# Data Mining of Air Traffic Control Operational Errors

Zohreh Nazeri, George *Mason University*
*znazeri@gmu.edu*
*4400 University Drive, Fairfax, VA 22030*

***Abstract*****—In this paper we present the results of applying data mining techniques to identify patterns and anomalies in air traffic control operational errors (OEs). Reducing the OE rate is of high importance and remains a challenge in the aviation safety community. Existing studies, which use traditional methods and focus on individual aspects of OEs, are limited to operations at a single facility, or events in a short period of time. A holistic study of historical data available on OEs has not been conducted. We have applied an attribute focusing technique to study 15 years of operational errors at all FAA Air Route Traffic Control Centers (ARTCCs) [1] in the National Airspace System (NAS) in the U.S. We have found 'interesting' patterns of common characteristics, anomalies, and changes in trends of operational errors. We interpreted the results with the help of domain experts and plan to do a similar analysis for OEs at other types of air traffic control facilities (towers and TRACONs) as well.***

## I. INTRODUCTION

A fundamental principle of aviation safety is maintaining a safe distance between the flying aircraft and other objects (other aircraft, terrain, obstructions, and airspace that is not designated for routine air travel). There are rules and procedures for maintaining this safe distance. Air traffic controllers employ these rules and procedures that define safe separation in the air and on the ground[2]. An operational error occurs when these rules and procedures are not followed due to equipment or human error.

For example, in April 2000, an operational error occurred at the Denver en route Center when a controller allowed two jet airliners to lose separation as they were approaching head on at flight level 390 (i.e., altitude of 39,000 feet). At about 6 miles apart, less than 20 seconds from a midair collision, an on-board Traffic Alert Collision and Avoidance System (TCAS) sounded an alert and prompted the pilots to take evasive action averting an accident. The two airplanes came within 1,100 feet vertically and 1 mile laterally [1].

1 See Appendix A for an explanation of air traffic control facilities and operations.

2 Standard separation is 5 miles laterally and 1,000 feet vertically up to 29,000 feet and 2,000 feet vertically above 29,000 feet in the en route environment.

Air traffic control operational errors can pose a very serious safety risk. Some of these errors can pose safety risks by directing aircraft onto converging courses and, potentially, midair collisions. At least four serious errors occur, on average, every 10 days in which aircraft collisions are barely averted [2]. Figure 1 shows historical data on total en route operations and errors in the U.S.

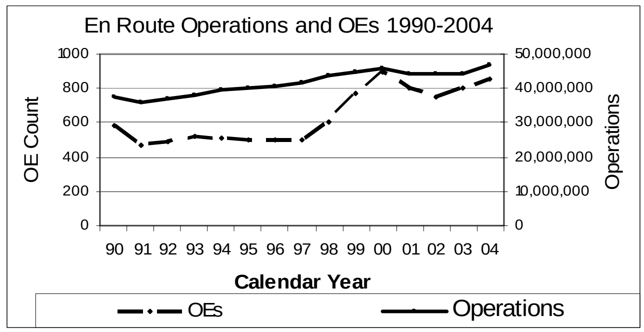


Fig. 1. Historical data on en route operations and errors in the U.S.

Operational errors are at risk of increasing even further because of the projected growth in air traffic. After going through a sharp reduction following September 11, 2001, air traffic volume has returned to its pre-9/11 level and is projected to exceed this level according to FAA Aerospace forecasts, fiscal years 2005 – 2016 [3] (see Fig. 2).

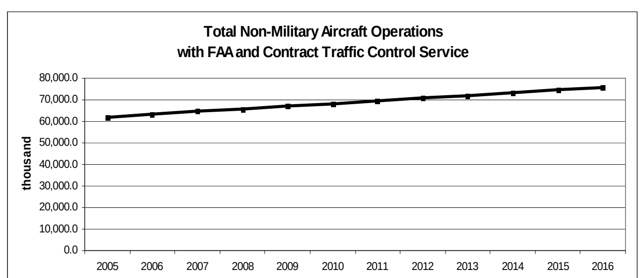


Fig. 2. Federal Aviation Administration (FAA) projection of air traffic growth in the U.S.

In addition to the challenge of expanding NAS capacity to accommodate the expected increase in the volume of air traffic over the next 10 years [4], a high rate of retiring air traffic controllers in the next few years is expected. Many air traffic controllers were hired in the post strike years of 1981 to 1985 and are either retiring now or will retire soon [5] [6]. There is a strong demand for new controllers to fill vacant positions in a 20 to 30 year career field. Hiring thousands of air traffic controllers to replace those expected to retire over the next decade means having less experienced controllers to handle operations at the facilities. This raises the potential for occurrences of errors.

Understanding characteristics of the operational errors and taking action to reduce them is a high priority for FAA and the aviation safety community as a whole [7].

## II. Related Studies

To understand the operational errors, different studies have been conducted. The past and current related studies use traditional analysis methods and have focused on individual aspects of the problem, or limit their research to a specific facility and/or to a short period of time. For example, Rodgers [8] focused on 1992-1995 data for the Atlanta airport. Kinney [9] analyzed the data from 1974 to 1976. Similarly, Majumdar [10] analyzed 1998 to 2000 data in New Zealand. Findings of each of these studies are useful in the scope of their analysis. However, none of them provides a complete picture and their individual results cannot be compared or linked together. Simply comparing the error rates are not enough for comparison of OEs and safety levels at different facilities.

By applying data mining techniques, we were able to look at collected OEs in all 21 Air Route Traffic Control Centers (ARTCCs) in the U.S. for the time period of the past 15 years. Our approach allowed us to identify common characteristics and anomalies in the data and compare them at the national level.

## III. Experiment

We applied an attribute focusing algorithm to the data from Operational Error reports for the 21 Air Route Traffic Control Centers in the U.S. The applied methodology is discussed below.

### A. Data Mining Technique

In this experiment, we used an attribute focusing technique [11] [12] to identify anomalies in distribution of different attributes of erroneous operations. The analyzed data consists of reports on erroneous operations (positive examples) only. There are no similar data collected for the flights that completed without any operational errors.

The attribute focusing algorithm was applied to the structured fields in data only. It examines distribution of each and every subset of attributes over different values of a selected attribute and compares each subset's distribution to an 'expected' reference for comparison. The 'expected' pattern is distribution of all data records over different values of the selected attribute, but scaled down to fit in the range of values for each subset. For example, for a particular attribute, AIRCRAFT, the algorithm calculates the overall distribution of AIRCRAFT in the data. Then it compares this distribution with the distribution of AIRCRAFT in various subsets of the data (e.g., different facilities). If a subset has a statistically different distribution of AIRCRAFT, then the condition that defines the subset is returned as an anomaly. Note that the overall distribution is our baseline rule, and the distributions for the subsets are the potential exceptions.

### B. Data

This experiment was conducted on data reported for en route Operational Errors in the U.S. Air Route Traffic Centers (ARTCCs), maintained in the National Airspace Incident Monitoring (NAIMS) database, for the period of 1990 through 2004. This data consists of various structured fields as well as unstructured text fields. For the purpose of this study, only selected structured fields were used. Fields with specifics on individual controllers and other sensitive information were not included.

The selected fields used in the experiment include date of the error, location of the facility performing the operation at the time of the error, make/model of the aircraft involved in the erroneous operation, airspace class, phase of flight, weather condition, causal factors, and factors contributing to the complexity of the situation at the time of error.

### C. Domain Knowledge

Findings of the experiment were reviewed with subject matter experts, i.e., former air traffic controllers and facility supervisors having 20 to 30 years of experience in air traffic control. The findings were interpreted with the help of these experts and marked as either obvious or interesting. Furthermore, the findings that revealed issues in the quality of data were identified separately.

Additional domain knowledge to explain the findings in a bigger context of time and space were obtained from the experts, available literature, FAA websites, and other relevant sources.

## IV. Results

We divide the discovered patterns into two categories. The first category contains obvious findings. These are findings that are not unusual or surprising to the subject matter experts. However, their discovery by the applied technique were reassuring to the experts since they know these patterns very well as typical characteristics of the domain. These discoveries earned the experts' trust and belief that the techniques work very well on their data. The second category of findings contains patterns that were interesting to the experts; some could be explained with the help of available domain knowledge after a quick research, others need further investigation before a final conclusion could be drawn. We show examples of each category below.

One factor contributing to the complexity of operations is the airspace design of the space where the errors occur. An interesting discovery shows this factor is much higher than expected for the Southern region (Fig. 3) for the 15-year period (1990-2004). One would expect Eastern region with a higher traffic and more complicated airspace have a high airspace design complexity factor. The discovered pattern is unexpected.

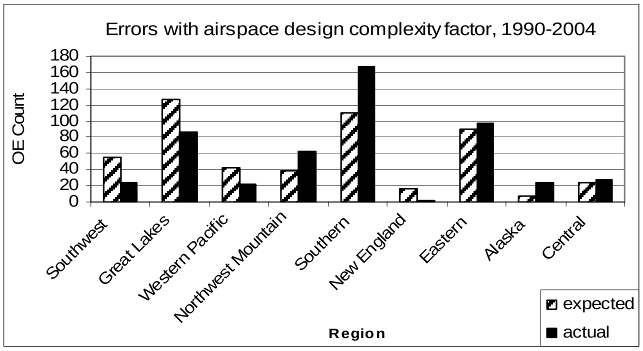


Fig. 3. Airspace design as a factor contributing to the complexity of operations at the time of error, is much higher than expected in Southern region.

Another interesting finding is the drop in the number of errors in the Eastern region after 2001 (Figure 4). To normalize the error rates, we considered number of errors per million Center operations in each region. The error rate for Eastern region remained lower than expected for the post 2001 years after normalization. (We found this true for Western Pacific and Central regions as well, although not with the same magnitude of decrease in error rate as for the Eastern region.)

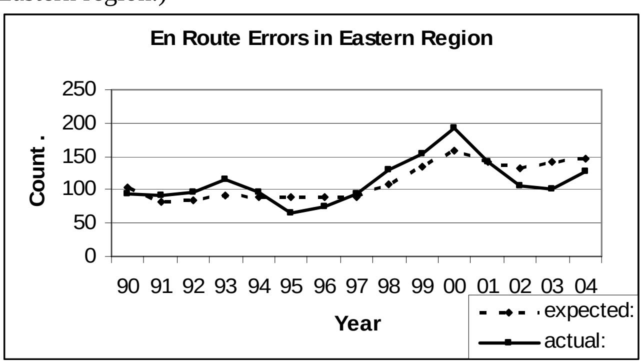


Fig. 4. Number of en route operational errors in Eastern region has dropped after 2001.

As mentioned before, some of our findings are obvious to the domain experts and did not add to the existing knowledge of the OEs. However, these findings served to show validity of the technique and some of them confirmed the experts' assumptions. Below are examples of these findings.

Flow Control (managing flow of multiple aircraft moving) is a factor that could contribute to the complexity of operations. Our findings show there are more OEs with Flow Control as a complexity factor when there are bigger volumes of traffic. This is not surprising, since managing the flow of aircraft operations becomes more complicated with bigger numbers of simultaneous aircraft in the airspace. Another obvious pattern in our findings is that the number of OEs involving small aircraft is higher in class E airspace compared to other airspace classes (see Appendix B). This finding is not surprising either. It is expected that smaller aircraft, mostly General Aviation, fly in class E airspace.

## V. Conclusion and Future Work

In this paper, we presented our experiment of applying data mining techniques to the en route operational errors. We discussed the technique we used, the data, and our findings in two categories: obvious and interesting. We also discussed that individual studies on subsets of data would not be able to provide the information found by a holistic study of the operational errors.

We would like to emphasize that subject matter experts' interpretation of the findings and further evaluation of the findings in the context of the domain knowledge is very important and absolutely necessary before drawing conclusions and taking any actions on the findings of any data mining application.

We are performing similar analyses on data reported for Towers and TRACONs operations in the U.S. and will publish the results in near future. We plan to perform an even higher level analysis of OEs looking at all three types of facilities (Towers, TRACONs, and Centers) together and incorporating the results. Another step we plan to do to enhance the results is to link the OE data to other relevant data available such as accidents and incidents data.

## Appendix A – FAA Facilities

The following information, from NASA's Air Traffic Management Tutorial website [13], describes different air traffic facilities and their function in the flight operations.

Towers control the air traffic at an airport and in its vicinity. They issue departure clearances, give instructions on aircraft push back from the gate and appropriate taxiways to take toward the takeoff runway, and issue takeoff clearances. Shortly after takeoff, the pilot is instructed to contact the TRACON.

TRACONs (Terminal Radar Approach Control) are facilities containing radar operations from which air traffic controllers direct aircraft during the departure, descent and approach phases of flight. TRACONs route the aircraft away from the airport and clear them to a new altitude. After departure, the aircraft (now moving into the en route phase of flight) is handed off to a Center.

Air Route Traffic Control Centers (ARTCCs), usually referred to as 'Centers', provide Air Traffic Service to aircraft within the controlled airspace, and principally during the en route phase of flight. There are 21 Air Route Traffic Control Centers (ARTCCs) in the United States. Any aircraft operating under Instrument Flight Rules (IFR) within the confines of an ARTCC's airspace is controlled by air traffic controllers at that Center. See Fig. 5 for a visualization of airspace assignments to different facilities. (The picture in Fig. 5 is obtained from NASA ATM Tutorial website [13].)

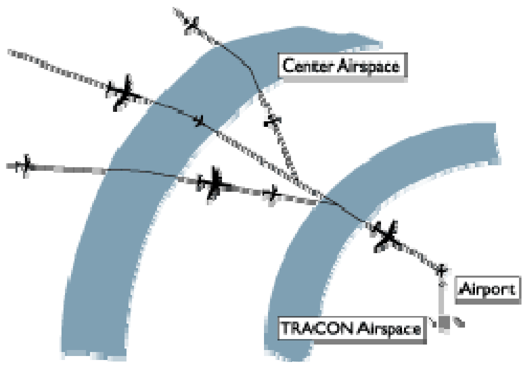


Fig. 5. Depiction of flights in controlled airspace.

When an aircraft is moving into its en route phase of flight, it is handed off to a Center controller who monitors the flight and gives instructions to the pilot in order to maintain the aircraft's separation from other aircraft. A flight path might go through more than one Center's space. Once the aircraft is close to its destination airport, it begins its descent phase (going from its cruising flight level to a lower altitude) and is handed off to a TRACON for approach phase of flight and then to a tower for landing.

## APPENDIX B – AIRSPACE CLASSES

Navigable airspace is divided into classes of three-dimensional segments. Each class is defined in terms of flight rules, interactions between aircraft and air traffic control, and pilot/equipment requirements. The following are brief definitions of different airspace classes [14]:

*Class A*: airspace from 18,000 feet MSL up to and including FL 600, including the airspace overlying the waters within 12 nautical miles of the coast.

*Class B*: airspace from the surface to 10,000 feet MSL surrounding the nation's busiest airports.

*Class C*: airspace from the surface to 4,000 feet above the airport elevation surrounding towered airports serviced by a radar approach control.

*Class D*: airspace from the surface to 2,500 feet above the airport elevation surrounding towered airports.

*Class E*: controlled airspace that is not Class A, Class B, Class C, or Class D. No specific pilot/equipment certification is required for flying in class E airspace.

*Class G*: small layer of uncontrolled airspace near the ground and in remote regions.

## ACKNOWLEDGMENT

The interpretation of the findings and explaining them in the context of air traffic control operations would not be possible without the help of the domain experts. The author would like to thank the air traffic controllers who reviewed the findings and provided explanation of many of the concepts and jargons in the air traffic control domain. While these domain experts helped with understanding and interpretation of the results, the responsibility of selecting examples for this paper and any conclusions made here goes solely to the author. The author would also like to thank Earl Harris who has developed the initial version of the attribute focusing algorithm used in this experiment.

## REFERENCES

[1] "Final report on review of the Federal Aviation Administration's (FAA) efforts to reduce air traffic control operational errors and deviations", report number: AV-2003-040, April 3, 2003. http://www.oig.dot.gov/.
[2] Department of Transportation Aviation Safety website, http://www.dot.gov/PerfPlan2004/safety_aviation.html.
[3] FAA Aerospace Forecasts, Fiscal years 2005-2016, March 2005. http://www.api.faa.gov/forecast05/Forecast_for_2005.htm'
[4] Aviation Capacity Enhancement (ACE) report, March 2003.
[5] National Airspace System - Progress and Ongoing Challenges for the Air Traffic Organization 2005.
[6] Denver (ZDC) Air Route Traffic Control Center website, http://www.nw.faa.gov/ats/zdvartcc/careers.htm.
[7] Aviation Safety: Information on FAA's Data on Operational Errors At Air Traffic Control Towers, GAO-03-1175R, September 23, 2003.
[8] Rodgers, M., Mogford, R., and Mogford, L., "Relationship of Sector Characteristics to Operational Errors", Report No. DOT/FAA/AM-98/14, FAA, U.S. Department of Transportation, 1998.
[9] Kinney, G.C. et al., "Human Element in Air Traffic Control: Observations and Analyses of the Performance of the Controllers and Supervisors in Providing ATC Separation Services", Report No. MTR-7655, METREK division, MITRE Corp., McLean, VA, 1977.
[10] Majumdar, A., Ochieng, W., Nalder, P., "Trend Analysis of Controller-Caused Airspace Incedents in New Zealand, 1994-2002", Journal of the Transportation Research Record, No. 1888.
[11] Bhandari, I., Colet, E., Parker, J., Pines, Z., Pratap, R., and Ramanujam, K. "AdvancedScout: Data Mining and Knowledge Discovery in NBA Data", 1996. Data Mining and Knowledge Discovery, Volume 1, #1, Kluwer Academic Publishers, pp 121-24.
[12] Harris, E., Bloedorn, E., Rothleder, N., "Recent Experiences with Data Mining in Aviation Safety", SIGMOD-DMKD98 Workshop, Seattle, WA, June 5, 1998.
[13] NASA Air Traffic Management System website, http://quest.arc.nasa.gov/projects/aero/virtualskies/ATM/tutorial/tutorial8.html
[14] Federal Aviation Administration (FAA) website http://www.faa.gov/ATPUBS/AIM/Chap3/aim0302.html#O8C3b6ROBE