Chapter 4
Conflict Resolution and Sector Workload Formulations

4.1. Occupancy Workload Constraint Formulation

We start by examining the approach used by Sherali, Smith, and Trani [47] to generate workload constraints for the APM model. Essentially, this approach determines the maximum number of aircraft that will simultaneously occupy each sector, given the selection of any set of surrogate flight plans, and imposes a corresponding penalty structure in the objective function.

Consider the total collection of flight plans $\bigcup_{f=1,...,F} P_f$. These plans involve traversals between certain pairs of fixes that might intersect in some sector $s = 1,...,S$ under present consideration. Define the *occupancy workload* for any such sector s at any moment in time to be the number of aircraft resident within that sector at that given instant in time. To characterize this type of workload for each sector $s = 1,...,S$, we can examine the occupancy durations of the various flights within s over the time horizon H. The model AOM of Sherali et al. [48] provides this information by constructing a Gantt chart of flight plan occupancy intervals for each sector. In practice, ATC operators routinely monitor several aircraft that are simultaneously traversing their respective sectors. Naturally, when the occupancy workload (maximum simultaneous occupancy of a number of aircraft) becomes too high, a potentially dangerous or untenable situation can arise.

For each sector $s \in S$, let i = 1,…, I_s index the collection of maximal overlapping sets C_{si} of flight plans (f,p), where an overlapping set of flight plans is called maximal if it is not a strict subset of another overlapping set. For example, examining Figure 4-1, we have I_s = 4 maximal sets. Hence, we have,

$$C_{si} = \left\{ (f,p): \begin{array}{l} \text{flight plan } (f,p) \text{ belongs to the } i^{th} \\ \text{maximal overlapping set for sector } s \end{array} \right\}, \forall i = 1,...,I_s, s \in S. \quad (4.1)$$

An efficient algorithm for determining these sets is described in Sherali and Brown [45]. Note that is possible that (f_1,p_1) and (f_2,p_2) $\in C_{si}$ for some s and i, with f_1=f_2 (i.e. the pair corresponds to the same flight), although in this case, the plans would pertain to two distinct surrogates.

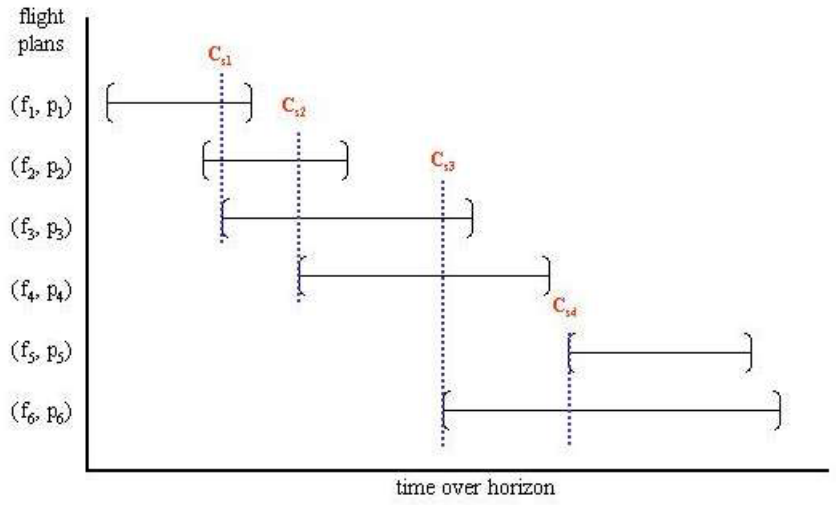


Figure 4-1: Gantt Chart for Formulating Workload Constraints

Now define the variable n_s to represent the maximum number of overlapping flight plans within each sector $s = 1,...,S$. Note that n_s is given by the largest number of flight plans selected from any of the maximal overlapping sets C_{si}, i = 1,…, I_s, i.e.,

$$n_s = \max_{i=1,\ldots,I_s} \left\{ \sum_{(f,p)\in C_{si}} x_{fp} \right\}, \tag{4.2}$$

because any other overlapping set is a subset of some maximal overlapping set. In the model formulation, the variable n_s is bounded on a suitable interval $[1, \overline{n}_s]$, and furthermore, its value is penalized in the objective function using a penalty factor that increases nonlinearly in an appropriate fashion with an increase in this type of workload. The motivation here is that if the maximum number of aircraft being simultaneously monitored in a sector increases from one to three, for example, the associated penalty should likely more than triple. Moreover, there should be some absolute maximum number $\overline{n}_s$ of overlapping flights at any point in time, as determined by the capacity of sector s.

This workload measure and the associated penalty structure can be modeled without the need to discretize time. Towards this end, define the binary variables

$$y_{sn} = \left\{ \begin{matrix} 1 & \text{if the occupancy workload in sector } s \text{ is } n \\ 0 & \text{otherwise} \end{matrix} \right\}, \forall s \in S, n = 1,\ldots,\overline{n}_s \tag{4.3}$$

and let μ_{sn} be the associated penalty for having y_{sn}=1. We assume that

$$\mu_{s2} \geq \mu_{s1}, \text{ and } \mu_{sj} \geq 2\mu_{s(j-1)} - \mu_{s(j-2)} \quad \forall j = 3,\ldots,\overline{n}_s . \tag{4.4}$$

Note that the condition (4.4) implies that

$$0 \leq (\mu_{s2} - \mu_{s1}) \leq (\mu_{s3} - \mu_{s2}) \leq \ldots \leq (\mu_{s\overline{n}_s} - \mu_{s(\overline{n}_s - 1)}) . \tag{4.5}$$

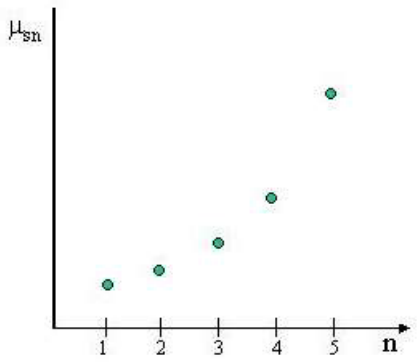


Figure 4-2: Illustration of a Convex Sector Workload Penalty Structure

Figure 4-2 illustrates the implied convex nature of this penalty structure. Observe that by enforcing $n_s \geq 1$, we always incur a workload cost of at least μ_{s1}, even when no aircraft are being monitored over the horizon. This is appropriate since there always exists a fixed monitoring cost. More importantly, by avoiding a cost of zero corresponding to $n_s=0$, there is greater flexibility in considering practical workload costs that would satisfy (4.4). For example, we might have $0 < \mu_{s1} = \mu_{s2} = ... = \mu_{s\tau}$ for some threshold number τ of aircraft being monitored, after which the costs might increase at an increasing rate as in (4.5). This is the cost structure that arises in practice, and the model assumes that this holds true. The penalty structure is then incorporated into the constraints as follows, noting (4.2).

$$\sum_{(f,p)\in C_{si}} x_{fp} - n_s \leq 0 \quad \forall i = 1,...,I_s, \ s = 1,...,S \tag{4.6a}$$

$$n_s = \sum_{n=1}^{\bar{n}_s} n y_{sn}, \quad \forall s = 1,...,S \tag{4.6b}$$

$$\sum_{n=1}^{\bar{n}_s} y_{sn} = 1, \quad \forall s = 1,...,S \tag{4.6c}$$

$$y_{sn} \geq 0, \quad \forall n = 1,...,\bar{n}_s, \ s = 1,...,S\,. \tag{4.6d}$$

The following term is included in the objective function:

$$(\min) \ \ldots + \sum_{s \in S} \sum_{n=1}^{\bar{n}_s} \mu_{sn} y_{sn} \ . \tag{4.6e}$$

Note that both n_s and y_{sn}, $\forall\ s$, n, have been declared as continuous variables in (4.6). Sherali et al. [47] prove that the binary restrictions on y hold automatically at optimality in the model, and hence, so do the integrality and bounding restrictions on the variables n_s, $\forall\ s = 1,...,S$.

4.2. Alternative Occupancy Workload Constraint Formulation

The foregoing approach focuses on the maximum sector monitoring workload for ATC personnel. However, this does not capture the total, or average workload requirements, or the persistence of peak periods. Hence, we offer the following additional modeling constructs.

The model AOM [48] determines the sector occupancies and the respective occupancy durations for each flight plan. Define Ω_s as the set of flight plans (f,p) that occupy the sector $s = 1,...,S$ during the horizon H, where the total occupancy time interval for flight plan p of flight f in sector s is given by t_{fp}^s.

The *average workload* for sector s can then be described as

$$w_s = \frac{1}{|H|} \sum_{(f,p) \in \Omega_s} t_{fp}^s x_{fp}, \quad \forall s = 1,...,S\ , \tag{4.7}$$

where $|H|$ is the length of the horizon being considered. Observe that if t_{fp} is the total airborne time corresponding to flight plan (f,p), we must have

$$\sum_{s \in S} t_{fp}^s = t_{fp}, \quad \forall (f,p)\ . \tag{4.8}$$

The average occupancy workload is penalized in the objective function in a linear fashion. The rationale here is that the average occupancy workload represents the nominal state of monitoring activity. As such, personnel and equipment can be

scheduled as a direct function of the expected amount of work to be performed (assuming, of course, that the expected workload is within the sector's capacity). Hence, we include in the objective function

$$(\min)\ \ldots + \sum_{s \in S} \gamma_s w_s\ , \tag{4.9}$$

where γ_s is a suitable constant penalty factor.

Recalling (4.2), and letting n_s be bounded above by some maximum number, $\overline{n}_s$, of simultaneously occupying aircraft, as before, we can characterize the *peak occupancy workload* in each sector s via

$$n_s \ \geq \sum_{(f,p) \in C_{si}} x_{fp}, \quad \forall i, \forall s = 1, \ldots, S, \ \text{and} \ \ n_s \leq \overline{n}_s, \forall s = 1, \ldots, S\ . \tag{4.10}$$

With respect to workload, an operations tempo that is constant and predictable is preferred to one that is either erratic or not predictable. When the ATC workload varies significantly, additional personnel and equipment resources are required that might remain idle during non-peak periods.

To accommodate this feature, we shall assign a penalty in the objective function corresponding to the maximal variability defined as the difference between the peak occupancy workload and the nominal (average) occupancy workload over the horizon. This maximal variability, $n_s - w_s$, is computed via

$$\left(n_s - w_s\right) \ = \ \sum_{n=0}^{\overline{n}_s} n\, y_{sn}, \quad \forall s = 1, \ldots, S\ , \tag{4.11}$$

where the y–variables represent convex combination weights, satisfying

$$\sum_{n=0}^{\overline{n}_s} y_{sn} \ = \ 1, \quad y_{sn} \geq 0\ , \forall n = 0, \ldots, \overline{n}_s, \ \text{for each}\ s = 1, \ldots, S\ . \tag{4.12}$$

Observe that the lower bound for n in (4.12) is zero. This corresponds to the case where the occupancy workload is constant over the horizon H (i.e. the peak workload equals the average workload). Furthermore, note that the quantity $(n_s - w_s)$ is not necessarily integral.

Remark: If $n(t)$ is the number of aircraft overlapping at time t, note that

$$w_s = \frac{1}{|H|}\int_0^H n(t)dt \le \frac{n_s}{|H|}\int_0^H dt = n_s, \tag{4.13}$$

and so, we have $(n_s - w_s) \ge 0$ as expected.

If we define the penalty structure for μ_{sn} as a function of $(n_s - w_s)$ in a similar manner as in (4.4) and (4.5), we can rely on the resulting implied convex structure to ensure that, at optimality, at most two y_{sn}-variables will be non-zero, and, if y_{sn_1} and y_{sn_2} are two such non-zero variables, then n_1 and n_2 are adjacent, so that the associated penalty is a convex combination of μ_{sn_1} and μ_{sn_2}.

Accordingly, we modify the constraints (4.6) as follows:

$$w_s = \frac{1}{|H|}\sum_{(f,p)\in\Omega_s} t^s_{fp} x_{fp}, \quad \forall s = 1,...,S \tag{4.14a}$$

$$\sum_{(f,p)\in C_{si}} x_{fp} - n_s \le 0, \quad \forall i = 1,...,I_s, \ s = 1,...,S \tag{4.14b}$$

$$n_s - w_s = \sum_{n=0}^{\bar{n}_s} n\, y_{sn}, \quad \forall s = 1,...,S \tag{4.14c}$$

$$\sum_{n=0}^{\bar{n}_s} y_{sn} = 1, \quad \forall s = 1,...,S \tag{4.14d}$$

$$n_s \le \bar{n}_s, \forall s = 1,...,S, \text{ and } y_{sn} \ge 0, \ \forall n = 0,...,\bar{n}_s, \ \forall s = 1,...,S. \tag{4.14e}$$

The related occupancy workload terms in the objective function are given by:

$$(\min)\ \ldots + \sum_{s \in S} \gamma_s w_s + \sum_{s \in S} \sum_{n=0}^{\bar{n}_s} \mu_{sn} y_{sn}\ , \tag{4.14f}$$

where

$$\mu_{s1} \geq \mu_{s0},\ \text{and}\ \mu_{sj} \geq 2\mu_{s(j-1)} - \mu_{s(j-2)},\ \ \forall j = 2, \ldots, \bar{n}_s\ . \tag{4.15}$$

Observe that the penalty terms of (4.14f) correspond to the "stress" placed on the ATC system as a result of the selection and execution of a particular set of flight plans. In a CDM environment, these system costs will be traded off against airline costs (e.g. fuel and delay costs) to select an optimal set of flight plans. These airline costs are discussed in detail in Chapter 5.

4.3. Conflict Resolution Workload Formulation

Resolving conflicts between aircraft traversing a sector imposes a *conflict resolution workload* that is in addition to the occupancy workload discussed in the foregoing section. For example, an ATC controller must contact conflicting aircraft, direct new vectors for them, and subsequently monitor compliance, to ensure that the required separation between aircraft pairs is maintained.

Accordingly, we ascribe a suitable penalty φ_{PQ} in the objective function for each conflict that must be resolved, corresponding to the flight plans P and Q, i.e.,

$$(\min)\ \ldots + \sum_{(P,Q) \in A} \varphi_{PQ} z_{PQ}\ , \tag{4.16}$$

where A is the set of pairs of flight plans that potentially conflict in the overall airspace under consideration, during any point in time within the horizon.

Remark: Using the structure of (4.16), we can ascribe a unique workload penalty for each conflict, based upon the geometry of the conflict itself, as a measure of the difficulty, or intensity, of the required conflict resolution actions. For example, two aircraft that are approaching each other head-on might require a quicker ATC response

than two aircraft traveling near parallel trajectories. Recall that the PAEM computes conflict geometry information for each conflicting pair of flight plans, (P,Q), including their relative trajectories and the minimum projected separation distance.

4.4. Discretized Time Slot Workload Formulation

We summarize the strategy employed by Sherali, Smith and Trani [47] to formulate conflict constraints for the APM model. First, the time horizon is discretized into uniform-length segments, $t = 1,...,T_s$, for each sector $s = 1,...,S$, where the duration of the segment is dependent on the sector's conflict resolution capacity. Based on its conflict analysis, APM then generates a *conflict graph* $G_{st}(N_{st}, A_{st})$ for each sector s and each time segment t, where N_{st} is the set of nodes representing all flight plans traversing sector s during segment t, and A_{st} is the set of edges such that if two flight plans are in conflict, then A_{st} includes an edge connecting the corresponding nodes. The graph G_{st} is typically a collection of disjoint components. Figure 4-3 shows an instance of such a conflict graph where the aircraft pairs (P,Q), (P,R), (Q,R), and (Q,W) are in conflict. To facilitate the model generation process, an *overall conflict graph* $G(N, A)$ is also constructed, where N is the set of nodes representing all flight plans, and A is the set of edges corresponding to pairs of flight plans that are in conflict in any sector during any point in time within the time horizon.

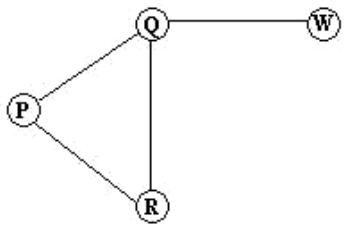


Figure 4-3: Example Conflict Graph $G_{st}(N_{st},A_{st})$

The APM model imposes the workload restriction that no more than one conflict may occur in each sector s during any time segment t. The algorithm examines the edges of A_{st}, in a pairwise fashion, and constructs constraints of the form,

$$\sum_{P \in S_k} x_P \le |S_k| - 1, \tag{4.17}$$

where S_k is the set of nodes at which the k^{th} pair of edges is incident. Observe that $|S_k|$ equals three or four, depending on whether the pair of edges is adjacent. The sequential process adopted by APM to generate constraints of the type (4.17) automatically precludes the creation of redundant constraints.

In an alternative formulation APM2, Sherali, Smith, and Trani [47] defined z_{PQ} as a binary variable that takes a value of one whenever conflicting flight plans P and Q are selected (by including the constraints $z_{PQ} \ge x_P + x_Q - 1,\ z_{PQ} \ge 0$), and prescribed the following restriction to represent the above conflict workload constraint, in lieu of (4.17):

$$\sum_{(P,Q) \in A_{st}} z_{PQ} \le 1, \quad \forall t = 1, ..., T_s,\ \forall s = 1, ..., S \tag{4.18}$$

One advantage of using (4.18) is that it readily generalizes to the case where some $r_s \ge 1$ simultaneously occurring conflicts are permitted, as discussed in the next section.

4.5. Continuous Time Workload Formulation with Conflict Buffers

Suppose that while running PAEM, whenever we encounter a conflict between a pair of flight plans P and Q, $P < Q$, over some duration $[t_1, t_2]$, such that this conflict is to be resolved in sector s, we record this interval along with the conflicting pair (P,Q) in a *Conflict Gantt Chart* for sector s, denoted $\boldsymbol{CGC_s}$. Note that if P and Q lie in different sectors over the duration $[t_1, t_2]$, the resolution of this conflict is assigned to the sector that contains the focal aircraft at time t_1, with (unlikely) ties being broken arbitrarily by assigning the conflict to the sector that contains the smaller-indexed (P) flight plan.

In addition, depending on the sector s and the severity and characteristics of the identified conflict, we define a *prep-buffer* duration $b^s_{PQ} > 0$ that serves to represent a preparatory duration required by the air traffic controller in sector s to address the imminent conflict between flight plans P and Q. We augment the interval $[t_1, t_2]$ represented in the CGC_s by adding the prep-buffer, i.e., by replacing t_1 with $t_1 - b^s_{PQ}$.

Having constructed the CGC_s for all sectors s, we now impose a conflict resolution workload constraint that the

$$\#\text{ of simultaneous augmented conflicts to be resolved in } s \text{ at any time } \leq r_s, \qquad (4.19)$$

where $r_s \geq 1$ is some designated workload parameter for sector s, $s = 1,...,S$. Using the algorithm prescribed by Sherali and Brown [45], suppose that we identify the entire collection of maximal overlapping sets M_{sk}, $k = 1,...,K_s$, for CGC_s, where each M_{sk} is comprised of pairs of conflicting flight plans (P,Q), P < Q, in CGC_s that overlap at some time, and is maximal in the sense that it is not a strict subset of any other set of simultaneously occurring pairs of conflicting flight plans. Figure 4-4 illustrates a hypothetical CGC_s and its associated maximal overlapping sets M_{sk}, $k = 1,...,K_s$. Observe that, as in the case of the pair (A, B) in Figure 4-4, disconnected conflict intervals for any pair of flight plans are treated as separate conflicts in this definition.

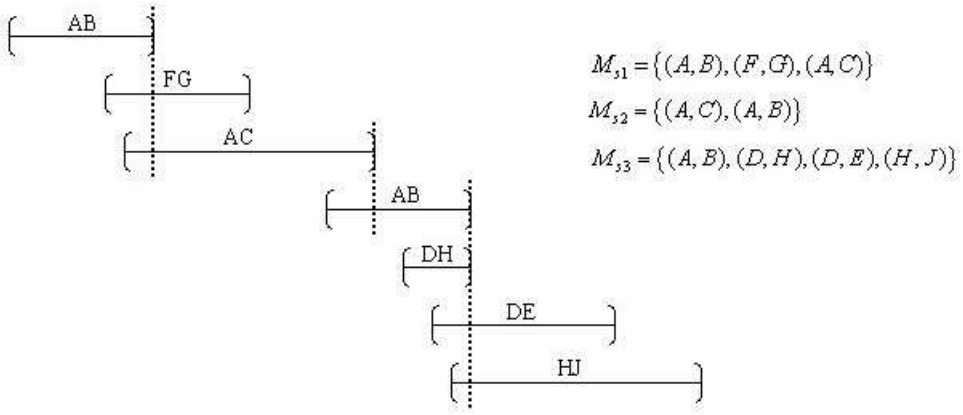


Figure 4-4: Conflict Gantt Chart CGC_s and its Associated Maximal Overlapping Sets

Accordingly, we can formulate (4.19) by imposing

$$\sum_{(P,Q)\in M_{sk}} z_{PQ} \le r_s, \quad \forall k = 1,...,K_s, \forall s = 1,...,S, \tag{4.20}$$

where recall that z_{PQ} equals one if flights plans P and Q are selected, and equals zero otherwise. Observe that (4.20) is a valid representation of (4.19), since any set of overlapping conflicts must be a subset of some M_{sk}, $k \in \{1,...,K_s\}$, where the sets M_{sk} themselves contain overlapping conflicts. Let us refer to the collection of constraints (4.20) for each $s = 1,...,S$ as $M_s - inequalities$.

Recall from Section 4.1 that in the previous formulation of conflict resolution workload constraints, given any sector s, we discretized the horizon into slots $t = 1,...,T_s$, each of duration t_s, and we constructed the conflict graphs $G_{st}(N_{st}, A_{st})$ for each slot $t = 1,...,T_s$, essentially based on conflicts that (partially) overlapped slot t in CGC_s (without any prep-buffers). Similar to (4.20), we then imposed the restrictions (using the same generalized parameter $r_s \ge 1$):

$$\sum_{(P,Q)\in A_{st}} z_{PQ} \le r_s, \quad \forall t = 1,...,T_s, \forall s = 1,...,S. \tag{4.21}$$

Let us refer to the collection of constraints in (4.21) for each s as $D_s - inequalities$, for $s = 1,...,S$.

Proposition 4-1

(a) Suppose that $b_{PQ}^s \equiv 0 \quad \forall (P,Q) \in A_s$. Then for any value of $t_s > 0$, the D_s-inequalities imply the M_s-inequalties.

(b) Conversely, suppose that given a value of $t_s > 0$ for each corresponding slot $t = 1,...,T_s$, we compute

$$\tau_s(t) = \max\left\{0,\ t_2(t) - t_1(t)\right\} \tag{4.22}$$

where

$$t_1(t) = \min\left\{t'' :\ [t', t''] \text{ is a conflict interval for some } (P,Q) \in A_{st}\right\} \tag{4.23}$$

$$t_2(t) = \max\left\{t' :\ [t', t''] \text{ is a conflict interval for some } (P,Q) \in A_{st}\right\}. \tag{4.24}$$

Then, $\tau_s(t) \le t_s \quad \forall t = 1,...,T_s$, and moreover, for

$$b_{PQ}^s \ge \tau_s(t), \quad \forall (P,Q) \in A_{st}, \quad \forall t = 1,...,T_s, \tag{4.25}$$

we have that the corresponding M_s-inequalities imply the D_s-inequalities.

Proof

(a) Consider an arbitrary M_{sk} and its corresponding inequality in (4.20). Since all the conflicts between pairs $(P,Q) \in M_{sk}$ occur simultaneously at some point in time (with $b_{PQ}^s = 0, \quad \forall (P,Q) \in A_s$), there exists some time slot t, for which the associated graph G_{st} contains all these conflicts, i.e., $A_{st} \supseteq M_{sk}$. Hence, we have that

$$\sum_{(P,Q)\in A_{st}} z_{PQ} \le r_s \quad \Rightarrow \quad \sum_{(P,Q)\in M_{sk}} z_{PQ} \le r_s \,, \tag{4.26}$$

i.e., the corresponding D_s-inequality in (4.21) implies this M_s-inequality in (4.20).

(b) Conversely, given the hypothesis of part (b), consider any slot $t \in \{1,...,T_s\}$ and the corresponding inequality (4.21) for the accompanying graph G_{st}. If $t_2(t) \le t_1(t)$ as depicted in Figure 4-5(a), then $\tau_s(t) \equiv 0 \le t_s$. Moreover, all the conflicts in A_{st} overlap at some point in time, because if not, then some conflict in A_{st} ends strictly before another begins, whence we would have $\tau_s(t) > 0$. Consequently, there exists an $M_{sk}, k \in \{1,...,K_s\}$, such that $M_{sk} \supseteq A_{st}$, and so,

$$\sum_{(P,Q)\in M_{sk}} z_{PQ} \le r_s \quad \Rightarrow \quad \sum_{(P,Q)\in A_{st}} z_{PQ} \le r_s \,. \tag{4.27}$$

On the other hand, if $t_2(t) > t_1(t)$ as depicted in Figure 4-5(b), let $t_1(t)$ and $t_2(t)$ respectively correspond to some conflicting pairs of flight plans (G,H) and (U,V). Since the conflict (U,V) begins after (G,H) ends, and since they both overlap time slot t, we have that both these events occur within time slot t, and so, $\tau_s \equiv t_2(t) - t_1(t) \le t_s$. Moreover, examining the time $t_1(t)$, for any conflict $(P,Q) \in A_{st}$ that occurs over some duration $[t',t'']$, say, we have by (4.22) that (i) if $t' < t_1(t)$, as occurs with conflict (A,B) in Figure 4-5(b), then since $t'' \ge t_1(t)$, we get

$$t' - \tau_s(t) \le t_1(t) \le t'' \tag{4.28}$$

and (ii) if $t' \ge t_1(t)$, as occurs with conflict (C,D) in Figure 4-5(b), then since $t' \le t_2(t)$, we again get that (4.28) holds true because $t' - \tau_s(t) = t' - t_2(t) + t_1(t) \le t_1(t) \le t' < t''$. But (4.28) asserts that under (4.25), all the conflict pairs in A_{st} would overlap at time $t_1(t)$ in

CGC_s, and so, there exists an $M_{sk}, k \in \{1,...,K_s\}$, for which $M_{sk} \supseteq A_{st}$. Therefore, we again have that (4.27) holds true, and this completes the proof. □

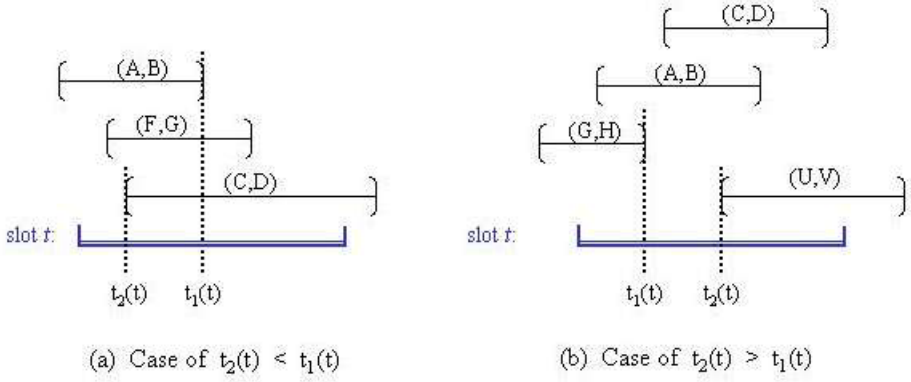


Figure 4-5: Illustration for Proposition 4-1

Observe from Proposition 4-1 that if we use a prep-buffer duration of zero, i.e., we do not augment the conflict intervals, then the D_s-inequalities impose a more stringent set of conflict resolution workload constraints (for any slot duration) than the M_s-inequalities. In this case, the slot-based restriction constrains not only the simultaneously occurring conflicts, but also recognizes that non-overlapping conflicts that might occur in relatively quick succession impose a stressful workload on the ATC controller. Indeed, as Proposition 4-1 illustrates, as the prep-buffer begins to increase, the condition (4.19) related to the M_s-inequalities begins to accommodate the consideration of conflicts that occur in relatively quick succession within the workload formulation, and the M_s-inequalities all imply the D_s-inequalities once the prep-buffer becomes sufficiently large. Proposition 4-1 demonstrates that this occurs at or before a prep-buffer value of t_s, the slot duration for the D_s-inequalities.

The foregoing discussion serves a two-fold purpose. First, it lends further insight into the slot-based conflict resolution constraints, and second, it offers an alternative modeling of such constraints within the framework of APM. Observe that each inequality of the type (4.20) is induced by an underlying conflict graph $G_{sk}(N_{sk}, A_{sk})$, where the edge set $A_{sk} \equiv M_{sk}$, and where N_{sk} is the set of nodes at which the edges

M_{sk} are incident. This is analogous to the conflict graphs $G_{st}(N_{st}, A_{st})$ for the slot-based formulation. Moreover, the overall conflict graph is given by $G(N, A)$ in either case. Consequently, replacing "(s,t)" with "(s,k)" appropriately, or vice versa, the results obtained using one of these modeling constructs can be transported to the other in an analogous fashion. In the sequel, we shall focus on developing further results in the context of the M_s-inequalities framework.

4.6. Valid Inequalities

4.6.1. Alternative Conflict Modeling Strategies

For each sector s, let $G_{sk}(N_{sk}, A_{sk})$ be the conflict graph that is constructed for the overlapping set $M_{sk}, k = 1,...,K_s$, where N_{sk} is the set of nodes representing the k^{th} overlapping set of flight plans traversing sector s during the time horizon, and $A_{sk} \equiv M_{sk}$ is the set of edges connecting the corresponding nodes in N_{sk} representing simultaneously occurring conflicts between pairs of flight plans within M_{sk}. Recall from Sections 3.4 and 4.2 that these conflict intervals are based upon some threshold probabilities, p_1 and $p_2 < p_1$, for conflicts of severity levels one and two, respectively, and include the respective prep-buffer durations, $b^s_{PQ} > 0$. As in Sherali, Smith, and Trani [47], we define $\boldsymbol{FC}$ as the set of flight plan pairs P and Q that pose a fatal conflict risk and we impose the constraint

$$x_P + x_Q \le 1, \quad \forall (P,Q) \in FC. \tag{4.29}$$

Accordingly, the graph G_{sk} and the entire discussion in the sequel below is concerned with the set of non-fatal or resolvable conflicts. As derived in Sherali, Smith, and Trani [47], and working with $r_s = 1, \forall s$ for now, let C_1 denote the set of conflict constraints in the x-space as given by

$$C_1 = \left\{ x: \quad \sum_{P \in S_k} x_P \le |S_k| - 1 \quad \forall k \in K_{NR}, \quad x \text{ binary} \right\}, \tag{4.30}$$

where each S_k is a set of nodes in the conflict graph (representing flight plans) at which designated pairs of edges (in M_{sk}) are incident (whence $|S_k|$ equals three or four), and where K_{NR} records the collection of all such non-redundant constraints over the entire set of conflict graphs G_{sk} $\forall (s,k)$.

Recall that Sherali, Smith, and Trani [47] also propose an alternative representation in a higher-dimensional space (x,z), where each z-variable z_{PQ} represents the quadratic product $x_P x_Q$. This results in the conflict constraint set

$$C_2 = \left\{ \begin{array}{lll} (x,z): & \sum\limits_{(P,Q) \in M_{sk}} z_{PQ} \le 1 & \forall (s,k) \\ & z_{PQ} \ge x_P + x_Q - 1 & \forall (P,Q) \in A \\ & z_{PQ} \ge 0 & \forall (P,Q) \in A \\ & x \text{ binary} & \end{array} \right\}, \tag{4.31}$$

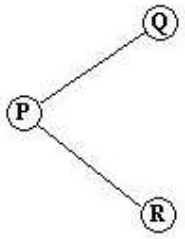
where A is the arc set of the overall conflict graph. The experiments conducted by Sherali, Smith, and Trani reveal that C_1 is preferable for sparse conflict situations, while C_2 is preferable for relatively more dense conflict graphs.

We shall now augment C_2 in order to derive a provably stronger representation than that given by C_1, and motivate our consideration of this revised representation, denoted C_3, defined below. Subsequently, we will further augment C_3 with some additional classes of valid inequalities.

To motivate the derivation of C_3, consider the following simple example.

Example 4-1

Consider the following conflict graph.



The constraints (4.31) defining C_2 for this instance are as given below.

$$\begin{aligned} & z_{PQ} + z_{PR} \le 1, \\ & z_{PQ} \ge x_P + x_Q - 1, \\ & z_{PR} \ge x_P + x_R - 1, \\ & z \ge 0, \quad x \text{ binary}. \end{aligned} \tag{4.32}$$

If we maximize $\{x_P + x_Q + x_R\}$ subject to the continuous relaxation of (4.32), we obtain the fractional extreme point solution $z_{PQ} = z_{PR} = \frac{1}{2}, \quad x_P = \frac{1}{2}, \quad x_Q = x_R = 1$. However, observe that the constraint

$$x_P + x_Q + x_R \le 2 \tag{4.33}$$

that would be inherent in the definition of C_1 deletes this fractional solution. □

Motivated by this example (as well as by Proposition 4-2 below) let us define T_{NC} to be a set of triplets (P,Q,R) that do not admit a clique in any conflict subgraph G_{sk}. More precisely, let

$$T_{NC} = \left\{ \begin{array}{l} (P,Q,R),\ P<Q<R: \quad \text{for some } (s,k),\ \text{a subgraph of } G_{sk} \text{ that is} \\ \text{induced by the nodes } P,Q, \text{ and } R \text{ contains precisely two edges,} \\ \text{but no such subgraph for any } (s,k) \text{ contains three edges} \end{array} \right\}. \tag{4.34}$$

Accordingly, let us define

$$C_3 = \left\{ (x,z): \begin{array}{lll} \sum_{(P,Q)\in M_{sk}} z_{PQ} \le 1, & \forall (s,k) & \text{(4.35a)} \\ x_P + x_Q + x_R \le 2, & \forall (P,Q,R) \in T_{NC} & \text{(4.35b)} \\ z_{PQ} \ge x_P + x_Q - 1, & \forall (P,Q) \in A & \text{(4.35c)} \\ z \ge 0, \quad x \text{ binary} & & \text{(4.35d)} \end{array} \right\}.$$

Consider the following result, where for any set C_i, the set $\overline{C_i}$ denotes its continuous relaxation, obtained by replacing x binary with $0 \le x \le e$, where e will always denote a (conformable) vector of ones.

Proposition 4-2

C_3 yields a tighter representation of the conflict constraints than does C_1 in the sense that the constraints defining $\overline{C_3}$ imply those defining $\overline{C_1}$.

Proof

Consider any constraint in $\overline{C_1}$ where $|S_k| = 3$, with $S_k = \{P,Q,R\}$, say, and where $P < Q < R$. If $(P,Q,R) \in T_{NC}$, then the corresponding constraint in (4.30) is directly present in (4.35b). Otherwise, there exists some (s,k) for which G_{sk} contains a clique on the node set $\{P,Q,R\}$. Accordingly, (4.35c) contains the constraints

$$z_{PQ} \ge x_P + x_Q - 1, \quad z_{QR} \ge x_Q + x_R - 1, \text{ and } z_{PR} \ge x_P + x_R - 1 \qquad (4.36)$$

while (4.35a) for this (s,k) implies

$$z_{PQ} + z_{QR} + z_{PR} \le 1. \qquad (4.37)$$

Summing the three constraints in (4.36) and using (4.37), we get

$$1 \geq z_{PQ} + z_{QR} + z_{PR} \geq 2\left(x_P + x_Q + x_R\right) - 3 \tag{4.38}$$

which implies $x_P + x_Q + x_R \leq 2$. Hence again, the corresponding constraint in $\overline{C_1}$ is implied in the continuous sense.

Now consider a constraint in $\overline{C_1}$ for which $|S_k| = 4$, and is given by

$$x_P + x_Q + x_R + x_S \leq 3 \tag{4.39}$$

based on edges (P,Q) and (R,S) in M_{sk} for some graph G_{sk}. For this (s,k), (4.35a) implies the constraint

$$z_{PQ} + z_{RS} \leq 1 \tag{4.40}$$

while (4.35c) contains the relationships

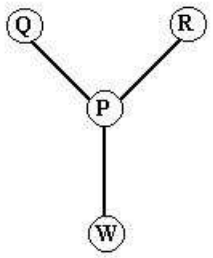
$$z_{PQ} \geq x_P + x_Q - 1 \text{ and } z_{RS} \geq x_R + x_S - 1. \tag{4.41}$$

Summing (4.40) and using (4.41), we have as before that (4.39) is implied. This completes the proof. □

The next example illustrates that $\overline{C_3}$ can yield a strictly tighter representation than $\overline{C_1}$.

Example 4-2

Consider the star conflict graph shown below.



The corresponding constraints of $\overline{C_1}$ are given by

$$x_P + x_Q + x_R \le 2, \quad x_P + x_Q + x_W \le 2, \quad x_P + x_R + x_W \le 2, \text{ and } 0 \le x \le e. \tag{4.42}$$

The solution $(x_P, x_Q, x_R, x_W) = (1, 1/2, 1/2, 1/2)$ is a vertex of (4.42) as evidenced by the four linearly independent constraints, given by $x_P = 1$, and the three structural inequalities in (4.42) being active. However, note that the constraints of $\overline{C_3}$ imply

$$\begin{aligned} &1 \ge z_{PQ} + z_{PR} + z_{PW} \ge (x_p + x_Q - 1) + (x_P + x_R - 1) + (x_P + x_W - 1), \\ &i.e. \quad 3x_P + (x_Q + x_R + x_W) \le 4. \end{aligned} \tag{4.43}$$

The above fractional solution is deleted by (4.43), and so in general, $\overline{C_3}$ can provide a strictly tighter representation than $\overline{C_1}$. However, to generate the star-graph facet $2x_P + (x_Q + x_R + x_W) \le 3$ for the underlying set $\overline{C_1}$ (see Sherali and Smith [46]), we would need to add $z_{PQ} + z_{PR} + z_{PW} \le x_P$ to $\overline{C_3}$. This extension is considered in Section 4.4 below. □

4.6.2. An Efficient Scheme for Generating C_3

In order to devise an efficient scheme for generating the conflict constraints C_3 defined in (4.35), suppose that we commence by constructing the overall conflict graph $G(N,A)$. Here, N is the set of nodes representing all the flight plans, and A is the set of arcs such that if flight plans P and Q conflict at any point in time over the horizon (with probability exceeding p_1 for severity level one conflicts and $p_2 < p_1$ for severity level two conflicts), then we have an edge connecting P and Q. Let us index the edges $(P,Q) \in A, (P<Q)$, as $c = 1,...,\overline{c}$. Let us also define an indexing scheme for pairs of conflicts as $1,...,\overline{\varsigma}$, where $\overline{\varsigma} = \dfrac{\overline{c}(\overline{c}-1)}{2}$, and where the two-tuple for any given pair of conflicts (l,l'), for $l < l'$, with $\{l,l'\} \subseteq \{1,...,\overline{c}\}$, is represented by the function $\varsigma(l,l'):\ \{1,...,\overline{c}\} \times \{1,...,\overline{c}\} \to \{1,...,\overline{\varsigma}\}$ given by

$$\varsigma(l,l') = \sum_{i=1}^{l-1}(\overline{c}-i) + (l'-l) \equiv \frac{(l-1)(2\overline{c}-l)}{2} + (l'-l). \tag{4.44}$$

Let us now define a one-dimensional vector E having $\overline{\varsigma}$ components that are initialized at zero, and where for any pair of edges $\{l,l'\} \subseteq \{1,...,\overline{c}\}, l < l'$, if l and l' are nonadjacent, then we leave $E[\varsigma(l,l')] \equiv 0$ permanently, while if l and l' are adjacent edges involving the nodes P, Q, and R, say, then $E[\varsigma(l,l')]$ might be revised to a value of 1 or 2 according to the following cases. Whenever we detect (P,Q,R) as a potential candidate for the set T_{NC} given by (4.34) based on some conflict graph G_{sk}, we set $E[\varsigma(l,l')] = 1$, and if (P,Q,R) is known not to belong to T_{NC} (perhaps at some later stage after being identified as a candidate for T_{NC}) then we set $E[\varsigma(l,l')] = 2$. Note that for any triplet (P,Q,R), if all the three possible edges l, l', and l'' induced by these nodes exist in the edge set M_{sk} for any conflict graph G_{sk}, then we enforce

$E\left[\varsigma\left(\overline{l,l'}\right)\right]=E\left[\varsigma\left(\overline{l,l''}\right)\right]=E\left[\varsigma\left(\overline{l',l''}\right)\right]=2$, where the notation $\left(\overline{l_1,l_2}\right)$ denotes that this two-tuple is $\left(l_1,l_2\right)$ if $l_1<l_2$ and it is $\left(l_2,l_1\right)$ if $l_2<l_1$.

Consider any G_{sk} being examined in turn, and suppose that we order the edges of M_{sk} according to their assigned indices in $\{1,...,\overline{c}\}$. For each $l\in M_{sk}$ examined in this order, consider each $l'>l$ in M_{sk} and do the following

- If l and l' are not adjacent, then reiterate.

- Else, let the incident nodes for this adjacent pair (l,l'), $l<l'$ be given by $\left(P,Q,R\right)$. Consider the following three possible cases:

Case (i)

If $E\left[\varsigma\left(l,l'\right)\right]=2$, then reiterate.

Case (ii)

If $E\left[\varsigma\left(l,l'\right)\right]=1$, then we have previously identified this $\left(P,Q,R\right)$ as a candidate for inclusion in T_{NC}. Hence, we check if this property still holds true as follows. If the third edge induced by $\left(P,Q,R\right)$ exists in M_{sk} and is given by $l''>l'$ (this edge triplet has not as yet been considered), then put

$$E\left[\varsigma\left(l,l'\right)\right]=E\left[\varsigma\left(l,l''\right)\right]=E\left[\varsigma\left(l',l''\right)\right]=2, \tag{4.45}$$

(i.e. $\left(P,Q,R\right)$ forms a clique and is hence no longer part of T_{NC}), and reiterate.

Case (iii)

If $E\left[\varsigma\left(l,l'\right)\right]=0$ and if the third edge induced by $\left(P,Q,R\right)$ exists in A_{sk} and is given by $l''>l'$, then execute (4.45) and reiterate. Otherwise, put $E\left[\varsigma\left(l,l'\right)\right]=1$ and reiterate.

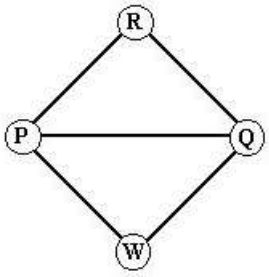
Once we have examined G_{sk} for all (s,k), we may then generate the constraint (4.35b), i.e., $x_P+x_Q+x_R\le 2$, for each $\left(l,l'\right)$ where $E\left[\varsigma\left(l,l'\right)\right]=1$. As we proceed sequentially for $\varsigma=1,...,\bar{\varsigma}$, we generate only non-redundant constraints by checking against previously generated constraints.

4.6.3. Additional Valid Inequalities

We now present an example to motivate a further tightening of $\overline{C_3}$ via the generation of some additional valid inequalities.

Example 4-3

Consider the following conflict graph, where $P<Q<R<W$, say,



Noting that the triplets $\left(P,R,W\right)$ and $\left(Q,R,W\right)\in T_{NC}$ as defined by (4.34), the constraints defining $\overline{C_3}$ (see (4.35)) are given as follows:

$$
\begin{aligned}
& z_{PQ} + z_{PR} + z_{PW} + z_{QR} + z_{QW} \le 1, \\
& z_{PQ} \ge x_P + x_Q - 1, \quad z_{PR} \ge x_P + x_R - 1, \quad z_{PW} \ge x_P + x_W - 1, \\
& z_{QR} \ge x_Q + x_R - 1, \quad z_{QW} \ge x_Q + x_W - 1, \\
& x_P + x_R + x_W \le 2, \quad x_Q + x_R + x_W \le 2, \\
& z \ge 0, \quad 0 \le x \le e .
\end{aligned} \tag{4.46}
$$

The problem

$$
\text{maximize } \{x_P + x_Q + x_R + x_W \ : \ \text{constraints (4.46)}\} \tag{4.47}
$$

yields the solution

$$
x_P = x_Q = x_R = 1/2, \ x_W = 1, \ \ z_{PQ} = z_{PR} = z_{QR} = 0, \ \text{ and } \ z_{PW} = z_{QW} = 1/2 \tag{4.48}
$$

with objective value equal to 2.5. However, since the x-variables are binary, we can impose the constraint

$$
x_P + x_Q + x_R + x_W \ \le \ \lfloor 2.5 \rfloor \ = \ 2 , \tag{4.49}
$$

which deletes the fractional solution. In fact, incorporating (4.49) within (4.47) now yields a binary extreme point solution. In this case, (4.49) is actually an even-hole facet-defining constraint for the corresponding set $\overline{C_1}$ (see Sherali and Smith [46]). □

Although we could solve a problem of type (4.47) to generate a cut of type (4.49) for each conflict graph G_{sk}, this might be burdensome. Hence, we propose the generation of the following two types of valid inequalities.

First, similar to (4.47), but using the total set of conflict constraints $\overline{C_3}$, we solve the linear programming problem

$$\nu_{C1} = \max \ \{ \ \sum_P x_P \ : \ \text{constraints } \overline{C_3} \text{ in (4.35)}, \ \sum_P x_{iP} = 1 \ \ \forall i \ \} . \tag{4.50}$$

If ν_{C1} is fractional, we impose the valid inequality

$$\sum_P x_P \ \leq \ \lfloor \nu_{C1} \rfloor \tag{4.51}$$

within C_3.

Second, we consider the linear program

$$\nu_{C2} = \min \ \{ \ \sum_P \overline{c}_P x_P \ : \ \text{constraints } \overline{C_3} \text{ in (4.35)},$$

$$\sum_P x_{iP} = 1, \forall i, \ \text{constraint (4.51) (if generated)} \ \} \tag{4.52}$$

where $\overline{c}_P = \lceil c_P \rceil \ \ \forall P$, or $\overline{c}_P$ is an integerized (rounded up, say) coefficient of x_P containing some of the other model objective terms as well. In any case, if ν_{C2} is fractional, we impose

$$\sum_P \overline{c}_P x_P \ \geq \ \lceil \nu_{C2} \rceil \tag{4.53}$$

within C_3. We can further tighten (4.53) in one of two ways. First, if $\Delta \geq 2$ is a (highest) common factor of $\overline{c}_P \ \ \forall P$, where $\lceil \nu_{C2} \rceil / \Delta$ is fractional, we can replace (4.53) by the stronger valid inequality

$$\sum_P \left(\frac{\overline{c}_P}{\Delta} \right) x_P \ \geq \ \left\lceil \frac{\lceil \nu_{C2} \rceil}{\Delta} \right\rceil . \tag{4.54}$$

Alternatively, we can derive a valid inequality (stronger than even (4.54)) by solving the following 0-1 knapsack problem

$$\nu_{C3} = \min \left\{ \sum_P \overline{c}_P x_P \ : \ \pi x \geq \pi_0, \, x \text{ binary} \right\} \tag{4.55}$$

where $\pi x \geq \pi_0$ is the strongest surrogate constraint for (4.52) derived by surrogating the structural constraints in (4.52) using the optimal dual multipliers obtained while solving this linear program.

Based on (4.55) we can impose

$$\sum_P \overline{c}_P x_P \ \geq \ \nu_{C3} \tag{4.56}$$

within C_3. We will experiment with the use of C_3 augmented with the foregoing valid inequalities in order to prescribe a viable strategy.

4.7. Alternative Conflict Constraint Formulation

Motivated by Example 4-2, let us further tighten the representation of C_3 by replacing the T_{NC}-cuts by higher-dimensional underlying star graph convex hull constraints as explained below.

For any (s,k), consider the conflict graph $G_{sk}(N_{sk}, M_{sk})$. Examine the edges in M_{sk} that are incident at any given node $i \in N_{sk}$, and define

$$J_{sk}(P) = \{ j \in N_{sk} : \ \ (P \sim Q) \in M_{sk} \}, \tag{4.57}$$

where $(P \sim Q)$ denotes (P,Q) if $P < Q$, and (Q,P) if $Q < P$, noting our convention that for any edge $(P,Q) \in M_{sk}$, we have $P < Q$. Likewise, we denote by (PQ) either PQ if $P < Q$, or QP otherwise. Consider the following result:

Proposition 4-3

Let $G_{sk}(N_{sk}, M_{sk})$ be the k^{th} conflict graph for sector s, and for any $P \in N_{sk}$, let $J_{sk}(P)$ be given by (4.57). Then, the following is a valid inequality for the conflict constraints:

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \leq x_P . \tag{4.58}$$

Proof

If $x_P = 0$, then (4.58) implies that $z_{(PQ)} = 0 \ \forall Q \in J_{sk}(P)$ (since $z \geq 0$), which is valid because $z_{(PQ)}$ represents the product $x_P x_Q$. If $x_P = 1$, then this is again valid because the conflict constraint asserts that we must have $\sum_{Q \in J_{sk}(P)} x_Q \leq 1$ when $x_P = 1$. This completes the proof. □

Based on Proposition 4-3, consider the following collection of constraints of the type (4.57), where we have imposed the condition $|J_{sk}(P)| \geq 2$ because, as we shall show in the sequel, the constraints for the case $|J_{sk}(P)| = 1$ are inconsequential with respect to the projection onto the original X-space.

$$\sum_{Q \in J_{sk}(P)} z_{(PQ)} \leq x_P, \quad \forall P \in N_{sk} \text{ such that } |J_{sk}(P)| \geq 2, \ \forall (s,k) . \tag{4.59}$$

Note that if we consider the overall conflict graph $G(N, A)$, and accordingly define, similar to (4.57),

$$J(P) = \{Q : \ (P \sim Q) \in A\} \quad \forall P \in N , \tag{4.60}$$

then several subsets of $J(P)$ will be used in the different graphs G_{sk} to generate constraints of the type (4.59) for any $P \in N$. Naturally, if $J_{s_1k_1}(P) \subseteq J_{s_2k_2}(P)$, then the constraint (4.59) that is based on $J_{s_1k_1}(P)$ may be dropped, since it is implied by that for $J_{s_2k_2}(P)$. Hence, let us define

$$I^* = \{P \in N: \quad \text{at least one constraint of the type (4.59) is generated}\} \tag{4.61}$$

and for each $P \in I^*$, let $J_r(P)$, for $r = 1,...,r_P$, be the collection of sets of type $J_{sk}(P)$ that yield a nonredundant system of constraints in (4.59). Accordingly, let us define the conflict constraint set

$$C_4 = \left\{ (x,z): \begin{array}{ll} \displaystyle\sum_{(P,Q)\in M_{sk}} z_{PQ} \le 1, & \forall (s,k) \qquad (4.62a) \\ \displaystyle\sum_{Q\in J_r(P)} z_{(PQ)} \le x_P, & \forall r = 1,...,r_P, \forall P \in I^* \qquad (4.62b) \\ z_{PQ} \ge x_P + x_Q - 1, & \forall (P,Q) \in A \qquad (4.62c) \\ z \ge 0, \ x \text{ binary} & \qquad (4.62d) \end{array} \right\}.$$

Consider the following result, which justifies the omission of constraints (4.59) corresponding to $|J_{sk}(P)| = 1$ in (4.62).

Proposition 4-4

Let C_4^+ be defined as in C_4 with the additional constraints

$$z_{(PQ)} \le x_P, \quad \forall P \in N_{sk} \text{ such that } J_{sk}(P) \equiv \{Q\}, \forall (s,k). \tag{4.63}$$

Accordingly, define $X = \left\{x: \quad (x,z) \in \overline{C_4}\right\}$ and $X^+ = \left\{x: \quad (x,z) \in \overline{C_4}^+\right\}$, where $\overline{C_4}$ and $\overline{C_4}^+$ respectively denote the continuous relaxations of C_4 and C_4^+. Then $X = X^+$.

Proof

It is clear that $X^+ \subseteq X$ because of the additional constraints (4.63) defining C_4^+. Hence, suppose that $\bar{x} \in X$ as evidenced by $(\bar{x}, \bar{z}) \in C_4$, and let us show that $\bar{x} \in X^+$ by demonstrating that there exists a $\hat{z}$ for which $(\bar{x}, \hat{z}) \in C_4^+$. Toward this end, consider

$$\hat{z}_{PQ} \equiv \max\{0,\ \bar{x}_P + \bar{x}_Q - 1\}, \quad \forall (P,Q) \in A. \tag{4.64}$$

Observe from (4.62c) and (4.62d) that we must have

$$\bar{z}_{PQ} \ \geq\ \max\{0,\ \bar{x}_P + \bar{x}_Q - 1\} \ =\ \hat{z}_{PQ} \quad \forall (P,Q) \in A. \tag{4.65}$$

Hence, since $(\bar{x}, \bar{z}) \in C_4$, we also have that $(\bar{x}, \hat{z}) \in C_4$ by virtue of (4.65). Furthermore, considering any constraint of type (4.63), we have that $(\bar{x}, \hat{z})$ satisfies this constraint as well because

$$\bar{x}_P \geq 0 \text{ and } \bar{x}_P \geq \bar{x}_P + \bar{x}_Q - 1 \quad \Rightarrow \quad \bar{x}_P \geq \max\{0,\ \bar{x}_P + \bar{x}_Q - 1\} = \hat{z}_{(PQ)}. \tag{4.66}$$

Therefore, we have that $\bar{x} \in X^+$, or that $X \subseteq X^+$. This completes the proof. □

We now exhibit that C_4 provides a tighter representation than C_3.

Proposition 4-5

C_4 provides a tighter representation for the conflict constraints that does C_3 in the sense that $\overline{C_4} \subseteq \overline{C_3}$.

Proof

Let $(\overline{x},\overline{z}) \in \overline{C_4}$. Noting the structure of $\overline{C_3}$, in order to show that $(\overline{x},\overline{z}) \in \overline{C_3}$, it is sufficient to demonstrate that $\overline{x}$ satisfies any T_{NC}-inequality. Toward this end, consider any such inequality

$$x_P + x_Q + x_R \;\leq\; 2 \tag{4.67}$$

where, without loss of generality, suppose that $\{P,Q,R\} \subseteq N_{sk}$ with $\{Q,R\} \subseteq J_{sk}(P)$ for some (s,k). Hence, there exists a constraint of the type (4.62b) defining C_4 for which $\{Q,R\} \subseteq J_r(P)$. From this constraint, using (4.62c), and $z \geq 0$, we get

$$x_P \;\geq\; \sum_{Q \in J_r(P)} z_{(PQ)} \;\geq\; z_{(PQ)} + z_{(PR)} \;\geq\; \left(x_P + x_Q - 1\right) + \left(x_P + x_R - 1\right), \tag{4.68}$$

which implies (4.67). This completes the proof. □

To illustrate that C_4 can provide a strictly tighter representation than C_3, consider the conflict graph of Example 4-2. For this graph, the set $\overline{C_3}$ is given by

$$\overline{C_3} = \left\{ \begin{array}{ll} (x,z): & z_{PQ} + z_{PR} + z_{PW} \leq 1 \\ & x_P + x_Q + x_R \leq 2, \quad x_P + x_Q + x_W \leq 2, \quad x_P + x_R + x_W \leq 2, \\ & z_{PQ} \geq x_P + x_Q - 1, \quad z_{PR} \geq x_P + x_R - 1, \quad z_{PW} \geq x_P + x_W - 1, \\ & z \geq 0, \quad 0 \leq x \leq e \end{array} \right\}. \tag{4.69}$$

The seven structural inequalities defining $\overline{C_3}$ are linearly independent, and their intersection yields the feasible fractional vertex

$$z_{PQ} = z_{PR} = z_{PW} = 1/3, \quad x_P = x_Q = x_R = x_W = 2/3. \tag{4.70}$$

However, note that $\overline{C_4}$ has the constraint

$$z_{PQ} + z_{PR} + z_{PW} \leq x_P \tag{4.71}$$

of type (4.62b) that deletes this fractional solution. Observe that (4.71) along with (4.62 c) implies the valid inequality

$$x_P \geq z_{PQ} + z_{PR} + z_{PW} \geq \left(x_P + x_Q - 1\right) + \left(x_P + x_R - 1\right) + \left(x_P + x_W - 1\right), \tag{4.72}$$

that is,

$$2x_P + \left(x_Q + x_R + x_W\right) \leq 3, \tag{4.73}$$

which also deletes the solution (4.70). In fact, as shown by Sherali and Smith [46], (4.73) is a facet for the underlying star conflict graph, and that in general for star graphs, the constraints of $\overline{C_4}$ characterize the complete convex hull representation for the conflict constraints, implying an exponential collection of facets in the original projected x-space representation. Hence, this representation C_4 for the conflict constraints might be particularly useful.

We shall conduct experiments using the alternative representations C_3 or C_4 (or both) of the conflict constraints, and examine their relative performance.

4.8. Generalized M_s-inequalities

Let us revisit the case where the number of conflicts to be allowed simultaneously in any sector s is restricted to be no more than some $r_s \geq 1$ in general. The extended form of C_4 is then given as follows:

$$C_4' = \left\{ (x,z): \begin{array}{ll} \displaystyle\sum_{(P,Q)\in M_{sk}} z_{PQ} \le r_s, & \forall (s,k) \\ \displaystyle\sum_{Q\in J_{sk}(P)} z_{(PQ)} \le r_s x_P, & \forall P \in N_{sk} \text{ such that } \left|J_{sk}(P)\right| \ge r_s + 1, \ \forall (s,k) \\ z_{PQ} \ge x_P + x_Q - 1, & \forall (P,Q) \in A \\ z \ge 0, \ x \text{ binary} & \end{array} \right\} \begin{array}{l} \text{(4.74a)} \\ \text{(4.74b)} \\ \text{(4.74c)} \\ \text{(4.74d)} \end{array}.$$

Note that as in Proposition 4-4, we can show that the projection of C_4' onto the x-space remains unaltered if we include the constraints of the type (4.62b) that correspond to $\left|J_{sk}(P)\right| \le r_s$, even if we tighten (4.62b) in that case to

$$\sum_{Q\in J_{sk}(P)} z_{(PQ)} \le \left|J_{sk}(P)\right| x_P . \tag{4.75}$$

As in the proof of Proposition 4-4, we obtain $\overline{x}_P \ge \hat{z}_{(PQ)} \ \ \forall Q \in J_{sk}(P)$, which upon summing over all $Q \in J_{sk}(P)$, implies that (4.75) is satisfied by the $(\overline{x}, \hat{z})$ used in that proof. Furthermore, as in (4.62b), we can identify and eliminate redundant constraints (4.74b). In particular, if for some (s_1, k_1) and (s_2, k_2) (where possibly $s_1 = s_2$), we have for some $P \in I^*$ defined similar to (4.61) that

$$J_{s_1 k_1}(P) \subseteq J_{s_2 k_2}(P) \ \text{ and } \ r_{s_1} \ge r_{s_2} , \tag{4.76}$$

then the constraint in (4.74b) corresponding to (s_1, k_1) may be dropped for this P. Hence, scanning (4.74b) in this fashion for each $P \in I^*$ for which at least two constraints of type (4.74b) are generated, we can eliminate the implied inequalities from this set. Note again that as shown by Sherali and Smith [46], $\overline{C_4}$ yields a convex hull representation for the case of star graphs.

4.9. Additional Conflict Modeling Constructs

An advantage of using the C_4 (or alternatively, C_3) conflict constraints defined over the (x,z)-space, aside from the tighter representation that it provides as elucidated in Proposition 4-2 and Proposition 4-5, is that it permits the consideration of more detailed modeling strategies. We describe two such additional modeling constructs below.

First, note that since we have a specific variable z_{PQ} that denotes the existence of a conflict between selected flight plans P and Q whenever $z_{PQ}=1$, we can ascribe a specific cost φ_{PQ} to this variable that reflects a penalty based on the (expected) duration and severity of this conflict. This provides a greater flexibility in filtering the set of conflicts that would need resolution, in addition to simply limiting the number of simultaneous (generic) conflicts.

Second, we can modify the constraints (4.62a) by instead imposing

$$\sum_{(P,Q)\in M_{sk}} z_{PQ} \le 1+\eta_s, \quad \text{where } 0\le\eta_s\le(r_s-1), \tag{4.77}$$

where η_s equals the additional conflicts permitted beyond 1. Accordingly, in the objective function, we can accommodate the term

$$\sum_s \bar{\mu}_s\eta_s\,, \tag{4.78}$$

where the coefficients $\bar{\mu}_s$ penalize the excessive workload imposed via admitting more than one simultaneous conflict in sector s. (Recall that the coefficients φ_{PQ} defined above penalize the specific conflicts themselves based on duration and severity.)

Including (4.77) softens (4.62a) and avoids blatant infeasibilities in tight situations. Observe also that in lieu of the linear penalty function (4.78), we can impose an increasing-rate penalty function by representing η_s in terms of its specific possible

values over the interval $[0, r_s - 1]$ using binary variables, and then by associating specific penalties to each of these value-representing binary variables.

We will investigate the foregoing two strategies outlined in this section at a subsequent stage, once the primary model has been analyzed.