Digital Avionics Systems Conference (DASC), October, 2003, Indianapolis, IN

METHODOLOGY FOR THE GENERATION OF AIR TRAFFIC SCENARIOS BASED ON RECORDED TRAFFIC DATA

Robert Oaks, Shurong Liu, David Zhou, Veridian ITS

Mike Paglione, Federal Aviation Administration

William J. Hughes Technical Center, Atlantic City International Airport, NJ 08405

Abstract

This paper presents a methodology to generate air traffic scenarios for the support of the testing and analysis of Air Traffic Management decision support tools. This methodology was developed by and is currently being used by the Conflict Probe Assessment Team, a workgroup within the Federal Aviation Administration's Simulation and Modeling Group located at the Federal Aviation Administration's William J. Hughes Technical Center. The individual flights within the scenarios generated by this methodology follow realistic flight routes, yet the air traffic in these scenarios contains aircraft-to-aircraft conflicts and encounters that do not exist in the field. Scenarios with these characteristics are necessary to evaluate decision support tools that predict conflicts. The paper describes each of the three steps that comprise the methodology. The first step is data extraction during which traffic data, available from various sources and recorded in different formats, is extracted and placed into a set of relational database tables. The second step is data modification, where the data in the tables may be manipulated for test purposes. This manipulation may consist of simply culling undesirable flights, or it may involve using a genetic algorithm to time-shift the flights to induce conflicts or encounters. The third step is scenario generation, where scenarios are created based on the traffic data retrieved from the modified database tables. During this final step the scenarios may be formatted for various target systems. The paper then describes how the Conflict Probe Assessment Team has used this methodology to generate scenarios that have been used for accuracy and risk reduction testing and for a study assessing the effect of weather forecast errors.

Introduction

Traffic flow management decision support tools (DSTs) use simulation as a tool for development, technical assessment, and field evaluation. The air traffic scenarios, used by these simulations, are data files describing the flow of aircraft traffic through an airspace over a period of time. For traffic flow management DSTs, these airspaces are generally those defined for Terminal Radar Approach Control facilities (TRACONs), such as those that manage arrivals and departures around New York and Dallas/Fort Worth, and Air Route Traffic Control Centers (ARTCCs), that manage air traffic as it crosses the country.

The scenario data files contain planning/advisory information and track data. The time-stamped planning/advisory data describe the aircraft's planned flight; it includes its flight plan and flight plan amendments, interim altitude clearances, and hold information. The track data or smoothed radar surveillance data represents the aircraft's actual flight path. It consists of a series of 4-dimensional components: two defining the aircraft's position (using either an XY-coordinate or latitude/longitude pair), another component defining its altitude, and the fourth component being an associated time.

One form of these scenario files is the playback scenario, which is based on the format of intercomputer messages recorded at the field site. Playback scenarios composed of test data specifically designed to test its internal algorithms are used to develop the DSTs. But in order to ensure operational capability, the tools must be tested with playback scenarios based on recorded field data containing the situations they are designed to protect against. Playback scenarios are also used to evaluate DSTs because they provide the most realistic simulation environments.

In 1996 the Federal Aviation Administration's Traffic Flow Management Branch (ACT-250) established the Conflict Probe Assessment Team (CPAT) to evaluate the accuracy of the conflict probes in DSTs. In 2002, CPAT became a part of the Simulation and Modeling Group (ACB-330). CPAT has measured the conflict prediction accuracy of the User Request Evaluation Tool (URET) [Cale et al., 1998], measured the trajectory modeling accuracy of both URET and the Center TRACON Automation System (CTAS) [Paglione et al., 1999], and assisted in the accuracy testing of URET Current Capability Limited Deployment (CCLD) [CPAT, 2000][CPAT, 2001], which is the operational implementation of URET. CPAT is developing risk reduction scenarios that are being used during the deployment of URET CCLD into the ARTCCs, has conducted a study that measured the sensitivity of the URET conflict probe to weather forecast errors [Paglione et al., 2002], and has collaborated with other researchers to develop acceptable methods of altering traffic data for use in evaluating conflict probe performance. Each of these projects required the generation of playback scenarios based on recorded field data. This paper describes the scenario generation process that CPAT developed to support these projects and provides some specific examples of its use.

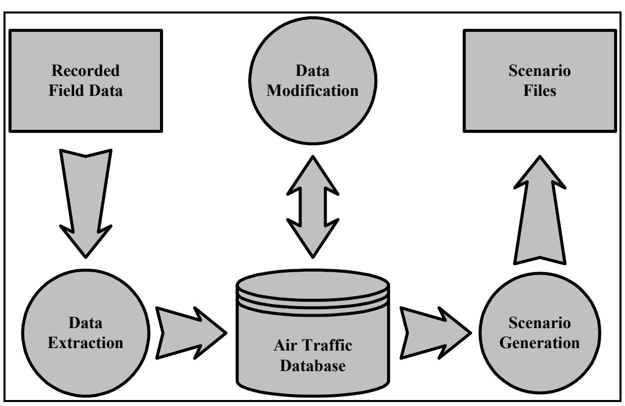

Figure 1: Scenario Generation Process

Scenario Generation Process

As depicted in Figure 1, the scenario generation process developed by CPAT consists of three basic steps: data extraction, data modification, and scenario generation. The goals of this process are (1) to be able to extract data from a variety of field data sources and (2) to be able to provide scenarios in a variety of formats.

Data Extraction

The first step of the scenario generation process uses software developed by CPAT to extract the data from the field recordings and populate an Air Traffic Database on a relational database management system. This database consists of tables grouped into three categories: the environmental tables, the bookkeeping tables, and the flight-centric tables.

- The environmental tables contain center specific and operational data. This includes coordinate conversion constants and preferential routing, sector assignment, and route status information.
- The bookkeeping tables contain data set and run identification information used internally by the software.
- The flight-centric tables contain flight message data such as flight plan information, track data, and controller information messages.

The concept of flight-centricity is key to the scenario generation process. During data extraction a flight is inserted into the database for each unique flight encountered in the recorded field data that had a flight plan message. Each of these extracted flights has a start time, which is the time of the flight's first recorded track message. All other events (e.g., flight plan messages, hold messages, interim altitude messages, and individual track messages) associated with the flight are relative to this start time. Each flight also has a delta time that is used to time-shift the flight. During the data modification step an entire flight can be shifted in time but retain its flight profile merely by adding this delta time to the flight's start time.

Data Manipulation

The second step in the scenario generation process is data modification. This may consist of simply culling undesirable flights, but more typically involves time-shifting the flights to induce conflicts or encounters. Time-shifting consists of three independent techniques: time compression, random time adjustment, and the use of a genetic algorithm. These are described in [Paglione et al., 2003a].

Scenario Generation

The final step in the scenario generation process is the actual generation of the scenarios using the time-shifted data. These scenarios are created in multiple formats, including both binary and ASCII formatted files.

Field Data Processing

The scenario generation process was implemented as a subsystem known as FD_PROC (Field Data Processing). The software contained in FD_PROC has evolved over the years, but can be classified into two distinct efforts: the original implementation and an upgrade.

Original Implementation

A brief overview of CPAT's original implementation of this scenario generation process was presented in [Oaks and Paglione, 2001]. The software consists of UNIX scripts and a number of C/C++ programs. CPAT used g++, the GNU C/C++ compiler, and libg++, the GNU C/C++ libraries. The C++ programs that access the Oracle database use the Oracle Pro*C/C++ Precompiler.

In the original implementation there are twelve database tables. There are four environmental tables: the *fd_airspace* table, which contains the constants required for x-y to lat-long coordinate conversion for a specific air traffic control facility; the *fd_rtix* table, which contains the preferential routes names, indices, and types; the *fd_sector_asgn* table, which contains information associated with the sector assignment messages; and the *fd_route_status* table, which contains information associated with the route status messages. There are two bookkeeping tables: the *fd_data_id* table, which contains information about a specific data extraction and the *fd_run* table, which contains information identifying the data sets used for a specific run. There are six flight-centric tables: the *fd_flight* table, which contains the static information related to a flight; the *fd_flight_plan* table, which contains a history of the flight plans for a flight; the *fd_track* table, which contains the individual track points for a flight; the *fd_int_alt* table, which contains a history of the interim altitude messages for a flight; the *fd_hold* table, which contains a history of the hold messages for a flight; and the *fd_pref_route* table, which contains the converted route information for a flight.

- <u>Data extraction</u> may be performed by either of two application programs: *ext* or *extscn*. *ext* processes a Host Interface Device (HID) recorder file, which is recorded by the URET Daily Use system, which was developed by MITRE. The messages recorded in these files use the Host Computer System 3.20 Patch, which is described in [FAA, 1998].

  extscn processes files in what CPAT calls P320-formatted scenario files. This is an ASCII file in a format compatible with a number of MITRE tools. This format is documented in [Lindsay, 1997] and in Appendix A of [Lindsay, 1998]. This format is also one of the scenario outputs available in the original implementation of the scenario process.

  These two application programs populate the two bookkeeping tables, all of the flight-centric tables, and two of the environment tables: *fd_sector_asgn* and *fd_route_status*.

  Prior to running either of the extraction programs the *fd_airspace* and *fd_rtix* tables must be populated. The *fd_airspace* table is updated manually, while the *fd_rtix* table is updated using the programs *getRTIX* and *getAlpha*. These refer to data that is obtained from the RTIX environment file on the URET DU system and to the alphanumeric data obtained from the ACES (Adapted Controlled Environment System) data. These programs extract and populate the tables with the required preferential route information.

- <u>Data manipulation</u> may be performed by either of two programs: *Odo* or *Cat*. The program *Odo* is a C++ program that accesses the Air Traffic Database to time shift the flights in the *fd_flight* table by altering their start times. The program *Odo* uses scenario time compression and random time adjustment to time shift the flights. Time compression means to apply a constant compression multiplier (C_m) times the difference between the start times of the flight's track and some base time. Random time adjustment means to modify the start time of the flight's track by adding a random delta time.

The program *Cat* is also a C++ program that accesses the Air Traffic Database to time shift flights in the *fd_flight* table. This program uses an implementation of the genetic algorithm to alter the flights' start times. This implementation is documented in [Oaks and Paglione, 2002].

- Scenario generation is performed by the program *sgp*, which is a C++ program that accesses the Air Traffic Database. This process writes ASCII scenario files in either of two formats: the P320 format discussed earlier as an input format or the CMS ASCII format, which is an ASCII interpretation of the Common Message Set (CMS) format used for intercomputer communications, which is documented in [FAA, 2000].

Upgrade

The original FD_PROC implementation described above evolved over a number of years. As a result, it is a very elaborate process involving a number of individual programs and numerous manual processes. In order to improve its maintainability and therefore its utility, it was decided to upgrade the FD_PROC system. This upgrade needed to resolve a number of specific problem areas; these include:

- Since the programs used generic table names it was difficult to shift from one project to another project. It was decided to resolve this problem by appending suffixes to the generic table names upon each project run.
- The existing system had a number of program trouble reports that needed to be resolved. It was decided that it would be better to resolve these problems through an upgrade rather than continue to modify the existing code.
- The existing system is maintained under CVS (Concurrent Versions System); however, the procedure in which it is maintained caused the application data to intermix with the source code to the programs.
- The programs in the existing system required too much manual intervention and had some inherent design flaws. This problem drove the upgrade and resulted in the decision to recode the system in Java, using its robust GUI and Java Database Connectivity (JDBC) capabilities.

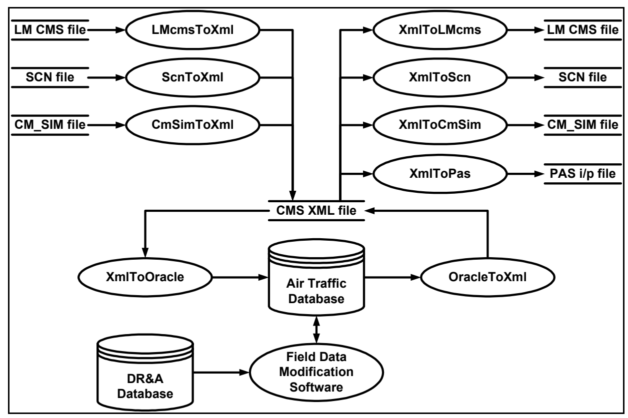

Figure 2: FD_PROC Upgrade Data Flow

Figure 2 shows the data flow design for the upgraded FD_PROC system.

- Data extraction in the upgraded system will be performed by a number of Java programs (shown as *LMcmsToXml*, *ScnToXml*, and *CmSimToXml*) that convert files from various sources (viz. Lockheed Martin CMS, MITRE SCN, and CTAS CM_SIM) and convert them to an XML formatted file. This XML file format is based on and consistent with the CMS Specification [FAA, 2000]. A single Java program (*XmlToOracle*) extracts the data from this XML file and inserts the data in the Air Traffic Database.
- Data manipulation in the upgraded systems in unchanged from the original implementation.
- Scenario generation in the upgraded system consists of a single Java program

(*OracleToXml*) that accesses the Air Traffic Database and creates an XML file. Numerous Java programs (*XmlToLMcms*, *XmlToScn*, *XmlToCmSim*, *XmlToPas*) will be available to convert the XML-formatted file to the various scenario formats.

Examples of Use

URET CCLD Accuracy Testing

The Federal Aviation Administration's Free Flight Phase One Program Office tasked CPAT to supply Lockheed Martin Air Traffic Management (LMATM) Division with scenarios supporting the accuracy testing of the User Request Evaluation Tool Core Capability Limited Deployment (URET CCLD) system. URET CCLD is the operational implementation of the URET DST, which was developed by MITRE and was in daily use at the Indianapolis and Memphis En Route Air Traffic Control Centers (ARTCCs). The scenarios were based on field data recorded at these ARTCCs.

The objective of this accuracy testing was to quantify the missed and false aircraft-to-aircraft and aircraft-to-airspace conflict predictions made by URET CCLD. This required test scenarios containing a significant number of conflicts. However, scenarios generated directly from the recorded field data did not contain conflicts, since the air traffic controllers had managed the airspace to avoid them. Therefore, it was necessary for CPAT to modify the recorded field data to induce the required number of conflicts, while at the same time retaining the realism of the air traffic. CPAT used the FD_PROC system to create these scenarios. In fact, this accuracy testing was the driving force for development of FD_PROC and it was during the development of these scenarios that the original implementation of FD_PROC evolved.

URET CCLD Risk Reduction

During the development of the ZME and ZID scenarios for the URET CCLD accuracy testing, the FAA and LMATM saw a need for additional scenarios for the operational deployment and for other ARTCCs at which URET CCLD was to be deployed. The purpose of these scenarios was to ensure that the URET CCLD was properly adapted for these other sites. CPAT used the original implementation of FD_PROC to generate risk reduction scenarios for the ZID, ZME, Kansas City ARTCC (ZKC), Washington ARTCC (ZDC), and Atlanta ARTCC (ZTL). CPAT is currently working with LMATM to develop risk reduction scenarios for the Jacksonville ARTCC (ZJX), Minneapolis ARTCC (ZMP), and Fort Worth ARTCC (ZFW).

URET Weather Prediction Error Impact Study

In 2002 CPAT designed and performed a quantitative experiment to evaluate the impact of weather forecast errors on URET trajectory and conflict predictions. The experiment used about two hours of traffic data recorded at the Indianapolis en route center in May 1999. The flights were time shifted to generate a sufficient number of test conflicts using a genetic algorithm technique developed by CPAT. This time-shifted scenario was used as input to the URET Prototype. Weather forecast errors were induced into the weather input file (Rapid Update Cycle, RUC) by adding 20 or 60 knots to the wind magnitude, 45 or 90 degrees to the wind direction, and 5 or 15 degrees Kelvin to the air temperature. This produced seven URET runs for the experiment – the unaltered control run and six treatment runs. The analysis compared the control run against the treatment runs. A methodology was developed to compare the trajectory and conflict prediction accuracy of these runs. The results of the study are documented in [Paglione et al., 2002]. In regards to CPAT's scenario development, the study illustrates how the field data was altered to produce a significant number of conflicts in a rather brief two-hour traffic sample. This capability of generating high levels of conflicts with relatively short traffic samples allowed the scenario to be used efficiently by being run repeatedly. This was a requirement for the study with one control run and over six treatment runs through the URET DST.

Methodology for Generating Conflict Scenarios by Time Shifting Recorded Traffic Data

NASA Ames Research Center worked with CPAT to develop a methodology for generating conflict scenarios that can be used as test cases to estimate the operational performance of a conflict probe. Recorded air traffic data was used to preserve real-world errors that affect the performance of conflict probes. However, due to controller actions to separate traffic, such data generally does not contain actual violations of legal separation standards. Therefore, the track data was time shifted to create traffic scenarios featuring conflicts with characteristic properties similar to those encountered in typical air traffic operations. First, a reference set of conflicts was obtained from trajectories that were computed using birth points and nominal flight plans extracted from recorded traffic data. Distributions were obtained for several primary properties (e.g., encounter angle) that are most likely to affect the performance of a conflict probe. A genetic algorithm was then utilized to determine the values of time shifts for the recorded track data so that the primary properties of conflicts generated by the time shifted data match those of the reference set. This methodology was successfully demonstrated using recorded traffic data for the Memphis Air Route Traffic Control Center; a key result was that the required time shifts were less than 5 min for 99% of the tracks. It was also observed that close matching of the primary properties used in this study additionally provided a good match for some other secondary properties. This study is documented in [Paglione et al., 2003b]. It illustrates how CPAT utilized the FD_PROC tools and developed an acceptable methodology of time-shifting traffic for later evaluation of conflict probe systems.

Conclusion

This paper presents a methodology to generate air traffic scenarios for the support of the testing and analysis of Air Traffic Management decision support tools. The paper first describes what a scenario based on recorded data is. Next, it defines why it must be altered in order to create the conflicts required for testing a decision support tool.

The paper describes each of the three steps that comprise the methodology. The first step is data extraction during which traffic data, available from various sources and recorded in different formats, is extracted and placed into a set of relational database tables. The second step is data modification, where the data in the tables may be manipulated for test purposes. This manipulation may consist of simply culling undesirable flights, or it may involve using a genetic algorithm to time-shift the flights to induce conflicts or encounters. The third step is scenario generation, where scenarios are created based on the traffic data retrieved from the modified database tables. During this final step the scenarios may be formatted for various target systems.

The paper described the original implementation of this methodology and then describes an upgrade of this methodology that is currently in process.

The paper ends by describing how the Conflict Probe Assessment Team has used this methodology to generate scenarios that have been used for accuracy and risk reduction testing and for a study assessing the effect of weather forecast errors.

Acronyms

ACES	Adaptation Controlled Environment System
ARTCC	Air Route Traffic Control Center
ASCII	American Standard Code for Information Exchange
CCLD	Core Capability Limited Deployment
CMS	Common Message Set
CPAT	Conflict Probe Assessment Team
CTAS	Center-TRACON Automation System
CVS	Concurrent Versions Systems
DST	Decision Support Tool
GNU	GNU's Not UNIX
GUI	Graphical User Interface
FD_PROC	Field Data Processing
HID	Host Interface Device
LMATM	Lockheed Martin Air Traffic Management
RUC	Rapid Update Cycle
TRACON	Terminal Radar Approach Control
URET	User Request Evaluation Tool

ZDC	Washington ARTCC
ZFW	Fort Worth ARTCC
ZID	Indianapolis ARTCC
ZJX	Jacksonville ARTCC
ZKC	Kansas City ARTCC
ZME	Memphis ARTCC
ZMP	Minneapolis ARTCC
ZTL	Atlanta ARTCC

References

[Cale et al., 1998] Cale, Mary Lee, Michael Paglione, Dr. Hollis Ryan, Dominic Timoteo, Robert Oaks (1998), "URET Conflict Prediction Accuracy Report," DOT/FAA/CT-TN98/8, WJHTC/ACT-250.

[CPAT, 2000] Conflict Probe Assessment Team (2000), CD-ROM, "URET CCLD Final Accuracy Scenario Delivery Refresh Data, Revision 1," Federal Aviation Administration, Engineering and Integration Branch, ACT-250, William J. Hughes Technical Center, New Jersey, 08405.

[CPAT, 2001] Conflict Probe Assessment Team (2001), CD-ROM, "URET CCLD Final Accuracy Scenario Delivery Scenario Data, Revision G," Federal Aviation Administration, Engineering and Integration Branch, ACT-250, William J. Hughes Technical Center, New Jersey, 08405.

[FAA, 1998] Federal Aviation Administration (1998), "Interface Requirements Document ARTCC Host Computer System (HCS) / Center TRACON Automation System (CTAS) - User Request Evaluation Tool (URET)," NAS-IR-82178219, October 5, 1998.

[FAA, 2000] Federal Aviation Administration (2000), "Interface Requirements Document ARTCC Host Computer System (HCS) / Air Traffic Management (ATM) Applications," NAS-IR-82170001, September 1, 2000, Revision M.

[Lindsay, 1997] Lindsay, Kenneth S. (1997), Lindsay, *Xeval Users Manual*, MITRE Technical Report, MTR 97W0000030, June, 1997.

[Lindsay, 1998] Lindsay, Kenneth S. (1998), *Algorithmic Evaluation Process and Tools for the User Request Evaluation Tool*, MITRE Technical Report, MTR 98W0000029, June, 1998.

[Oaks and Paglione, 2001] Oaks, R. D. and M. M. Paglione, "Generation of Realistic Air Traffic Scenarios Based on Recorded Field Data," *46th Annual Meeting of the Air Traffic Control Association*, Washington, DC, November, 2001..

[Oaks and Paglione, 2002] Oaks, R. D. and M. M. Paglione, "Generation of Realistic Air Traffic Scenarios Using a Genetic Algorithm," 21st Digital Avionics Systems Conference: Air Traffic Management for Commercial and Military Systems, Irvine, CA, October, 2002.

[Paglione et al., 1999] Paglione, Mike, Dr. Hollis F. Ryan, Robert D. Oaks, J. Scott Summerill, Mary Lee Cale (1999), "Trajectory Prediction Accuracy Report User Request Evaluation Tool (URET)/Center-TRACON Automation System (CTAS)," DOT/FAA/CT_TN99/10, WJHTC/ACT-250.

[Paglione et al., 2002] Paglione, Mike M., Lori Charles, Shurong Liu, Robert D. Oaks, Hollis F. Ryan, Ph.D, J. Scott Summerill (2002), "User Request Evaluation Tool (URET) Conflict Probe Sensitivity to Weather Forecast Errors," Draft Report, October 2002.

[Paglione et al., 2003a] Paglione, Mike M., Robert D. Oaks, and J. Scott Summerill (2003), "Time Shifting Air Traffic Data for Quantitative Evaluation of a Conflict Probe," *AIAA Guidance, Navigation, and Control Conference*, Austin, TX.

[Paglione et al., 2003b] Paglione, Mike M., Robert D. Oaks, and Karl D. Bilimoria (2003), "Methodology for Generating Conflict Scenarios by Time Shifting Recorded Traffic Data," *AIAA 3rd Annual Aviation Technology, Integration, and Operations Technical Forum*, Denver, CO.