A Distributed Scheduler for Air Traffic Flow Management

Steven J. Landry*

Todd Farley

Ty Hoang

NASA Ames Research Center

M/S 210-6

Moffett Field, CA 94035

Brian Stein

Methods Corporation

Keywords: Air traffic flow management, distributed scheduler, air traffic arrival scheduling, decision support

---

*Corresponding author, currently with Purdue University, School of Industrial Engineering, 315 N. Grant St. West Lafayette, IN 47906. (765) 494-6256, slandry@purdue.edu, Fax (765) 494-6802

Abstract A system was developed to efficiently schedule aircraft into congested resources over long ranges and present that schedule as a decision support system. The scheduling system consists of a distributed network of independent schedulers, loosely coupled by sharing capacity information. This loose coupling insulates the schedules from uncertainty in long-distance estimations of arrival times, while allowing precise short-term schedules to be constructed. This "rate profile" mechanism allows feasible schedules to be produced over long ranges, essentially constructing precise short-range schedules that also ensure that future scheduling problems are solvable while meeting operational constraints. The system was tested operationally and demonstrated reduced airborne delay and improved coordination.

Acronyms

AMDT: allowable maximum delay time
ARTCC (or Centers): Air Route Traffic Control Centers
ATAs: actual times of arrival
CTAS: Center-TRACON Automation System
DP: "dynamic planner"
DS: distributed scheduler
ETAs: estimated times of arrival
FAA: Federal Aviation Administration
FCFS: first-come, first-served
McTMA: Multi-center Traffic Management Advisor
MiT: miles-in-trail
NASA: National Aeronautics and Space Administration
NextGen: next generation air transportation system
nmi: nautical miles
PHL: Philadelphia International Airport
RTA: required time of arrival
SESAR: Single European Sky ATM Research programme
SPDP "single point DP"
TMA: traffic management advisor
TMCs: traffic managers
TRACON: Terminal Radar Approach Control facilities
ZBW: Boston Center
ZDC: Washington Center
ZNY: New York Center
ZOB: Cleveland Center

1 Introduction

The government agency in the U.S. responsible for planning the next generation air transportation system (NextGen), the Joint Planning and Development Office, has identified “trajectory-based operations” as a fundamental technology required for substantially increasing the capacity of the air traffic system. One requirement for trajectory-based operations is that capacity restrictions be considered over long ranges – i.e., that real-time, long-range scheduling of aircraft into constrained resources be implemented. This paper describes the design, development, and testing of a system capable of such scheduling.

The system’s design is that of a loosely-coupled, distributed scheduler. Individual scheduling instances create short-range schedules that are enforced by controllers, while passing constraints to other instances of schedulers. When scheduling, a scheduler instance must consider both short-range constraints as well as those fed to it by other schedulers to which it is delivering aircraft. This design satisfies several important considerations: it is fast enough to produce schedules in near real-time for an extremely large problem, it is implementable within the current operational air traffic control system, it creates schedules whose sequences match the physical sequencing being used by controllers, and it minimizes reliance on long-range estimates of arrival times.

The resulting system does not create optimal schedules. Each scheduling solution is designed to create a feasible, not optimal, solution, where it is feasible in the local sense of having proper separation at the meter points and only passing along solvable problems to downstream schedulers. Since the schedule is done in a first-come, first-served manner, it is also produces an equitable solution for the different airlines.

The architecture, implemented within the Multi-center Traffic Management Advisor (McTMA) decision support tool, is called the “distributed scheduler,” and it has been tested operationally with successful results. The concept was accepted for implementation by the U.S. Federal Aviation Administration.

This paper first discusses the problem in detail, including past work. The solution is then introduced, followed by a description of the extensive field testing done in an operational air traffic control environment. The field testing was the primary means of establishing the acceptability of the system. However, as a final test, the system was used operationally over two days at four U.S. air route traffic control centers. The effect of the use of the system over those two days was quantified.

The system being described in this paper is a software system developed and implemented within an operational context. As such, there is no mathematical description of the problem, nor is a mathematical solution provided. The relationship of this system to such problems, however, is discussed. Instead, this paper is focused on disseminating what is believed to be an innovative solution to a difficult applied problem in a challenging operational environment, including the solution's architecture and algorithms. It is believed that this architecture, and to some extent the algorithms, can be applied to similar problems in many different application domains.

2 Background

2.1 Past work

There are currently several operational systems for scheduling aircraft into airports and the surrounding terminal airspace, such as the traffic management advisor (TMA) (Wong, 2000), OASIS (Ljungberg & Lucas, 1992), COMPAS (Voelckers, 1990), and OSYRIS. However, these systems are focused on scheduling into single resources consisting of an airport's runway complex and surrounding airspace, and work well only over distances of about 250 nautical miles (nmi). For example, numerous airports in the United States currently use TMA, which uses a scheduler called the "dynamic planner" (DP) to sequence and schedule arriving aircraft.

In order to achieve some of the goals of NextGen and Single European Sky ATM Research programme (SESAR), aircraft will need to be scheduled over ranges of up to thousands of nautical miles, a problem for which the current generation of scheduling systems cannot be scaled up. Several problems arise when the scale of the scheduling problem increases, including interactions between the individual airport problems due to sharing of upstream resources by aircraft bound for the same or different airports and increases in the uncertainty in the estimated arrival times of aircraft used in the scheduling process. The infrastructure and algorithms used in TMA and similar systems are incapable of dealing with these complexities.

A large number of mathematical approaches have also been proposed to handle such problems, including individual arrival problems and more general traffic flow management problems, generally focused on identifying more efficient or valuable sequences of aircraft in order to optimize some parameter such as throughput or delay (Barnhart, Belobaba, & Odoni, 2003). A good overview of past methods can be found in (Wu & Caves, 2002). A number of these approaches have been focused on the multiple-airport ground holding problem and the traffic flow management problem (Andreatta, Brunetta, & Guastalla, 2000; Bersimas, Lulli, & Odoni, 2011; Bertsimas & Stock-Patterson, 1998; Odoni, 1987; Richetta, 1995; Terrab & Odoni, 1993; Vranas, Bertsimas, & Odoni, 1994).

However, these mathematical methods have not been implemented within an operational system, with the possible exception of some recent work on collaborative decision making for releases from ground holds (Ball, Donohue, & Hoffman, 2006; Idris, Evans, Evans, & Kozarsky, 2004), for several reasons. Because of the size of a realistic optimization problem, consisting of hundreds of aircraft with numerous constraints, the solution time for such methods is often very high, well beyond a system’s likely requirement that the solutions be presented in real time. In addition, methods proposed to date can yield infeasible solutions because they do not consider all of the operational aspects of the problem such as trajectory prediction uncertainty, flow sequencing constraints, and holding queue capacity. More recent methods have been proposed to overcome these problems (Balakrishnan & Chandran, 2006; Bersimas, et al., 2011; Saraf & Slater, 2008), but none of these methods have ever been tested operationally.

2.2 The problem of air traffic management and scheduling

Prior to the development of systems such as TMA, no useful scheduling systems have been available to air traffic controllers or managers to assist them in coordinating or scheduling air traffic. Air traffic authorities have historically managed aircraft at the level of flows, applying local distance-based spacing constraints between aircraft, airborne holding, and ground holding on particular streams of aircraft to manage air traffic congestion. These “Eulerian” methods, while easy to implement and requiring little inter-facility communication, are uncoordinated at a regional and national level, are not tightly connected to actual demand profiles on congested resources, do not consider more efficient sequences, and can inordinately penalize the busiest streams of aircraft. Such methods are still used extensively today where TMA is not available or under conditions where it is not considered useful by the facilities involved.

The problem of air traffic management is complicated by the organizational structure required to manage air traffic over thousands of miles using radar systems to track aircraft and without significant communication infrastructure such as a data network. (The air traffic system has been very slow to modernize.) Most inter-facility coordination has been conducted using telephone, which has a high workload overhead and very limited bandwidth since the communication occurs between two human operators.

Specifically, the U.S. airspace structure is (roughly) broken into a number of Air Route Traffic Control Centers (ARTCCS or Centers) that control traffic between its origin and destination, and Terminal Radar Approach Control facilities (TRACONs) that control departing and arriving aircraft. These facilities are not physically co-located, with the 20 continental U.S. Centers typically located near the center of the airspace they control. (Airspace outside the U.S. is broken up in roughly the same way.) TRACONs are commonly located on the grounds of the airport for which they are controlling the arriving and departing aircraft.

Centers are sub-divided into numerous sectors, each of which circumscribes a three-dimensional volume of airspace. One air traffic controller is responsible for the air traffic within each sector, although during busy periods they will often have a “data controller” to assist them. Controllers for different sectors, but within the same Center, sit together in the same room, often within arm’s reach of one another. TRACONs are subdivided into arrival and departure sectors designed to handle particular problems; the arrival and departure routes are structured so as to keep these problems procedurally separated. Moreover, Center airspace is also commonly structured to manage arrival and departure problems into the busiest airport(s) in their Center.

As an example of a typical problem in a Center, consider the following, shown graphically in Fig. 1, where two “streams” of aircraft from separate sectors (B and C) within a Center need to be merged into one stream, and where the merge happens in a separate Center (sector A in Fig. 1). Controllers must, by regulation, keep aircraft in these streams separated by at least 5 nmi horizontally and 1,000 feet vertically. However, in order to keep this spacing after being merged at point B1, controllers would place a restriction of 10 nmi on each of the unmerged streams between B1 and C1 and B1 and C2. This is meant to ensure that the controller in sector A would be able to merge two full streams to obtain one stream with 5 nmi between aircraft.

This restriction would be placed on points C1 and C2, so that the controllers in sectors B and C would need to manipulate the aircraft on this stream to obtain the proper spacing. This manipulation is done manually by slowing one aircraft in comparison to the preceding aircraft; this slowing can be accomplished by reducing the speed of the aircraft or by “path stretching” – increasing the distance the aircraft has to fly to get to C1 or C2 by creating deviations from the direct path.

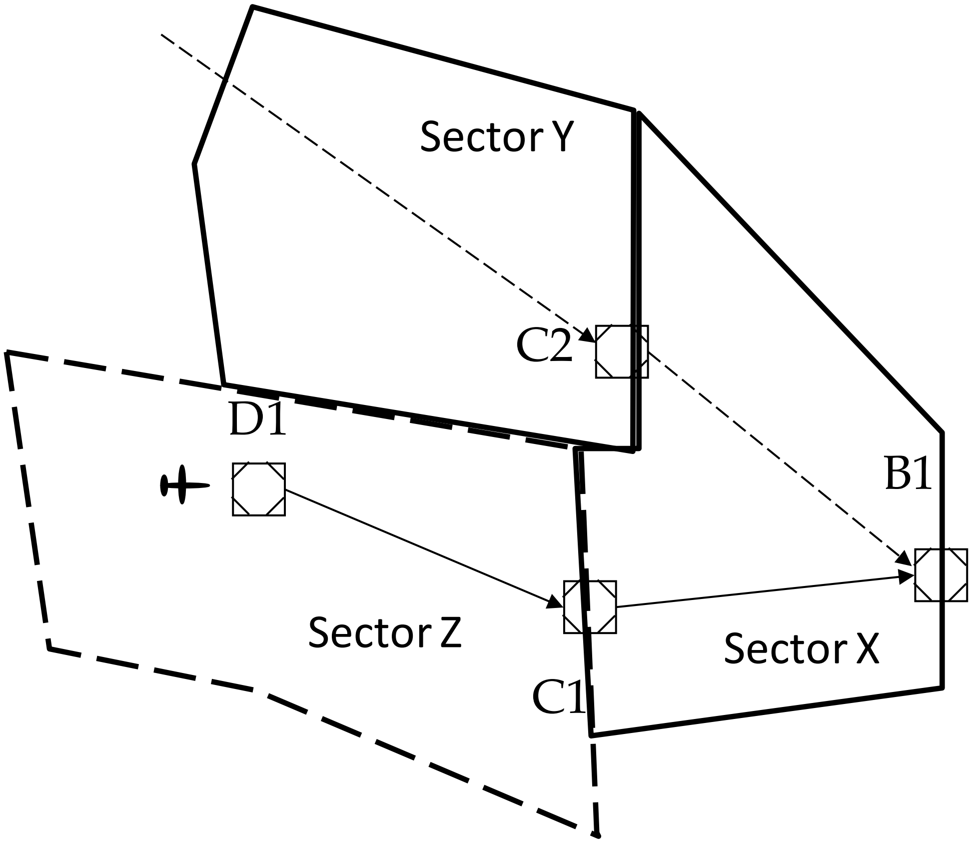


Fig. 1 Merging streams example

These restrictions, however, are static and constructed without much knowledge of actual demand. (Typically, they are based on "historically validated restrictions" applied after analyzing gross demand.) If little or no demand materializes on the stream in sector C, the 10 nmi spacing in sector B is excessive or even unnecessary. The busier stream (in sector B) is inordinately and unnecessarily delayed. This results in extra fuel burned by the aircraft and under-utilized air traffic capacity. Such methods do not explicitly require scheduling since controllers are simply keeping aircraft separated by a preset distance.

Time-based metering, which does require scheduling, is an alternative (or supplement) to the traditional methods of managing air traffic, and attempts to address several of these sources of inefficiency: coordination at a regional and national level, connecting schedules to actual demand, and allocating demand equitably across streams. Time-based metering is a "Lagrangian" approach, where control is applied at an aircraft level as opposed to an Eulerian approach that would apply control at the trajectory level, by assigning meet times over defined points along each aircraft's route of flight. Time-based metering has been shown to be superior to Eulerian methods of managing air traffic theoretically (Sokkappa, 1989), through modeling (Idris, et al., 2004; Moreau & Roy, 2005), and in practice (Knorr, 2003; Mann, Stevenson, Futato, & McMillan, 2002). Its disadvantage is that it requires a great deal of communication and some centralized coordination.

The barriers to implementing a practical and efficient time-based metering system are extremely high. It is necessary but not sufficient for a scheduling system to demonstrate improvements in throughput, delay, or other measures of efficiency. Such a system must also be widely deployable, must work in real time, must be compatible with the technology used within the air traffic control infrastructure, must have no adverse effect on controller workload, and must present implementable solutions. Overall, the case must be made that the benefits of the system exceed its costs. It is not strictly necessary that the solutions presented by such a system be "optimal" with respect to throughput or delay.

Current implementations of time-based metering, such as TMA, utilize a first-come, first-served (FCFS) approach, although TMA utilizes a two-stage coordinated schedule between "meter fixes" located at the entry points into the airport's terminal airspace and the runway (Wong, 2000). Although this approach is known to be less than optimal in terms of throughput, it has nearly universal agreement that it is at least fair to all parties. The use of other objective functions, such as minimizing delay per aircraft, maximizing throughput, and minimizing delay per passenger may be unfair to some particular class of aircraft operator, although some recent work has attempted to utilize constrained position shifting to address fairness (Balakrishnan & Chandran, 2006; Saraf & Slater, 2008).

2.3 A discussion of aspects of the complexity of this problem

There are a number of aspects of the air traffic problem that are atypical in comparison to more conventional scheduling problems. For one, there are constraints on both the minimum and maximum time between resources due to operational speed restrictions on aircraft and its fuel quantity. The positions of the aircraft are continuous and not discrete, as are the dynamics of the aircraft, and there are minimum separation constraints between any two aircraft. Most importantly, delaying aircraft between resources is a manual process and the aircraft cannot be stopped. There are also restrictions on the number of aircraft in a sector at any given time due to controller workload limitations. Lastly, there are strong precedence constraints within a given sector, reflecting the high workload required to have one aircraft "pass" another aircraft under most conditions.

2.3.1 Uncertainties in arrival times

Another difficulty faced by an operational system is how to deal with uncertainties in estimating arrival times to resources when that prediction is made very far in advance of the expected arrival time. A practical system must be robust with respect to these uncertainties, and the architecture described in this paper was driven in large part by this particular problem. That is, it is insufficient to incorporate uncertainty into the scheduling decisions, but rather an entirely different architecture is necessary to handle the uncertainty.

Specifically, the uncertainties in the airspace system and the complexity of predicting numerous trajectories in real time create significant problems for estimating arrival times over long distances (Paelli & Erzberger, 1997; Slattery & Zhao, 1997; Warren, 2000). Such estimated times of arrival (ETAs) are subject to "errors" that generally increase with distance. These "errors" are really only differences between predicted and actual arrival times, which result from a variety of unpredictable factors, such as wind forecast errors, aircraft performance modeling errors, airline/aircrew procedure modeling errors, and unforeseen air traffic control actions (e.g. controllers altering the course of certain aircraft to avoid separation problems). Many of these long-range trajectory prediction errors are impossible to predict; a scheduling system must be robust to such uncertainties.

A fairly comprehensive study was done on prediction time errors was conducted in 2000 (Heere & Zelenka, 2000). Outside of 200 nmi the accuracy remained approximately constant, but the precision decreased. That is, there were more frequent large differences between actual times of arrival (ATAs) and ETAs as distance increased beyond 200nmi.

One consequence of such differences is that the predicted sequences may be wrong, which can make the resulting schedule infeasible. Due to limitations such as position and speed determination errors, trajectory prediction produces accurate estimates only over about 200 – 250 nmi.

Uncertainty in the estimate of arrival times, using the best possible trajectory estimation techniques, is on the order of 10 minutes at distances of over 300 nmi (Heere & Zelenka, 2000). If we make sequence/schedule decisions at 350 nmi and the schedule is fixed based on this information, aircraft may be unable to meet the scheduled times, causing disruptions.

In addition, because we have no knowledge of the true open-loop behavior of the system, sequence optimization based on deterministic trajectory models can lead to suboptimal schedules. Furthermore, it may not be evident that a suboptimal schedule has been selected or implemented. Dynamic scheduling methods (Beasley, Krishnamoorthy, Sharaiha, & Abramson, 2004; Ouelhadj & Petrovic, 2009) are not designed to handle this problem, where errors are manifest slowly over time, but are rather focused on sudden disruptions (such as additions, deletions, or shifts in the ETAs of aircraft).

Such uncertainties mean that schedules produced over long ranges may be inaccurate or infeasible. Methods based on releasing aircraft from ground holding typically assume that the unimpeded flight times are deterministic. What seems to be needed instead is an adaptive approach, where schedules can be adjusted as the uncertainties are resolved. An alternative is to use a stochastic approach, where the uncertainties are modeled into a distribution and decisions made to control the probability of particular undesirable conditions from occurring (Ren & Clarke, 2007).

2.3.2 Holding stacks and ground holds

The significant use of airborne holding stacks, where aircraft fly in racetrack patterns at an assigned altitude to absorb large delay, is extremely disruptive and therefore undesirable. Such stacks take up airspace in a sector and require monitoring by air traffic controllers. Large holding stacks require significant volumes of airspace and can completely occupy a sector. In such cases, no aircraft can be allowed to enter the sector, causing ripple-effect holding upstream. In extreme cases, holding stacks can result in gridlock where outbound aircraft cannot takeoff from airports due to sector congestion and inbound aircraft cannot land due to ground congestion caused by the inability of the outbound aircraft to depart.

In addition, airborne holding patterns allow only somewhat large, discrete intervals of delay. That is, one cannot impart 30 seconds of delay in a holding pattern, because the turn rates of aircraft do not allow it. Typically, holding patterns are no shorter than a racetrack with 5nmi straight-aways (legs) and 180° turns, but are more frequently found with 10 nmi legs. Each turn takes no less that 60 seconds (180° at 3° per second), and each leg takes approximately 1 minute at 5 nmi/minute. This means the minimum delay is 4 minutes, and can only be imparted (roughly) in multiples of 4 minutes.

Ground holding areas, required so that aircraft awaiting departure do not block gates needed for arriving aircraft, are less disruptive and more fuel efficient, but such space for ground holding is highly variable by airport (Atkin, Burke, Greenwood, & Reeson, 2008). Moreover, reserve requirements dictate the minimum amount of fuel an aircraft must carry for a given flight. Significant indeterminate ground holding with engines running requires fuel beyond the minimum required for the flight and represents a substantial expense if ground delay is significantly different than expected. In addition, when snow or other freezing precipitation is present, ground hold times are limited by de-icing time constraints.

3 Scheduling problem and algorithm

3.1 Problem and solution overview

Consider a network of resources (A1-A4, B1, C1 and C2, and D1), such as shown in Fig. 3. Traditionally, scheduling problems have been confined to the airport and sometimes the arrival fixes into the airspace surrounding the airport (A1 – A4). The arrival fixes mark the entry into the TRACON, where the aircraft are under the control of specialized arrival controllers. Outside of this area, the aircraft are managed by air traffic controllers in one or more Centers.

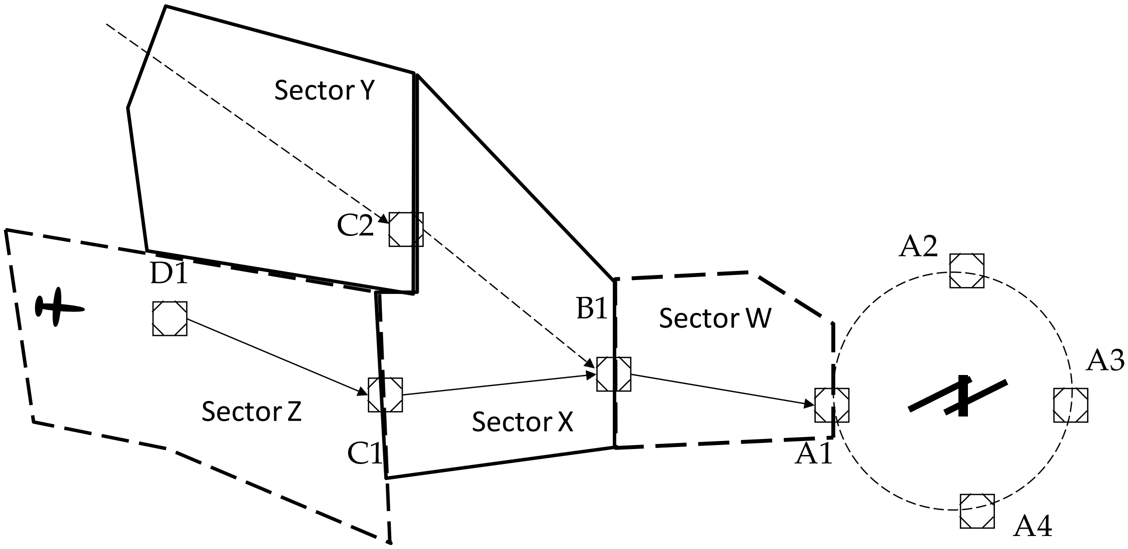


Fig. 3 Example scheduling topology

Centers are divided into sectors, such as sectors A, B, C, and T1 shown in Fig. 3. Each scheduler is designed to operate (roughly) within one Center sector, and can be considered as creating schedules for the exit point from the sector. More specifically, each scheduler creates a local schedule to one resource, and there is no limit to how many schedulers can be running in one sector or Center.

An algorithm is also run for inside the TRACON, although controllers inside the TRACON are not expected to adhere to that schedule – the system is designed to only present a manageable problem to the TRACON controllers, who are then free to control the aircraft as necessary. (Presenting unmanageable problems to TRACON controllers can result in short- or no-notice holding for the controllers in the Center sectors adjacent to the TRACON such as sector T1, which, as mentioned, is extremely disruptive.)

Given the ETAs of each aircraft to each resource, the minimum requirement is that the overall scheduling system generates deconflicted required times of arrival (RTAs) that meet all FAA separation constraints while presenting manageable problems to the controller. “Manageable” means, in part, that the scheduler assigns no more delay to be absorbed in a sector than that sector is capable of absorbing. This allowable maximum delay time (AMDT) is sector-specific and depends on the spatial characteristics/constraints of the sector and the complexity of the traffic flow. Most likely, it is dynamic, but for the purposes of simplicity it is treated as a static value defined by air traffic experts. In addition, the schedule should be dynamic, robust to disturbances, produce feasible sequences of aircraft within each sector, and not rely strictly on the most distant and least accurate estimates of arrival times.

If the distance from D1 to A1 is several hundred nautical miles, the predictions of arrival times at A1 will be poor. (Predictions to the runway will be even worse.) Predictions of arrival times at B1, C1, and D1, for example, get increasingly more accurate. As such, schedules constructed at the airport when the aircraft is beyond D1 utilize the least reliable arrival time of any available. An alternative is to schedule over short ranges (where the arrival time predictions are accurate) in such a way that downstream capacity is not exceeded. For this reason, the algorithm is focused on scheduling a particular problem within each sector, then linking to other instances of the algorithm by providing information on the capacity of the scheduler to produce feasible schedules.

Each resource is assigned a scheduler that operates independently from every other scheduler, while accepting capacity information from other schedulers. (For this example, there would therefore be several schedulers running for the network shown in Fig. 3 – one scheduling into A1, including the airport, one scheduling for B1, and one each scheduling for C1 and C2.) Each scheduler assigns RTAs for each aircraft to its resource, referred to as a "meter point," which are typically located at either exit points from the sector and/or merge points.

An example architecture is shown in Fig. 4. In Fig. 4, ETAs are computed by NASA's Center-TRACON Automation System (CTAS) using wind information from the U.S. National Weather Service Rapid Update Cycle system and input from the "HOST" computer system used by the FAA to process radar track information and flight plans and present that information to air traffic controllers. The schedulers subscribe to pertinent ETAs and capacity information ("rate profiles"), compute schedules, then send RTAs to the Center's mainframe computer and publish capacity information for other schedulers. Schedulers do not subscribe to RTA information from other schedulers.

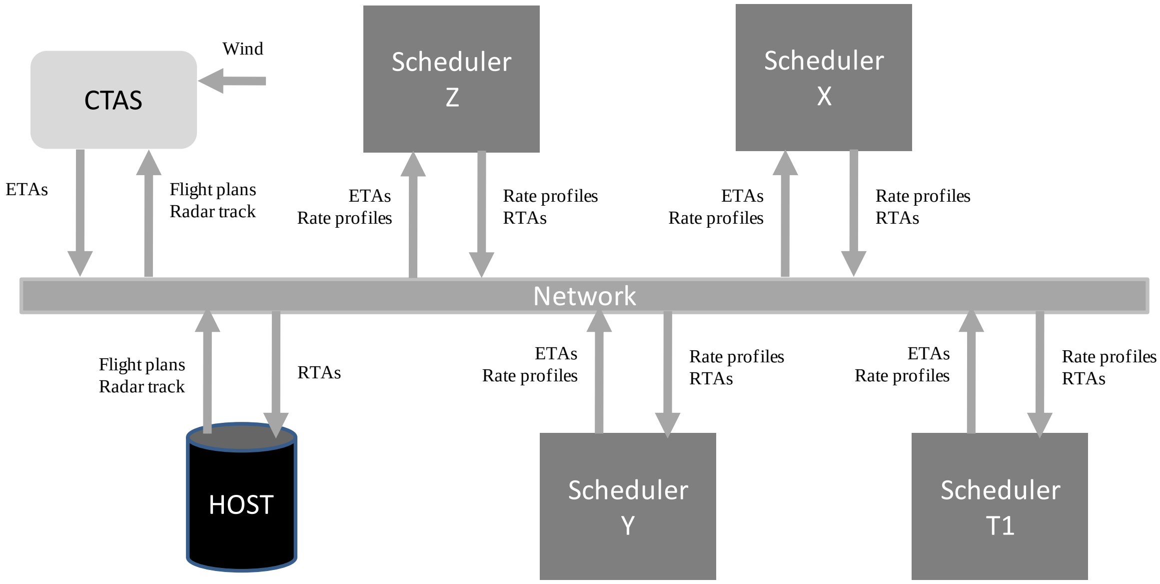


Fig. 4. Example scheduler architecture.

In addition to meeting local constraints, such as FAA requirements for lateral separation, each scheduler considers capacity information sent to it from other schedulers. The scheduler uses this information to construct a schedule that, in addition to keeping aircraft properly separated within the sector, also ensures that the problem it passes on to the next sector can be solved without exceeding the sector’s AMDT.

The sending of capacity information, instead of schedule information, is critical. If schedule information were sent, the most constrained resource, typically the runway, would drive sequence decisions. As mentioned, however, estimates of arrival time at the runway are the least accurate. As such, such sequences are liable to be infeasible far from the runway where the aircraft is currently being controlled. By passing capacity information only, the upstream schedulers can assign sequences based on the short-range schedule, which has accurate estimated times of arrival to its meter point, while ensuring only that downstream capacity is not exceeded.

3.2 Rate profiles

The coupling between schedulers in a system where short-range schedulers must also ensure that downstream capacity is not exceeded can be accomplished by exchanging capacity information in the form of “rate profiles.” Each rate profile consists of a time sequence of short-term (on the order of one minute) rate limitations placed on a meter point by a destination resource. Rate profiles define whether one resource can send an aircraft to another resource such that the aircraft arrives at a particular time; if capacity does not exist in the desired time period for that particular aircraft, then it must be delayed so that it arrives in a time period which does have capacity.

For example, a scheduler schedules traffic into meter point B1. After doing so, it calculates the capacity it has “set aside,” i.e. scheduled, for aircraft from C1 and from C2, considering both its local constraints and the capacity available to it at A1 and the runway. It then transmits information to C1 and C2 indicating that capacity. When the scheduler at C1 schedules its traffic, it ensures that the resulting schedule does not exceed the capacity available to C1 at B1. C1 also obtains such capacity information from A1 and ensures it does not exceed capacity at A1 or the runway as well.

Capacity information is computed and recorded in rate profiles as follows. At each meter point, arrival fix, and airport, a scheduler calculates the capacity that has either been assigned, or is unused and therefore available, to each upstream point that is supplying aircraft to it. Assigned capacity can be determined once a schedule has been created. Available capacity, where capacity exists but is not being used, is identified as available for use to all upstream points.

Schedulers recalculate rate profiles each time a new schedule is created, such as when a new set of radar track data is received. As new aircraft enter the system and are assigned to available time slots, or as arrival time estimates change, those previously-available time slots are subsequently marked as assigned to a particular upstream resource and are no longer identified as available in the next set of transmitted rate profiles.

Different scheduling entities in a network of schedulers interact with each other only through the use of rate profiles. Rate profiles are capacity information, generated by the scheduler at one meter point and indicating the capacity available at that meter point for aircraft coming from some other meter point. Each scheduler generates a rate profile for every other scheduler that is providing aircraft to it.

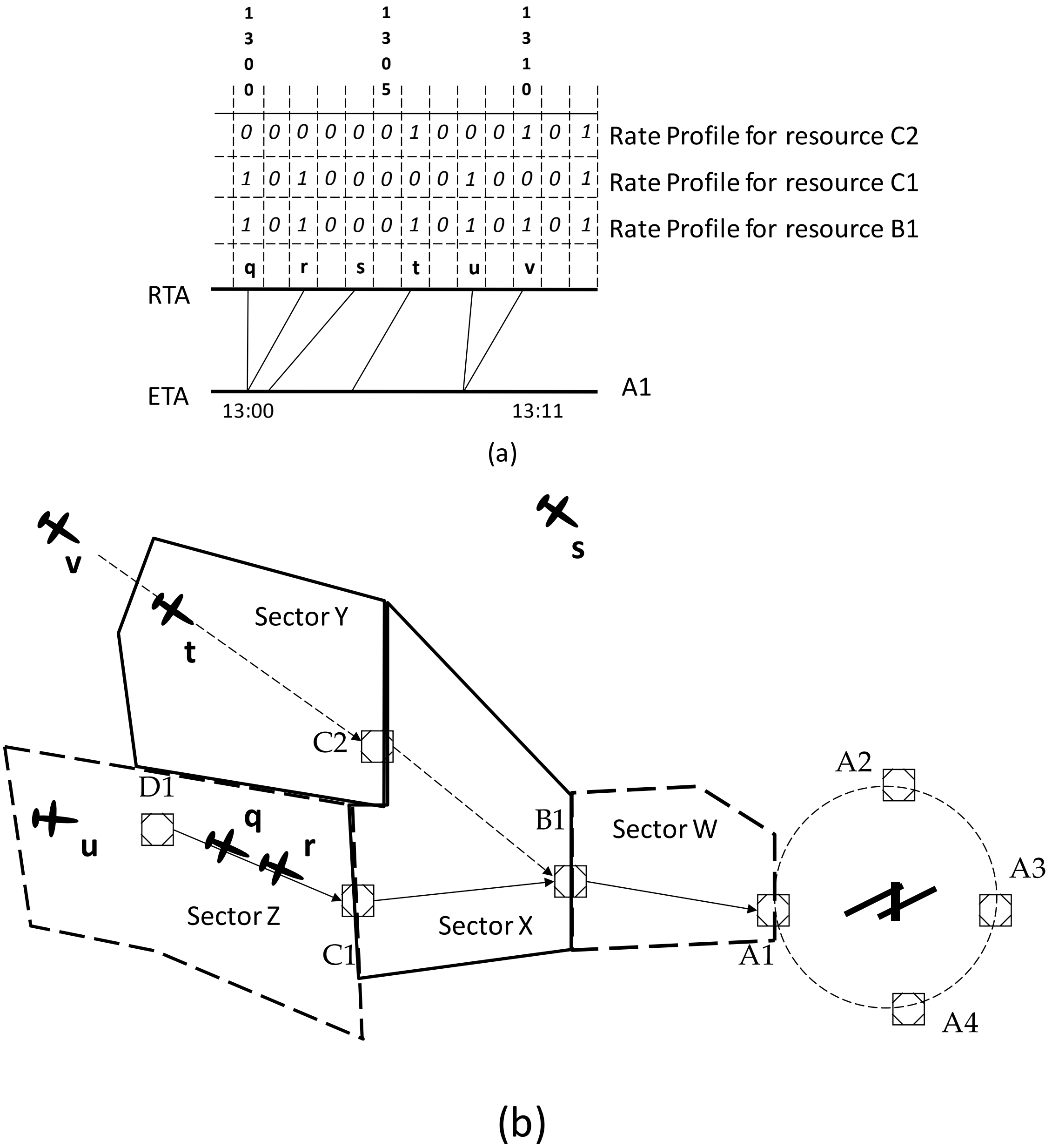
An example of a rate profile is shown in Fig. 5. In Fig. 5a, the scheduler at A1 has scheduled aircraft *q* through *v*, which are arriving from the locations indicated in Fig. 5b.

Aircraft *q* and *r* are grouped together, both expecting to arrive at 13:00. The scheduler at the runway spaces them out according to local constraints, as shown on the RTA line. (In this simplified example, lateral separation requirements are translated into a constant 2 minutes between aircraft; in practice the distance-based separations translate into variable time-based distances based on the ground speeds of the aircraft.)

An identification of possible positions for the aircraft is shown in Fig. 5b. Aircraft *q*, *r*, and *u* are along one stream from C1 to B1, and aircraft *t* and *v* are along one stream from C2 to B1; these streams join to form the stream from B1 to A1 and the airport. The position of aircraft *s* is notional; from the rate profile we only know it is not arriving through B1, C1, or C2.

Note that the sequence of *q* and *r* at A1 is reversed from their physical sequence between D1 and C1. A sequence change is generally impractical, so the sequence at A1 should reflect the sequenced at C1. If one monolithic schedule were created, such that the sequences were driven by the sequences at A1 (or the airport), the scheduler would enforce the sequence at A1 upstream, presenting an infeasible sequence to the controller in sector C. However, because the scheduler at C1 does not use the schedule, but instead only capacity information, it will be free to set the sequence according to its local constraints. As the aircraft get closer to A1, their estimates of arrival time will improve and the sequence at A1 will reverse to reflect the proper sequence. (The ability to accomplish this comes from another feature of the scheduling system, called the “rolling freeze,” which will be discussed in the next section.)

To compute the rate profiles, the RTA line is broken into one-minute bins, or “slots,” and the number of aircraft in the bin from a given resource is identified. (These bins are representative of the actual data structure used to hold rate profiles in the software.) For example, aircraft *q* is scheduled to arrive at 13:00, and its route is C1-B1-A1. Therefore, one of the slots is allocated to B1 in bin 1300, and a slot is allocated to C1 in bin 1300. No slot is allocated to C2. Aircraft *t* is arriving over (for example) C2, and so no slot is allocated to C1. After the arrival of aircraft *v* there is a gap. This capacity, at 13:12, is shown as being available to all points, indicating that it is available capacity rather than allocated capacity.



(a)

(b)

Fig. 5 The rate profile for a sample problem (a) and a geographic depiction of possible locations of the aircraft corresponding to that rate profile (b).

3.3 Rolling freeze

Controllers, each of whom manages traffic in one sector, are presented with a "frozen" schedule and are responsible for imparting the assigned delay to each aircraft. It is necessary to freeze the schedule so that the controller has a fixed target RTA to control the aircraft to meet. This “freeze horizon” is similar to a “commitment time horizon” (Bidot, Vidal, Laborie, & Beck, 2009) with which many readers may be more familiar.

However, an aircraft is typically frozen with respect to only one scheduler, which is the scheduler for the next meter point the aircraft will pass. As the aircraft is being controlled to meet the RTA to its next meter point, ETAs to meter points downstream of the next meter point may change. Since the schedulers for these downstream meter points are not frozen, these changing ETAs may induce schedule changes at downstream meter points. As the aircraft passes its next meter point, it (most likely) becomes frozen with respect to the subsequent meter point in its route. So as an aircraft traverses its route of flight, its RTA becomes fixed to the next, and only the next, meter point. This effect is known as the “rolling freeze” effect and is critical to the operation of the scheduling system.

Referring to Fig. 5, for example, aircraft *q*, *r*, and *u* would have frozen RTAs to C1, where the controller for sector C1 is actively imparting delay so that those aircraft can meet those RTAs. However, the RTAs for those aircraft to B1 and A1 are not frozen. As *q*, *r*, and *u* are delayed to meet their fixed RTAs to C1, the ETAs of those aircraft to B1 and A1 are being affected as well. Since the RTAs for those aircraft at A1 and B1 are not frozen, those RTAs may change as a result of changes to the estimated arrival times.

Recall that *q* and *r* have an improper sequence at A1 as compared to C1. The RTAs at A1 are not frozen until the aircraft enter sector T1. Therefore, the sequence at A1 is not set until the aircraft are very close to A1, where the estimated arrival times are more accurate.

Moreover, suppose B1 were highly congested. In that case, aircraft *q* and *r* would have significant delay at B1. If the sequence were set based on the sequence at A1, aircraft *s* would get “stuck” behind those aircraft, due to the delay at B1, which should not affect *s*. However, the rolling freeze means that *s* could be allowed to pass *q* and *r*, since its RTA at A1 is not frozen, assuming available, and attainable, capacity in front of *q* and *r* is available at A1. In this way, the distributed nature of the scheduler, and the rolling freeze effect, combine to better utilize available capacity.

3.4 Scheduler algorithm and implementation

To accomplish the scheduling task, the distributed scheduler (DS) was developed. For operational details of the system in which the DS was implemented, see Landry (Landry, 2008).

Scheduling starts with the generation of unimpeded ETAs for every aircraft over a set of points along its route of flight, including the resources for which schedules are being built. Such resources are currently defined in advance, but could also be dynamically defined. In the operational system, ETAs are generated by the trajectory synthesis module of CTAS, which uses algorithms that are similar to those used in the flight management computers onboard commercial aircraft (Denery & Erzberger, 1995; Heere & Zelenka, 2000). (The modeling of trajectories within the TRACON is greatly simplified, however.)

ETAs are recalculated when any one of a number of events occur, including radar track updates (approximately every 12 seconds in enroute airspace), flight plan updates, or the receipt of departure messages. This means that new information is generated at least every 12 seconds, requiring schedules to be recalculated with at least the same frequency. In practice, it is necessary for the algorithm to be able to recalculate schedules within a few seconds due to the asynchronous arrival of rescheduling event triggers.

CTAS, a suite of air traffic control automation software, creates estimates of future four-dimensional (three spatial dimensions plus time) trajectories for aircraft. To accomplish this, CTAS utilizes aircraft equations of motion, aircraft type performance characteristics, wind data, flight plan information, assigned altitude information, radar track information, models of climb and descent procedures, and knowledge of local constraint and expectation information. CTAS uses forward and reverse integration of these equations of motion subject to all these constraints to predict ETAs at numerous points along each aircraft’s route of flight.

3.5 Algorithm

Initially, the RTA of each aircraft is set equal to its ETA; an aircraft cannot be scheduled to arrive earlier than its ETA (i.e. it cannot be asked to speed up). This constraint differs from some previous work, and probably should change, but reflected the observation and comments of controllers that they do not typically ask pilots to increase speed. Controllers can assign speed reductions, but otherwise will only have pilots fly their "expected" speed as identified in their flight plan.

Depending on the particular phase of flight, trajectory modeling uses either actual groundspeed, a modeled groundspeed, or the expected groundspeed. Which groundspeed to use has significant implications for the operation of the system, but it requires detailed treatment that is beyond the scope of this paper.

The RTAs are then revised in chronological order using a FCFS method. (There are a few exceptions to the FCFS rule, such as medical evacuation or low fuel aircraft.) The next aircraft cannot arrive prior to the first aircraft's RTA (plus constraints), and so on. These constraints include longitudinal separation (miles-in-trail or MiT) that reflect procedurally-mandated minimum separation, restrictions on arrival rates (acceptance rates) such as no more than 10 aircraft per 10 minutes, and blocked intervals (periods in which no aircraft are allowed to arrive). In the operational system, any or all of these constraints can be set by a facility traffic manager.

Except for the generation of rate profiles to be used at points upstream of the TRACON, scheduling inside the TRACON is accomplished by the DP. Upstream points in the DS are scheduled by a process called the "single point DP" (SPDP). The SPDP differs from the DP in that it produces a schedule for only one resource, and it must comply with downstream rate profiles. Specifically, for an aircraft i (sorted by ETA) in a stream for meter point q that also has a total of n resources along its route of flight, the algorithm can be approximated with the following pseudo-code:

```
$x_i^{q,*} = x_i^q$
if $x_{i-1}^{q,*} + \frac{MIT_q}{\text{Max}(S_i, S_{i-1})} > x_i^{q,*}$ then $x_i^{q,*} = x_{i-1}^{q,*} + \frac{MIT_q}{\text{Max}(S_i, S_{i-1})}$
if $x_{i+1}^{q,*} - \frac{MIT_q}{\text{Max}(S_i, S_{i+1})} < x_i^{q,*}$ then $x_i^{q,*} = x_{i+1}^{q,*} + \frac{MIT_q}{\text{Max}(S_i, S_{i-1})}$
for $r = q + 1$ to $n$
        if $x_i^q + RP\_Delay_{q,r}(x_i^q + (x_i^r - x_i^q) - AMDT_{q,r}) > x_i^{q,*}$
        then $x_i^{q,*} = x_i^q + RP\_Delay_{q,r}(x_i^q + (x_i^r - x_i^q) - AMDT_{q,r})$
next $r$
Do
        if $AR\_DELAY_q(x_i^{q,*}) > 0$ then $x_i^{q,*} = x_i^{q,*} + AR\_DELAY_q(x_i^{q,*})$
        if $BI\_DELAY_q(x_i^{q,*}) > 0$ then $x_i^{q,*} = x_i^{q,*} + BI\_DELAY_q(x_i^{q,*})$
until $AR\_DELAY_q(x_i^{q,*}) = 0$ and $BI\_DELAY_q(x_i^{q,*}) = 0$
```

where:

$x_i^{q,*} =$ the RTA of aircraft i in stream for meter point q

$x_i^q =$ the unimpeded ETA of aircraft i in stream for meter point q

$MIT_q =$ the miles-in-trail separation requirement at meter point q

$S_i =$ the groundspeed of aircraft i

$RP\_Delay_{q,r} =$ the delay required at meter point q due to rate profile constraints at resource r

$x_i^r - x_i^q =$ the estimated time enroute of aircraft i from meter point q to resource r, where $r > q$

$AR\_Delay_q =$ the delay required at meter point q due to acceptance rate constraints.

$BI\_Delay_q =$ the delay required at meter point q due to blocked interval constraints

$AMDT_{q,r} =$ the allowable maximum delay time between meter point q and resource r

The first line sets the RTA of aircraft i to its ETA, reflecting the constraint that the RTA cannot be earlier than the ETA. The second and third lines ensure that the RTA does not violate the MiT constraint on the stream at the meter point, either with the preceding aircraft or the succeeding aircraft (if it was scheduled first due to its being a priority aircraft). The for-next loop ensures that the rate profile constraints are not violated after considering the AMDT of the intervening sectors. Finally, the do-until loop ensures that the resulting time does not violate acceptance rate or blocked interval constraints.

(Acceptance rate constraints are constraints on maximum flow per time period, and blocked intervals are periods for which the acceptance rate is zero.)

3.5.1 Program flow

Although the sequence of scheduling events can be somewhat asynchronous without affecting performance to a large degree, it is easier to understand as an ordered sequence of events. First, the A1 scheduler schedules the runway and arrival fixes in a coordinated, centralized fashion. The scheduler at A1 then generates rate profiles for all upstream meter points that are delivering aircraft to it. These rate profiles reflect capacity at the runway and arrival fixes available to be used by the particular upstream resource.

A simple example of the output of the scheduler at A1 is shown in Fig. 5a above, where the schedule at A1 has been completed and the rate profiles for B1, C1, and C2 have been computed. The rate profile for B1 is then sent to that scheduler using a publish-subscribe mechanism. C1 and C2 obtain their rate profiles in the same way.

In practice the rate profile calculations must consider the continuous range of time, rather than discrete bins. In the example, if a one-minute bin has no aircraft in it, and if the bin on either side of it has no aircraft in it, it is considered available. In practice the gap between each pair of aircraft would have to be evaluated to determine whether any capacity exists.

This computation is fairly straightforward, and involves examining the size of the gap between aircraft. Pseudo-code for the computation is shown below, where a gap between two aircraft is converted into a number of “holes,” where one hole means one aircraft can be scheduled in the gap. Any one-minute bin that contains a non-zero number of holes has available capacity.

$$\underline{g} = x_i^{q,*} + \frac{MIT_q}{\max \;\; S_i, S_{i+1}}$$

$$\overline{g} = x_{i+1}^{q,*} - \frac{MIT_q}{\max \;\; S_i, S_{i+1}}$$

$$if \;\; \underline{g} > \overline{g} \;\; then \;\; h = 0 \;\; else$$

$$h = \frac{\underline{g} > \overline{g}}{MIT_q \big/ \max \;\; S_i, S_{i+1}}$$

where:

$\underline{g} =$ start of available gap

$\overline{g} =$ end of available gap

$h =$ number of holes in gap

An example of the computation, transfer and use of the rate profiles is shown in Fig. 6. The rate profile shown in Fig. 5a calculated by the scheduler at A1 for B1 is published to the network, and subscribed to by the scheduler at B1. In scheduling its aircraft, it ensures that not only are local constraints met, but that capacity is available for the aircraft downstream at A1 given their RTA at B1. For example, although local constraints would not require aircraft *t* to be delayed at B1, it is being delayed an additional two minutes so as to conform to the rate profile at A1.

Note that the sequence for aircraft *r* and *q* are different at B1 and A1, as discussed previously. The flexibility of passing a rate profile instead of a time allows the scheduler at B1 to allocate *r* to the earlier slot and delay *q*, rather than the other way around, which would yield an infeasible solution at B1.

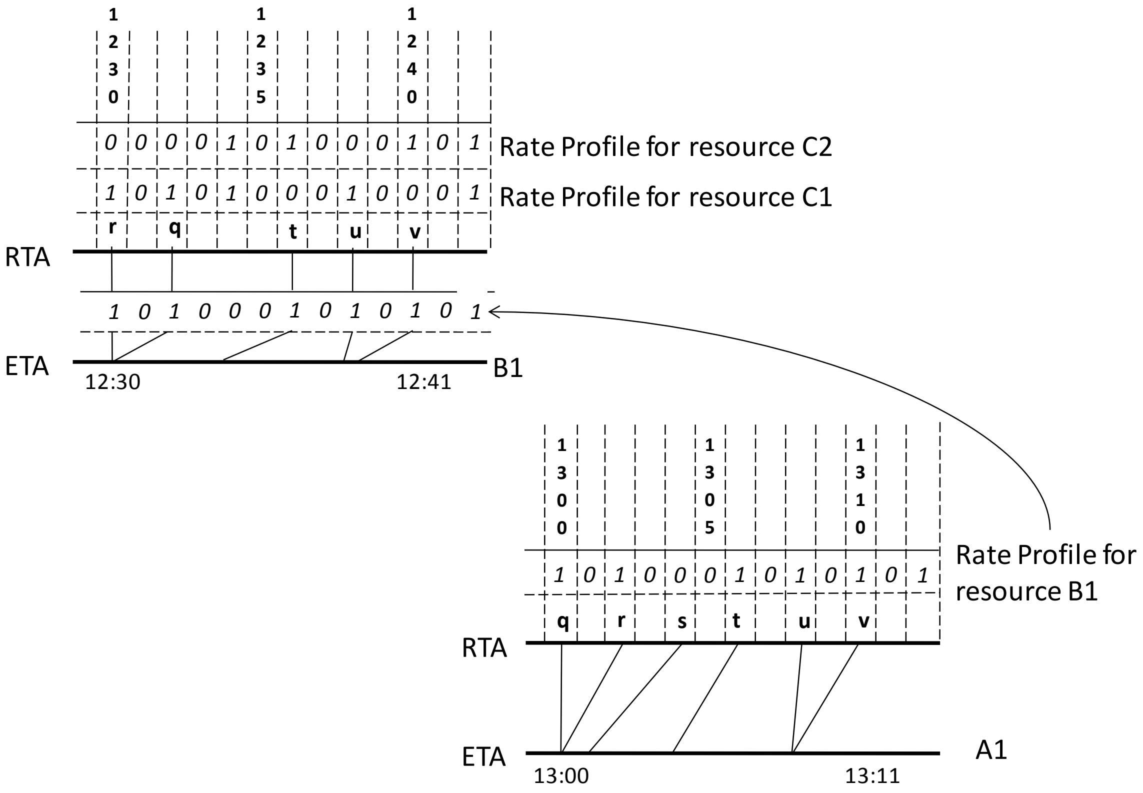


Fig. 6 Scheduling for resource B1 using rate profile from A1.

4 System test and results

The system described was implemented within the Multi-Center Traffic Management Advisor tool and deployed to FAA facilities for testing. The system was configured to schedule arrivals into Philadelphia International Airport (PHL). This problem requires coordination between at least four Centers in addition to the PHL TRACON. Those Centers are Boston Center (ZBW), New York Center (ZNY), Cleveland Center (ZOB), and Washington Center (ZDC). The range over which the scheduler functioned was approximately 400 nmi. Additional information on the configuration for the test and the operational environment can be found in Landry (Landry, 2008). An example of the output of the scheduler is shown in Fig. 7. (The left-hand side of each timeline shows the estimated arrival times and the right-hand side shows the scheduled times of arrival, with the required delay in minutes shown next to the aircraft identification.)

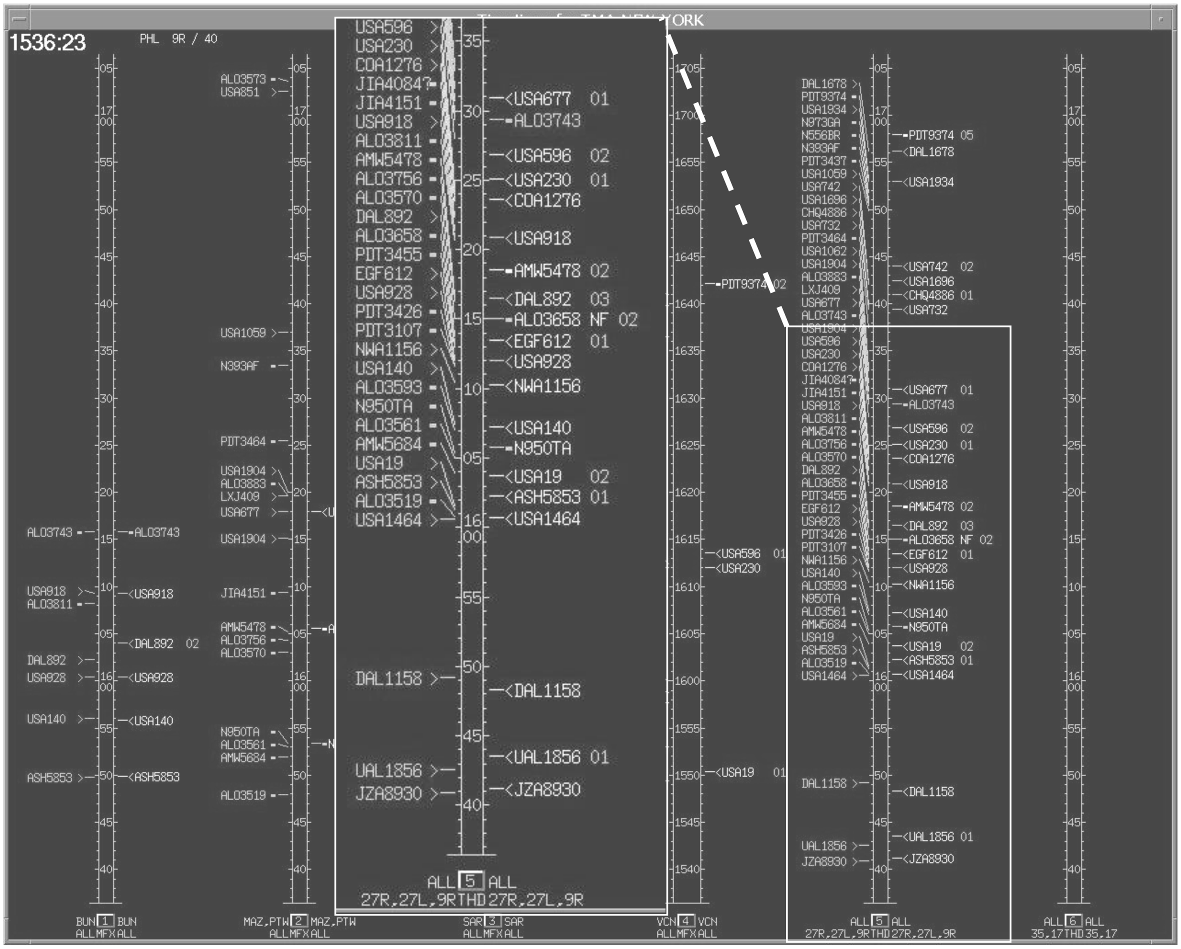


Fig. 7. Scheduler output during operational tests.

The field test program consisted of twenty-eight separate events, or “field trials.” These field trials were conducted between January 2003 and November 2004 to evaluate system performance and refine the operational concept as the prototype matured. The primary objective of the field test program was to satisfy the transition criteria for FAA/NASA Technology Readiness Level 6, effectively establishing the technical case, the operational case, and the business case for transfer of the technology from NASA to the

FAA for deployment. These tests are documented more completely in other publications (Farley et al., 2003; Farley et al., 2005).

The field trials culminated in an actual operational test, where the system was used to control aircraft departure times. Actual airborne metering was not accomplished during the operational test due to the high cost of training the large number of controllers that would be needed to conduct such a test. Departure metering required only the traffic management personnel.

Due to the complexity of the overall system, DS algorithm execution times were not recorded. However, the DS must compute and display scheduling solutions within 10 seconds to avoid a backlog of scheduling requests. Moreover, any delays of more than about 3 seconds would be noticeable to system operators. No such delays or backlogs were recorded. Certain previous versions of the software had iterative checks that did result in such delays, and the iterative checks were removed for that reason.

Departure metering accomplished by using the DS caused no problems that were apparent in the FAA's Aviation System Performance Metrics data, the logs of the sessions made by the NASA field personnel, nor the anecdotal comments of air traffic control personnel gathered in the debrief sessions. As shown in Fig. 8, the periods during the operational trials had, on average, 30% lower average airborne delay at PHL as compared to the average airborne delay in each of six control periods that experienced nearly identical weather and demand conditions. (This difference has a confidence interval of 0.744 to 4.736 minutes per flight.) This low airborne delay was accompanied by no significant penalty in average gate arrival delay during the system evaluation periods. (The average difference was 26% less gate arrival delay during McTMA periods, but this difference was not statistically significant, with a confidence interval of -3.64 to 7.60 minutes.) The FAA has corroborated these results, including comparisons with similar tests at Houston International Airport (Office of Operations Planning, 2005).

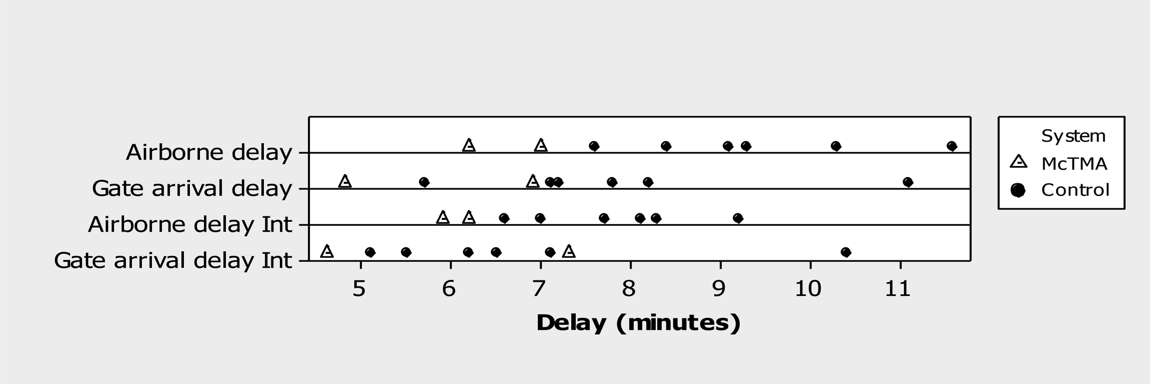


Figure 8. Dotplot of airborne and gate arrival delay for all aircraft and internal departures only (Int) for McTMA test periods and control periods.

Airborne holding in the Center sectors surrounding the TRACON and within the TRACON was reduced 90% as compared to the control period, and traffic flow in the Center airspace was more predictable, orderly, and efficient during the evaluation period. It is especially noteworthy that such favorable results were observed merely with departure metering and no airborne metering. Airborne metering will close the distributed scheduling control loop and therefore is expected to produce even more significant delay savings. There are many variables that influence air traffic efficiency, and only a limited number of test periods were conducted. Because of this, one cannot conclusively state that the system was solely responsible for the positive results, although there was strong similarity between the control periods and the test periods. The most likely explanation is that the McTMA system was primarily responsible, but additional tests and simulations should be conducted to validate these findings.

The system's departure advisories were generally more conservative than the clearances that traffic managers (TMCs) reported they would have issued without the system. That is, the system advised as much or more ground delay. This was not unexpected, because the system considers additional restrictions beyond those with which the TMCs must ordinarily comply. Whereas TMCs usually consider only their internal restrictions, the system also considers the effect of the proposed departure on congestion downstream at other resources and at the runway. The observed data, where significantly less airborne delay and holding was seen along with little or no increase in average ground delay, support this interpretation.

Based on the results, the FAA has incorporated the ability to schedule to individual resources from the system into its En route Automation Modernization System, which has been developed to replace the HOST, currently being deployed. The distributed scheduling algorithm, including the rate profiles, and networking capability of the system, are under consideration for implementation. Such integration would thereby implement all the capabilities of the system, enabling time-based metering scheduling over long distances.

5 Extended use of the system

Additional tests were conducted on the system to test its ability to schedule to multiple airports and for multiple purposes. The conclusiveness of this test was limited by the need for a more detailed implementation of the airports in each of the test Centers. However, the system was set up to meter simultaneously into Boston-Logan and Philadelphia within the Boston Center's system. The system was also configured to meter westbound traffic for Chicago-O'Hare and southbound traffic to Dulles, although schedulers for O'Hare and Dulles were not running during these tests.

In this configuration, Boston Center traffic managers could see schedules built for aircraft inbound to Logan, as well as meter aircraft outbound to Philadelphia, O'Hare, and Dulles. Traffic managers utilized this capability in "shadow" mode, where the system was not driving actual decisions, but was evaluated against the actual decisions made without the use of automation.

While no official results were recorded for these tests, traffic managers expressed approval for the capability. Despite the lack of official testing, the capability to meter departures from a Center has already been incorporated into the deployed TMA system.

6 Discussion – can the system be scaled up?

The successful testing of the distributed scheduler demonstrates that the system is capable of generating feasible schedules over ranges up to 500 nmi, and suggests that it is capable of generating feasible schedules to multiple destinations and for multiple purposes over that range.

Since each scheduler utilizes only capacity information from other schedulers, there is no reason to suspect that it will not scale to much larger problems. The sequence is set based on local information; this ensures that the sequences produced by the scheduler are feasible. Since downstream errors in ETAs generally manifest themselves as sequence differences, the system should scale indefinitely. Additional studies are being undertaken to ensure this scalability exists.

If, as expected, the system is scalable to thousands of nautical miles, and if fully deployed and networked, the DS can enable time-based metering between any departure airport and any arrival airport for all aircraft in the U.S. national airspace system. There is currently no basis to suspect that this cannot occur, and there are no alternatives under consideration for this purpose.

If deployed in this manner, it can be used for several important purposes. First, “ground stops,” where aircraft are held indefinitely, with no definitive departure time, can be significantly reduced or perhaps eliminated by using the system to provide each aircraft a scheduled departure time. Of course, with significantly reduced capacity this time may be operationally equivalent to a ground stop, but would at least eliminate the uncertainty associated with indefinite ground stops, and would be applied individually rather than in blocks. When given scheduled departure times, air carriers can start creating recovery plans from such disruptions earlier and with more accuracy.

Second, the distributed scheduler assigns delay such that no sector along the route of flight of a scheduled aircraft is given a problem that exceeds its capability to incur delay. In routine operations, this would translate into the absence of the need for airborne holding. However, there may still be cases of sudden, unplanned drops in capacity that would require delay on aircraft beyond the capability of the remaining sectors. In such cases airborne holding or extended vectors would need to be used, but the use of the system should significantly reduce unplanned airborne holding.

The distributed scheduler, as incorporated within the overall system, also provides the capability for better regional or NAS-wide decision-making, including improved inter-facility collaboration on traffic problems. With the distributed scheduler and the visualization tools available within the system, traffic managers at the local, regional, and national level will be able to identify sources of congestion and control restrictions that can throttle the air traffic as necessary. Moreover, traffic managers at all levels can see these sources of congestion even if they are not within their own Center, a capability that does not exist now and can lead to better collaborative decisions.

7 Conclusions and future work

The distributed scheduler is capable of producing schedules for aircraft arrivals over distances of up to 500 nmi that are operationally robust to the uncertainty inherent in long-range estimates of arrival times. This is the first algorithm of its kind tested in air traffic control facilities using operational air traffic. As such, the distributed scheduler is a critical development toward constructing the capability needed to meet the vision of NextGen and SESAR.

The algorithm enforces schedules over short ranges at distributed resources, where estimates of arrival time are accurate, while simultaneously ensuring that downstream capacity is not exceeded. This is done through the use of “rate profiles,” which are sets of capacity information targeted for upstream resources. In using rate profiles instead of schedule information, the local schedulers can set sequences based on the most accurate (short range) information rather than the least accurate (long range) information.

The system tests described in this document occurred in 2003-2004. Since that time, the system’s technology has been transferred to the FAA. Incorporation of the technology is planned, but has been delayed as other developments have had priority.

Since the system does not provide any optimization per se, future work should include the incorporation of optimization routines to improve throughput and reduce delay. Such routines will need to operate in real-time, which will be a challenge given the size of the problems and the number of constraints involved. However, the system provides a framework within which optimization can be accomplished while still ensuring feasible solutions are presented to controllers and traffic managers.

References

Andreatta, G., Brunetta, L., & Guastalla, G. (2000). From ground holding to free flight: An exact approach. *Transportation Science, 34*(4), 394-401.

Atkin, J. A. D., Burke, E. K., Greenwood, J. S., & Reeson, D. (2008). On-line decision support for take-off runway scheduling with uncertain taxi times at London Heathrow airport. *Journal of Scheduling, 11*, 323-346.

Balakrishnan, H., & Chandran, B. (2006, 21-24 August). *Scheduling Aircraft Landings under Constrained Position Shifting.* Paper presented at the AIAA Guidance Navigation, and Control Conference, Keystone, CO.

Ball, M., Donohue, G. L., & Hoffman, K. (2006). Auctions for the safe, efficient, and equitable allocation of airspace system resources. In P. Cramton, Y. Shoham & R. Steinberg (Eds.), *Combinatorial Auctions* (pp. 507-538). Cambridge, MA: MIT Press.

Barnhart, C., Belobaba, P., & Odoni, A. (2003). Applications of operations research in the air transport industry. *Transportation Science, 37*(4), 368-391.

Beasley, J. E., Krishnamoorthy, M., Sharaiha, Y. M., & Abramson, D. (2004). Displacement problem and dynamically scheduling aircraft landings. *Journal of the operational research society, 55*(1), 54-64.

Bersimas, D., Lulli, G., & Odoni, A. (2011). An integer optimization approach to large-scale air traffic flow management. *Operations Research, 59*, 211-227.

Bertsimas, D. J., & Stock-Patterson, S. (1998). The air traffic flow management problem with enroute capacities. *Operations Research, 46*(3), 406-422.

Bidot, J., Vidal, T., Laborie, P., & Beck, J. C. (2009). A theoretic and practical framework for scheduling in a stochastic environment. *Journal of Scheduling, 12*, 315-344.

Denery, D. G., & Erzberger, H. (1995). *The Center-TRACON Automation System: Simulation and Field Testing. (Report No. TM-110366).* (No. TM-110366). Moffett Field, CA: National Aeronautics and Space Administration.

Farley, T., Hoang, T., Landry, S., Nickelson, M., Levin, K., & Rowe, D. (2003). *Multi-Center Traffic Management Advisor Initial Field Test Results.* Paper presented at the AIAA 3rd Aviation Technology, Integration and Operations (ATIO) Forum, Denver, CO.

Farley, T., Landry, S., Hoang, T., Nickelson, M., Levin, K., Rowe, D., et al. (2005). *Multi-Center Traffic Management Advisor: Operational Test Results.* Paper presented at the 5th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Arlington, Virginia.

Heere, K. R., & Zelenka, R. E. (2000). *A comparison of Center-TRACON Automation System and airline time of arrival predictions. (Report No. TM-2000-209584).* (No. TM-2000-209584). Moffett Field, CA: National Aeronautics and Space Administration.

Idris, H., Evans, A., Evans, S., & Kozarsky, D. (2004, September 20-22, 2004). *Refined benefits assessment of multi-center traffic management advisor for Philadelphia and New York.* Paper presented at the AIAA 4th Aviation Technology, Integration and Operations Forum, Chicago, IL.

Knorr, D. (2003). *Free Flight Program Performance Metrics to Date: December 2003*. Washington, DC: Federal Aviation Administration.

Landry, S. (2008). The design of a distributed scheduling system for multi-center time-based metering of air traffic into congested resources. *Air Traffic Control Quarterly, 16*(1), 69-97.

Ljungberg, M., & Lucas, A. (1992). *The OASIS air traffic management system: Technical note 28*. Carlton, Australia: Australian Artificial Intelligence Institute.

Mann, J., Stevenson, C., Futato, S., & McMillan, K. (2002). *Introducing time-based metering at Los Angeles Air Route Traffic Control Center*. Paper presented at the 2002 Air Traffic Control Association Annual Conference.

Moreau, D. C., & Roy, S. (2005). *A stochastic characterization of en route traffic flow management strategies.* Paper presented at the AIAA Guidance, Navigation, and Control Conference and Exhibit, San Francisco, CA.

Odoni, A. (1987). The flow management problem in air traffic control. In A. Odoni, L. Bianco & G. P. Szego (Eds.), *Flow Control of Congested Networks* (pp. 269-288). Berlin: Springer-Verlag.

Office of Operations Planning. (2005). *Analysis of Potential TMA-MC Airports*.

Ouelhadj, D., & Petrovic, S. (2009). A survey of dynamic scheduling in manufacturing systems. *Journal of Scheduling, 12*, 417-431.

Paelli, R., & Erzberger, H. (1997). Conflict probability for free flight. *Journal of guidance, control, and dynamics, 20*(3), 597-604.

Ren, L., & Clarke, J.-P. B. (2007). Separation analysis methodology for designing area navigation arrival procedures. *Journal of Guidance, Control, and Dynamics, 5,* 1319-1330.

Richetta, O. (1995). Optimal algorithms and a remarkably efficient heuristic for the ground-holding problem in air traffic control. *Operations Research, 43*(5), 758-770.

Saraf, A., & Slater, G. L. (2008). Optimal dynamic scheduling of aircraft arrivals at congested airports. *Journal of Guidance, Control and Dynamics, 31*(1), 53-65.

Slattery, R., & Zhao, Y. J. (1997). Trajectory synthesis for air traffic automation. *Journal of guidance, control, and dynamics, 20*(2), 232-238.

Sokkappa, B. G. (1989). *Impact of metering methods on airport throughput. (Report No. MP-89W000222)* (No. MP-89W000222). McLean, VA: MITRE Corporation.

Terrab, M., & Odoni, A. (1993). Strategic flow management for air traffic control. *Operations Research, 41*(1), 138-152.

Voelckers, U. (1990, 23-25 May). *Arrival planning and sequencing with COMPAS-OP at the Frankfurt ATC-Center.* Paper presented at the 9th American Control Conference, San Diego, CA.

Vranas, P. B., Bertsimas, D. J., & Odoni, A. (1994). The multi-airport ground-holding problem in air traffic control. *Operations Research, 42*(2), 249-261.

Warren, A. (2000, June, 2000). *Trajectory prediction concepts for next generation air traffic management.* Paper presented at the 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy.

Wong, G. L. (2000). *The dynamic planner: Sequencer, scheduler, and runway allocator for air traffic control automation. (Report No. TM-2000-209586).* (No. TM-2000-209586). Moffett Field, CA: National Aeronautics and Space Administration.

Wu, C., & Caves, R. (2002). Research review of air traffic management. *Transport Reviews, 22*(1), 115-132.