

Formalizing the ATM problem

Bryan O’Gorman, Tobias Stollenwerk

May 18, 2016

These notes attempt to formalize the ATM problem: given a set of flights, schedule them in a way that minimizes cost.¹

1 Configuration space

At the highest level, the configuration space is a set of 4D trajectories, one for each flight. By flight we mean an ideal trajectory (times and spatial points) between a departure city and arrival city, including an ideal departure time. Even with relatively coarsely discretized spacetime, that space is infeasably large. Therefore we parameterize each trajectory by its deviations from the ideal, conflict-ignorant one. Such deviations are of two types: temporal and spatial. A temporal deviation is simply a delayed start (typically on the order of up to 30 minutes). We assume that within the time-scale of such delays, the ideal conflict-ignorant trajectory is time-independent. We consider local spatial deviations, which consist of a set of maneuvers that change relatively small parts of the spatial trajectory². Such maneuvers may introduce a delay, but, as with the delays due to temporal deviations, we assume that the subsequent ideal trajectory is unaffected other than simply being delayed.

Henceforth, we focus on origination delays (temporal deviations) and maneuvers (local spatial deviations). We assign an index to each flight

$$i \in \{1, \dots, n\}$$

For each flight i , the wind-optimal trajectories are given as a sequence of M_i 4-dimensional space-time points

$$\tau_i = \{p_1^i, \dots, p_{M_i}^i\}$$

with

$$p_j^i = (\varphi_j^i, \lambda_j^i, z_j^i, t_j^i)$$

where $(\varphi_j^i, \lambda_j^i)$ are the latitude and longitude. For simplicity, we will assume a constant altitude z_j^i for all flights.

$$p_j^i = (\varphi_j^i, \lambda_j^i, t_j^i)$$

2 Potential conflicts

Conflicts may occur if multiple flight trajectories

$$\{\tau_i : i \in I_k \subseteq \{1, \dots, n\}\}$$

become too close to each other in time and space.

In the following we will define spatial, potential, and real conflicts. A *spatial conflict* occurs if two trajectories become too close to each other in space. The existence of a *real conflict* depends on the arrival

¹It remains to be proven that this problem is hard, either in general or for realistic values of its parameters, e.g. cost functions.

²There could also be global spatial deviations, which change the entire trajectory, such as are used in Olga’s thesis. For example, one parameterization of global spatial deviations treats each trajectory as an arc in a plane perpendicular to the ground and considers smooth one-parameter curves to the projection of the trajectory onto the ground.

time of both flights and at the spatial conflict. A *potential conflict* exists, if a real conflict is possible considering a maximal flight delays upon arrival at the spatial conflict. For details, see appendix A.

In any case, we index the potential conflicts by k . A classical precalculation gives us the mapping to the involved flights

$$k \mapsto I_k \subseteq \{1, \dots, n\}$$

as well as the set of conflicts the flight i is involved in

$$i \mapsto \{k_1^i, \dots, k_{N_i}^i\} =: K_i \quad (1)$$

Moreover, we can order the N_i conflicts the flight i is involved in by their temporal appearance:

$$i \mapsto (k_1^i, \dots, k_{N_i}^i)$$

In the following it will prove useful to introduce the set of conflicts the flight i was involved in before it arrives at a potential conflict k

$$K_{ik}^< = \{\tilde{k} : \tilde{k} \in \{k_0^i, \dots, k_{N_i}^i\}, \tilde{k} < k\}$$

Note that the conflicts may not be point conflicts, but may extend in time and space. See the appendix.

3 Cost function

In reality, the cost function results from factors such as fuel, labor, time, etc.; here we treat it as a unitless quantity that appropriately weights these factors. Because we are interested in the relative costs of different configurations, we treat the cost of the configuration in which all flights are routed along their independently optimal trajectories, ignoring conflicts, and are exactly on time, as zero. We further assume that the total cost is the sum of two different types of costs:

- for each flight i , the cost c_i^{late} of being late on arrival,
- for every conflict k , the cost c_k^{avoid} of avoiding it.

The total cost therefore is

$$c = \sum_i c_i^{\text{late}} + \sum_k c_k^{\text{avoid}},$$

where each contribution to the cost function depends (implicitly as written so far) on the configuration space. Because the lateness and avoidance costs are associated with flights and conflicts, respectively, we write simply $c_i = c_i^{\text{late}}$ and $c_k = c_k^{\text{avoid}}$ where the distinction is clear from context. Henceforth, we simplify things by ignoring the direct cost of the local maneuvers. However, they will still indirectly contribute to the cost function via their introduction of delays. The quantity to be minimized can then be written

$$c = \sum_i c_i(D_i),$$

where $D_i = d_i + \sum_{K_i} d_{ik}$ is the delay of flight i at its destination, which will be a function of its origination delay d_i and delays $\{d_{ik}\}$ introduced by avoiding the potential conflict k by a local maneuver.

Lastly, we simplify things further by assuming that all destination delays are equally important, (which comes by necessity from the absence of information to the contrary), so that the quantity to minimize is simply the total delay of all flights: ³

$$c = D = \sum_i D_i. \quad (2)$$

³The lack of information with respect to the relative importance of flights in reality only requires that all c_i be the same function. One could imagine minimizing the L_2 norm (which would help avoid having a few very late flights) rather than the L_1 norm as we do here, or indeed any other function. The simple sum here is the one currently used by Olga et al., and has the advantage of being most conducive to mapping to QUBO.

4 Classical precalculation

We assume the availability of the following classical routines, whose resource requirements we regard as negligible:

- Given the trajectories of all flights, return the potential conflicts

$$\{\tau_1, \dots, \tau_n\} \mapsto \{k_1, \dots, k_m\}$$

and the corresponding parameters attached to a potential conflict k :

- The set of involved flights I_k
- The trajectory points involved $\{p_j^i : j \in J \subseteq M_i, i \in I_k, \}$
- Given the above potential conflict parameter, return the delays introduced by a conflict avoiding maneuver

$$k \mapsto d_{ik} \quad \forall i \in I_k$$

This delay may depend on a set of maneuver parameters.

$$\mathbf{a}_k = (a_k^1, \dots, a_k^{L_k})$$

5 Assumptions made

- For each flight, the optimal route is time-independent (both spatially and in duration) over the timescale of possible delays. For example, if the flight has no potential collisions, its lateness on arrival will exactly equal its departure delay.
- Actively avoiding a collision (rather than avoiding via departure delays) does not change the cost of actively avoiding other collisions on the same flight path. In reality, one could reason, e.g. that the little bit of fuel used in each active collision avoidance could add up to a disastrously empty tank, but in general we would like the effect of active collision avoidances to be as little as possible. Similarly, one could argue that the plane burns fuel while idling.

6 Potentially helpful assumptions

- (Sparseness) All conflicts are pairwise. That is, when two planes have a potential conflict, there are no other planes nearby. Furthermore, when this conflict is avoided, it does not create a conflict with another flight. (This may be the most unreasonable assumption, e.g. near major airports or on crowded jetstreams.)
- Collisions are symmetric. That is, the cost of avoiding them is a function only of the magnitude of the difference in delays of the two flights up until that point. For example, the cost of avoiding a collision between i and j is the same whether i is delayed by d or j is.
- The lateness cost functions are non-decreasing. This is obviously satisfied for the linear cost functions with unit slope used here, and hard to imagine relaxing.

7 Models

The arrival time of flight i at a potential conflict k reads

$$T_{ik} = t_{ik} + d_i + \sum_{p \in K_{ik}^<} d_{ip}$$

where t_{ik} is the arrival time in the wind-optimal case. The delay it has picked up upon arrival at the potential conflict k , is

$$\begin{aligned} D_{ik} &= T_{ik} - t_{ik} \\ &= d_i + \sum_{p \in K_{ik}^<} d_{ip} \end{aligned}$$

7.1 General Model

In general, the delay introduced by conflict avoiding maneuver d_{ik} is a function of the arrival delay of each involved flight $D_{jk}, \forall j \in I_k$ and the maneuver parameters \mathbf{a}_k

$$d_{ik} = d_{ik}(\mathbf{D}_k, \mathbf{a}_k) \quad (3)$$

where $\mathbf{D}_k = (D_{jk})_{j \in I_k}$. This function is given as a classical routine mentioned in section 4. The total delay (2) is then given by

$$\begin{aligned} D &= \sum_i \left(d_i + \sum_{k \in K_i} d_{ik} \right) \\ &= \sum_i \left(d_i + \sum_{k \in K_i} d_{ik}(\mathbf{D}_k, \mathbf{a}_k) \right) \end{aligned}$$

7.2 Sparse case

First, we describe a method for solving the “sparse” version of the problem. By sparsity, we mean that all potential conflicts are between only two trajectories and that no other trajectories are nearby in spacetime, so that potential maneuvers for avoiding a conflict between two flights do not introduce additional potential conflicts with other flights.⁴

Making the assumption that the maneuvers depend only on the difference between the arrival times of the flights, we can write d_{ik} as a function of the difference of the delays for the two flights, where i can be either of the two indices j and l , with $l < j$, involved in conflict k .

$$\Delta_k := D_{lk} - D_{jk} \quad l < j \in I_k$$

and the maneuver parameters \mathbf{a}_k . It’s plausible to assume that it furthermore depends only on the magnitude of this difference $|D_{lk} - D_{jk}|$, though it’s not yet clear that that will help. With this (3) becomes

$$d_{ik} = d_{ik}(\Delta_k, \mathbf{a}_k)$$

and the total delay reads

$$D = \sum_i \left(d_i + \sum_{k \in K_i} d_{ik}(\Delta_k, \mathbf{a}_k) \right).$$

Three special cases are as follows.

7.2.1 No Delay Model

In the (totally unrealistic) case where maneuvers do not introduce any further delays, we have simply

$$\Delta_k = d_i - d_j$$

⁴This could be extended to an m -sparse version, where each conflict can affect $m - 1$ others.

and

$$d_{ik} = d_{ik}(d_l, d_j) \quad (l, j) \in I_k$$

$$= \begin{cases} 0, & \text{potential conflict } k \text{ can be avoided given delays } d_l \text{ and } d_j, (l, j) \in I_k \\ \infty, & \text{otherwise.} \end{cases}$$

For the assumption that maneuvers do not introduce further delays to make any sense at all, we must bound their size, which is why some conflicts simply cannot be avoided within this approximation. Of course, we define d_{ik} as piece-wise infinite for consistency with the rest of these notes, but in practice this would be implemented as a hard constraint. (Interestingly, this notational device has the morose interpretation that if the flights do collide, the further delay introduced is indeed infinite.)

7.2.2 Independent Delay Model

In this model we assume that the delay introduced by a conflict-avoiding maneuver is independent of the other flight. Each potential conflict k is an actual conflict when the difference between the total delays of the constituent flights are in some range $[\Delta_k^{\min}, \Delta_k^{\max}]$. However, it is not clear how to penalize real conflicts.

7.2.3 Delay Frontier Model

For a given Δ_k and $(l, k) \in I_k$ we assume that there is a minimal delay of each flight d_{lk}^* and d_{jk}^* assuming that the other is unchanged, and there is (presumably) a concave “delay frontier” in the region $[0, d_{lk}^*] \times [0, d_{jk}^*]$ connecting $(d_{lk}^*, 0)$ and $(0, d_{jk}^*)$ that gives the possible delays (d_{lk}, d_{jk}) . We parameterize this curve by the scalar $a_k \in [0, 1]$. More generally, there is a “delay surface” when various values of Δ_k are considered. Discretizing a_k into only two values $\{0, 1\}$ is equivalent to requiring that exactly one flight maneuvers around the other, whose trajectory remains unchanged.

In order to formulate the problem as a QUBO, we discretize and introduce bits for each discrete value for the variables d_i , d_{ik} , $\Delta_k = D_{lk} - D_{jk}$, and a_k :

$$C = D + p_1 C^{\text{one}} + p_2 C^{\text{equal}},$$

where

$$D = \sum_i \sum_{\alpha} \alpha d_{i\alpha} + \sum_{ik} \sum_{\beta} \beta d_{ik\beta}$$

is the delay to be minimized,

$$C^{\text{one}} = \sum_i \left(\sum_{\alpha} d_{i\alpha} - 1 \right)^2 + \sum_{ik} \left(\sum_{\beta} d_{ik\beta} - 1 \right)^2 + \sum_k \left(\sum_{\gamma} \Delta_{k\gamma} - 1 \right)^2 + \sum_k \left(\sum_{\delta} a_{k\delta} - 1 \right)^2$$

is the penalty function ensuring that each variable has a well defined value encoded by a single bit,

$$C^{\text{equal}} = \sum_k \left[\left(\sum_{\alpha} \alpha d_{l\alpha} + \sum_{p \in K_{lk}^<} \sum_{\beta} \beta d_{lp\beta} \right) - \left(\sum_{\alpha} \alpha d_{j\alpha} + \sum_{q \in K_{jk}^<} \sum_{\beta} \beta d_{jq\beta} \right) - \left(\sum_{\gamma} \gamma \Delta_{k\gamma} \right) \right]^2 +$$

$$\sum_{ik} \left[\left(\sum_{\beta} \beta d_{ik\beta} \right) - \left(\sum_{\gamma\delta} d_{ik}(\gamma, \delta) \Delta_{k\gamma} a_{k\delta} \right) \right]^2$$

is the penalty function ensuring that $\Delta_k = D_{lk} - D_{jk}$ and $d_{ik} = d_{ik}(\Delta_k, a_k)$, the Greek indices represent the discrete values that the corresponding quantities can assume, and p_1 and p_2 are penalty weights.⁵ Note, that the last term introduces cubic and quartic terms.

⁵In practice we could have varying penalty weights for the various penalties.

We could make a further approximation and for each conflict pick some optimal-in-expectation a_k , thereby removing that variable from the formulation. This would decrease the degree of the above cost function resulting in a QUBO. Alternatively, we could allow the delays (d_{ik}, d_{jk}) to take on any value at or beyond the delay frontier (i.e. somewhere in between the independent delays and delay frontier cases). We could also suitably parameterize the “delay surface” so that the cost function could be written in a functional form.

7.2.4 Variations of the Delay frontier model

We can reduce the number of variables by

1. Both flights will always *split the delay*: Set a_k to a fixed value inside $[0, 1]$, presumably $a_k = 0.5$. This would eliminate all variables $a_{k\beta}$.
2. One flight makes a detour, the other remains on the wind-optimal route: Reduce the possible values of a_k to either 0 or 1. Hence a_k itself becomes a binary variable.

7.3 Dense case

The sparsity condition assumed above is in general false, even when ignoring near-airport trajectories. The other extreme is the “dense” case, in which all flights are restricted to a “highway” as in Olga’s thesis. An appropriate QUBO for that problem is relatively straightforward, and can be written explicitly if needed.

7.4 General case

General instances fall into neither of the above extremes. One potential solution is to use some sort of master-slave subproblem decomposition to combine solvers for each extremal case into a general solution.

8 Mapping to Job Shop Scheduling

We can map the ATM problem to job shop scheduling by imagining the space at a potential conflict k as a resource, the flights I_k have to compete for.

R_k = spatial trajectory points of the flights $i \in I_k$ which are in (spatial) conflict with each other

In the language of job shop scheduling, we have a ordered set of tasks for each flight $i \in I_k$:

$$\mathcal{T}_i = (R_i^{t_0}, \dots, R_i^{t_M})$$

where t_0 and t_M are the earliest and the latest time in the whole trajectory data set. R_i^t is the resource assigned to flight i at time t . If a flight i at time t' occupies a spatial trajectory point which is not in spatial conflict with any other flight, it is viewed as using the resource R_i , i.e. $R_i^{t'} = R_i$. The role of conflict avoiding maneuvers is replaced by the a task waiting for a resource occupied by a conflicting flight to free up. In contrast to the above models where the cost function was the sum of all delays, the cost function of this approach is the maximal delay among all flights.

8.1 Example

Assume three flights and two potential conflicts between flight 1 and 2 as well as between 2 and 3 (see figure 1). Let the set of tasks for each flights read

$$\begin{aligned}\mathcal{T}_1 &= (R_1, R_1, R_{k_1}, R_1, R_1, R_{k_2}, R_1) \\ \mathcal{T}_2 &= (R_2, R_2, R_{k_1}, R_2, R_2, R_2, R_2) \\ \mathcal{T}_3 &= (R_3, R_3, R_3, R_3, R_{k_2}, R_3, R_3)\end{aligned}$$

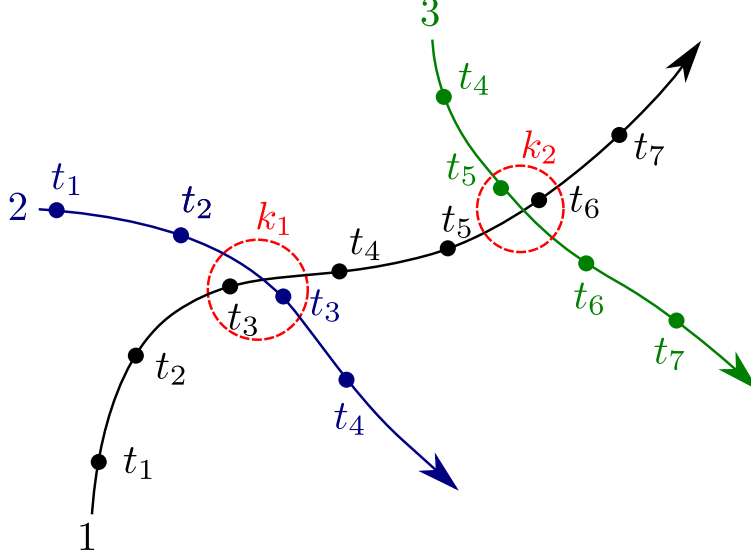


Figure 1: Example for mapping to job shop scheduling

A Potential conflicts

A.1 Spatial conflicts

In the following we will restrict ourselves to pairwise conflicts. Given a spatial conflict between the flights i and j , a real conflict exists if the arrival times at the spatial conflict k , T_{ik} and T_{jk} (see figure 2) obey certain conditions. The corresponding departure times from the spatial conflict are denoted by

$$T'_{ik} = \frac{L_k}{v_i} + T_{ik}$$

where L_k spatial length of the spatial conflict k and v_i is the constant velocity of flight i .

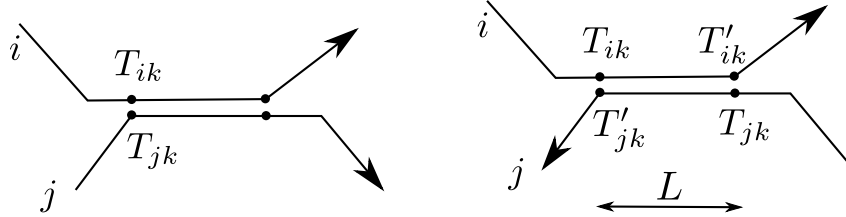


Figure 2: Parallel and antiparallel spatial conflict

- **Parallel conflict:** A real conflict exists if the difference between the arrival times is smaller than a threshold Δ_p which depends on the trajectories, in particular the speed and position of both flights.

$$\underbrace{|T_{jk} - T_{ik}|}_{\Delta_k} < \Delta_p$$

with

$$\Delta_p = \Delta_p(v_i, v_j, r_i, r_j)$$

This can also be written as

$$-\Delta_p < \Delta_k < \Delta_p \quad (4)$$

- **Antiparallel conflict:** A real conflict exists if

$$T'_{ik} - T_{jk} < \Delta_{ap}^1(v_i, v_j, r_i, r_j)$$

and

$$T'_{jk} - T_{ik} < \Delta_{ap}^2(v_i, v_j, r_i, r_j)$$

Hence

$$\frac{L_k}{v_i} + \underbrace{T_{ik} - T_{jk}}_{-\Delta_k} < \Delta_{ap}^1$$

$$\frac{L_k}{v_j} + \underbrace{T_{jk} - T_{ik}}_{\Delta_k} < \Delta_{ap}^1$$

This yields

$$\frac{L_k}{v_i} - \Delta_{ap}^1 < \Delta_k < \Delta_{ap}^2 - \frac{L_k}{v_j} \quad (5)$$

A.2 Potential conflicts

The arrival time T_{ik} is given by the arrival time in the wind optimal case t_{ik} plus the delay it picked up due to earlier conflict avoiding maneuvers or a departure delay D_{ik} .

$$T_{ik} = t_{ik} + D_{ik} \quad \forall i, k$$

and the difference of the arrival time of the two flights i and j at spatial conflict k is given by

$$\begin{aligned} \Delta_k &= T_{ik} - T_{jk} \\ &= t_{ik} - t_{jk} + D_{ik} - D_{jk} \end{aligned}$$

The maximal value of D_{ik} , $\forall i, k$ is bounded from above by some D^{\max} . Hence, Δ_k is bounded by

$$(t_{ik} - t_{jk} - D^{\max}) < \Delta_k < (t_{ik} - t_{jk} + D^{\max}) \quad (6)$$

Therefore, we identify a spatial conflict k as a **potential conflict** if for a Δ_k which is bounded by (6), one of the inequalities (4) or (5) is fulfilled.

From now on, all indices k will denote potential conflicts.