Crew Assignment with Duty Time Limits for Transport Services: Tight Multi-commodity Models

A. Balakrishnan, P. Mirchandani, and S. Lin

Electronic Companion, Part 1
Appendices B to E

Appendix B: Proofs of Propositions

Proposition 1: The CSP2DT problem is polynomially solvable, in $O\big(n(m + n\,logn)\big)$ time.

Proof of Proposition 1: Consider a feasible single-depot CA-DT problem instance with two crew members, denoted as the CSP2DT problem. We index the n trip nodes $I = \{1,2,\dots,n\}$ in non-decreasing order of start times s_i, and include in the underlying graph an origin node o and a destination node d (with appropriate incident arcs) representing the start and end of the duties for the crew members. Let $LRT = \max_{i:(i,d)\in A}(f_i + \tau_{id})$ denote the latest time at which a crew member returns to the depot when all trips are completed. We refer to the trip with this maximum value as the 'last' trip. Define $EST = \min_{i:(d,i)\in A}(s_i - \tau_{di})$ to be the earliest time at which a crew member needs to leave the depot to operate any of the n trips. We refer to the trip with this minimum value as the 'first' trip. Without loss of generality, we can shift all trip start and end times so that $EST = 0$.

For each trip i, let ERT_i denote the earliest depot return time after a crew member completes trip i. This time may be different from $f_i + \tau_{id}$ if the arc set A does not contain the depot-return arc (i,d) (i.e., a crew member cannot directly return to the depot after completing trip i) or if the transit (plus appropriate trip times) times do not satisfy the triangle inequality. Let LST_i denote the latest time that a crew member must start from the depot in order to operate trip i (again, this time may be different from $(s_i - \tau_{di})$ if either the network does not contain arc (d,i) or transit times (together with appropriate trip times) do not satisfy the triangle inequality).

Let crew 1 denote the person who is assigned to the first trip; crew 2 is the other crew member. We first partition the set of trips I into the following three subsets:

- $I_1 = \{i \in I: (LRT - LST_i) > DT_{\max}\}$;
- $I_2 = \{i \in I: ERT_i > DT_{\max}\}$; and
- $I_3 = I\backslash(I_1 \cup I_2)$.

The set I_1 represents the set of trips that cannot be assigned to the crew member who operates the last trip due to the maximum DT restriction. Conversely, the set I_2 is the set of trips that cannot be assigned to the crew member who operates the first trip due to the DT limit $DT_{\max}$. I_3 is the remaining set of trips that can be assigned to either crew member. Note that I_1 will be empty if and only if $LRT \le DT_{\max}$, i.e., the DT limits are not binding for any valid assignment of trips to a crew member. In this case, I_2 must also be empty, and all the trips can be assigned to either crew member, subject to assignment validity requirements. Otherwise, if I_1 is non-empty (and so I_2 is also non-empty), then crew 1 must necessarily operate all the trips in I_1, and cannot operate any of the trips in I_2. Conversely, crew 2 must necessarily operate the last trip and all other trips in I_2, and cannot operate any trip in I_1.

For the remainder of this discussion, we will assume that I_1 and I_2 are non-empty (the method also extends when these sets are empty. To ensure that the problem has a feasible solution, it must satisfy the following necessary conditions:

(i) For all trips $i \in I_1$ arranged in order of non-decreasing start times, the network contains a path from the depot to each of these trips in sequence (specifically, the trips must be non-overlapping and transfers between adjacent trips in the sequence are permitted); and,

(ii) For all trips $j \in I_2$ arranged in order of non-decreasing start times, the network contains a path through each of these trips in sequence, starting from the first trip in this set, and ending at the depot.

We first assign all trips in I_1 (in sequence, starting from the depot) to crew 1 and all trips in I_2 (in sequence, returning to the depot after the last trip) to crew 2, and compute the associated crew usage, assignment, and transfer costs. Let i^* denote the last trip in I_1 assigned to crew 1, and j^* the first trip in I_2 assigned to crew 2. To determine which trips in I_3 to assign to crews 1 and 2, we solve a minimum cost assignment problem defined over the following bipartite network BG. This approach is a variant of the approach discussed in Bertossi et al. (1987). The network contains two nodes corresponding to each trip $i \in I_3$, one labeled i_e in the left-hand side layer and the other labeled i_s in the right-hand side layer. These two nodes represent, respectively, the end and start of trip i. In addition, the network contains two pairs of nodes corresponding to the two crew members. Nodes $d1_e$ and $d1_s$, in the left- and right-hand sides of the network respectively represent the ending and starting 'events' of crew 1. Similarly, nodes $d2_e$ and $d2_s$ denote the ending and starting of crew 2. Network BG contains the following arcs:

- Arc $(d1_e, j_s)$ for all $j \in I_3$ such that $(i^*, j) \in A$, with cost $c_{i^*,j}$, and arc $(i_e, d1_s)$ for all $i \in I_3$ such that $(i, d) \in A$, with cost c_{id}. Selecting these arcs in the assignment respectively correspond to scheduling trip j as the next trip after i^* for crew 1, and assigning trip i as the last trip for this crew member before returning to the depot;
- Arc $(i_e, d2_s)$ for all $i \in I_3$ such that $(i, j^*) \in A$, with cost c_{i,j^*}, and arc $(d2_e, j_s)$ for all $j \in I_3$ such that $(d, j) \in A$, with cost c_{dj} Selecting these arcs in the assignment respectively correspond to scheduling trip i as the trip just preceding trip j^* for crew 2, and assigning trip j as this crew member's first trip; and
- Arcs (i_e, j_s) for all $(i, j) \in A$, with $i, j \in I_3$, having cost c_{ij}. Choosing this arc corresponds to assigning trip j immediately after trip i to a crew member.

The optimal solution to the minimum cost assignment problem on network BG provides the optimal assignment (and sequencing) of trips in I_3 to the two crew members.

The minimum-cost assignment problem on a bipartite graph can be solved in $O(n(m + nlogn))$ (Fredman and Tarjan, 1987), where n and m respectively denote the number of trips and number of arcs in the graph. So, the overall complexity of the solution method (including the preprocessing stage to identify and assign the trips in I_1 and I_2) for the CSP2DT problem is $O(n(m + nlogn))$. ♦

Proposition 2: For crew assignment with DT restrictions, the optimal LP value of the multi-commodity model [*CA-DT*] equals the optimal LP value of the set partitioning model [*SP-CA-DT*].

Proof of Proposition 2. For each depot $d \in D$, let Π^d denote the set of all the feasible o_d-to-e_d paths for crew members from this depot. Set $\delta_{ij}(p)$ =1 if arc (i, j) belongs to path p, and 0 otherwise. Let $\boldsymbol{y} = \{y_p^d\}$ denote a feasible solution to the LP relaxation of [*SP-CA*], and let ϕ_p^d, for $d \in D$, denote the first trip node on path p. Since each path $p \in \Pi^d$ satisfies the trip assignment, transfer time and DT constraints, every arc (i, j) in G' satisfying $\delta_{ij}(p)$ has a corresponding arc in G^{kd} (in which node d corresponds to e_d) except arc (o_d, ϕ_p^d). The solution $\{\boldsymbol{x}, \boldsymbol{z}\} = \{x_{ij}^{kd}, z^{kd}\}$ constructed by setting, for each $d \in D$:

$$z^{kd} = \sum_{p \in \Pi^d : k = \phi_p^d} y_p^d \qquad \text{for all } \langle k, d \rangle \in K \text{, and}$$

$$x_{ij}^{kd} = \sum_{p \in \prod^d : k=\phi_p^d} y_p^d \, \delta_{ij}(p) \qquad \text{for all } (i,j) \in A^{kd}, \langle k,d \rangle \in K$$

is feasible to formulation [*CA-DT*] and costs the same as solution $\boldsymbol{y}$. The feasibility of solution $\{\boldsymbol{x}, \boldsymbol{z}\}$ follows: Constraints (3.2) are met since for each $p \in \prod^d$ there is a corresponding path from $k = \phi_p^d$ to d in G^{kd}, and constraints (3.3) and (3.4) follow from (4.2) and (4.3) and the construction of $\{\boldsymbol{x}, \boldsymbol{z}\}$ above.

Conversely, given a feasible solution $\{\boldsymbol{x}, \boldsymbol{z}\} = \{x_{ij}^{kd}, z^{kd}\}$ to model [*CA-DT*], construct a flow decomposition F^{kd} consisting of paths p each with flow $\theta(p)$ for each $\langle k,d \rangle \in KD$ with $z^{kd} > 0$. Observe that since G^{kd} is acyclic, there are no cycles in F^{kd}, and because all paths from k to d in G^{kd} are feasible, all paths in F^{kd} are feasible (that is, satisfy the trip assignment, transfer time and DT constraints). By the flow decomposition property, for each $\langle k,d \rangle \in KD$, $z^{kd} = \sum_{p \in F^{kd}} \theta(p)$ and $x_{ij}^{kd} = \sum_{p \in F^{kd}} \delta_{ij}(p)\theta(p)$. For each $p \in F^{kd}$, construct a corresponding path p' in G' with arc (o_d, k) appended to p with flow $\theta(p)$. Note that node d in G^{kd} corresponds to node e_d in G'. Since $p \in F^{kd}$ is feasible in G^{kd}, p' is feasible in G'. Construct a solution $\boldsymbol{y} = \{y_{p'}^d\}$ for model [*SP-CA*] by setting $y_{p'}^d = \theta(p)$. Feasibility of (4.2) and (4.3) follows from (3.3) and (3.4), and the cost of solution $\boldsymbol{y}$ is the same as that of $\{\boldsymbol{x}, \boldsymbol{z}\}$. ♦

Proposition 3: For single-depot crew assignment problems with DT restrictions, the optimal LP value of the multi-commodity model [*CA-DT-EQ*] is at least as high as BC's initial Lagrangian dual value.

Proof of Proposition 3. Let model [*CA-DT-LPR*] denote the LP relaxation of model [*CA-DT-EQ*] obtained by relaxing the binary requirements (3.5) and Z_{CS-DT}^{LPR} the corresponding objective function value. Let [*CA-DT-LAG*($\boldsymbol{\mu}$)] denote the Lagrangian relaxation of model [*CA-DT-EQ*] obtained by dualizing the trip coverage constraints (3.4) using the multipliers $\boldsymbol{\mu} = \{\mu_i\}$. Let

$$Z_{CS-DT}^{LAG(\mu)} = \min\left(\sum_{\langle k,d \rangle \in KD} \left(\sum_{(i,j) \in A^{kd}} c_{ij}^{kd} x_{ij}^{kd} + c_{dk}^{kd} z^{kd} \right) + \sum_{i \in I} \mu_i \left(\sum_{d \in DF(i)} z^{id} + \sum_{\langle k,d \rangle \in KD, k \neq i} \sum_{(j,i) \in A^{kd}} x_{ji}^{kd} - 1 \right) \right)$$

subject to (3.2), (3.3) and (3.5). Let $Z_{CS-DT}^{LAG} = \max_{\boldsymbol{\mu}} Z_{CS-DT}^{LAG(\boldsymbol{\mu})}$. Let $\boldsymbol{A}$ denote the constraint matrix corresponding to constraints (3.2) and (3.3) such that these two sets of constraints are equivalent to $\boldsymbol{A}\binom{\boldsymbol{x}}{\boldsymbol{z}} \leq \boldsymbol{b}$ where $\boldsymbol{x}$ is the vector corresponding to the commodity flow variables, $\boldsymbol{z}$ is the vector corresponding to the crew selection variables, and $\boldsymbol{b}$ corresponds to the right-hand side of (3.2) and (3.3). Note that the elements of $\boldsymbol{b}$ are integer-valued (0 or M_d).

Table B1: Constraint matrix structure for Proposition 3

$\boldsymbol{A}_{k_1 d_p}$	**0**	**0**	**0**	$\boldsymbol{e}(k_1, d_p)$	**0**	**0**	**0**
0	$\boldsymbol{A}_{k_2 d_p}$	**0**	**0**	**0**	$\boldsymbol{e}(k_2, d_p)$	**0**	**0**
0	**0**	…	**0**	**0**	**0**	…	**0**
0	**0**	**0**	$\boldsymbol{A}_{k_{n_p} d_p}$	**0**	**0**	**0**	$\boldsymbol{e}(k_{n_p}, d_p)$
0	0	0	0	1	1	1	1

We will first show that model [*CA-DT-LAG*($\boldsymbol{\mu}$)] satisfies the integrality property (Geoffrion's 1974), implying that $Z_{CS-DT}^{LPR} = Z_{CS-DT}^{LAG}$. To show that [*CA-DT-LAG*($\boldsymbol{\mu}$)] has integer optimal solutions, we prove that the coefficient matrix $\boldsymbol{A}$ corresponding to constraints (3.2) and (3.3) is totally unimodular. This matrix is block decomposable by depot. So, it suffices to show that the block corresponding to each depot is totally unimodular. Let $\langle k_q, d_p \rangle$, $q = 1, 2, \ldots, n_p, n_p \geq 1$ be the feasible commodities for depot

$d_p \in D$. For each q, let $\boldsymbol{A}_{k_q d_p}$ denote the node arc incidence matrix for graph $G^{k_q d_p}$, and $\boldsymbol{e}(k_q, d_p)$ denote a $V^{k_q d_p}$-dimension vector with value -1 in position k_q, 1 in position d_p, and zero elsewhere. **0** represents an appropriately dimensioned zero matrix (or vector). The block in matrix $\boldsymbol{A}$ for depot d_p has the structure shown in Table B1.

The block consists of n_p sets of rows, one for each commodity; the block's last row corresponds to the crew availability constraint. Each column of each $\boldsymbol{A}_{k_q d_p}$ has one -1 and one +1 coefficient, and all other entries in the column are zero. Similarly, each $\boldsymbol{e}(k_q, d_p)$ has one -1 and one +1, and all other entries are zero. For any subset of rows R, construct a partition R_1 and R_2 as follows. For $q = 1, 2, \ldots, n_p$, if the row corresponding to the -1 entry of $\boldsymbol{e}(k_q, d_p)$ is in R and the row corresponding to the +1 entry is not in R, then include all the rows of R corresponding to q in R_1, else include them in R_2. Include the row corresponding to the crew availability constraint in R_1. For this partition, the absolute value of the difference of entries in R_1 and R_2 is no more than one for every column. Hence, the block corresponding to depot d_p is totally unimodular, implying that $\boldsymbol{A}$ is totally unimodular and $Z_{CS-DT}^{LPR} = Z_{CS-DT}^{LAG}$.

Let $(\boldsymbol{f}^*, \boldsymbol{\mu}^*)$ denote the optimal Lagrangian dual solution for BC's approach, that is, the Lagrangian dual solution to the single depot version of [*SC-CA-DT*] with (4.7) set as equalities, (4.9) incorporated implicitly, and (4.8) dualized using multipliers $\boldsymbol{\mu}^*$. In addition, the BC procedure relaxes the binary constraints (4.10) to general integer constraints, $f_{gh} \geq 0$ and integer for all $(g, h) \in A'$. Let d denote the single depot (recall that BC study the single depot problem). To generate the flow vector $\boldsymbol{f}^*$, the BC procedure finds the lowest (Lagrangian cost) path from o_d directly to i and then from i to e_d for each $i \in I$. By construction, these paths all satisfy the DT constraints. Their algorithm then selects the M_d lowest cost paths, and sets f_{gh}^* equal to the total number of times arc (g, h) appears on these paths. Note that the f_{gh}^* values may exceed one.

Let $(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{\mu}^*)$ denote the optimal solution to the Lagrangian problem *CA-DT-LAG*($\boldsymbol{\mu}$*) (with a single depot o_d). Set $\bar{f}_{o_d k} = (z^{kd})^* \; \forall k \in I$, and $\bar{f}_{ij} = \sum_{\langle k,d \rangle} (x_{ij}^{kd})^* \;\; \forall (i, j) \in A', i \neq o_d$. Since $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ corresponds to a feasible flow (which satisfies DT constraints), $\bar{\boldsymbol{f}}$ is feasible to the Lagrangian dual in BC. Hence, $Z_{CS-DT}^{LAG(\boldsymbol{\mu}^*)}$ is at least the optimal value of the Lagrangian relaxation solution value of the BC approach. Proposition 3 follows since $Z_{CS-DT}^{LPR} = Z_{CS-DT}^{LAG} \geq Z_{CS-DT}^{LAG(\boldsymbol{\mu}^*)}$. ♦

Proposition 5: (i) For any E-E subgraph *EE* of G' with an odd number ρ of conflict relationships, the following inequality is valid for the CA-DT problem and strengthens model [*CA-DT*]:

$$\sum_{q=1}^{Q} X_{i_q, j_q}^{KD_q} - \sum_{q \in Cond(\sigma_{EE})} X_{g_q, h_q}^{KD'_q} \leq \frac{(\rho - 1)}{2}. \tag{5.1}$$

Proof of Proposition 5: (i) By construction, every pair of consecutive indices, q and $q + 1$ in the sequence $q = 1, 2, \ldots Q$ (we set $Q + 1$ to equal 1) have either a conflict or a conditional relationship. Recall that $Conf(\sigma_{EE})$ denotes the subset of indices q for which we have a conflict relationship between flows on arcs (i_q, j_q) and (i_{q+1}, j_{q+1}) and the corresponding inequality is

$$X_{i_q, j_q}^{KD_q} + X_{i_{q+1}, j_{q+1}}^{KD_{q+1}} \leq 1 \qquad \forall\, q \in Conf(\sigma_{EE}). \tag{A.1a}$$

The remaining set of indices, that is, $\{1, 2, \ldots, Q\} \backslash Conf(\sigma_{EE}) \equiv Cond(\sigma_{EE})$ define a conditional relationship of one of the following types:

$$X_{i_q, j_q}^{KD'_q} \leq X_{i_{q+1}, j_{q+1}}^{KD'_q} \equiv X_{g_q, h_q}^{KD'_q} \quad \text{or} \quad X_{i_{q+1}, j_{q+1}}^{KD'_q} \leq X_{i_q, j_q}^{KD'_q} \equiv X_{g_q, h_q}^{KD'_q} \qquad \forall\, q \in Cond(\sigma_{EE}). \tag{A.1b}$$

Adding inequalities (A.1a), (A.1b), and the identities $X_{g_q, h_q}^{KD'_q} = X_{g_q, h_q}^{KD'_q}, \forall\, q \in Cond(\sigma_{EE})$ and dividing

both sides by two, and rounding down the right side, leads to inequalities (5.1). ♦
(The example in the paper, following Proposition 5, establishes part (ii) of this proposition.)

Proposition 7: (i) For given mutually exclusive subsets of trip nodes I_1, I_2, I_3, and depots $D_1 \subset D$ and $D_2 = D \backslash D_1$, the following Multi-commodity Partition inequality is valid for the CA-DT problem:

$$Y^{D_1}(D_1, I_{13}) + Y^{D_1}(I_4, I_{13}) + Y^{D_2}(I_{13}, D_2) + Y^{D_2}(I_{13}, I_4) + 2Y^D(I_1, I_1) + 2Y^D(I_3, I_3) + Y^D(I_2, I_2) + Y^D(I_1, I_2) + 2Y^D(I_1, I_3) + Y^D(I_2, I_3) \le 2|I_1| + |I_2| + 2|I_3| - 2, \quad (5.3)$$

where $I_{13} = I_1 \cup I_3$, and $I_4 = I \backslash \{I_1 \cup I_2 \cup I_3\}$.

(ii) Inequality (5.3) is at least as tight as, and can be strictly tighter than, FLMT's Lifted Path Inequality.

Proof of Proposition 7. The proof of the validity of (5.3) relies upon a linear combination of the EPEC, Strengthened Cut-IN and Cut-OUT inequalities. Let $I_{123} = I_1 \cup I_2 \cup I_3$. Then, we have the EPEC Constraints:

$$Y^{D_1}(V \backslash I_{123}, I_{123}) + Y^{D_1 \cup D_2}(I_{123}, I_{123}) + Y^{D_2}(I_{123}, V \backslash I_{123}) = |I_{123}| = |I_1| + |I_2| + |I_3|, \text{ and} \quad (A.2a)$$

$$Y^{D_1}(V \backslash I_{13}, I_{13}) + Y^{D_1 \cup D_2}(I_{13}, I_{13}) + Y^{D_2}(I_{13}, V \backslash I_{13}) = |I_{13}| = |I_1| + |I_3|. \quad (A.2b)$$

Strengthened Cut-IN and Strengthened Cut-OUT inequalities:

$$Y^{D_1 \cup D_2}(I_1, V \backslash I_1) + Y^{D_1 \cup D_2}(I_1, I_1) = |I_1|, \text{ and} \quad (A.2c)$$

$$Y^{D_1 \cup D_2}(V \backslash I_3, I_3) + Y^{D_1 \cup D_2}(I_3, I_3) = |I_3|. \quad (A.2d)$$

Subtour elimination constraints:

$$Y^{D_1 \cup D_2}(I_1, I_1) \le |I_1| - 1, \quad (A.2e)$$

$$Y^{D_1 \cup D_2}(I_2, I_2) \le |I_2| - 1, \text{ and} \quad (A.2f)$$

$$Y^{D_1 \cup D_2}(I_3, I_3) \le |I_3| - 1. \quad (A.2g)$$

Using a weight of 1/3 for inequalities (A.2a), (A.2e) and (A.2g), and a weight of 2/3 for the remaining four inequalities, adding the weighted inequalities, and rounding down the right hand side gives inequality (5.3). In this derivation, we have expanded terms. For example,

$$Y^{D_1 \cup D_2}(I_{13}, I_{13}) = Y^{D_1 \cup D_2}(I_1, I_1) + Y^{D_1 \cup D_2}(I_3, I_3) + Y^{D_1 \cup D_2}(I_1, I_3) + Y^{D_1 \cup D_2}(I_3, I_1).$$

Now, note that

$$Y^{D_1}(V \backslash I_{123}, I_{13}) = Y^{D_1}(D_1, I_{13}) + Y^{D_1}(I_4, I_{13}) \text{ and } Y^{D_2}(I_{13}, V \backslash I_{123}) = Y^{D_2}(I_{13}, D_2) + Y^{D_2}(I_{13}, I_4),$$

and thus we can write (5.3) as

$$Y^{D_1}(D_1, I_{13}) + Y^a(I_4, I_{13}) + Y^{D_2}(I_{13}, D_2) + Y^{D_2}(I_{13}, I_4) + 2Y^{D_1 \cup D_2}(I_1, I_1) + 2Y^{D_1 \cup D_2}(I_3, I_3) + Y^{D_1 \cup D_2}(I_2, I_2) + Y^{D_1 \cup D_2}(I_1, I_2) + 2Y^{D_1 \cup D_2}(I_1, I_3) + Y^{D_1 \cup D_2}(I_2, I_3) + Y^a(I_2, I_1) + Y^a(I_3, I_1) + Y^b(I_3, I_1) + Y^b(I_3, I_2) + Y^b(I_3, I_4) \le 2|I_1| + |I_2| + 2|I_3| - 2. \quad (A.3)$$

Expressing the above inequality using FLMT's single-commodity variables, we get

$$f(D_1, I_{13}) + Y^a(I_4, I_{13}) + f(I_{13}, D_2) + Y^b(I_{13}, I_4) + 2f(I_1, I_1) + 2f(I_3, I_3) + f(I_2, I_2) + f(I_1, I_2) + 2f(I_1, I_3) + f(I_2, I_3) + Y^a(I_2, I_1) + Y^a(I_3, I_1) + Y^b(I_3, I_1) + Y^b(I_3, I_2) + Y^b(I_3, I_4) \le 2|I_1| + |I_2| + 2|I_3| - 2$$

which is equivalent to

$$Y^a(I_4, I_{13}) + Y^b(I_{13}, I_4) + Y^a(I_2, I_1) + Y^a(I_3, I_1) + Y^b(I_3, I_1) + Y^b(I_3, I_2) + Y^b(I_3, I_4) + f(D_1, I_{13}) + f(I_1, I_2 \cup D_2) + f(I_2, I_{23}) + f(I_3, D_2) + 2f(I_1, I_{13}) + 2f(I_3, I_3) \le 3 + 2(|I_1| - 1) + 2(|I_2| - 1) + 2(|I_3| - 1). \quad (A.4)$$

Inequality (A.4) is at least as tight as, and can be strictly tighter than, the Lifted Path Inequality (Inequality 17) in FLMT because of the first seven terms on the left hand side all of which are nonnegative. ♦

Proposition 8: Finding a feasible solution for a single-depot WT-constrained crew assignment problem with two or more crew members, even without DT limits, is NP-hard.

Proof of Proposition 8. We will prove this assertion by transforming the Partition problem, denoted as Problem PP, which is NP-hard, in polynomial time to the WT-constrained crew assignment with two crew members, denoted as Problem CSP2WT.

Problem PP is defined as follows.

Problem PP. Given R integers $S_1, S_2, \dots, S_R$, whose sum is Q, partition the integers into two disjoint sets U and V such that $\sum_{i \in U} S_i = \sum_{i \in V} S_i = Q/2$.

We reduce Problem PP to Problem CSP2WT as follows. The CSP2WT problem instance has R trips, indexed in the same order as the R integers. The start time of trip 1, $s_1 = 0$, and the start time s_i of trip i, $2 \leq i \leq R$, is $\sum_{j=1}^{i-1} S_j$. The finish time, f_i of trip i is $s_i + S_i$. All transfer times, including those to and from the depot are zero. All costs are also zero. (The start and finish locations of the trips are not important and can be anywhere.) The working time limit for both crew members is $Q/2$. Then, any feasible solution to Problem CSP2WT gives a duty of total working time of at most $Q/2$ for each of the two crew members. Since the total working time of all trips is Q, the working time of each crew member must be $Q/2$. The subsets of integers corresponding to the trips in each of the two duties gives a partition of the R integers into two subsets each with sum equal to $Q/2$. If the CSP2WT problem is infeasible, then so is the given Partition problem. Since the Partition problem is NP-hard, this reduction shows that verifying the feasibility of a CSP2WT problem instance is also NP-hard.

This proof can be extended to $J > 2$ crew members as follows. Start with the construction above. Next, add $J - 2$ trips. For each such trip j $s_j = 0$ and $f_j = Q/2$ (that is, the trip duration of each trip j is $Q/2$). Clearly, we will need $J - 2$ of the J crew members for these newly added trips, and the remaining two crew members will service the original R trips. ♦

Proposition 9: If $\bar{c}_{ij}^{kd} > (\hat{Z} - Z_{LP}^R)$, then x_{ij}^{kd} must be zero in every optimal solution to the original problem.

Proof of Proposition 9: Since $\bar{c}_{ij}^{kd} > \left(\hat{Z} - Z_{LP}^R\right) \geq 0$, the LP primal and dual solutions to the restricted CA-DWT model are optimal even if we include the variable x_{ij}^{kd} in the model. Further, since the reduced cost for all omitted arc flow variables are non-negative, Z_{LP}^R equals the LP value Z_{LP} for the original problem. Suppose we add the constraint $x_{ij}^{kd} \geq 1$ to this model (with variable x_{ij}^{kd}) in order to force the commodity $\langle k, d \rangle$ to flow on arc (i, j). Then, by the definition of this arc flow variable's reduced cost, we know that the LP relaxation value of the augmented model together with the added constraint must be at least $Z_{LP}^R + \bar{c}_{ij}^{kd} = Z_{LP} + \bar{c}_{ij}^{kd} > \hat{Z}$, where the latter inequality follows from the condition $\bar{c}_{ij}^{kd} > \left(\hat{Z} - Z_{LP}^R\right)$. In other words, the LP value of the augmented model with the added constraint exceeds a known upper bound for the problem, implying that x_{ij}^{kd} cannot be one in the optimal solution to the original problem. ♦

References

Bertossi, A. A., P. Carraresi, and G. Gallo. 1987. On some matching problems arising in vehicle scheduling models, *Networks*, **17**, 271-281.

Fredman, M. L., and R. E. Tarjan. 1987. Fibonacci heaps and their uses in improved network optimization algorithms, *J. Assoc. Comput. Mach.*, **34**, 596-615.

Appendix C: Minimum DT Requirement

Adding the minimum DT requirement can make the trip assignments more equitable. Likewise, excluding this requirement can decrease equitability. Specifically, we show that without the minimum DT constraint, the difference between the highest and lowest DTs of a schedule as a proportion of the lowest DT can be arbitrarily large.

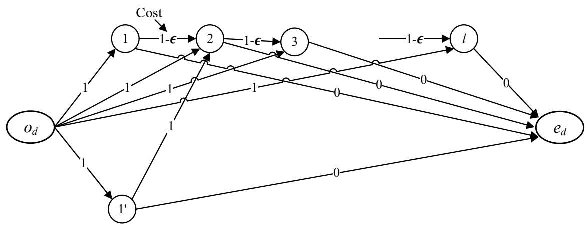

Figure C1(a): Problem Instance

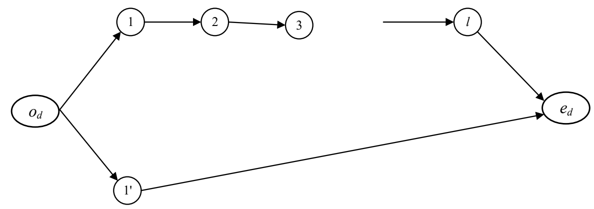

Figure C1(b): Solution without Minimum DT Constraint

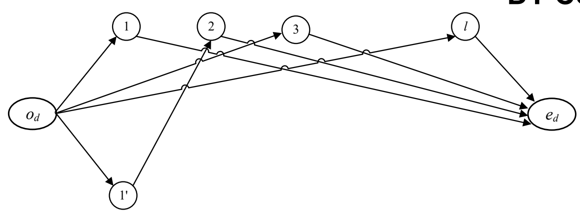

Figure C1(c): Solution with Minimum DT Constraint

Figure C1(a) describes the problem instance. The nodes o_d and e_d denote the depot node, and the instance has $l + 1$ trip nodes: $1, 1', 2, \dots, l$. The numbers on the arcs denote the costs, where ϵ denotes a small number. Missing arcs have high cost. Each trip duration is M (a large number), except $1'$ which has a duration of one. Figure C1(b) describes the lowest cost solution without the minimum DT constraint with one trip having a DT of lM and the other trip having a DT of one. Thus, the ratio of the difference between the highest and the lowest DT as a function of the lowest DT is $lM - 1$. If we impose a minimum DT of M, then we use l crew members and one trip has a DT of $M + 1$ while all the others have a DT of M. In this case, the ratio of the difference between the highest and the lowest DT as a function of the lowest DT becomes $1/M$, which corresponds to a more equitable solution.

Appendix D: E-E Inequalities

D.1 Conditional Relationships in E-E Inequalities

To show the power of conditional inequalities in an E-E inequality, consider the structure in Figure D1. In this structure, there are three trip nodes (2, 3, and 4) and node 1 is a depot node. There are two commodities—denoted by a solid line (commodity A), and a dashed line (Commodity B). The fractions written on each arc denote the flow value on that arc. One conditional inequality ($x_{31}^A \le x_{12}^A$) and three conflict inequalities ($x_{12}^A + x_{24}^B \le 1, x_{24}^B + x_{23}^B \le 1$ and $x_{23}^B + x_{31}^A \le 1$) form the E-E inequality $x_{31}^A + x_{24}^B + x_{23}^B \le 1$, which is violated by the current solution.

Observe that for this example, the thorny submultigraph inequality is not effective for three reasons: (i) the total flow on the 3-node cycle is one, (ii) adding an extension edge to any of the cycle nodes violates the second condition in the definition of a thorny submultigraph in HMS, and (iii) one of the nodes in the set that defines the inequality is a depot node.

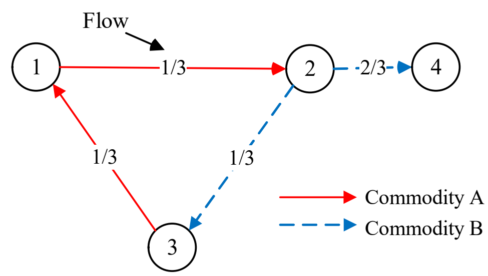

Figure D1: Effectiveness of using Conditional Inequalities

D.2 Separation Procedure for E-E Inequality

We now describe the separation procedure for the E-E inequality. Let $G_S: (I, A_S)$ denote the subgraph of G in which the nodes correspond to trips and the arcs correspond to arcs that carry positive flow corresponding to the current LP solution. We refer to this graph as the support graph. In general, there can be an exponential number of E-E inequalities, and we develop a heuristic procedure to identify violated E-E inequalities.

Given the support graph $G_S: (I, A_S)$, our heuristic to identify a violated E-E inequality runs as follows. We start with a single arc and then extend it into a path or simple cycle by adding arcs that are in conflict with the last arc added. Extending the current subgraph from the last added arc $(i, j) \in A_S$, we identify and add arc $(j, l) \in A_S$ or arc $(l, j) \in A_S$ that is in conflict with arc (i, j). To ensure that the computational time is not excessive, we add up to nine arcs in this fashion. Specifically, we considered the following Eulerian subgraph structures, and permitted adding at most one extension: (i) three-node cycle (ii) four-node cycle, and (iii) two cycles, each with three or four nodes, attached at a common node. To each of the above Eulerian cycles, we added up to one arc as an extension. So, the largest among these subgraphs—consisting of two four-node cycles with a common node together with one extension—contains a total of nine arcs. If the E-E inequality corresponding to any of these structures is violated, we add it to the formulation. We chose not to consider larger subgraphs since the time for the search procedure can increase exponentially with the number of nodes spanned by the subgraph.

Appendix E: Duty-splitting method for omitting arcs in the Restricted problem

To determine whether to eliminate the flow variable x_{ij}^{kd} for commodity $\langle k, d\rangle$ on arc (i, j) from the CA-DWT model when constructing the restricted problem instance, we apply the following *a priori* (before solving the model or its LP relaxation) cost-based test.

Consider a solution that routes commodity $\langle k, d\rangle$ on arc $(i, j) \in A^{kd}$. Figure E1(a) shows a schematic of this commodity's given path through the network. Node $l \in L^{kd}$ is a possible last trip node for the crew member that is reachable from trip j in network G.

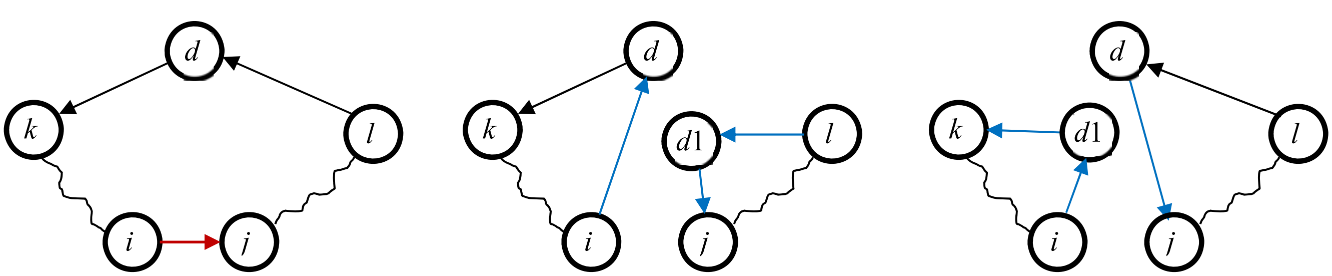

Figure E1(a): Original duty **Figure E1(b): Split duties 1** **Figure E1(c): Split duties 2**

We consider two alternate solutions that split the original duty into two separate duties, but without traversing arc (i, j). In the first alternate solution (Figure E1(b)), the crew member who was assigned the original duty starts at k from depot d, and returns to the depot from trip i (instead of operating trip j and subsequent trips), while a second crew member possibly from a different depot $d1 \in DF(j) \cap DL(l)$ starts with trip j, operates all the trips from j to l and returns to $d1$. In the second alternate solution (Figure E1(c)), the original crew member from depot d operates the latter half of the original duty (from trip j to l), while another crew member from a depot $d1 \in DF(k) \cap DL(i)$ operates trip k to trip i. Observe that, under the mild assumption that the transfer times from depot $d1$ (if different from depot d) to trip j, or from trip i to depot $d1$ are not excessive, both the split tour solutions satisfy the DT and WT limits since each crew member operates only a subset of the trips assigned to the crew member in the original solution. Note, however, that both these solutions use an additional crew member, assumed to be available at some depot $d1$. We wish to determine the maximum possible cost difference between the original tour and the two tours with split duties. For this purpose, we will identify the trip l and depot $d1$ that maximize the cost difference. For simplicity, we assume that a crew member from any depot can start or end with any trip, i.e., $DF(k) = DL(l) = D$ for all $k, l \in I$. Since we are focusing on a particular commodity $\langle k, d \rangle$, we will omit the commodity superscript for the cost parameters.

First, comparing the costs of the Split duties 1 (Figure E1(b)) and the original duty, the maximum cost difference is:

$$\Delta 1 = Max_{d1,l}\left(c_{id} + \gamma^{j,d1} + c_{l,d1} - c_{ld}\right) - c_{ij},$$

where l is the set of all nodes in L^{kd} that are reachable from node j. Next, the maximum cost difference between Split duties 2 (Figure(c)) and the original duty is:

$$\Delta 2 = Max_{d1}\left(\gamma^{k,d1} + c_{i,d1} + \gamma^{jd} - \gamma^{kd}\right) - c_{ij}.$$

So, the better of the two solutions (one that minimizes the maximum cost difference) has a cost difference of $\Delta = Min\{\Delta 1, \Delta 2\}$. If this value is non-positive, then omitting arc (i, j) for commodity $\langle k, d \rangle$ does not increase the total cost of the solution. For the model CAU-DT with Uniform costs, we can develop a similar cost-based test to determine if arc (i, j) can be dropped for each commodity k. Observe, however, that these tests are based on the assumption that when we attempt to locally change the solution by replacing the current tour (containing arc (i, j) with split tours), we have an additional (not currently used) crew member available at one of the depots. Since we do not know a priori if an additional crew member will be available, we need to apply the optimality test (based on dual prices) discussed in Section 7.2 to confirm if dropping arc (i, j) is valid, i.e., does not increase the optimal value of the original problem.

Appendix F. Summary Statistics for all BC Instances

(Single-depot CSP with only DT limits, max. DT = 480 minutes)

# Trips	# Crew Available	# Variables	# Constr.	Initial LP Gap (%)	LP Time (secs)	# B&B nodes	Total Time (secs)
	27	266	122	0.000%	0	0	0.02
	28	266	122	0.000%	0	0	0.02
50	29	266	122	0.000%	0	0	0.02
	30	266	122	0.000%	0	0	0
	31	266	122	0.000%	0	0	0.01
	44	1172	353	0.000%	0	0	0.05
	45	1172	353	0.000%	0	0	0.05
100	46	1172	353	0.000%	0	0	0.06
	47	1172	353	0.000%	0.02	0	0.06
	48	1172	353	0.000%	0.02	0	0.08
	69	2438	665	0.398%	0.02	0	0.47
	70	2438	665	0.167%	0.06	0	0.64
150	71	2438	665	0.000%	0.02	0	0.06
	72	2438	665	0.000%	0.02	0	0.11
	73	2438	665	0.000%	0	0	0.06
	93	5106	1122	0.000%	0.01	0	0.16
	94	5106	1122	0.000%	0.03	0	0.16
200	95	5106	1122	0.000%	0.03	0	0.17
	96	5106	1122	0.000%	0.02	0	0.17
	97	5106	1122	0.000%	0.02	0	0.16
	108	10091	1947	0.000%	0.08	0	0.48
	109	10091	1947	0.110%	0.06	0	0.56
250	110	10091	1947	0.000%	0.08	0	0.76
	111	10091	1947	0.000%	0.06	0	0.53
	112	10091	1947	0.000%	0.05	0	0.42
	130	18785	3073	0.000%	0.14	0	0.89
	131	18785	3073	0.000%	0.11	0	0.84
300	132	18785	3073	0.000%	0.17	0	0.9
	133	18785	3073	0.000%	0.11	0	0.69
	144	22353	3766	0.000%	0.22	0	1.08
	145	22353	3766	0.000%	0.19	0	1.79
350	146	22353	3766	0.000%	0.19	0	1.03
	147	22353	3766	0.000%	0.2	0	1.03
	148	22353	3766	0.000%	0.22	0	1.48
	159	33199	5041	0.081%	0.44	0	2.93
	160	33199	5041	0.000%	0.44	0	1.75
400	161	33199	5041	0.000%	0.38	0	2.48
	162	33199	5041	0.000%	0.39	0	1.65
	163	33199	5041	0.000%	0.36	0	1.78
	182	44689	6410	0.023%	0.64	0	3.24
	183	44689	6410	0.024%	0.62	0	4.57
450	184	44689	6410	0.024%	0.64	0	3.48
	185	44689	6410	0.000%	0.62	0	2.34
	186	44689	6410	0.000%	0.59	0	3.07
	204	62407	8259	0.000%	0.79	0	3.29
	205	62407	8259	0.015%	0.83	0	4.31
500	206	62407	8259	0.000%	0.75	0	3.29
	207	62407	8259	0.000%	0.76	0	3.21
	208	62407	8259	0.000%	0.76	0	4.93
Average				**0.017%**	**0.23**	**0**	**1.25**