6.43.28.8 CASE STUDY: AIR TRAFFIC MANAGEMENT SYSTEMS[1]

Alexandre M. Bayen, Department of Aeronautics and Astronautics, Stanford University, Stanford CA 94305-4035, USA.
Claire J. Tomlin, Department of Aeronautics and Astronautics, Stanford University, Stanford CA 94305-4035, USA.

Keywords: Air Traffic Management, Hybrid Systems, Safety, Polynomial Algorithms, Scheduling.

Contents

Glossary

Airspace congestion: Traffic pattern potentially leading to a situation in which the number of aircraft in a given area exceeds a threshold fixed by regulations.

Hybrid system: A system which combines continuous-state and discrete-state dynamics, in order to model systems which evolve both continuously and according to discrete jumps.

Lagrangian model of traffic: A model of traffic which tracks the trajectories of each vehicle individually through the domain of interest and accounts for property changes of vehicle variables (speed, acceleration, heading) along the trajectories.

Model validation: Process of verifying on a recorded data set that a mathematical model (abstraction) of a physical system agrees with the data measured.

Polynomial algorithm: An algorithm that has a running time which is a polynomial function of the input size.

[1]This research is supported by DARPA under the Software Enabled Control Program (AFRL contract F33615-99-C-3014), by NASA under Grant NCC 2-5422, and by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research under Grant N00014-02-1-0720.

Summary

The National Airspace System is a large scale, hybrid, dynamic system, with an Air Traffic Control (ATC) authority which is organized hierarchically. In this chapter, certain subsystems within the current ATC system are presented as case studies for analysis and full or partial automation through hybrid control design. A brief history of ATC is first presented; and its organization and structure are described. Then, some case studies of hybrid modeling and control for prototype automation are presented. In particular, a control theoretic model of sector-based air traffic flow using hybrid automata theory is detailed. This model is Lagrangian, meaning that it models the properties of the system along its trajectories. A subset of this model is used to generate analytic predictions of air traffic congestion. A *dynamic sector capacity* which is used to predict the time it takes to overload a given portion of airspace is described. The design and validation of an air traffic flow simulator, used to assess the accuracy of these predictions, is presented. Some existing automation tools are then described, and analysis results based on optimization and game theory for hybrid systems are used to derive results in congestion control, routing and sequencing, and collision avoidance.

1. Introduction

The U.S. *National Airspace System* (NAS) is a large scale, nonlinear dynamic system, with a control authority which is organized hierarchically. A single *Air Traffic Control System Command Center* (ATCSCC), in Herndon VA, supervises the overall traffic flow, and this is supported by 22 (20 in the continental US or CONUS) *Air Route Traffic Control Centers* (ARTCCs, or simply, Centers) organized by geographical region and controlling the airspace up to 60,000 feet [Nolan, 1999, Perry, 1997, Jackson and Green, 1998, Kahne and Frolow, 1996, Gazit, 1996, Pujet and Feron, 1996, Menon et al., 2002]. Each Center is sub-divided into about 20 sectors, with at least one air traffic controller responsible for each sector. Each sector air traffic controller (ATC) may talk to 20-25 aircraft at a given time (the maximum allowed number of aircraft per sector depends on the sector itself). The controller guides the aircraft through the sector using a set of standard commands (over voice channels). In general, the controller has access to the aircraft's flight plan and may revise the altitude and provide temporary heading assignments, amend the route, speed, or profile, in order to attempt to optimize the flow and to keep aircraft *separated*.

The history of U.S. ATC dates back to the 1920's, when congestion and safety issues made an organization of aircraft flow necessary. Since then, the major updates to ATC have been in the addition of new technologies and an overall structure incorporating each geographical region. Throughout its history, it has been critical to maintain ATC as a human-controlled system with exacting levels of safety. Recently, the system is experiencing a new phenomenon: airspace saturation – there are too many aircraft for the NAS and ATC to handle without forcing backups and resulting delays. In addition, new safety challenges have arisen in the past year and a half. One major question that has received much recent research attention is: can the problems of airspace saturation be alleviated by automation or partial automation of some of ATC functionality, while still maintaining, or improving, the levels of safety in the system.

One of the most important, and time consuming, controller tasks is to prevent *losses of separation* (LOS), between aircraft; for high altitude sectors (above 29,000 ft), this means that the controller must keep each pair of aircraft in the sector separated by more than 5 nautical miles (nm) horizontally, and 2000 feet vertically. The terminology *protected zone* is

used to represent the 5 nm radius, 2000 foot height cylinder around an aircraft, that another aircraft must not penetrate. For any pair of aircraft, their relative configuration or *state* (relative position and orientation), is referred to as *unsafe* if there is a *rational* process of actions which leads one aircraft to penetrate the protected zone of another.

This chapter presents some recent research results which could lead to partial automation of ATC functionality. Portions of the NAS and ATC are modeled as a hybrid dynamical system, and analysis results based on optimization and game theory for hybrid systems are used to derive results in congestion control, routing and sequencing, and collision avoidance. For some of these results, the models and methods pertain to the hybrid system verification methodology presented in [TOMLIN et al., 2003] in this volume. First, a brief history of ATC is presented, and its organization is overviewed. The chapter then presents a mathematical model for a controlled sector, based on a hybrid system model for each aircraft which encodes simple aircraft dynamics under the discrete action of the controller. It is observed that the set of commands used by controllers, while large, is finite, and consists of simple actions such as: "turn to heading of x degrees", "hold current heading", "fly direct to jetway y", "increase speed to z knots". The model is then analyzed and the concept of sector dynamic capacity is defined, combined with analysis to predict the time it takes to overload given sectors of airspace, and thus predict delays, assuming controllers use a subset of their available control actions. These results are validated using a simulator of the system, which is implemented in C++ interfaced with MATLAB, and uses data from the *Enhanced Traffic Management System* (ETMS), which is a set of time stamped data for all aircraft in the NAS, at intervals of a few minutes between time stamps.

The data presented in this chapter pertains to several sectors within the Oakland ARTCC, located in Fremont, CA.

2. A Short History of Air Traffic Control

Air Traffic Control in the United States began in the late 1920s, pioneered by airport employees using red and green flags, and lights to signal their instructions to pilots. The Air Commerce Act of May 20, 1926, was the first step of the Federal government towards regulation of civil aviation. This legislation was pushed by the leaders of the aviation industry, who were convinced that the airplane could not reach its full commercial potential without Federal action to define, improve and maintain safety standards. The Air Commerce Act defined several tasks, including issuing and enforcing air traffic rules, licensing pilots, certifying aircraft, establishing airways, and operating and maintaining aids to air navigation. The first city to have a radio-equipped control tower was Cleveland (1930). The first three centers for ATC were established by an airline consortium, encouraged by the Federal government, between 1935 and 1936. Maps, blackboards, and mental calculations were the first tools used by early Air Traffic Controllers to ensure the safe separation of aircraft traveling between cities along designated routes.

In 1938, the Civil Aeronautics Act transferred the Federal civil aviation responsibilities from the Department of Commerce to a new independent agency, the Civil Aeronautics Authority. The legislation also expanded the government's role by giving the Authority the power to regulate airline fares and to determine the routes that airlines would serve. The Authority was split in 1940, giving birth to the Civil Aeronautics Administration (CAA) and the Civil Aeronautics Board (CAB) placed under the Department of Commerce. The CAA was responsible for ATC, airman and aircraft certification, safety enforcement, and airway development. The CAB's task was safety rulemaking, accident investigation, and economic

regulation of the airlines.

The increasing airspace congestion triggered by the growing traffic in the 1940s, the introduction of jet airliners and a series of midair collisions motivated passage of the Federal Aviation Act of 1958, which generated a new agency: Federal Aviation Agency. Even though a special committee had already recommended the use of radar in 1947, it was not until the late 1950s that a civilian radar system was installed by the CAA. The Federal Aviation Agency was given sole responsibility to develop and maintain a common civil-military system of air navigation and ATC. The Act also transferred safety rulemaking from the CAB to the Federal Aviation Agency. On April 1, 1967, the Federal Aviation Agency became one of several organizations within the Department of Transportation (DOT) and became the Federal Aviation Administration (FAA).

In the mid-1970s, the FAA achieved a semi-automated ATC system based on a combination of radar and computer technology. By automating certain functionalities of ATC, the system allowed controllers to concentrate more efficiently on the vital task of aircraft separation, which is still not automated today. The controller graphical display encompassed technology able to visualize numerous information about aircraft (identity, altitude, and ground speed of aircraft carrying radar beacons), while controlling the airspace. Despite its effectiveness, this system was not able to keep up with the growth of traffic and increasing congestion. The NAS Plan, created in January 1982 by FAA aimed at finding solutions to the congestion problem, defined more advanced systems for En Route and Terminal ATC, modernized flight service stations, and improved in ground-to-air surveillance and communication. Several other levels of automation were introduced until the events of September 11, 2001. Short after, Congress created the Transportation Security Administration (TSA), whose principal responsibility is civil aviation security.

Despite the negative impact of September 11 and the concurrent economic recession, which made passenger demand fall initially by more than 20%, airspace congestion is a continued problem. Recent studies [Shaver, 2002] have shown that the overall impact of September 11 on air traffic congestion was just a two year delay in previous estimates. In 40 years, the delays have increased by 50% in the United States. From 1995 to 1999, the average delay grew from 42 minutes to 50 minutes. Flight cancellations increased by 68% between 1995 and 1999. Annual traffic growth is still 2.3% and airlines have increased their flight times on 80% of all busy routes, up to 27 minutes. Such a list of alarming numbers could be extended almost endlessly. It vehemently speaks for automation and optimization of the current ATC system, in order to satisfy the ever-increasing amount of traffic.

3. Organization of Air Traffic Control

This section explains how the airspace is currently divided geographically, and how it is organized hierarchically. Then, the onboard and ground navigation infrastructure is briefly presented. Finally, the communications between all agents in the network is presented. The information enclosed in this section pertains for the US airspace. We mention the European airspace at the end and highlight some specificities of this airspace.

3.1 Airspace Structure

The airspace is divided in different *classes*, which correspond to regions under different regulations and use. An exhaustive classification of airspace is available in [Nolan, 1999], and is only briefly summarized here.

Class A airspace exists from 18,000 to 60,000 feet. All operations in this airspace must

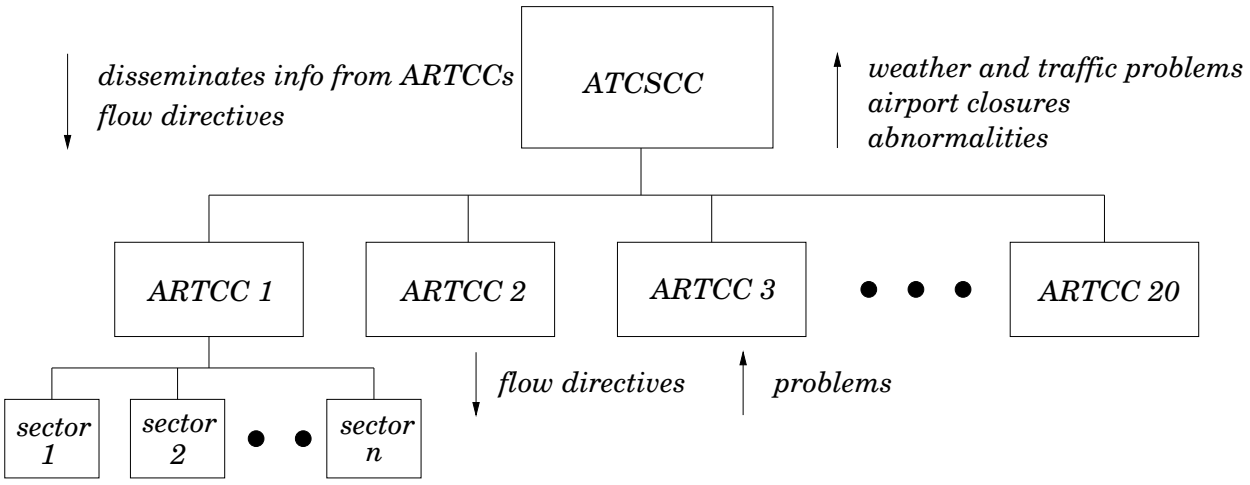

Figure 1: Control hierarchy in the current structure of NAS.

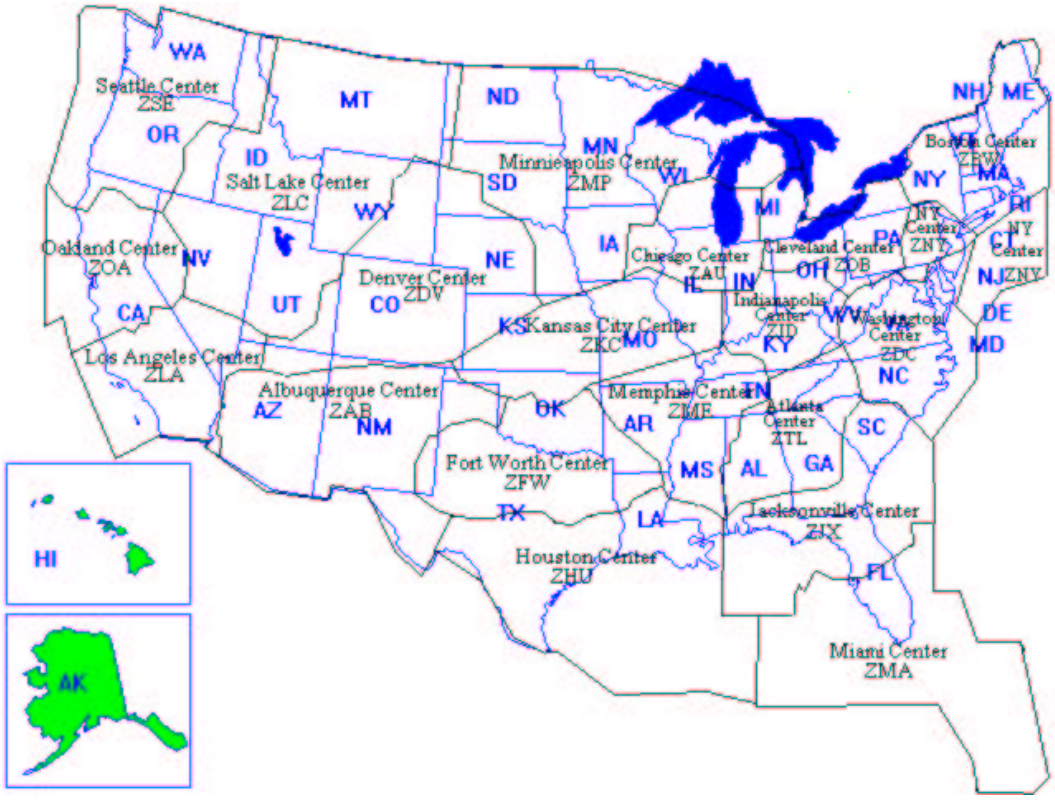

Figure 2: Map of the 22 ARTCCs in the U.S. (map courtesy of `www.seaartcc.org`)

be under *instrument flight rules* (IFR) (pilots must be rated to fly according to the rules governing the procedures for conducting instrument flight) and are subject to air traffic control clearances and instructions. *Class B airspace* surrounds "busy" airports in the US. Each Class B area is individually tailored and consists of a surface area and two or more surrounding layers (see Figure 3) (most Class B airspace would look like an inverted wedding cake if viewed in profile). Again, pilots must receive an ATC clearance to enter class B airspace. *Class C airspace* generally surrounds "smaller" airports with an operating control tower, a radar approach control facility, and a certain number of IFR operations. The area encompassed by this airspace is delimited by two circles with the inner circle extending 5 nautical miles from the airport starting at the surface and extending up to 4000 feet above airport elevation. The outer circle extends to 10 nautical miles from the airport and consists of a shelf from 1200 feet to 4000 feet above airport elevation. The rest of civilian airspace is divided in further categories (*Classes D, E, G*), not relevant for the description in this chapter. Airspace also includes *special use airspace*, which encompasses prohibited areas, restricted areas, warning areas and military operations area, which will not be detailed here.

The *National Airspace System* (NAS) is a large scale, layered, hybrid dynamic system: its control authority is currently organized hierarchically with a single *Air Traffic Control System Command Center* (ATCSCC), in Herndon VA, supervising the overall traffic flow. This is supported by 22 (20 in the continental US or CONUS) *Air Route Traffic Control Centers* (ARTCCs, or simply, Centers) organized by geographical region up to 60,000 feet [Nolan, 1999, Kahne and Frolow, 1996]. Each Center is sub-divided into about 20 sectors, with at least one air traffic controller responsible for each sector. Each sector controller may talk to 25-30 aircraft at a given time (the maximum allowed number of aircraft per sector depends on the sector itself). The controller is in charge of preventing *losses of separation* between aircraft, keeping them separated by more than 5 nautical miles horizontally, and 2000 feet vertically. In general, the controller has access to the aircraft's flight plan and may revise the altitude and provide temporary heading assignments, amend the route, speed, or profile, in order to attempt to optimize the flow and to keep aircraft *separated*, as well as to provide weather reports and winds. An illustration of the current control structure is presented in Figure 1.

There are about 17,000 controllers in the NAS infrastructure, each controlling a zone with rough diameter from 20 to 200 miles [Nolan, 1999]. There are about 19,000 landing facilities, with about 400 of these major airports with ATC towers. The acceptance rate of each airport is usually 1 aircraft/minute per runway in normal operations (if the runway is used for both take off and landing); this capacity is doubled if the runway is used for landing only.

Within the Center airspace, the low traffic density region away from airports is known as the *en route airspace* and is under jurisdiction of the ARTCC. The high traffic density regions around urban airports are delegated to *Terminal Radar Approach Control* (TRACON) facilities. The TRACONs generally control this airspace up to 15,000 feet. There are more than 150 TRACONS in the United States: one may serve several airports. For example, the San Francisco Bay Area TRACON includes the San Francisco, Oakland, and San Jose airports along with smaller airfields at Moffett Field, San Carlos, and Fremont. The regions of airspace directly around an airport as well as the runway and ground operations at the airport are controlled by the familiar *Air Traffic Control Towers*.

3.2 Navigation and Surveillance

Surveillance is performed by ATC through the use of radar: a primary radar system which

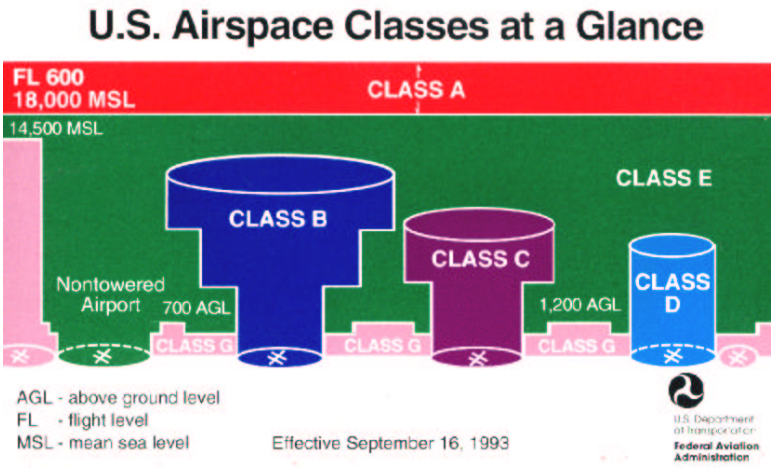

Figure 3: Airspace Classes (courtesy of U.S. Department of Transportation)

processes reflected signals from the aircraft skin, and a secondary radar system, which triggers a transmitter in the aircraft to automatically emit an identification signal. The range of the radars depends on the type of airspace being served: in the En Route airspace the long-range *Air Route Surveillance Radar* (ARSR) is used, while in the TRACON the shorter range *Automated Radar Terminal System* (ARTS) is used. The accuracy of the radars, and their slow (12 second) update rates (6 seconds in TRACON, 12 seconds in En Route airspace), contribute to the FAA standards for aircraft separation, which are 5 nautical miles horizontal separation, 1000 feet (2000 feet above 29,000 feet) vertical separation in the Center airspace, and 3 nautical miles horizontal separation, 1000 feet vertical separation in the TRACON. Each ATC facility is equipped with a computer system which takes the radar signals as input and provides a very limited amount of flight data processing, including a rudimentary conflict alert function. This information is displayed to controllers in two-dimensions on the black and green *plan view displays* (PVDs). Controllers issue directives to pilots using two-way voice (radio) channels.

ATC currently directs air traffic along predefined *victor airways* (low altitude $< 18,000$ feet) and jetways (high altitude), which are "freeways in the sky", or straight line segments connecting a system of *beacons* (*non-directional beacons* (NDBs), *very high frequency omni-range receivers* (VORs), and *distance measuring equipment* (DME)). These beacons are used by pilots (and autopilots) as navigational aids, to update and correct the current position information provided by the *inertial navigation systems* (INS) on board each aircraft.

New systems for navigation and surveillance are currently in the process of certification for use in the NAS.

The *Global Positioning System* (GPS) and its *Wide Area* and *Local Area Augmentation Systems* (WAAS and LAAS) provide 3D position information worldwide using signal information from a constellation of 24 satellites. A single GPS receiver can determine its position to an accuracy of a few meters, using signals from at least 4 out of these 24 satellites; if this information is augmented with differential corrections from another receiver (*differential GPS* or DGPS), this accuracy can be increased to a few centimeters. Many factors make the use of GPS in the cockpit a desirable alternative to the current ATM navigation methods [Gazit, 1996]: the accuracy is uniform from aircraft to aircraft whereas with the currently used INS, the accuracy decreases in time due to sensor drift rates; each GPS receiver acts

like an atomic-accurate clock, thus making it possible for many aircraft to coordinate among each other over a communication link; a GPS receiver is much cheaper than an INS system, and orders of magnitude cheaper than a VOR beacon. Fueled by the success of GPS and its augmentations, the EU started a European Satellite Navigation system, called Galileo. Galileo will be built around 30 satellites (27 operational and 3 reserve craft), and will offer features similar to the GPS. Aside from the evident industrial benefits of such a system (industry will now provide the same type of equipment for GPS and Galileo), the advent of such a system is also very important for the ATC community, since it will enable the use a redundant guidance system relying on both technologies.

Automatic Dependent Surveillance (ADS) is a communication protocol by which aircraft would transmit over digital satellite communication their GPS position information, velocity, as well as information about their intended trajectory, to the ground ATC. ADS-B (for broadcast) is a protocol for broadcasting this information to neighboring aircraft. Its major advantage over the current ATM surveillance methods is its ability to provide very accurate information for trajectory prediction, without relying on the radar system. Two immediate benefits of such a communication link are a huge improvement in surveillance over oceanic airspace, which is not covered by radar, and the possibility of reducing the separation standards between aircraft in all airspace.

Traffic Alert and Collision Avoidance System (TCAS) is an instrument integrated into an aircraft cockpit, which consists of hardware and software providing the pilot with information about traffic in its direct vicinity. In case of a potential upcoming collision with another aircraft, TCAS will sound an alarm and will provide the pilot with an escape maneuver to follow, coordinated with the other involved aircraft.

User Request Evaluation Tool (URET) is a tool which automatically predicts upcoming conflicts between the aircraft and notifies the Air Traffic Controllers. It is currently in use in six Centers in the US, and is in the process of being implemented in 14 Centers. URET also works as an advisory, resulting in greater efficiency in airspace use, in particular regarding direct routing and restrictions at sector boundaries. URET has also the potential of helping guidance of free through its advisory capability.

3.3 Communication and Procedures

All IFR pilots must file a *flight plan* at least 30 minutes before pushing back from the gate. The pilot reviews the weather along the intended route, maps the route and files the plan. The flight plan includes: flight number (which includes the airline identification), the aircraft type, the intended airspeed and cruising altitude, the route of flight (departure airport, Centers that will be crossed and destination airport). It also include additional information, such as *waypoints, navaids*, or *fixes*, which will be used by the aircraft to navigate through sectors of airspace. The flight plan also contains the *arrival*, which is a set of closely spaced waypoints, navaids or fixes leading to an airport. An example of arrivals into the Oakland airport (in California) is shown in Figure 4, with corresponding infrastructure. The pilot transmits the desired flight plan information to ATC, where a controller called a *flight data person* reviews the weather and flight plan information and enters the flight plan into the FAA main, or "host" computer. The computer generates a set of flight progress *strips* that are sent electronically from sector controller to controller across the flight plan; these strips, and flight plans, may be updated by each controller throughout the flight. The flight progress strip contains all of the necessary data for tracking the aircraft.

After the pilot has filed the flight plan, ATC may modify the flight plan according to con-

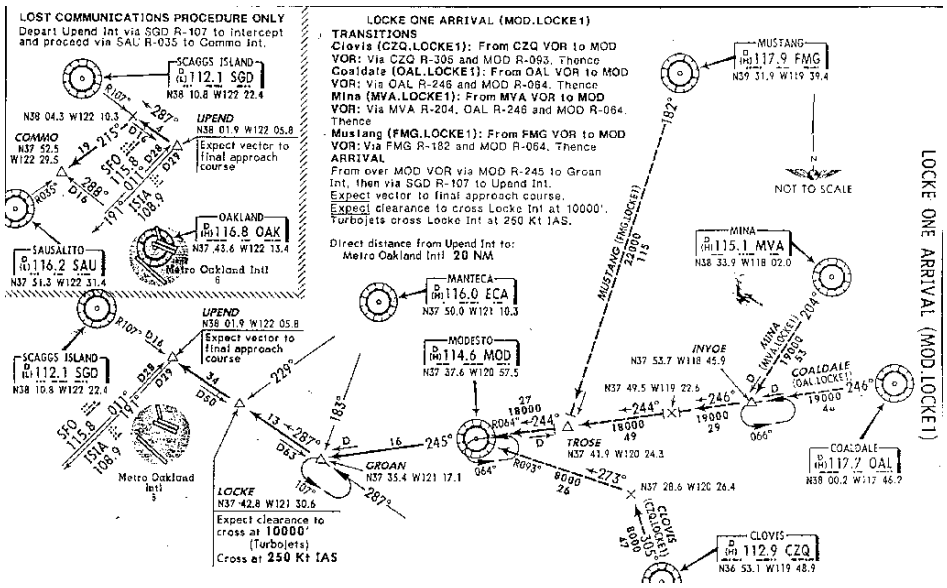

Figure 4: Example of arrivals into the Oakland (OAK) airport : LOCKE 1 (MOD.LOCKE1). Aircraft enter this airspace through the waypoints: MUSTANG, MINA, COALDALE, CLOVIS, whose acronyms are FMG, MVA, OAL, CZQ. Note the tracks for holding patterns (shown as loops at various merge points). Source: [JEPPESEN, 2000].

straints of the NAS and other aircraft (information which is available to each controller from conversations with the ARTCC and ATCSCC controllers), and issues a clearance to the pilot. After take-off, the control of the aircraft is passed through the Tower, TRACON, and possibly several Center facilities until the destination TRACON is reached.

Each sector controller may talk to 25-30 aircraft at a given time [NOLAN, 1999]. When an aircraft crosses the boundary from one sector to the next, there is a "hand-off" in which the communication is transferred from one controller to the next. Potential conflicts must be resolved before hand-off occurs. The controller directs the aircraft according to a set of simple *control directives*, voiced sequentially. One of the most important, and time consuming, controller tasks is to prevent LOS, between aircraft.

Radio communications are a critical link in the ATC system. The most important aspect in pilot-controller communications is mutual understanding of the command and response. Therefore, pilots acknowledge each radio communication with ATC by using the appropriate aircraft call sign; contacts are kept as brief as possible. For example, a contact procedure is codified as follows: name of facility being called, full aircraft identification as filed in the flight plan, and eventually, the request or type of message to follow. Each procedure is codified in a similar way. Each sector is handled by one key controller, each controller has his own radio frequency over which the communication with pilots in his sector takes place. As a flight progresses from one sector to another, the pilot is requested to change to the appropriate frequency. The *International Civil Aviation Organization* (ICAO) phonetic alphabet is used by FAA personnel when communications conditions are such that the information cannot be readily received without their use. The grammar and phraseology used in the current system is available in [NOLAN, 1999] and has been the focus of recent studies [HISTON and HANSMAN, 2002]. In general, the commands given to the aircraft by ATC are very precise and can be easily categorized in a discrete set of functions, parameterized by real numbers indicating speed, heading, or other flight variables. This very procedural com-

mand environment facilitates the task of modeling human ATC action, communication and aircraft behavior, as will be shown in this chapter. A sample command given by a human ATC to an aircraft might be: *"achieve flight level 290, turn to a heading of 130, reduce airspeed to 120 knots ..."*. In addition, the procedures differ from TRACON to Center control: in the TRACON, the controller is responsible for taking the aircraft from *Climbout* $\rightarrow$ *En Route* (the control actions must meet impromptu flow restrictions, hand-off to En Route control, clear to join filed route); in the Center, the controller may revise the altitude and provide temporary heading assignments, amend the route, speed, profile, and provide weather reports and winds.

In this way, the control is *distributed*, since it is applied locally in each sector. There is loose *coupling* within the ATC hierarchy: the ATCSCC controllers talk to the ARTCC controllers several times a day to provide updates and receive feedback about the flow control in each Center/sector. In the case of bad weather, airport closures, or other large disturbances, this feedback tightens, and the directives and updates among the levels of the hierarchy becomes more frequent.

Air Traffic Control regulations and infrastructure in other parts of the world differ from the US airspace. In particular, in Europe, unlike other large countries such as Russia or China, the skies are not unified, despite a similar size infrastructure. Eurocontrol, a European agency unifies the different national entities participating in ATC under a single organization, but the different sovereign States remain responsible for their airspace, which sometimes are of small geographical dimensions, leading to several handover procedures for short flights. The European skies includes 75 centers in charge of En-Route traffic, with around 18,000 Air Traffic Controllers (13,000 for the 15 states of the EU). Even if regulations differ from airspace to airspace, a Regulatory Committee and an accompanying Regulatory Unit have been created within Eurocontrol to ensure that regulations are properly observed by member states. An example of differences in standards with the US is vertical separation. Since January 2002, the *Reduced Vertical Separation Minimum* (RVSM) standards have provided six additional flight levels between 29,000ft and 41,000ft in the airspace of 41 European and North African countries, by reducing the separation minima from 2,000ft to 1,000ft. The midair crash above Überlingen mentioned above

4. Levels of Automation in the Current System

The main goal of ATC is to maintain safe separation between aircraft while guiding them to their destinations. However, the tight control that it has over the motion of every aircraft in the system frequently causes bottlenecks to develop. Avoiding these bottlenecks is also a goal ATC, but the priority of this task, however, is less than that of maintaining separation. Uncertainties in the positions, velocities, and wind speeds, as well as the inability of a single controller to handle large numbers of aircraft at once tend to lead to overly conservative controller actions and procedures to maintain safety.

Recently, there has been increasing research in the development of analysis tools and methods to automate or partially automate some of what is today manually performed by ATC. Here, two key problems arise: understanding the characteristics of the current system, in order to be able to predict its reaction under stress or severe conditions, and identifying the portions of the system which can be automated, in order to optimize traffic and relieve ATC from a part of its workload. Section 4.1 presents existing models of the current ATC system, including infrastructure and human control. It is shown how models can be used to analyze the system and their potential predictive capabilities are explained. Section 4.2 lists some

of the functionalities of ATC which are already automated or partially automated today. Section 4.3 demonstrates some of the algorithms which are focus of current research for automation of some other functionalities of ATC.

4.1 NAS and ATC Models

There are many elements which make a model of the NAS useful for the FAA to understand and assess the behavior of the current ATC system. Once the model is validated against the actual system, it may be used according to different functions. It can be used in *simulation* as an evaluation tool: simulations are useful to anticipate consequences of certain decisions made by ATC. The model can also be used for its *predictive capabilities*, for example assessing the feasibility of certain maneuvers or flows. Finally, it can be used for *optimization*, that is to modify certain parameters of the system and assess the value that these changes have on the system performance, such as aircraft count per sector, arrival rates at airports, or airborne encountered delays.

Existing NAS modeling tools [Plaettner-Hochwarth et al., 2000] have functionality which spans the modeling of runway and airport capacity and operations, through airspace operations [Erzberger et al., 1993] and conflict resolution, to human factors and man-machine integration. See [Odoni et al., 1996, Kuchar and Yang, 2000] for detailed surveys of NAS modeling and conflict detection and resolution methods. A recent tool, FACET [Bilimoria et al., 2001, Bilimoria and Lee, 2001], represents the first accurate NAS simulation tool, with the additional capability of a "playback" mode using actual traffic flow data. Instead of providing an extensive list of the different models with respective functionalities, we analyze one model in depth and explain how it can be used to analyze the system and generate predictions. Further details for this model can found in [Bayen et al., 2002b].

The goal of the model in [Bayen et al., 2002b] is to complement existing tools, by providing a control theoretic component which models the influence of ATC on the traffic. While the additional logic required to model the actions of the controller does not pose a significant computational problem if the aircraft density in the airspace is low, it becomes an issue as the density increases. The long term goal of increasing capacity as well as safety in the NAS cannot be achieved without an in-depth analysis of the applied control logic. If such a system were shown to model the current airspace with sufficient accuracy, then a wide array of applications would become feasible, including providing additional support for ATC in predicting delay.

The structure of the NAS is complex, for it involves a multitude of interacting agents and technologies: aircraft monitoring, flow management, communication, and human-in-the-loop. For the present work, whose goal is in predicting delay, only the features which are important for this purpose are extracted. A portion of the Oakland ARTCC is modeled, which contains five *sectors*. These sectors surround the Oakland TRACON, which controls the aircraft on their approaches into San Francisco, San Jose and Oakland airports. The TRACON is the final destination of the traffic considered here.

A sector is modeled by a portion of airspace containing aircraft under the local control of the responsible air traffic controller (Figure 5). The interior of this domain is the controlled area (in which the local controller can actuate the flow). Within each sector, navigation infrastructure, including jetways, waypoints and navigation aids, is used to help the flow follow desired patterns; they are therefore included in the model and used, even if it is observed that more than 40% of the aircraft may fly off the jetway at any given time. The model allows for aircraft to fly at different altitudes, but not to climb or descend. Altitude

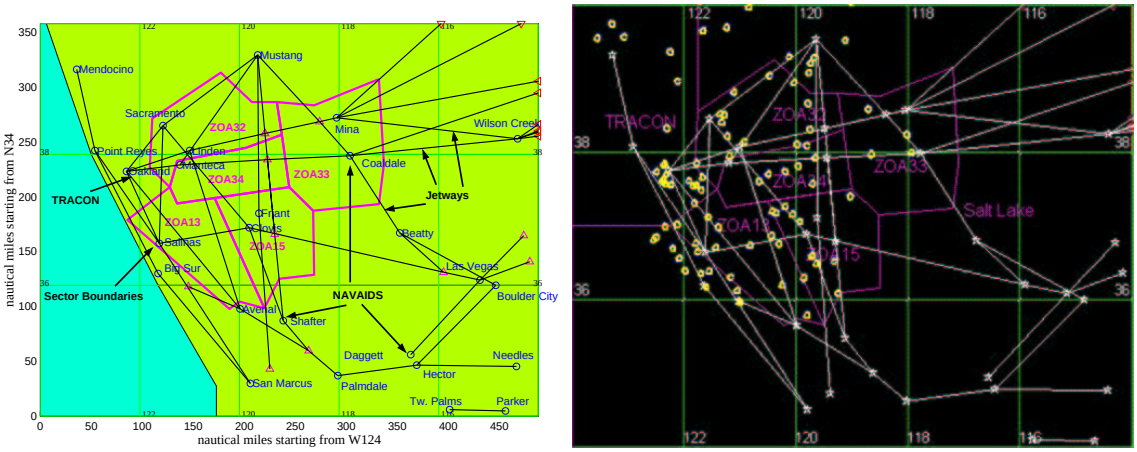

Figure 5: **Left:** ATC sectors modeled for this study: 32, 33, 34, 13 and 15 within the Oakland ARTCC (labeled above as ZOA32, etc.). Most crucial jetways and waypoints for a San Francisco approach are shown here. The data modeled comes from FACET [Bilimoria et al., 2001] as well as JEPPESEN high altitude en route charts [Jeppesen, 2000]. **Right:** Visual display of the simulator used in [Bayen et al., 2002b]. Traffic in the Oakland ARTCC (aircraft not in this Center have been filtered out). This plot has been generated using ETMS data.

changes are not crucial for the effects of interest: the type of sector overloads of interest results mainly from aircraft acceptance rates at destination airports.

4.1.1 Aircraft Model

A *hybrid model* is used for each aircraft. It describes the evolution of a system by a set of *discrete modes*, each associated with a *continuous dynamical system*, and *discrete switches* which enable the system to jump from one mode to another instantaneously. In mathematical terms, the motion of aircraft i is described by:

$$\dot{\vec{x}}_i = \frac{d\vec{x}_i}{dt} = \vec{v}_{\text{current heading}} \tag{1}$$

where $\vec{v}_{\text{current heading}} \in \mathbf{R}^2$ is a constant velocity vector held by the aircraft until the next discrete switch: a heading or speed change which changes $\vec{v}_{\text{current heading}}$. $\vec{x}_i \in \mathbf{R}^2$ is the planar position of aircraft i. Integration of equation (1) over time produces a continuous piecewise affine trajectory. Such a model is preferred over a continuous dynamical model for two reasons. First, the time scale of a "change in aircraft behavior" (for example a turn or slow down) is on the order of 30 seconds, whereas the time scale of a straight line portion of the flight is usually much longer, sometimes half an hour or more, thus the dynamics of such maneuvers is ignored in favor of their effects only (the set of resulting straight lines). Second, the update rate of ATC monitoring is in general not more than 30 seconds, which makes the details of these maneuvers inaccessible to the ATC. This approximation is widely accepted in the literature [Dugail et al., 2001, Pallottino et al., 2001, Bilimoria, 2000, Mao et al., 2000, Nilim et al., 2001, Bertsimas and Stock Patterson, 1998].

Monitoring ATC shows that a finite set of maneuvers, which depends on local parameters, is used. Combinations of these maneuvers result in a conflict-free flight environment in which the constraints of the air traffic flow are met. The maneuvers shown in Figure 6 are modeled (and consist of changing the right hand side of (1) according to certain rules which are

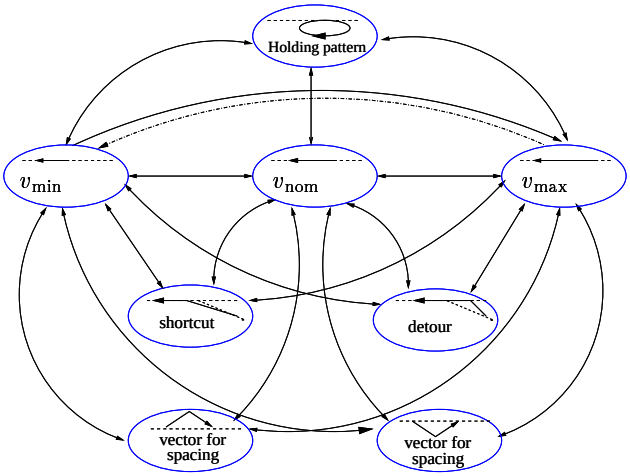

Figure 6: Hybrid automaton representing the action of one controller on a single aircraft. Each of the eight modes represents one possible state of the aircraft. The arrows joining these states are the mode switches, initiated by the controller. The dash-dotted transitions are used for the predictive analytical models. The complete set of arrows is used for the simulation and models of human ATC behavior.

now made explicit). The validity of models similar to this has been confirmed by statistical studies realized at MIT ICAT [Histon and Hansman, 2002].

1. *Speed change:* ATC may decelerate or accelerate the aircraft along its flight plan:

$$\vec{v}_{\text{modified speed}} := \lambda \cdot \vec{v}_{\text{current heading}} \tag{2}$$

where $\lambda \in \mathbf{R}^+$ defines the magnitude of the velocity change. The model will allow a finite set of speeds (which means λ has a finite number of acceptable values). This encodes the fact that the ATC generally has a finite set of possibilities in the choice of speeds, because the aircraft flies at its optimal speed per altitude and ATC will speed up or slow down the aircraft by not more than 10% of the current value.

2. *Vector-for-spacing (VFS):* This maneuver consists of a deviation of the aircraft from its original flight plan for a short time (part 1 of the maneuver), and a second deviation, bringing it back to its original flight plan (part 2 of the maneuver). This stretches the path that the aircraft must follow, and therefore generates a delay. The length of this maneuver depends on the geometry of the sector. Calling R_ψ the rotation matrix by angle ψ:

$$\begin{array}{ll} \vec{v}_{\text{part 1}} := R_\psi \cdot \vec{v}_{\text{current heading}} & \text{(First half of the maneuver)} \\ \vec{v}_{\text{part 2}} := R_{-2\psi} \cdot \vec{v}_{\text{part 1}} & \text{(Second half of the maneuver)} \end{array} \tag{3}$$

3. *Shortcut / Detour:* In certain situations, the ATC will have the aircraft "cut" between two jetways, a maneuver which could either shorten or lengthen the flight plan. The decision to command such a maneuver is often dictated by conflict resolution, but could also be used to shorten the overall flight time if sector occupancy allows it (sometimes called "direct-to" by pilots):

$$\vec{v}_{\text{shortcut}} := R_\psi \cdot \vec{v}_{\text{current heading}} \quad \text{for the duration of the maneuver} \tag{4}$$

until the next ATC action is taken. Here again ψ is the angle by which ATC turns the

aircraft to achieve the shortcut.

4. *Holding pattern:* In some extreme conditions, holding patterns are used to maintain an aircraft in a given region of space before eventually letting it follow its original flight plan. This is modeled by assigning the aircraft to a predefined zone and keeping it there while preventing other aircraft from entering that zone.

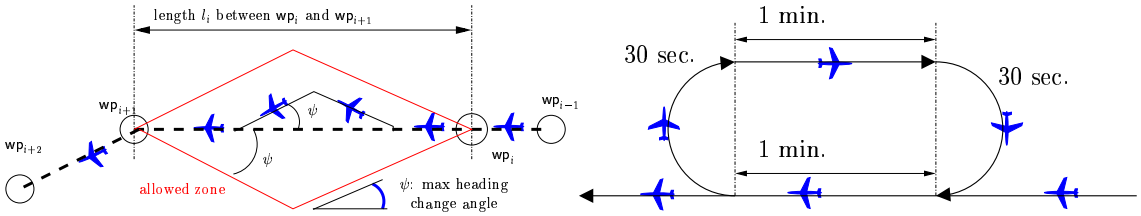

Figure 7: Two examples of maneuvers corresponding to modes of the automaton in Figure 6. **Left:** Deviation ψ from flight plan using a *Vector For Spacing* (VFS) available in a given sector. More precise models are also available in [DUGAIL et al., 2002]. **Right:** *Holding Pattern* (HP). The prescribed "time to lose" is given to the aircraft in minutes: one minute in each straight portion and 30 seconds in each half circle. For this scenario, $T_{\mathsf{hp}} = 1\,\text{min}. + 30\ \text{sec}. + 1\,\text{min}. + 30\ \text{sec}. = 3\,\text{min}.$

4.1.2 Lagrangian Delay Propagation Model

A large proportion of air traffic jams, i.e. portions of airspace saturated by aircraft, is generated by restrictions imposed at destination airports, usually themselves driven by weather or airport congestion. These restrictions are imposed as either miles-in-trail or minutes-in-trail, representing the distance (or time) required between aircraft in a flow incoming to the TRACON, and are referred to as *metering constraints*. Figure 8 illustrates the topology of the flows incoming to San Francisco Airport (SFO) which are often subject to this type of constraint. These constraints tend to propagate backwards from the airport into the network, and result in miles-in-trail constraints imposed at the entry points of each sector. For example, in the case of Figure 8, typically, the backpropagation of these metering conditions is as follows: TRACON $\rightarrow$ sector 34 $\rightarrow$ sector 33 $\rightarrow$ Salt Lake Center $\cdots$ and similarly for the two other flows.

In the current system, these restrictions are imposed empirically. The model presented here is useful for understanding *(i)* how the traffic jams propagate; and *(ii)* what the optimal control policy should be under these restrictions, in order to ensure maximal throughput into the TRACON.

A simple Lagrangian model of merging flows is presented, introduced in [BAYEN et al., 2002a]. This model of merging flows predicts the backpropagation of a traffic jam from a destination airport into the network. It is Lagrangian, because it models the trajectories of all agents in the considered space, rather than averaged quantities in a control volume, such as the number of aircraft in a given sector. This model can be applied to the merging flows of the type shown in Figure 8. It is used here to derive the dynamic capacity of a sector. Given prescribed inflow and outflow conditions, the *dynamic capacity* of a sector is defined to be the number of aircraft which can be actuated[2] in this sector, until ATC action has to occur upstream from this sector in order to not violate the outflow conditions. It is assumed in this definition that the sector is initially empty, but it may be applied to cases in which the

[2]ATC *actuation* of an aircraft means any alteration of its original flight plan, in order to meet desired conditions in the region of interest.

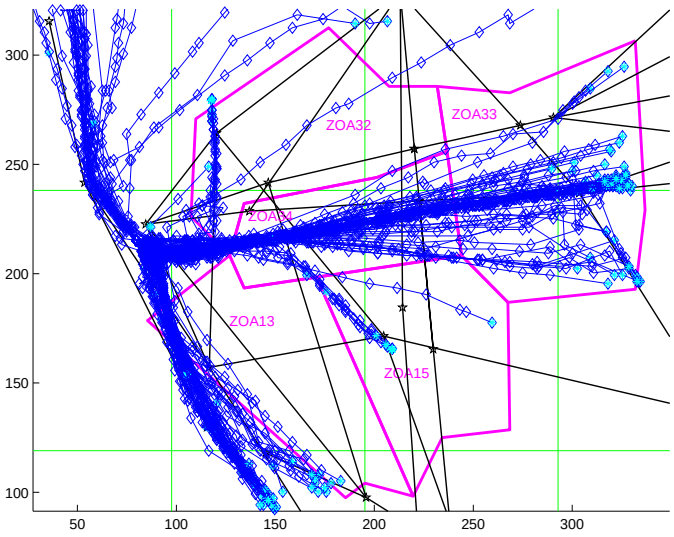

Figure 8: Overlay of trajectories merging into San Francisco (11 hours of traffic). The data modeled comes from ETMS and FACET [Bilimoria et al., 2001]. The Lagrangian approach adopted in this work models each of the trajectories included in this plot.

sector is initially non-empty. The dynamic capacity is a concept which appears naturally in the following problem:

Given a required spacing between the aircraft (metering constraint) of ΔT_{out}*, compute a controller policy which will force groups of aircraft to* exactly *satisfy the metering constraint at the sector exit point (each aircraft separated by exactly* ΔT_{out}*) while maintaining separation at all times.*

A very simple version of the problem is first considered, in which the controller uses only two modes (fast and slow). This model is not overly restrictive: several methodologies have been developed to map the full automaton of Figure 6 to this model: see for example [Bayen and Tomlin, 2003, Dugail et al., 2002].

The initial arc length distance $a_i^0 \in \mathbf{R}$ of aircraft i along its arrival route to the airport is introduced. For example, $a_i^0 = -200$ means that aircraft i has to fly 200 nm before landing in the airport. The location at which the metering condition is imposed is called $x_{\text{ex}} \in \mathbf{R}$. For example, $x_{\text{ex}} = -50$ means that the metering is applied 50 nm from the airport. It is possible to assume without loss of generality that the aircraft are numbered in order of arrival (the a_i^0 are indexed in increasing order of magnitude).

The following situation is investigated: all aircraft are initially at maximum speed $v_{\max}$, and in order to enforce metering, ATC slows down aircraft i to its minimum speed $v_{\min}$ (see Figure 8), at a location x_i^{switch} and time t_i^{switch} which will be determined (see Figure 10). This scenario is represented as a dash-dot line in Figure 6. The following condition is imposed: each aircraft crosses the metering point x_{ex} at exactly $t_{\text{block}} + (i-1)\Delta T_{\text{out}}$ where t_{block} is the time at which the metering condition is initiated. For simplicity, the origin of time is taken at $t_0 = 0$. This leads to the following kinematic equations of the aircraft under this

actuation:

$$\begin{aligned} x_i(t) &= a_i^0 + v_{\max} t \quad t \in [0, t_i^{\text{switch}}] \\ x_i(t) &= b_i + v_{\min} t \quad t \in [t_i^{\text{switch}}, t_{\text{block}} + (i-1)\Delta T_{\text{out}}] \end{aligned}$$

The assumption of continuity of $x_i(t)$ enables one to solve for b_i, from which the switching time is computed: $t_i^{\text{switch}} = (b_i - a_i^0)/(v_{\max} - v_{\min})$. Under the following feasibility conditions:

$$a_i^0 \in [x_{\text{ex}} - v_{\max}(t_{\text{block}} - (i-1)\Delta T_{\text{out}}), x_{\text{ex}} - v_{\min}(t_{\text{block}} - (i-1)\Delta T_{\text{out}})] \tag{5}$$

the propagation speed of the traffic jam may be computed analytically, by solving for the location of the edge of the traffic jam in space and time:

$$\begin{aligned} t_i^{\text{switch}} &= \frac{x_{\text{ex}} - v_{\min} t_{\text{block}} - (i-1)\Delta L - a_i^0}{v_{\max} - v_{\min}} \\ x_i^{\text{switch}} &= a_i^0 + \frac{v_{\max}[x_{\text{ex}} - v_{\min} t_{\text{block}} - (i-1)\Delta L - a_i^0]}{v_{\max} - v_{\min}} \end{aligned} \tag{6}$$

where $\Delta L := v_{\min}\Delta T_{\text{out}}$ is the metered spacing at the outflow of the sector. At a given time t, the set of aircraft such that $t_i^{\text{switch}} \leq t$ is called *traffic jam* or *metered platoon*. These aircraft have already been actuated (see Figure 8). It follows directly from (6) that the traffic jam will not grow if the two following conditions are met:

$$\begin{aligned} t_i^{\text{switch}} < t_{i+1}^{\text{switch}} &\quad \Leftrightarrow \quad \Delta L < a_i^0 - a_{i+1}^0 \\ x_{i+1}^{\text{switch}} < x_i^{\text{switch}} &\quad \Leftrightarrow \quad \left(\frac{v_{\min}}{\Delta L}\right) < \left(\frac{v_{\max}}{a_i^0 - a_{i+1}^0}\right) \end{aligned} \tag{7}$$

Condition (6) above is a sufficient condition for the length of the traffic jam to decay (which can be observed by inspection of the slope of the switching curve or *shock wave* of points $(x_i^{\text{switch}}, t_i^{\text{switch}})$ displayed in Figures 10 and 11). It is a local property of the problem: the conditions depend only on $a_i^0 - a_{i+1}^0$, not on all aircraft considered here. The second equation in (6) is in fact a one-dimensional discretized steady Lighthill-Whitham-Richard (LWR) equation, which appears naturally in highway congestion problems [Newell, 1993]. This fact links this Lagrangian approach, based on aircraft trajectory analysis, with Eulerian approaches such as [Menon et al., 2002], which are based on conservation equations; it relates local properties of the flow (local monotonicity of a variable, here the x-location of the wavefront) to global quantities (here the trajectories of the aircraft). This result is illustrated in Figure 10.

Using the analysis of the previous section, it is fairly easy to predict the *dynamic capacity* of a sector. Consider the worst case scenario: an incoming flow of aircraft, each at $v_{\max}$, separated in time by ΔT_{in} chosen to violate the second condition in (7). This will create a traffic jam originating at SFO (equated with entrance to SFO TRACON), which "piles up" and progressively fills sector 34. The arc length distance of the portion of the arrival jetway contained in sector 34 (see Figure 9) is called l. It is assumed that the sector is initially empty. Using equations (6), it is possible to compute the maximum number N_{limit} of aircraft which can be stacked along the length l of the jetway in the sector, before space is no longer available on this jetway. These aircraft are labeled as *metered platoon* in Figure 9; for example, for the situation shown in this Figure, approximately half of l is occupied by the metered platoon at the time considered, so the number of metered aircraft is approximately $N_{\text{limit}}/2$. When the number of metered aircraft reaches N_{limit}, after a time called T_{limit}, the actuation shown in Figure 9 (ATC slows aircraft down) has to occur upstream (in sector

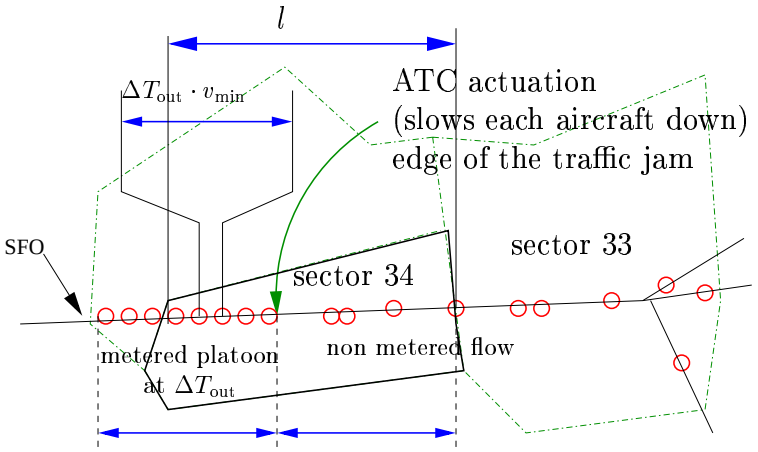

Figure 9: ATC control action on the merging flow. The traffic jam extends from SFO to the edge of the platoon with $\Delta T_{\text{out}} \cdot v_{\text{min}}$ miles-in-trail (one aircraft every ΔT_{out} time units). Once the actuation point (edge of the traffic jam) has moved upstream (into sector 33), sector 34 is called saturated.

33). N_{limit} and T_{limit} are respectively given by:

$$N_{\text{limit}} = \frac{l(v_{\text{max}} - v_{\text{min}})}{v_{\text{max}} v_{\text{min}} (\Delta T_{\text{out}} - \Delta T_{\text{in}})} \tag{8}$$

Several comments can be made regarding the two previous results: *(i)* As $v_{\text{min}} - v_{\text{max}} \to 0$, $N_{\text{limit}} \to 0$: no aircraft can be handled in the sector because no actuation is possible (it is not possible to "make the aircraft lose time" in this sector); *(ii)* As $\Delta T_{\text{out}} - \Delta T_{\text{in}} \to 0$, $N_{\text{limit}} \to \infty$ and $T_{\text{limit}} \to \infty$: if the incoming flow is such that it is almost metered as imposed at the exit of the sector, the number of aircraft required to saturate this airspace becomes large and the time it takes to saturate this sector grows accordingly.

The construction of the switching curve $(x_i^{\text{switch}}, t_i^{\text{switch}})$ described previously, can be used to compute the maximal extent of a traffic jam, or the portion of jetway affected by a traffic jam. Using (6), one can trace the shock location in the (x, t) plane (Figure 11). The edge of the traffic jam, called x_m, obtained at t_m gives the worst situation obtained from the initial configuration a_i^0 of the aircraft: at t_m, the extent of the traffic jam is at its maximum. In the case of Figure 11, one can see that the traffic jam does not propagate more than 300 nm upstream from the destination of the aircraft, called x_{ex}. Therefore no metering conditions need to be applied upstream from that point. In the current system, such information is not available to the ATC, thus leading to extra buffers taken by the controllers, which in turn leads to non optimal operating conditions as well as backpropagation of "virtual overload", a set of conservative precautions.

4.1.3 Human Air Traffic Controller Model

The models of the previous section rely on a mathematical analysis of metering, based on geometry. In order to understand how realistic these models are, one needs to validate them against the real system. Since the real system is not available as a testbed, an abstraction of it which reproduces its behavior adequately was designed. In this section, the design of a simulator which mimics true ATC behavior is presented, as well as the validation of this simulator against ETMS data. The graphical interface of the simulator is shown in Figure 5. This section explains how analytical predictions of the previous section can be validated against the real system using a simulator.

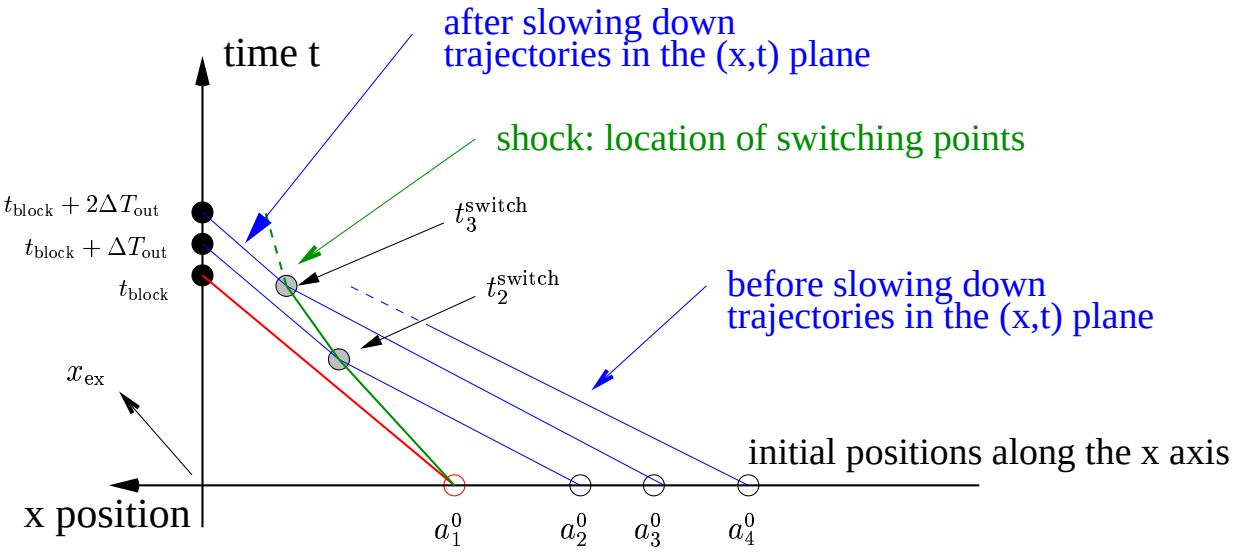

Figure 10: Shock construction. The aircraft trajectories are represented in the (x, t) plane. They originate at $t = 0$ from the horizontal axis (white circle on each trajectory). After some amount of time, the aircraft may be switched to speed $v_{\min}$ at location $(x_i^{\text{switch}}, t_i^{\text{switch}})$ (shaded circle on each trajectory). Ultimately they reach x_{ex}, the entrance of TRACON (black circle).

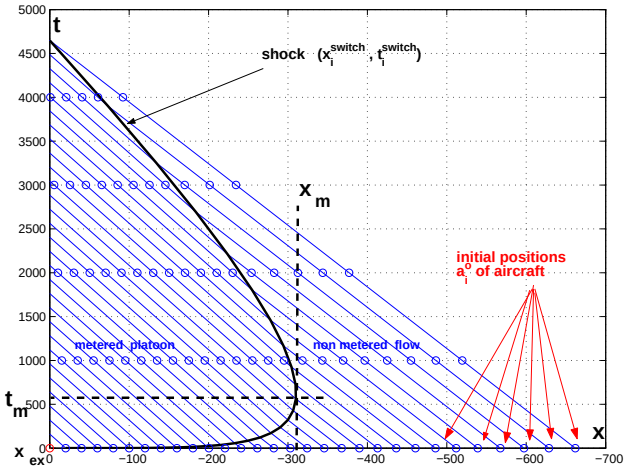

Figure 11: Switching curve (shock) for a vanishing traffic jam. x denotes the distance to the metering point (SFO). The lines are the trajectories of the aircraft in the (x, t) space. The positions of aircraft are represented every 1000 sec. as dots. Once they have passed through the shock, they are separated by $v_{\min}\Delta T_{\text{out}}$. The point (x_m, t_m) is the furthest reachable point by this traffic jam. Note that the slope of the lines changes through the shock. The slope difference can hardly be seen visually, because the speed change is small.

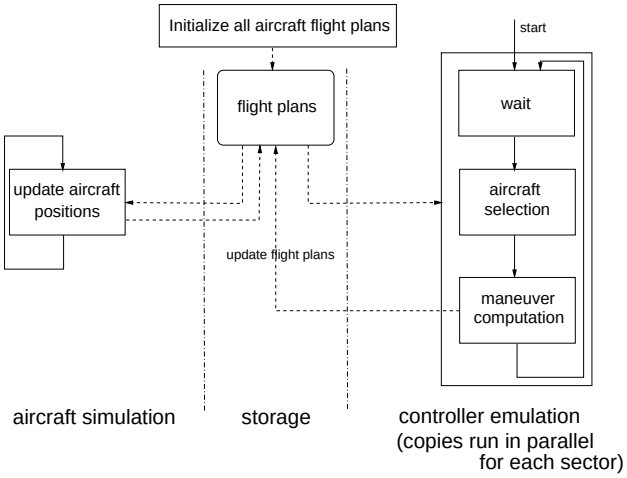

Figure 12: Program flow of the simulator used in [BAYEN et al., 2002b].

The simulator is based on empirical studies that were realized at the Oakland ARTCC, its core is based on observed behavior. Figure 6 (all transitions enabled) summarizes the behavior model observed at the ARTCC. The switching logic behind the transitions is the object of this section, and is implemented in the form of a cost function, which is described in this section.

Overall program flow The overall program flow of the simulator is shown in Figure 12. The input to the code is a set of aircraft filed flight plans (Figure 12, middle column), that can either be user generated or taken from ETMS data (as in FACET). As in the true system, these flight plans are not conflict-free and usually do not satisfy metering conditions imposed on the network. Once the program is initialized, aircraft motion simulation follows these flight plans (Figure 12, left column). As time is advanced, conflict as well as metering constraints are dealt with on a sector by sector basis (with sector-wide look ahead, Figure 12, right column), according to the full automaton shown in Figure 6. The flight plans are updated accordingly.

Key Data Structures Aircraft dynamic equations (1) produce a set of segments; the knowledge of the points connecting the segments and of the aircraft velocity is thus enough to define an aircraft trajectory. This trajectory is thus implemented as a *linked list* of *points* $[x, y, z]$, with a prescribed velocity between the points. The linked list is modified by the simulated controller in the program. The output for each aircraft is the updated linked list. The sectors are implemented as sets of linked lists accessible to a controller. They also contain additional data such as metering conditions (number of aircraft through a given boundary per time unit).

Controller Emulation ATC behavior is modeled by three levels of priority:

- *Priority 1: No loss of separation.* The prevalent requirement for ATC is to ensure that any aircraft pair is always separated by more than 5 nautical miles.
- *Priority 2: Metering conditions.* The controller needs to ensure that the outflow from his sector is an acceptable inflow for the next sector (or TRACON). Metering conditions can be of various nature: admittance rate or separation at downstream junctions.
- *Priority 3: Best possible throughput.* ATC will try if possible to give direct routes to

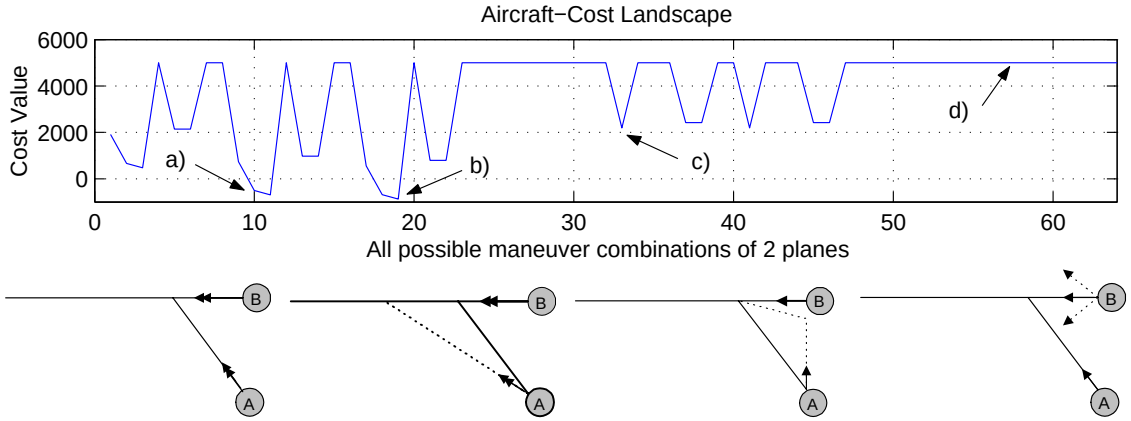

Figure 13: Top: cost values for all possible maneuver combinations in a two-aircraft intersection scenario, where the eight maneuvers of Figure 6 are enabled (thus generating $8^2 = 64$ possible values of J). Four out of 64 examples are extracted and illustrated (four lower pictures). (a) Both aircraft maintain same speed; (b) Aircraft A takes a shortcut maintaining aircraft B at max speed; (c) A makes a VFS at low speed; (d) A does nothing, B is not able to prevent the loss of separation. In this case, the simulated controller would choose solution (b) since the lowest cost is associated with that maneuver. The cost J has been truncated at $5 \cdot 10^3$ for readability. Thus a LOS cannot be visually differentiated from a breach in this plot, though it can in our data. Thus, too large breaches as well as LOS do not appear on this plot.

aircraft in order to minimize their flight times.

These priorities may be modeled using the cost function J:

$$J = \text{cost}_{\text{LOS}} + \text{cost}_{\text{BC breach}} + \text{cost}_{\text{delay}} + \text{cost}_{\text{aircraft actuation}} + \text{cost}_{\text{maneuver}} + \text{cost}_{\text{min dist}} \quad (9)$$

Each term of the cost is a weighted function:

$$J = \sum_{i=1}^{N} \frac{n^i_{\text{LOS}} \cdot w_{\text{LOS}}}{\Delta T^i_{\text{LOS}}} + \sum_{i=2}^{N} (T^i_{\text{breach}})^2 \cdot w_{\text{breach}} + \sum_{i=1}^{N} (TOA^i_{\text{pred}} - TOA^i_{\text{real}}) \cdot w_{\text{delay}}$$
$$+ N_{\text{moved}} \cdot w_{\text{single move}} + \sum_{i=1}^{N} J^i_{\text{maneuver}} + \sum_{i=1}^{N_{\max}} f(d^i_{\min}) \cdot w_{\text{dist}}$$

1. *Loss of separation* (LOS) *cost:* n^i_{LOS} is the number of losses of separation involving aircraft i in the current sector with its current flight plan. ΔT^i_{LOS} is the time until the first loss of separation for aircraft i.

2. *Boundary condition (BC) breach cost:* T^i_{breach} is the time by which an aircraft violates the ΔT time separation constraint from its predecessor (set to zero if the two aircraft are separated by more than ΔT).

3. *Delay cost:* $TOA^i_{\text{pred}} - TOA^i_{\text{real}}$ accounts for the difference between predicted and actual time of arrival (TOA) at the last waypoint of the flight. Positive delays are penalized; earlier arrivals are favored. TOA^i_{pred} and TOA^i_{real} are computed by integration of the flight plans for each aircraft.

4. *Aircraft actuation cost:* N_{moved} accounts for the number of flight plan modifications chosen in the current solution. Large N_{moved} are penalized (the solution chosen by the ATC is often the simplest).

5. *Maneuver cost:* J^i_{maneuver} accounts for the cost of the maneuver selected for aircraft i. Not all maneuvers are of equal preference and therefore have different costs. It is as easy for a controller to prescribe a 10% speed change, a VFS or a shortcut. A holding pattern is the least preferred option, for it requires constant monitoring of the aircraft. This reflects in the weight choice: $J^i_{\text{speed change}} \sim J^i_{\text{shortcut}} \sim J^i_{\text{VFS}} \leq J^i_{\text{holding pattern}}$. The ratio $J^i_{\text{holding pattern}}/J^i_{\text{speed change}}$ is on the order of 10.

6. *Minimal distance cost:* $f(d^i_{\min})$ penalizes aircraft distributions in which aircraft are closely spaced (but do not lose separation) against more sparse distributions. Here, $\text{dist}_{\max} = 7\text{nm}$.

$$f(d^i_{\min}) = \frac{1}{d^i_{\min}} \cdot w_{\text{dist}} \quad \text{if } d^i_{\min} < \text{dist}_{\max}$$
$$f(d^i_{\min}) = 0 \quad \text{otherwise.}$$

In order to reflect the three levels of priority of the human ATC stated earlier, the weights shown in the cost function J are: $w_{\text{LOS}} \sim 10^{300} \gg w_{\text{breach}} \sim 10^4 \gg$ other weights ~ 10. Thus a computation which tries to minimize J will first deal with losses of separation, then metering conditions, and finally optimization of the flow. An example of a cost landscape for a given topology of two aircraft is illustrated in Figure 13.

In order to reduce the computational time, N_{choice} is defined as the maximum number of aircraft considered for maneuvering by the simulated controller in each time iteration. N_{choice} is restricted to be generally less than the actual number of aircraft per sector. This term is a trade-off between run-time and control-quality and in the simulations it was set in the range of 4 to 8, depending on the targeted goals. Aircraft are selected according to the following rule: aircraft involved in LOS are selected first, then aircraft breaching boundary conditions, and finally remaining aircraft until the selection list has reached N_{choice} aircraft, or until there are no more aircraft to select. In practice, $4 \leq N_{\text{choice}} \leq 8$, where $N_{\text{choice}} = 8$ enables more complicated situations but makes the code run more slowly. The set of all maneuver combinations for the N_{choice} aircraft is called the maneuver set.

At each iteration of the controller emulation loop, an exhaustive recursive search on the maneuver set is run for the chosen aircraft in order to find a set of N_{choice} maneuvers which minimizes J. The computational complexity of finding the optimal J for N_{choice} aircraft subject to n_{maneuver} possible discrete maneuvers is $O((n_{\text{maneuver}})^{N_{\text{choice}}})$. This cost can be reduced to $O((n_{\text{maneuver}} - 2)^{N_{\text{choice}}})$: (i) the cost of the current maneuver has already been computed at the previous step and thus does not need to be recomputed; (ii) two maneuvers are mutually exclusive (shortcuts), therefore only one needs to be called. Including the cost of checking for conflicts, the total cost of a time iteration becomes:

$$O\left(N^2_{\max} \cdot (n_{\text{maneuver}} - 2)^{N_{\text{choice}}}\right)$$

where $N_{\max}$ represents the total number of aircraft in the sector. Due to both the discretization of time, and the restriction of the search space to a manageable number of aircraft, the search is not guaranteed to find the global optimum. However, as it was shown in [Bayen et al., 2002b] that the search does provide a reasonable approximation of the controller's behavior. By adjusting the two key control parameters, the number of selected

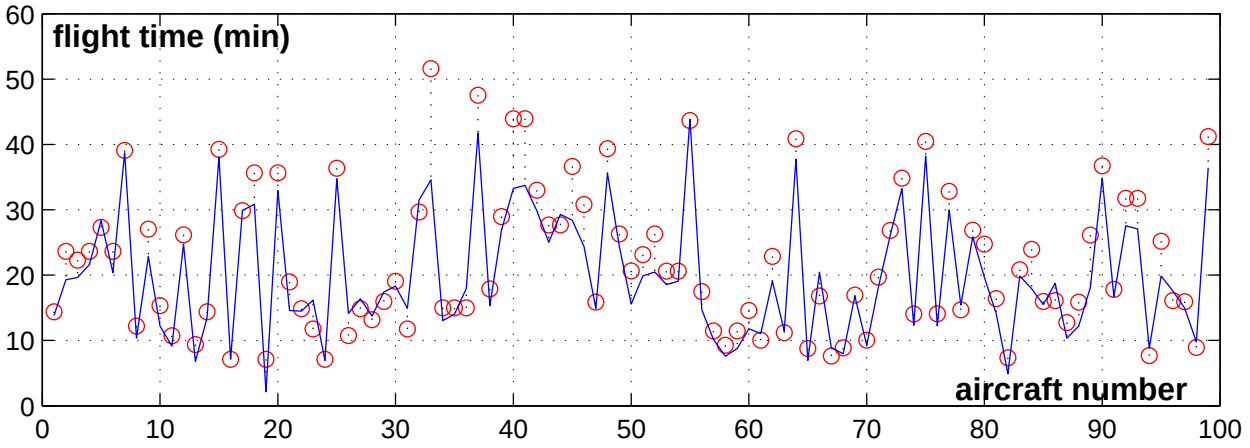

Figure 14: Flight time comparisons for the first 100 aircraft going through sector 33 in the ETMS data set we used. The dots are the flight times for the ETMS recorded points. The solid curve is the result of simulating the model presented here.

aircraft N_{choice} and time between controller activation ΔT_{act}, a transparent trade-off between run-time and control quality can be made.

4.1.4 Validation of the Models

The models of Section 4.1.2 and 4.1.3 were validated against real data in [Bayen et al., 2002b]. A few key points are important in this validation.

Validation of the model of the human Air Traffic Controller The controller logic presented in the previous section is the result of numerous observations made in the Oakland ARTCC, monitoring the work of air traffic controllers. The fact that one can classify ATC action into a set of preferred directives was experimentally validated for a different airspace (see [Histon and Hansman, 2002]). However, even if the automaton of Figure 6 and the cost function of the previous section implemented in the simulator are consistent with observations realized at the Oakland ARTCC, there is no a priori guarantee that these would produce the same effects on the system as a human controller. For this reason, it is assessed how well the cost function describes the decision making of a human controller by validating the code against recorded aircraft trajectories. ETMS data provided by NASA Ames is used.[3]

The data extraction process which enables us to convert ETMS data to a readable format for the simulator is first explained. Three types of validations realized are then presented.

Two types of information can be extracted from the ETMS data: the actual flown aircraft trajectories, and the filed flight plans for each aircraft (eventually updated if modifications are made during the flight). The position is given in latitude / longitude and in terms of navaids, fixes, and jetways represented in Cartesian coordinates, an approximation valid for the portion of airspace of interest to us. The filed flight plan is given in terms of navaids, fixes, jetways, which are also convert into Cartesian coordinates using a public database

[3]Data is collected from the entire population of flights with filed flight plans in the NAS. ETMS data is sent from the Volpe National Transportation System Center (VNTSC) to registered participants via the Aircraft Situation Display to Industry (ASDI) electronic file server. A file containing all recorded data is generated. It displays for each aircraft the current flight data (time, position, speed, heading), as well as the filed flight plan (in terms of navaids, jetways, fixes, etc.). The update rate of the measurements is of the order of one minute.

(`http://www.airnav.com`). Future versions of the simulator might use recently developed ETMS analysis tools such as [LIN and GIFFORD, 2002] to perform these tasks automatically. This study is limited to sector control, and the *Traffic Management Unit* (TMU) action was not implemented. TMU operates at the Center level and makes strategic flow scheduling decisions, which go beyond the range of a single sector controller. It is therefore needed to validate the simulator at a scale at which TMU actuation is already incorporated in the flight plans (typically one or two sectors). Since sectors 32, 33, 34, 15, and 13, are of interest, the actual flight plans are truncated, and only points in that sector are kept. The filed flight plans are truncated similarly. Their estimated time of entrance in the sector is set to the actual time of entrance of the closest actual recorded point. The entrance point in the computational domain is taken to be the closest conflict-free point to the intersection of the flight plan and the boundary of the corresponding sector. The altitude assigned to the filed flight plans is set to the average altitude of the actual trajectory in that sector.

(i) Comparison of flight times. The first 100 aircraft of the ETMS data set which are flying above 33000ft for more than 6 minutes are selected. Their recorded trajectories are extracted as sequences of waypoints which are used as initialized flight plans for the simulations. One then compares the flight times in the simulation to the actual ones. The experiment is run for the following set of speeds: $M \in \{0.6, 0.7, 0.8\}$ (M is the Mach Number), which matches observations in the data for this altitude. In the run, the controller is activated every $\Delta T_{\text{act}} = 10$sec. The results are shown in Figure 14. Two main conclusions can be made: *(i)* the simulator is able to recreate the flow without major modification, and eventually resolves apparent conflicts in the data – these conflicts can be due to inaccuracy of the measurements or transmission (two aircraft separated by less than 5 nm at the same altitude), or due to problems of interpolation when speed changes in time; *(ii)* the time comparison (Figure 14) shows relatively good matching. The flight times provided by the simulator are usually shorter than in reality because by default the simulator will always try to maximize the throughput in the sector. The mean deviation is 120 sec. for flight plans with an average duration of 1300 sec. (9.2% error).

(ii) Validation of conflict resolution. Aircraft flying through sector 33 in a time frame of 10 hours are selected (a total set of 314 aircraft) and the corresponding flights are simulated. The filed flight plans are not conflict-free. It is needed to show that the simulator is able to provide a conflict-free environment. For this run, the activation time of the controller is $\Delta T_{\text{act}} = 20$sec.[4] The set of speeds allowed is $M \in \{0.55,\ 0.75,\ 0.89\}$ (since the full range of altitude is considered, one needs to consider the full range of speed). The simulator is able to provide a conflict free environment. During the simulation, it has to actuate 50 different aircraft. The number of resolved conflicts can be assumed to be on the same order, since a single intervention will usually resolve not more than one conflict.

(iii) Validation of maneuver assignments. The core of the simulator is the model of the human controller by a decision procedure based on the cost function described previously. The validation so far has shown the correlation of flow patterns generated by the code and these observed in reality. It is also needed to assess the validity of the decision procedure. For this, one identifies conflict resolution maneuvers, which are typically obtained by identifying deviations from the filed flight plan. For these maneuvers, the following input data is generated: all aircraft are assigned their actual flight plan (recorded trajectories), and the

[4] $\Delta T_{\text{act}} = 10$sec. or $\Delta T_{\text{act}} = 20$sec. is on the order of the maximal actuation rate of a controller. ΔT_{act} is chosen to be 20sec. in this particular case because of the duration of the computation (10 hours of real time are simulated).

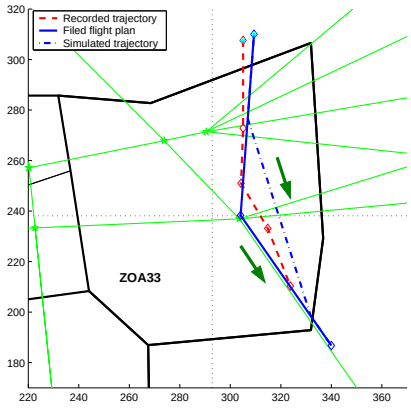

Figure 15: An example of maneuver caused by conflict resolution, reproduced by the simulator. The recorded data (dashed) exhibits an actual shortcut from the filed flight plan (solid). The simulated trajectory (dashed-dotted) is a shortcut of the same type.

aircraft for which the maneuver is identified is assigned its filed flight plan (a set of way points). The simulator is thus put in the same situation as the human controller, in which it has to make the decision that was actually taken. For sector 33, 20 distinct maneuvers out of 300 examined flight plans were identified. The simulator reproduced 16 of them.[5] An example is shown in Figure 15.

Validation of the analytical predictions In this section, the analytical predictions are validated against the real system, by comparing them with simulations. An example of two backpropagating shocks is presented, and solved with an extension of the method explained in Section 4.1.2. Two platoons, each at 10 miles-in-trail, are respectively subjected to 15 miles-in-trail and 20 miles-in-trail outflow boundary conditions (the boundary conditions for platoon 2 start after all aircraft of platoon 1 have reached the TRACON at time $t = 4300$). The speeds are $M \in \{0.59,\ 0.75,\ 0.89\}$. The resulting aircraft flows for this analytic solution are shown in Figure 16. The results and interpretations are shown in Figures 18 and 17. Two shocks appear successively. From Figure 18, one can see that within the second platoon, the first twelve aircraft need to be actuated within the Oakland Center, whereas the last eight need to be actuated upstream (Salt Lake Center). Since in general, no knowledge of the required boundary conditions is propagated upstream, one can predict that in the real system, the last eight aircraft would not be actuated until they enter the Oakland Center and that no solution to this metering problem would be found without putting the aircraft on hold.

This is verified by simulating this flow. In Figure 17, one can see that for the last eight aircraft, the first activation time in the simulator is higher than the predicted (upper plot): the simulated controller is not able to actuate the aircraft on time, because they are not in its airspace. In the middle plot, one can see that these aircraft are breaching the boundary

[5]Small-scale maneuvers are less likely to be executed correctly by the simulator because the probability of selecting the respective aircraft at exactly the 'right' time is small, which explains the small discrepancy between the results. Also, even if the maneuver is executed correctly by the simulator, the resulting flight plan will look different from the ETMS data, since the simulator is restricted to a single angle of deviation ($\theta = 22.5°$).

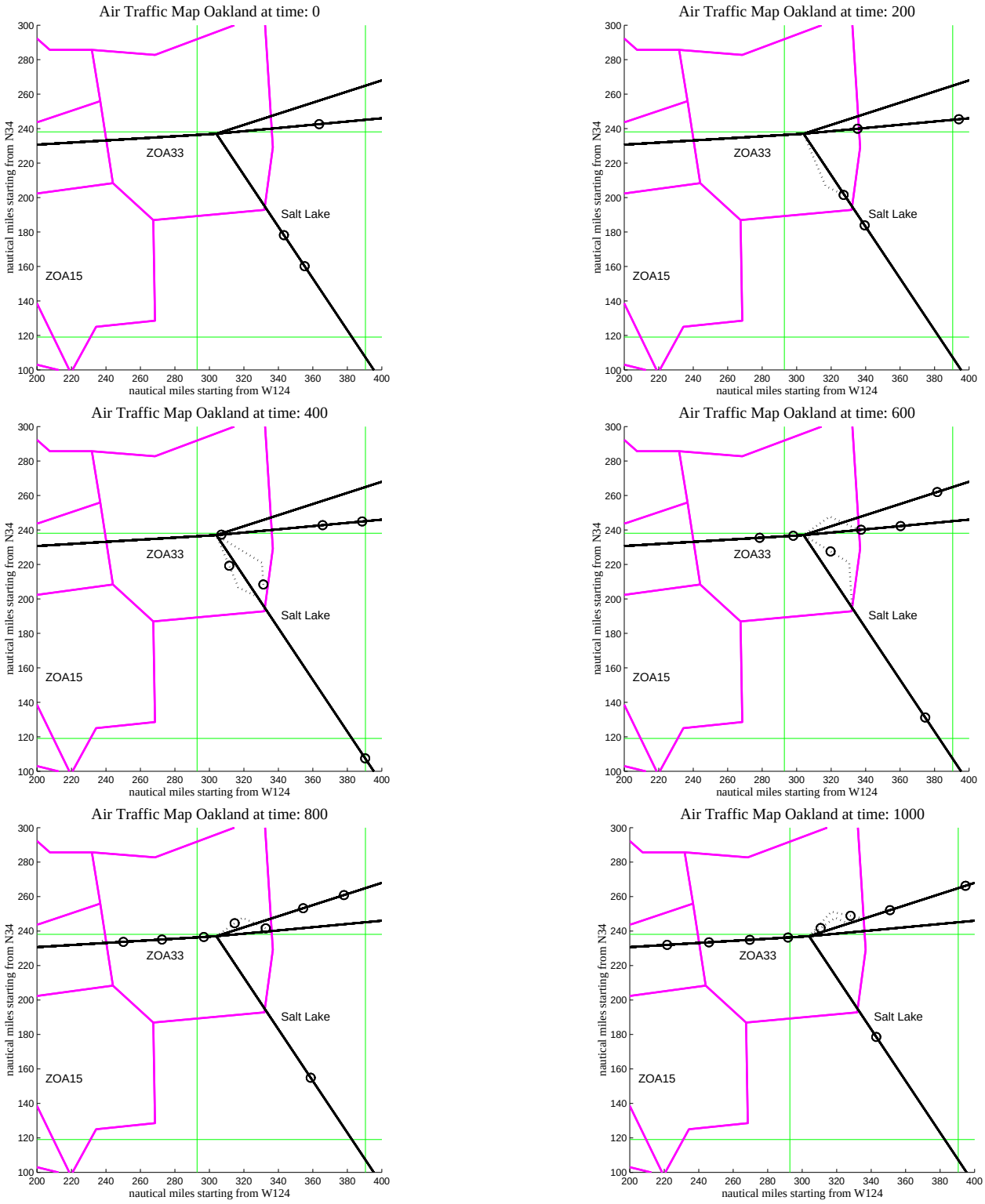

Figure 16: Sector 33: traffic flow for the merging traffic simulation of Figures 18 and 17. The radius around the aircraft is 2.5 nm. The solid lines represent the aircrafts' flight plan. The dotted lines correspond to maneuvers assigned by the simulator. The simulator makes extensive use of VFS to meter the aircraft.

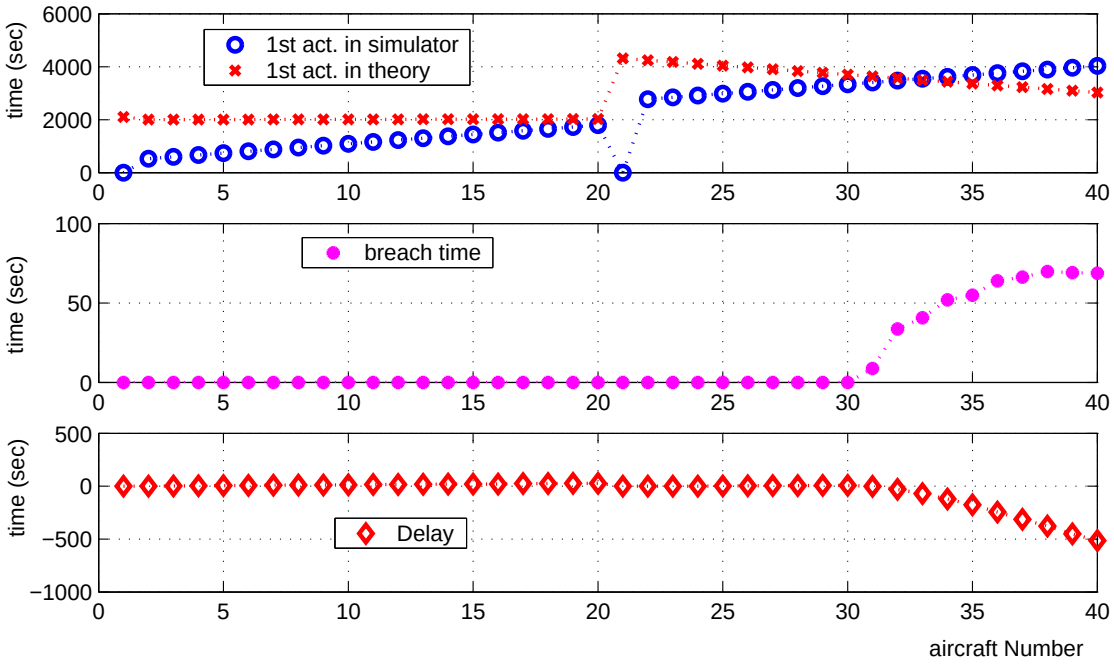

Figure 17: First actuation times of the aircraft (upper plot), simulated and predicted; breaches in metering conditions (middle plot), simulated; delays (lower plot), simulated, for the case of the two platoons of Figure 18.

conditions (by about one minute each), which can also be seen in the lower plot: their delays (the inverse of the cumulated breaches) become negative, i.e. they arrive in advance. This is an illustration of distributed and decentralized control: the actuation occurs in different sectors, and the only communication between the sectors is through the metering conditions. Obviously, the lack of centralized actuation (here communication and strategic TMU planning) disables efficient flow scheduling.

4.2 Onboard and Ground Automation

Several new technologies are under development and certification, or already partially used, and are fueling a change in the structure of ATM.

4.2.1 TCAS: Onboard Collision Avoidance System

The *Traffic Alert and Collision Avoidance System* (TCAS) is an instrument integrated into the avionics systems in the cockpits of civilian jetliners with more than 30 passenger seats. It consists of hardware and software that together provide a set of monitoring systems so the pilot is alerted when any other aircraft comes into the direct vicinity of his aircraft. Part of the TCAS capability is a display showing the pilot the relative positions and velocities of aircraft close to causing a LOS or collision. The instrument sounds an alarm when it determines that another aircraft will pass too closely to the subject aircraft. It issues a corresponding protocol to follow in order to avoid the conflict (or collision). TCAS provides a backup to the ATC process of separating aircraft. During the recent midair crash above Überlingen (Germany) between a DHL Being 757-200 (DHX 611) and a Bashkirian Airlines Tupolev 154 (BTC 2937), on July 1st, 2002. Short before the crash, TCAS issued an escape maneuver to both aircraft to avoid the collision. Due to the intervention of the an Air Traffic Controller, one of the aircraft did not follow TCAS, while the other was conforming which

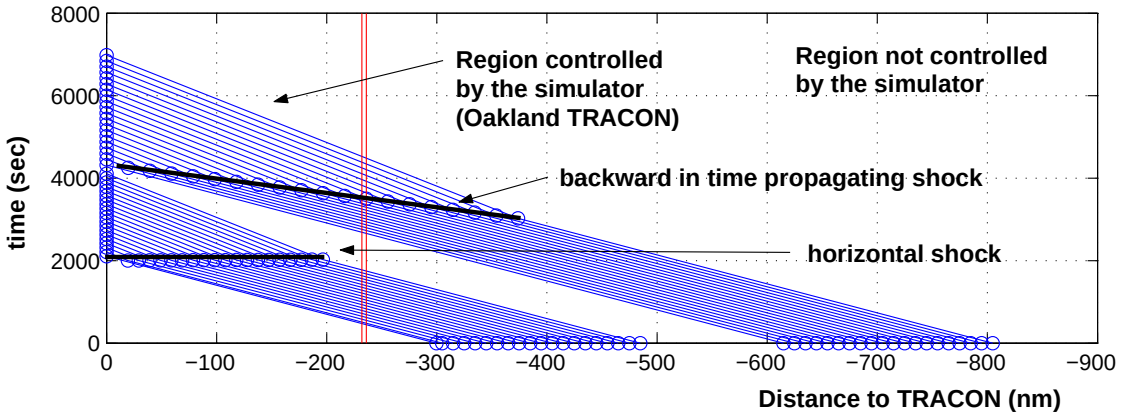

Figure 18: Shocks generated by two successive platoons. The first shock is steady in time (it only propagates backward in space). It corresponds to a piling up process on a highway where all vehicles slow down at the same time. The second shock propagates backward in space and time (which is much harder to handle in practice, because actuation must be performed upstream first).

resulted in the crash. This example is an illustration of difference of procedures between countries; the rule in the EU is that in case of conflict between TCAS and the Air Traffic Controller, the pilot is supposed to follow TCAS. However, one of the pilot was trained according to different rules originating from the former Soviet Union, which prescribe that ATC has precedence over TCAS.

4.2.2. Ground Automation Functionalities

An area of current activity is the development of decision support tools for air traffic controllers. One such tool is the *Center-TRACON Automation System (CTAS)* which is currently in use and a research focus at NASA Ames [Erzberger et al., 1993]. CTAS is software code which runs on computer workstations next to the air traffic controller; it uses radar data, current weather information, standard landing patterns, aircraft flight plans and simplified dynamic aircraft models to predict the aircraft trajectories, alert the controllers about potential conflicts, and provide advisories to the controller about landing sequences. A CTAS installation at the Dallas/Fort Worth Airport (DFW) has improved the sustained landing rate by 10%. CTAS contains two primary tools: the *Final Approach Spacing Tool* (FAST), used by low altitude controllers near airports, and the *Traffic Management Advisor* (TMA), for ATC managing flow further away from the airports.

The Center air traffic controllers and *Traffic Management Coordinators* (TMC) control arriving aircraft that enter the Center from an adjacent Center or depart from feeder airports within the Center. On the basis of the current and future traffic flow, the TMC creates a plan to deliver the aircraft, safely separated, to the TRACON at a rate that fully subscribes, but does not exceed, the capacity of the TRACON and destination airports. The TMC's plan consists of sequences and *scheduled times of arrival* (STA) at the meter fix, published points that lie on the Center-TRACON boundary. The Center air traffic controllers issue clearances to the aircraft in Center so that they cross the meter fixes at the STAs specified in the TMC's plan. Near the TRACON, the Center controllers hand the aircraft off to the TRACON air traffic controllers.

More information about the CTAS tools is available at `http://www.ctas.arc.nasa.gov/`.

4.3 Open Problems for Automation

4.3.1 Conflict Detection and Avoidance

Conflict detection and avoidance has received much attention in the last decade, motivating numerous research efforts. In the current system, conflict resolution is operated manually by ATC, using a simple flight plan extrapolation tool which operates on pairs of aircraft, as well as experience-based conflict avoidance protocols acquired after years of training and practice. Conflict detection and resolution is the main task of a sector controller, and there are many ways to solve a particular conflict. It is however possible to outline some trends observed among numerous controllers.

The most frequently encountered type of conflict consists of two aircraft with intersecting respective flight plans, (see Figure 19). The most popular maneuvers used to solve this type of conflicts are: heading change (ATC will alter the heading of one of the aircraft), or altitude change (ATC will change the cruising altitude of one of the aircraft). Most of the time, such a decision is taken five to ten minutes before the potential penetration in the protected zone of one of the aircraft.

A less frequent, but common conflict happens when one aircraft gains on another aircraft following the same course. This is often the case in *converging flows* (flows close to a destination airport, where aircraft are lined up before or during descent, in preparation for landing). For this type of conflict, ATC's preferred action is either speed change or VFS.

Numerous protocols have been proposed to automate this problem in the current system. In a recent survey [Kuchar and Yang, 2000], the authors classify the different protocols according to the following four criteria.

- *State propagation:* Conflict detection between two aircraft is based on anticipation (propagation) of the state (position, altitude and flight parameters) of the two aircraft. This propagation can be *nominal*: the two aircraft are assumed to follow their nominal course; *worst-case*: all possible trajectories of both vehicles are taken into account; and *probabilistic*: a probability is associated to each potential future state.
- *State dimension.*
- *Conflict detection:* Different metrics are available to assess if a situation is a conflict or not. Some example metrics include predicted minimum separation or estimated time to closest point of approach.
- *Conflict resolution:* The conflict resolution can be *prescribed*: when the conflict is detected, the aircraft follows a predefined protocol; *optimized*: the aircraft follows a path computed using an optimization criterion; *force field*: the aircraft follows a path generated by a force field method, which is a mathematical technique based on an analogy with electromagnetic forces; *manual*: the human generates his own resolution mechanism; or *no resolution*: the method only detects the conflict but does not provide a solution method.

In order to illustrate the previous classification, we show a particular protocol proposed for conflict avoidance in [Mitchell et al., 2002, Bayen et al., 2003, Tomlin et al., 2003] (the background for this example is presented in detail in [Tomlin et al., 2003] in this volume). This particular method may be classified as worst case / three dimensional /

predicted minimum separation / optimized. It models the conflict as a *differential game* [ISAACS, 1965], by accounting for the worst case, in which one of the aircraft (pursuer) tries to *cause* a LOS, while the other aircraft (evader) tries to escape from a LOS. The result of the differential game analysis, and some new computational algorithms presented in [MITCHELL et al., 2002], is a *reachable set* enclosing the evader's protected zone. If the pursuer is inside the reachable set, then it is always able to cause a LOS, no matter what the evader does; if the evader is outside the set, there does always exist an escape maneuver for the evader. Furthermore, this escape maneuver may be computed using optimal control techniques [MITCHELL et al., 2002]. This approach is illustrated in Figure 19, in which, for a pair of aircraft progressing through the Oakland center (obtained from ETMS data), the slice of the reachable set is shown around the evader (chosen arbitrarily to be one of the aircraft, though each aircraft tested is assigned both evader and pursuer roles in sequence). In this figure, one can see that this method may can be used for conflict detection. As soon as the pursuer enters the reachable set around the evader, the controller assesses the LOS threat. Even though the method of [BAYEN et al., 2003, MITCHELL et al., 2002] used for these computations enables one to solve the conflict in the horizontal plane, the particular choice of ATC facing this situation was to climb the evader.

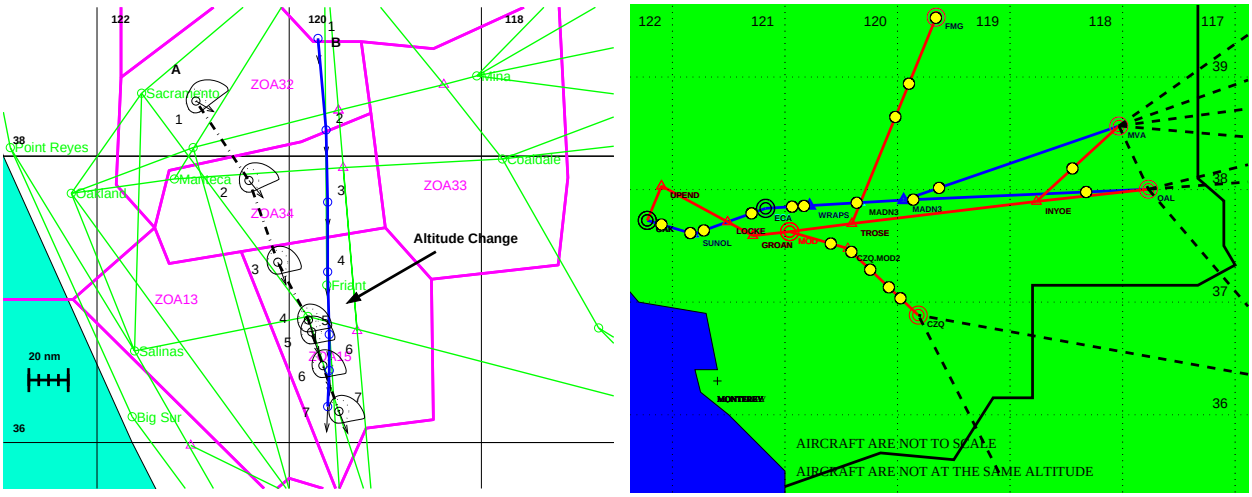

Figure 19: **Left:** Example of conflict detection realized using the algorithm presented in [BAYEN et al., 2003, MITCHELL et al., 2002]. The Figure shows the successive positions of the two aircraft and corresponding reachable sets at different times. The aircraft with the dash-dotted flight plan (labeled A) detects a threat from aircraft B (solid flight plan) for the position labeled 4. At this instant, aircraft B is inside the reachable set. The recorded flight data shows that the human ATC detected a potential LOS and chose to resolve the conflict by climbing aircraft A from 33,000 ft to 37,000 ft, and maintaining aircraft B at 33,000 ft. **Right:** Example of traffic optimization problem. All aircraft displayed in the map are heading to Oakland (OAK). A constraint is imposed at the airport, for example the airport will not accept more than one aircraft every $\Delta^* = 2$min. The problem is to find if there exist a maneuver assignment for each aircraft in the automaton of Figure 6 which satisfies the Δ^* constraint without putting the aircraft into a holding pattern. In the case of a too high inflow into the airport, holding patterns have to be used, and the problem is to minimize the number of holding patterns per aircraft.

4.3.2 Traffic Optimization

In the direct vicinity of airports, traffic concentration is high, which can lead to *airspace saturation*, meaning that the FAA-allowed number of aircraft per sector is exceeded. This is particularly crucial near airports such as San Francisco or Chicago, very often subject to sudden weather changes which may trigger the temporary closure of one (or all) runways.

Recent software tools exist which, given a configuration of many aircraft entering the vicinity of the airport airspace, compute the ordering and spacing of these aircraft so that the intervehicle spacing constraints are satisfied, and all aircraft satisfy their speed and path constraints. A description of the core of one such algorithm (in [BAYEN and TOMLIN, 2003]), which uses *Mixed Integer Linear Programming (MILP)* at its core, follows. In the example, the algorithm is described for the approach pattern and corresponding navigation beacons for the Oakland airport, detailed in Figure 4.

1. Map the automaton of Figure 6 to the automaton of Figure 20, as described in the caption of Figure 20. This task is computationally inexpensive.

2. Pose the arrival scheduling problem as an optimization problem. The details of the transformation are presented in [BAYEN and TOMLIN, 2003]. Consider a set of N aircraft converging to a single airport as in Figure 8. From the automaton of Figure 20, it is possible to compute the set of possible arrival time t_i for aircraft i (for all $i \in \{1, \cdots, N\}$) into the destination airport:

$$\begin{array}{lll} \textbf{Maximize:} & \Delta & \\ \textbf{Subject to:} & t_i \in S_i := \bigcup_{k=1}^{n_i} [a_k^i, b_k^i] & \forall i \in \{1, \cdots, N\} \\ & |t_i - t_j| \geq \Delta & \forall (i,j) \in \{1, \cdots, N\}^2, \text{ s.t. } i > j \end{array} \tag{10}$$

where $t_i,\ a_k^i,\ b_k^i\ \in \mathbb{R}$. S_i represents the union of achievable intervals of arrival times for aircraft i, and $n_i \in \mathbb{N}$ is the number of such intervals for aircraft i. For example, if the only achievable arrival time intervals for aircraft 3 are [7.23 min, 8.2 min] using a given route, and interval [10.23 min, 11.2 min] using the same route plus a 3 min loop, then $S_3 = [7.23, 8.2] \cup [10.23, 11.2]$ and $n_3 = 2$. The second constraint means that two aircraft are separated by at least Δ time units at the destination airport. The objective is to maximize Δ (since most of the time, the restriction is given at the destination airport). For example, if the airport cannot admit more than one aircraft every minute, the optimal value Δ^* of (10) must be at least 1 min. This problem is not convex, yet it can be transformed into a MILP:

$$\begin{array}{lll} \textbf{Maximize:} & \Delta & \\ & & \\ \textbf{Subject to:} & t_i \geq a_1^i & \forall i \in \{1, \cdots, N\} \\ & t_i \leq b_{n_i}^i & \forall i \in \{1, \cdots, N\} \\ & t_i \geq a_{k+1}^i - D d_{ik} & \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, n_i - 1\} \\ & t_i \leq b_k^i + D(1 - d_{ik}) & \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, n_i - 1\} \\ & d_{ik} \in \{0, 1\} & \forall i \in \{1, \cdots, N\}, \forall k \in \{1, \cdots, n_i - 1\} \\ & t_i - t_j \geq \Delta - C c_{ij} & \forall (i,j) \in \{1, \cdots, N\}^2, \text{ s.t. } i > j \\ & t_i - t_j \leq -\Delta + C(1 - c_{ij}) & \forall (i,j) \in \{1, \cdots, N\}^2, \text{ s.t. } i > j \\ & c_{ij} \in \{0, 1\} & \forall (i,j) \in \{1, \cdots, N\}^2, \text{ s.t. } i > j \end{array} \tag{11}$$

There is no unique way to transform (10) into (11) (some encodings of the optimization into a MILP might be more efficient computationally than others, and this encoding is not necessarily the best).

3. From the solution $(\Delta^*, t_1^*, \cdots, t_N^*)$ of the MILP, compute the set of modes of the automaton of Figure 20 (and then of Figure 6) which achieve these times of arrival and spacing.

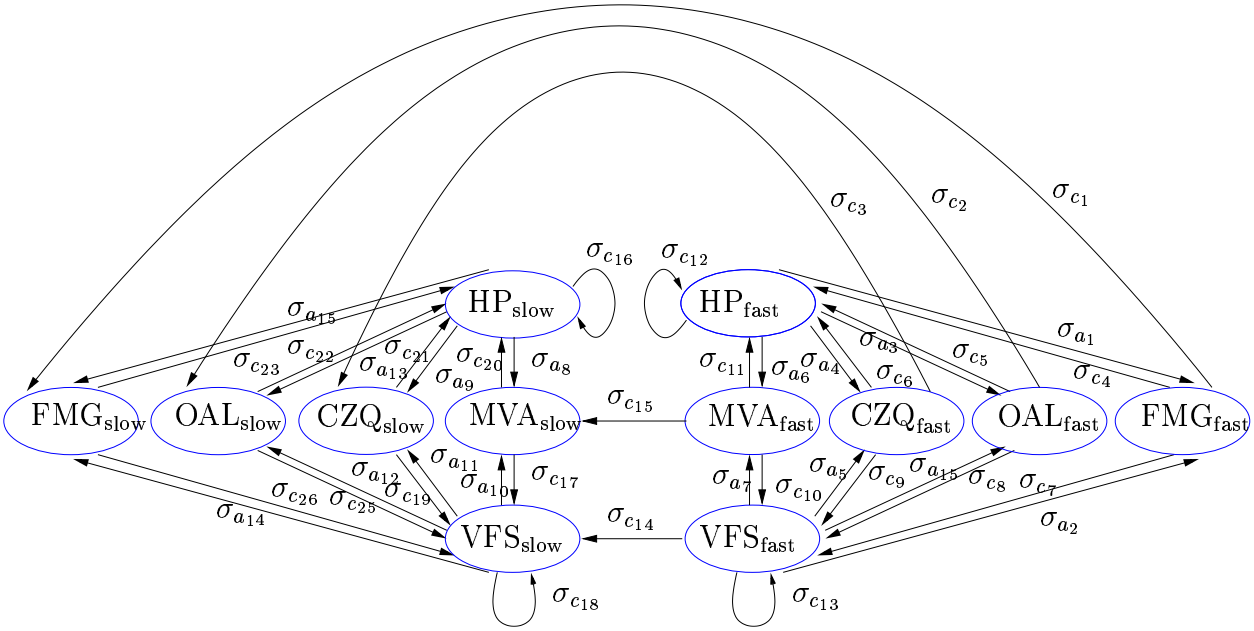

Figure 20: Hybrid automaton used in [Bayen and Tomlin, 2003] for timing each aircraft precisely along its last 200 mile s before landing. It is obtained from Figure 6 by duplicating each mode of that Figure as many times as there are ways to reach the destination airport (SFO) of Figure 8, accounting for the possible speed assignments available for this portion of the flight. The modes for the particular case of Figure 8 are: OAL$_{\text{slow}}$, MVA$_{\text{slow}}$, CZQ$_{\text{slow}}$, FMG$_{\text{slow}}$, OAL$_{\text{fast}}$, MVA$_{\text{fast}}$, FMG$_{\text{fast}}$, CZQ$_{\text{fast}}$, HP$_{\text{fast}}$, VFS$_{\text{fast}}$, HP$_{\text{slow}}$, VFS$_{\text{slow}}$. The first eight modes correspond to the dynamics of the aircraft incoming into the arrival from any entry waypoint (acronym), at a speed (fast or slow). The rest of the modes are HP (holding patterns) and VFS (vector for spacing), respectively, at fast and slow speed. Two types of switches are possible for this automaton. The first set (σ_{c_p}, where p is an integer) is generated by ATC: for example "slow down on arrival MOD.LOCKE1 from MVA" ($\sigma_{c_{15}}$), "hold at slow speed on arrival MOD.LOCKE1 from OAL" ($\sigma_{c_{22}}$). The second set (σ_{a_p}, where p is an integer) is generated by the airspace: for example at the end of a slow HP on MOD.LOCKE1 from CZQ, retrieve original course (σ_{a_9}).

In this method, all of the tasks except the resolution of the MILP are performed in *polynomial time*. This means that it is possible to bound the computation time required to complete these tasks with a polynomial of the variable N. In the case considered here, the degree of the polynomial is low enough that the number of aircraft considered is not a limiting factor in the method. However, the numerical solution of (11) is in general computable in *exponential time*, which means that in most of the cases, an optimization software trying to solve (11) might have to explore all cases before finding the optimum, and the number of such cases is exponential in the number of aircraft. It is crucial for an online implementation of such a method to have bounded computational time, because this gives a bound on the time a human ATC would have to wait for such a tool to prescribe a maneuver assignment. In [Bayen and Tomlin, 2003], several cases of (11) are identified. For certain cases, it is possible to prove that the solution can be computed in polynomial time. For example, if $\forall i \in \{1, \cdots, N\}$, $n_i = 1$, the MILP can be solved in polynomial time. If furthermore one assumes that the order of the aircraft is fixed, the following Δ^* solves the problem [Bayen and Tomlin, 2003], and the construction of a schedule may be performed in linear time:

$$\Delta^* = \min_{(i,j)\in\{1,\cdots,N\}^2,\ i<j} \left(\frac{b_1^j - a_1^i}{j-i} \right) \tag{12}$$

This result can be used to compute Δ^* for $n_i \geq 1$ in the case in which the order of the aircraft is fixed. The algorithm is still polynomial.

5. Conclusions

A brief description of the current ATC system has been presented, and areas where modeling and analysis may contribute to system improvements discussed. A control theoretic model of sector-based traffic flow using hybrid automata theory has been presented, and a subset of this model has been used to generate Lagrangian analytic predictions of the traffic flow (dynamic sector capacity, extend of traffic jams). These results in turn were used to predict the conditions under which airspace saturation cannot be treated at the level of a single sector, but requires centralized actuation (communication and strategic TMU planning). These predictions were validated against an abstraction of the real system (a simulator using the full model, which was validated against ETMS data). After a short description of some existing tools for automating aircraft collision avoidance and ground ATC functionality, open problems for automation of ATC were presented. Finally, recent results for collision detection and resolution using game theoretic methods for hybrid systems, and for automating the arrivals of aircraft into busy airports using mixed integer linear programming, were outlined.

Acknowledgments

The authors would like to acknowledge the help of the researchers at NASA Ames, in particular: George Meyer for his innovative ideas and support, Gano Chatterji for his initial suggestions concerning the construction of the shock wave, Banavar Sridhar for fruitful discussions about modeling this problem, Shon Grabbe for his help on FACET and ETMS data, and Karl Bilimoria for his recommendations concerning the converging flow. The help of Pascal Grieder and Shriram Santhanam in this work is also acknowledged.

References

[Bayen et al., 2002a] Bayen, A. M., Grieder, P., and Tomlin, C. J. (2002a). A control theoretic predictive model for sector-based air traffic flow. In *Proceedings of the AIAA Conference on Guidance, Navigation and Control*, Monterey, CA.

[Bayen et al., 2002b] Bayen, A. M., Grieder, P., Tomlin, C. J., and Meyer, G. (2002b). Control theoretic Lagrangian models for sector-based air traffic flow. Submitted Dec. 2002 to the *AIAA Journal on Guidance, Dynamics and Control.*

[Bayen et al., 2003] Bayen, A. M., Santhanam, S., Mitchell, I., , and Tomlin, C. J. (2003). A differential games formulation of alert levels in ETMS data for high altitude traffic. In *AIAA Conference on Guidance, Navigation and Control*, Austin, TX.

[Bayen and Tomlin, 2003] Bayen, A. M. and Tomlin, C. J. (2003). Real-time discrete control law synthesis for hybrid systems using MILP: applications to congested airspaces. In *Proceedings of the American Control Conference*, Denver, CO.

[Bertsimas and Stock Patterson, 1998] Bertsimas, D. and Stock Patterson, S. (1998). The air traffic flow management problem with enroute capacities. *Operations Research*, 46:406–422.

[Bilimoria, 2000] Bilimoria, K. (2000). A geometric optimization approach to aircraft conflict resolution. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Denver,CO.

[Bilimoria and Lee, 2001] Bilimoria, K. and Lee, H. (2001). Properties of air traffic conflicts for free and structured routing. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Montreal, PQ.

[Bilimoria et al., 2001] Bilimoria, K., Sridhar, B., Chatterji, G., Seth, K., and S.Graabe (2001). FACET: Future ATM concepts evaluation tool. In *Proceedings of the 3rd USA/Europe Air Traffic Management R&D Seminar*, Naples, Italy.

[Dugail et al., 2002] Dugail, D., Feron, E., and Bilimoria, K. (2002). Conflict-free conformance to En-Route flow-rate constraints. In *Proceedings of the AIAA Conference on Guidance, Navigation and Control*, Monterey, CA.

[Dugail et al., 2001] Dugail, D., Mao, Z.-H., , and Feron, E. (2001). Stability of intersecting aircraft flows under centralized and decentralized conflict avoidance rules. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Montreal.

[Erzberger et al., 1993] Erzberger, H., Davis, T. J., and Green, S. (1993). Design of Center-TRACON Automation System. In *Proceedings of the AGARD Guidance and Control Symposium on Machine Intelligence in Air Traffic Management*, pages 11.1–11.12, Berlin, Germany.

[Gazit, 1996] Gazit, R. Y. (1996). *Aircraft Surveillance and Collision Avoidance using GPS*. PhD thesis, Department of Aeronautics and Astronautics, Stanford University.

[Histon and Hansman, 2002] Histon, J. and Hansman, R. J. (2002). The impact of structure on cognitive complexity in air traffic control. Technical Report ICAT-2002-4, MIT.

[Isaacs, 1965] Isaacs, R. (1999 (1965)). *Differential Games*. Dover (reprint from John Wiley).

[Jackson and Green, 1998] Jackson, J. W. and Green, S. (1998). Control applications and challenges in air traffic management. In *Proceedings of the 1998 American Control Conference*, Philadelphia, PA.

[Jeppesen, 2000] Jeppesen (2000). High altitude enroute charts. http://www.jeppesen.com.

[Kahne and Frolow, 1996] Kahne, S. and Frolow, I. (1996). Air traffic management: Evolution with technology. *IEEE Control Systems Magazine*, 16(4):12–21.

[Kuchar and Yang, 2000] Kuchar, J. and Yang, L. (2000). A review of conflict detection and resolution modeling methods. *IEEE Transactions on Intelligent Transportation Systems*, 1(4):179–189.

[Lin and Gifford, 2002] Lin, J. C. and Gifford, D. K. (2002). VisualFlight: The air traffic control data analysis system. Technical report, MIT.

[Mao et al., 2000] Mao, Z., Feron, E., and Bilimoria, K. (2000). Stability of intersecting aircraft flows under decentralized conflict avoidance rules. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Denver,CO.

[Menon et al., 2002] Menon, P. K., Sweriduk, G., and Bilimoria, K. (2002). A new approach for modeling, analysis and control of air traffic flow. In *Proceedings of the AIAA Conference on Guidance, Navigation and Control*, Monterey, CA.

[Mitchell et al., 2002] Mitchell, I., Bayen, A. M., and Tomlin, C. J. (2002). Computing reachable sets for continuous dynamic games using level set methods. Submitted January 2002 to *IEEE Transactions on Automatic Control*.

[Newell, 1993] Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research*, 27(4):281–287.

[Nilim et al., 2001] Nilim, A., El-Ghaoui, L., Duong, V., and Hansen, M. (2001). Trajectory-based air traffic management (TB-ATM) under weather uncertainty. In *Proceedings of the 4rd USA/Europe Air Traffic Management R&D Seminar*, Santa Fe, NM.

[Nolan, 1999] Nolan, M. S. (1999). *Fundamentals of Air Traffic Control, 3rd Edition*. Brooks/Cole Publishing, Pacific Grove, CA.

[Odoni et al., 1996] Odoni, A., Bowman, J., Delahaye, D., Deyst, J., Feron, E., Hansman, J., Khan, K., Kuchar, J., Pujet, N., and Simpson, R. (1996). Existing and required modeling capabilities for evaluating ATM systems and concepts. Technical Report NAG2-997, MIT.

[Pallottino et al., 2001] Pallottino, L., Bicchi, A., and Feron, E. (2001). Mixed integer programming for aircraft conflict resolution. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Montreal.

[Perry, 1997] Perry, T. S. (1997). In search of the future of air traffic control. *IEEE Spectrum*, 34(8):18–35.

[Plaettner-Hochwarth et al., 2000] Plaettner-Hochwarth, J., Zhao, Y., and Robinson, J. (2000). A modularized approach for comprehensive air traffic system simulation. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Denver, CO.

[Pujet and Feron, 1996] Pujet, N. and Feron, E. (1996). Flight plan optimization in flexible air traffic environments. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, San Diego, CA.

[Shaver, 2002] Shaver, R. D. (2002). The growing congestion in the national airspace system (NAS) : How do we measure it? Are current plans sufficient to constrain its growth? If not, what else can we do? *Air Traffic Control Quarterly*, 10(3):169–195.

[Tomlin et al., 2003] Tomlin, C. J., , Mitchell, I., and Bayen, A. M. (2003). Verification of hybrid systems. In Gobaisi, D. A., editor, *Encyclopedia Of Life Support Systems*. UNESCO. Ref: 6.43.28.6.