A co-availability scheduling model for coordinating multi-disciplinary care teams

Serpil Mutha, James Benneyana*, John Terrellb, Victoria Jordanb and Ayten Turkcana

aHealthcare Systems Engineering Institute, Colleges of Engineering and Health Sciences, Northeastern University, Boston MA, USA; bMD Anderson Cancer Center, Houston TX, USA

(Received 30 April 2014; accepted 23 January 2015)

We introduce a co-availability scheduling problem that arises in various healthcare settings in which personnel from different disciplines work together as care teams and for which synchronisation of their availability impacts scheduling flexibility and procedure timeliness. Examples include breast cancer surgery involving oncologic and plastic surgeons, primary and specialty care integrated visits, and vascular interventions involving cardiac surgeons, radiologists and radiology technicians. We develop an integer programming model to help create optimal schedules that maximise the amount of co-available time across the scheduling templates of the desired team members, while still satisfying each of their clinic coverage, preference and extraneous responsibilities constraints. Application to breast surgery at a major cancer centre increased team co-availability by 94%, with sensitivity analysis in other scenarios producing 64–152%, increases in favourable team assignments, and without negatively affecting operating room neither utilisation nor surgery delays.

Keywords: healthcare; personnel scheduling; multi-disciplinary teams

1. Introduction

Coordinating multi-disciplinary healthcare teams is an increasingly common need in many medical contexts (Taylor et al. 2010). Examples include multiple morbidities (Goldstein et al. 2004; Kuzma et al. 2008; Stocker 2003), chronic care (Kasper et al. 2002; Levetan et al. 1995; Wagner 2000), cancer care (Forrest et al. 2005; Ruhstaller et al. 2006; Sidhom and Poulsen 2006), mental health (Burns 2004; Liberman et al. 2001; Van den Berg et al. 2005) and surgical care (Catchpole et al. 2008; ten Cate et al. 2004; Hu 2008). Forming the most effective teams is also important to efficiency, care quality, safety, provider and patient satisfaction and disease management (Bunnell et al. 2010; Carter, Garside, and Black 2003; Committee on Quality of Health Care in America, Institute of Medicine 2000; Davenport et al. 2007; Gabel, Hilton, and Nathanson 1997; Makary et al. 2006; Mazzocco et al. 2009; Sexton et al. 2006). Collectively synchronising schedules of a large number of staff in multiple departments, however, such that all individuals of each desired team subsequently are co-available when needed, subject to their other weekly responsibilities and scheduling constraints is a difficult problem, especially in complex settings involving multiple specialists (Kane, O’Byrne, and Luz 2010; Kane et al. 2007; Wagner et al. 2001).

The resulting inefficient team scheduling has been shown to negatively impact workflow in radiology and pathology (Kane et al. 2007), co-scheduling of dermatologic oncology specialists (Caudron et al. 2010), schedule matching in neonatal intensive care units (Brown et al. 2003), and care for diabetes (Ritholz et al. 2011), cancer (Wright et al. 2009), palliative radiotherapy (Fairchild et al. 2009) and psychiatry (Fiddler et al. 2010) patients. In contrast, the value of efficient teams has been well-studied in the literature, some via randomised studies, including the benefits on patient outcomes (Goldstein et al. 2004), readmissions (Kasper et al. 2002), length of stay (Levetan et al. 1995), survival (Forrest et al. 2005), quality of life (Van den Berg et al. 2005), patient safety (Catchpole et al. 2008), and timely care and patient and provider satisfaction (Bunnell et al. 2010; Gabel, Hilton, and Nathanson 1997).

Rather than schedule day-to-day cases directly (e.g. surgeries), we introduce a schedule template optimisation model that maximises the amount of time during which required or well-matched people are ‘co-available’, making subsequent procedure scheduling itself easier and less constrained. The general idea is that by maximising mutually available unscheduled times, such as by optimally scheduling other activities (e.g. teaching responsibilities, clinic coverage, release time, etc.), healthcare procedures can be scheduled more efficiently, and with more desirable team compositions. Whereas, many studies that primarily focus on scheduling rooms or procedures (e.g. Cardoen, Demeulemeester, and Belien 2010; Cayirli and Veral 2003) assume provider schedules are given, our problem is more of a staff timetable
template problem (Ernst et al. 2004). To our knowledge, no studies have considered this problem of simultaneously
coordinating individual scheduling templates in order to ease subsequent scheduling difficulties and increase team
composition quality. Personal preferences, historical data on high functioning teams and technical or personality
incompatibility also can be taken into account in this approach.

2. Optimisation model

We define co-availability as the per cent of time slots in which all required or desired individuals are simultaneously
available, such that at some later time they could be scheduled to work together on patient procedures. We formulate
this problem as an integer programming optimisation model, with a group \( G_i \) defined as the set of all providers from the
same discipline \( i \) who will be scheduled into teams that each perform one of \( T \) distinctive types of tasks (see Figure 1),
and with the objective function seeking to maximise within-team co-availability subject to coverage, obligatory tasks,
personal requests and preference constraints. Table 1 summarises model inputs, notation and definitions.

The planning period is divided into \( S \) time slots across which each individual \( m \in G_i \) must spend a fixed number of
time slots \( R_{mt} \) performing each task \( (t = 1, \ldots , T) \). Coverage requirements exist for each task in each time slot, in which
a fixed number of members \( C_{it} \) from skill group \( G_i \) must be assigned to perform task \( t \). The set \( O_i \) is defined as all
obligatory tasks for members of skill group \( G_i \) such that individual \( m \) must perform task \( t \) during time slot \( s \).

![Figure 1. Relationship of personnel skill groups and multi-disciplinary teams.](image)

Table 1. Input notation used in IP model.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I )</td>
<td>Number of groups that consist of people with the same skills, where index ( i \in {1, 2, \ldots , I} )</td>
</tr>
<tr>
<td>( R )</td>
<td>Number of teams that consist of people from different groups performing a task together, where index ( r \in {1, 2, \ldots , R} )</td>
</tr>
<tr>
<td>( m, n )</td>
<td>Subscripts denoting members of a group or a team</td>
</tr>
<tr>
<td>( T )</td>
<td>Number of types of tasks, where index ( t \in {1, 2, \ldots , T} )</td>
</tr>
<tr>
<td>( S )</td>
<td>Number of time slots, where index ( s \in {1, 2, \ldots , S} )</td>
</tr>
<tr>
<td>( k )</td>
<td>Subscript denoting number of members in partial or whole teams</td>
</tr>
<tr>
<td>( G_i )</td>
<td>Set of providers from the same discipline with the same skill set</td>
</tr>
<tr>
<td>( D_{int} )</td>
<td>Set of members with whom member ( m ) of group ( i ) works together for the common task ( t )</td>
</tr>
<tr>
<td>( U_{int} )</td>
<td>Set of members with whom member ( m ) of group ( i ) cannot work together for the common task ( t )</td>
</tr>
<tr>
<td>( O_i )</td>
<td>Set of obligations for members of group ( i ), ( O_i = {(m,t,s)} ), member ( m ) should perform task ( t ) in time slot ( s )</td>
</tr>
<tr>
<td>( Q' )</td>
<td>Set of teams, members of which prefer to work together for task ( t )</td>
</tr>
<tr>
<td>( Q'_r )</td>
<td>( r )th team in set ( Q' )</td>
</tr>
<tr>
<td>( R_{mt} )</td>
<td>Total number of time slots member ( m ) must allocate to task ( t )</td>
</tr>
<tr>
<td>( C_{it} )</td>
<td>Total number of members from group ( i ) that are required to cover each time slot for task ( t )</td>
</tr>
<tr>
<td>( w_{kt} )</td>
<td>Increase in objective function by scheduling ( k ) members of a given team for task ( t ) in the same time slot</td>
</tr>
<tr>
<td>( K_{tr} )</td>
<td>Number of team members of ( p_{th} ) team in set ( Q' )</td>
</tr>
<tr>
<td>( Goal_i )</td>
<td>Minimum number whole teams that must be assigned to task ( t ) in each time slot</td>
</tr>
</tbody>
</table>
Some providers additionally may have strong preferences or requirements regarding with whom they work, either
due to technical compatibility or personality dynamics. Preferred and not preferred individuals by member \( m \) of group \( i \) for task \( t \) are represented by the sets \( D_{int} \) and \( U_{int} \), respectively.

Defining \( Q' \) as the set of \( R \) teams \( Q'_r, r = 1, \ldots, R \), whose membership is desirable to perform task \( t \) for any of the above reasons, the co-availability problem then assigns providers to perform each task at each time slot in a manner that maximises the number of co-available desirable unscheduled teamings, while satisfying all other coverage, obligations, preference and compatibility constraints. Since partial team assignments may still create partial value, weight coefficients \( 0 \leq w_{kt} \leq 1 \) denote the value of \( k \) members of a desired team being co-available for task \( t \). This also causes the IP model to assign as many team members together as possible, where \( p_{ktrs} \) denotes \( k \) members of team \( r \) are assigned to task \( t \) in time slot \( s \), the weight factors can be chosen to reflect the relative value of fewer whole or partial teams, and the parameter \( \text{Goal}_t \) is defined as the minimum number of whole teams that must be assigned to task \( t \).

The integer programming model then is formulated as follows:

\[
\text{Max} \quad \sum_{k \leq K_r} \sum_{t} \sum_{r} \sum_{s} w_{kt} \cdot p_{ktrs} \\
\text{s.t.} \quad \sum_{t} A_{mts} \leq 1 \quad \forall i, s \text{ and } \forall m \in G_i \\
\sum_{s} A_{mts} = R_{mt} \quad \forall i, t \text{ and } \forall m \in G_i \\
A_{mts} = 1 \quad \forall (m, t, s) \in O_i \\
\sum_{n \in D_{int}} A_{nts} \geq A_{mts} \quad \forall i, t, s \text{ and } \forall m \in G_i, D_{int} \neq \emptyset \\
A_{nts} + A_{mts} \leq 1 \quad \forall i, t, s \text{ and } \forall m \in G_i, \forall n \in U_{int} \\
\sum_{m} A_{mts} \geq C_{it} \quad \forall i, t, s \text{ where } m \in G_i \\
\sum_{m} A_{mts} = \sum_{k \leq K_r} k \cdot p_{ktrs} \quad \forall t, r, s \text{ where } m \in Q'_r \\
\sum_{t} p_{ktrs} \geq \text{Goal}_t \quad \forall t, s \\
A_{mts} \in \{0, 1\} \quad \forall m, t, s \\
p_{ktrs} \in \{0, 1\} \quad \forall k, t, r, s
\]

where the decision variable \( A_{mts} = 1 \), if member \( m \) is assigned to task \( t \) for time slot \( s \) (and 0 otherwise) and the decision variable \( p_{ktrs} = 1 \) if \( k \) members of team \( r \) are assigned together to task \( t \) for time slot \( s \) (and 0 otherwise).

The objective function in Equation (1) maximises the total weighted number of unscheduled common time slots of compatible or required team members over all tasks. Constraints (2)–(3) are assignment constraints that ensure each group member is allocated to a time slot for at most one task, and that each assigned group member \( m \) spends \( R_{mt} \) time units for task \( t \), respectively. Constraint (4) ensures fulfilment of all obligatory tasks. Constraints (5)–(6) define compatibility requirements and provider preferences. Specifically, constraint (5) ensures that all time slots of each member \( m \) who is assigned to the common task \( t \) coincide with task \( t \) time slots of members that belong to his/her \( D_i \) set, meaning that at least one member with whom it is desirable to work on task \( t \) is available for each provider.
Constraint (6) ensures that tasks of each member \( m \) are scheduled to different time slots than those of the members that belong to his/her \( U_t \) undesirable set, so that none of the strictly-not-preferred members are scheduled to work together on task \( t \).

Constraint (7) is a coverage constraint to ensure that at least \( C_{it} \) members of group \( i \) are assigned to cover task \( t \) in each time slot. Note that the coverage parameter \( C_{it} \) could be extended to include an additional subscript, \( C_{its} \), if requirements change according to time slot. Constraint (8) determines if a whole \( (k = K_{tr}) \) or partial \( (k < K_{tr}) \) team is assigned to a time slot for any defined team with any number of members. Constraint (9) ensures that at least \( \text{Goal}_t \) whole teams are assigned to task \( t \) in every time slot and avoids team accumulation in some time slots. Constraints (10)–(11) are standard binary integer assignment constraints. To solve the below examples, this IP model was implemented in CPLEX 12.1 on an Intel Core 2 CPU with 4.00 GB RAM.

3. Case study: breast cancer surgery

We used the above model in a major cancer centre to schedule the co-availability of oncology and plastic surgeons performing both mastectomy and reconstruction during the same case, where removal of breast tissue by an oncologic surgeon is followed by cosmetic reconstruction by a plastic surgeon under the same episode of anaesthesia (Figure 2) (Scanlon 1991). Scheduling large number of such cases (roughly 110 per week across 8–10 operating rooms) requires sufficient time slots in which both types of compatible specialists are available.

The cancer centre has 13 oncology and 18 plastic surgeons, respectively, performing tissue removal and reconstruction (denoted here as O1, O2, …, O13 and P1, P2, … P18) each working morning and afternoon sessions 5 days a week in surgery or some other task (e.g. clinic appointments, research, administrative work and teaching). Since all teams in this application are composed of two members \( (K_{tr} = 2 \forall r) \), the weight factor used was \( w_{kt} = 1 \). Typically, most surgeons work two days a week in surgery, two days in clinic, and one day on other tasks. Plastic surgeons perform reconstruction of breast and other types of patients (e.g. throat, neck, ears, etc.), with on average 40% of their surgery time spent on composite breast cases, 25% on other composite cases, and 35% on elective operations (where only a plastic surgeon is involved), although these percentages are heterogeneous across surgeons. Of the current plastic surgeons in this cancer centre, six spend all their surgery time on composite breast cases, eight most of their surgery time, and the remaining three only a little time.

Figure 3 summarises weekly available operating rooms for breast surgeries. Since some complex cases require a full day, weekly schedules should include some days in which surgeons are assigned to surgery for both morning and afternoon sessions, referred to here as whole day assignments. This is represented in our model by forcing some per cent of team assignments to be consecutive.

On average, historically three breast surgeons are in clinic each day, so we assumed in the model that at least two breast surgeons must be in clinic each day, with other non-clinic tasks deferrable to days, the surgeon is not performing surgeries nor seeing patients in clinic. For preference relationships, three pairs of surgeons prefer not to work with each other (Surgeon P2 and O11, P12 and O2, P12 and O12), with 10 pairs historically performing better than others ((O2, P11), (O6, P7), (O6, P8), (O8, P11), (O8, P4), (O12, P12), (P3, O6), (P3, O4), (P3, O3) and (P3, O13)). In this particular application, no obligatory tasks were identified by the surgeons. Using these inputs, the resulting optimal weekly schedule is shown in Figure 4, with the optimal solution resulting in a 94% increase in total desired co-availability improving from 18 to 35 oncology-plastic surgeon pairings. This result represents a significant improvement in cancer team schedules.

Figure 2. Breast cancer oncology and reconstruction composite surgery.
Figure 3. Breast centre OR availability. Between 7 and 10 ORs are assigned to breast surgeries in each time slot. Black cells indicate unavailable ORs and coloured cells indicate ORs assigned to specific oncology surgeons (e.g. OR 9 is assigned to oncology surgeon O8 on Tuesdays).

Figure 4. Optimal co-available solution of breast cancer surgery case study (with set of task types A: administrative, C: clinic and S: breast or other surgery).
While Figure 4 shows the full schedule, to provide some insight to the solution Figure 5 illustrates the improved co-availability for certain pairs composed of four specialty surgeons (O4, P2, O6 and P8), with the ovals highlighting co-availability. The bottom row of each part of Figure 5 shows the resulting total number of desirable team assignments across all surgeons in the full model.

4. Performance

4.1 Breast cancer surgery case study

To evaluate how these optimised templates would perform in practice, we simulated the actual scheduling of breast cancer surgeries with random demand assumed to be Poisson estimated from 17 months of historical data (January 2011 through May 2012). To place these surgeons into scheduling templates some logic assumptions needed to be made. Cases were scheduled for the first available time slot of the assigned surgeon with an available OR, with two wait time scenarios assumed for composite cases, summarised in Figure 6. In the first scenario (referred to here as less waiting or LW), composite patients are scheduled for the first availability of the assigned breast surgeon, after which an attempt is made to co-schedule any compatible plastic surgeon; if none is available then any non-compatible plastic surgeon is scheduled.

The second scenario (referred to as more matching or MM) assigns a composite case to the first time slot at which both the assigned breast surgeon and any compatible plastic surgeon are available. In both scenarios, if there is no option in the given week then the case is scheduled in a subsequent week to the first availability of the assigned breast surgeon partnered with any available plastic surgeon.

While neither simulation logic exactly matches the actual process, they provide some sense for how the performance of current vs. proposed schedules compare. In practice, OR requests for composite cases are made by breast surgeon scheduling assistants after finding candidate dates with plastic surgeon scheduling assistants. Hence, the MM strategy is closer to the actual process, although randomly choosing between options that fit the requirements for the surgery.

The simulation was coded in the C programming language and run for five replications of 52 weeks each to compare the current and optimal schedules. Add-on, emergency and emergency return cases were ignored, since scheduled
cases account for over 95% of all volume. Historical frequencies and durations for breast, composite and other cases are summarised in Figure 7 and Table 2. Since no standard probability distribution was a good fit to actual surgery duration data, tabulated values were used to define an empirical distribution function.

Figure 8 summarises the increased number of surgeries assigned each week to desirable teams, with the LW and MM strategies resulting in 152 and 64% improvements, respectively, and with no negative consequence on OR utilisation (Figure 9) nor patient waits (Figure 10). (Note that the remaining utilisation typically is absorbed by the omitted emergency and add-on cases.) The difference in performance of the two simulation logics provide a general feel for the improvement that would result in actual practice.

4.2 Other examples

To develop insight into potential improvements in other settings, we studied the effect of several factors (size of personnel groups, number of obligatory tasks and balancing goal value) in a similar hypothetical scenario, as summarised in
Table 3, requiring the co-availability of a surgeon and an anaesthesiologist (K_tr = 2 ∀ r). Table 4 compares the number of hours that each class of personnel (surgeons, anaesthesiologists) allocate to various tasks (R_mt), with each clinician also classified by clinical or research orientation (Leitch and Walker 2000; Soh 1998; Winslow, Bowman, and Klingensmith 2004). The listed values for the balancing goal constraint Goal_t indicate the minimum number of available whole teams in each time slot. Although greater balanced team availabilities may be desirable, this may not always be

<table>
<thead>
<tr>
<th>Scheduled duration</th>
<th>Actual range</th>
<th>Breast (%)</th>
<th>Composite (%)</th>
<th>Other (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>20–60</td>
<td>6.49</td>
<td>0.03</td>
<td>18.63</td>
</tr>
<tr>
<td>90</td>
<td>65–90</td>
<td>6.04</td>
<td>0.16</td>
<td>6.56</td>
</tr>
<tr>
<td>120</td>
<td>95–120</td>
<td>6.52</td>
<td>0.23</td>
<td>6.03</td>
</tr>
<tr>
<td>150</td>
<td>125–150</td>
<td>3.43</td>
<td>0.12</td>
<td>3.32</td>
</tr>
<tr>
<td>180</td>
<td>155–180</td>
<td>3.53</td>
<td>1.29</td>
<td>3.54</td>
</tr>
<tr>
<td>210</td>
<td>185–210</td>
<td>0.60</td>
<td>0.53</td>
<td>0.93</td>
</tr>
<tr>
<td>240</td>
<td>215–240</td>
<td>1.30</td>
<td>2.99</td>
<td>5.12</td>
</tr>
<tr>
<td>300</td>
<td>245–300</td>
<td>0.63</td>
<td>1.55</td>
<td>3.58</td>
</tr>
<tr>
<td>360</td>
<td>315–360</td>
<td>0.94</td>
<td>2.19</td>
<td>4.11</td>
</tr>
<tr>
<td>480</td>
<td>375–480</td>
<td>0.48</td>
<td>1.41</td>
<td>5.39</td>
</tr>
<tr>
<td>600</td>
<td>500–780</td>
<td>0.09</td>
<td>0.34</td>
<td>1.88</td>
</tr>
</tbody>
</table>

Figure 8. Improvement in weekly number of desired team assignments for the (a) LW scheduling strategy and (b) MM scheduling strategy.

Figure 9. Weekly OR utilisation under the (a) LW scheduling strategy and (b) MM scheduling strategy (no statistical significance, paired t test, p > 0.90).
feasible; similarly the number of obligated time slots can limit the solution. Obligation levels indicate how flexible a setting is, since they are used to determine how many of the assignments are fixed while generating the set of obligatory tasks.

For the coverage requirements $C_{in}$, we assumed that at least one surgeon of each type is available throughout the week for clinic appointments and another for performing surgeries. We considered three types of pair relationships between surgeons and anaesthesiologists: Of all surgeons 40% must work with at least one of some individuals (type-1), 20% cannot work with some individuals (type-2) and 40% prefer to work with some individuals (type-3). These percentages were used to randomly choose members of the $D_{imp}$, $U_{imp}$ and $Q^f$ sets, respectively.

Five problem instances were randomly generated and solved for each of the 48 factor combinations, with the optimisation program terminated if an optimal solution was not found after 4 h. (Some factor–level combinations reflect settings with extensive requirements that are not possible to be satisfied as a whole.) Figure 11 summarises the impact of larger problem sizes, percentages obligatory tasks and goal levels on problem feasibility. Figure 12 summarises the (a) average computation times for problems that were solved optimally and (b) optimality gap from the upper bound for those that were not, both appearing reasonably good. As would be expected, larger percentages of obligatory tasks and
problem sizes result in more infeasible problems (11a), larger goal levels result in challenges to solve problems only to near-optimality within the time limit (11b), larger problem sizes and goal levels result in longer average computation time (12a) and smaller goal values result in smaller average gaps (12b). No XL problem could be solved to optimality (12a), and similarly no small-size problem was solved only to near optimality (12b). An unpredicted result is the decrease in the optimality gap when problem size changes from ‘large’ to ‘XL’, since one would expect the solutions to be further from the upper bound as the problem size increases, although this apparent improvement may be a result of using relative gap, where the difference is reported as a percentage. Additionally, further improvements from further search in branch-and-bound tree should be relatively smaller when the objective value is large.

5. Discussion

Multi-disciplinary teams are increasingly important in many healthcare and non-healthcare contexts, with co-scheduling often being a difficult problem due to asynchronous availability of necessary or desired team members. While many studies focus on scheduling specific cases to rooms and personnel, limited co-availability often causes even these ‘optimal’ schedules to be fairly poor or to require effort to reschedule conflicts. In such cases, focusing first on improving the more macro scheduling templates can allow for the eventual day-to-day schedules to improve significantly, as well as streamline scheduling efforts that can be largely manual and time-intensive. As demonstrated across a range of applications, optimising co-availability schedule templates can enable more efficient case scheduling, with significant increases in desired team compositions and without adversely affecting OR utilisation nor patient waits.

Moreover, although implementation of many healthcare optimisation models can be technically and culturally challenging to implement on an ongoing basis, off-line solution of models of the type described here may be easier to implement than those requiring integrating into daily practice. Finally, note that our simulation assumes patients are scheduled via a first-come-first-scheduled rule, assigning them to the first available time, whereas some patients in
realism may request later dates based on to their schedules. Results nonetheless give a fairly good indication of the improvement potential of this approach.

Funding
This research was supported in part by National Science Foundation (NSF) [grant number IIP-10341990]; Center for Medicare and Medicaid Innovation (CMMI) [grant number 1C1CMS331050]. Any statements or opinions expressed herein are those of the authors and do not necessarily reflect those of NSF nor CMMI.

References


