good as the OD-link-based model. In this section, we conduct numerical experiments to test the solution efficiency of the models. The network and demand are randomly generated, and the models are all solved by matlab calling CPLEX 12.2 on a 3.2 GHz Dual Core laptop with 4 GB of RAM.
We first carry out an experiment with 20 ports, 5 ship routes, and a total of 16,805 twenty-foot equivalent units (TEUs) to be shipped. Table 1 indicates that the difference of the four models lies in that the numbers of variables and constraints are associated with $|W|$ for the OD-link-based model, $|O|$ for the origin-link-based model, $|D|$ for the destination-link-based model, and $|O| + |D|$ for the hybrid-link-based model. Hence, we report in Table 3 the size each model in terms of the value of $|W|$, $|O|$, $|D|$, and $|O| + |D|$, respectively. Since in an optimization model, some variables and constraints may be eliminated by a simple pre-processing step, we report in Table 3 the numbers of variables and constraints for the four models after pre-processing. We also report the CPU time (ms) for solving the problem to optimality. Results show that all the four models obtain the same optimal solution (in terms of $y_{od}$) and the same optimal objective function value. This validates the correctness of the four models. Table 3 indicates that after pre-processing, the numbers of decision variables and constraints of the hybrid-link-based model are the smallest. In terms of CPU time, the OD-link-based model is significantly inferior to the other three models.

<table>
<thead>
<tr>
<th>Table 3 20 ports, 5 ship routes, 16,805 TEUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td>#variables after pre-processing</td>
</tr>
<tr>
<td>#constraints after pre-processing</td>
</tr>
<tr>
<td>CPU time (ms)</td>
</tr>
</tbody>
</table>

We conduct more numerical experiments to further compare the origin-link-based model, the destination-link-based model, and the hybrid-link-based model, and the results are shown in Table 4 to Table 7. These results demonstrate the superiority of the hybrid-link-based model over the origin-link-based model, the destination-link-based model with regard to the numbers of variables and constraints and the CPU time. Finally, we note that in some problems the hybrid-link-based model could not significantly reduce the number of origins and destinations, for example, in networks where almost every port pair has demand. In such situations, the computational performance could not be considerably improved by using the hybrid-link-based model.
Table 4 30 ports, 10 ship routes, 214 OD pairs, 41,437 TEUs

<table>
<thead>
<tr>
<th>Model</th>
<th>Origin (O)</th>
<th>Destination (D)</th>
<th>Hybrid (O + D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>29</td>
<td>30</td>
<td>16</td>
</tr>
<tr>
<td>#variables after pre-processing</td>
<td>4,760</td>
<td>4,917</td>
<td>2,693</td>
</tr>
<tr>
<td>#constraints after pre-processing</td>
<td>2,168</td>
<td>2,238</td>
<td>1,232</td>
</tr>
<tr>
<td>CPU time (ms)</td>
<td>218</td>
<td>187</td>
<td>177</td>
</tr>
</tbody>
</table>

Table 5 30 ports, 20 ship routes, 267 OD pairs, 55,427 TEUs

<table>
<thead>
<tr>
<th>Model</th>
<th>Origin (O)</th>
<th>Destination (D)</th>
<th>Hybrid (O + D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>30</td>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>#variables after pre-processing</td>
<td>17,217</td>
<td>17,217</td>
<td>11,013</td>
</tr>
<tr>
<td>#constraints after pre-processing</td>
<td>6,758</td>
<td>6,758</td>
<td>4,360</td>
</tr>
<tr>
<td>CPU time (ms)</td>
<td>515</td>
<td>905</td>
<td>468</td>
</tr>
</tbody>
</table>

Table 6 40 ports, 10 ship routes, 465 OD pairs, 90,490 TEUs

<table>
<thead>
<tr>
<th>Model</th>
<th>Origin (O)</th>
<th>Destination (D)</th>
<th>Hybrid (O + D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>40</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>#variables after pre-processing</td>
<td>8,250</td>
<td>8,242</td>
<td>5,082</td>
</tr>
<tr>
<td>#constraints after pre-processing</td>
<td>3,547</td>
<td>3,541</td>
<td>2,171</td>
</tr>
<tr>
<td>CPU time (ms)</td>
<td>350</td>
<td>422</td>
<td>296</td>
</tr>
</tbody>
</table>

Table 7 40 ports, 20 ship routes, 458 OD pairs, 91,963 TEUs

<table>
<thead>
<tr>
<th>Model</th>
<th>Origin (O)</th>
<th>Destination (D)</th>
<th>Hybrid (O + D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>40</td>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>#variables after pre-processing</td>
<td>22,328</td>
<td>22,326</td>
<td>14,138</td>
</tr>
<tr>
<td>#constraints after pre-processing</td>
<td>8,983</td>
<td>8,981</td>
<td>5,698</td>
</tr>
<tr>
<td>CPU time (ms)</td>
<td>1,467</td>
<td>2,153</td>
<td>905</td>
</tr>
</tbody>
</table>

6 Conclusions

This study has proposed a hybrid-link-based container routing model that is more compact than existing OD-link-based, origin-link-based, and destination-link-based models. The idea of the hybrid-link-based model is that an appropriate combination of origins and destinations could reduce the number of decision variables because not all port pairs have demand. We further formulated an integer linear programming model to identify the optimal choice of origins and destinations in the hybrid-link-based model. We rigorously proved that this integer linear programming model has the totally unimodular property and hence can be easily solved as a linear programming problem.
The contribution of the paper to the literature is the proposition of the novel hybrid-link-based container routing model. This model nests the origin-link-based and destination-link-based models as special cases, and is at least as compact as the other link-based model. The hybrid-link-based model can also be used to solve multi-commodity flow problems.

References


