Aircraft Operation Anomaly Detection Using FDR Data

Lishuai Li, Maxime Gariel, R. John Hansman
IAB/Airline Industry Consortium
Nov 4, 2010
Motivation

- Commercial aircraft accident rate has dropped significantly.
- Further improvement requires proactive safety management.
  - Identify risks in day-to-day operations
- Large amount of routine flight data “available”
  - Flight Operations Quality Assurance (FOQA) in US
  - Flight Data Monitoring (FDM) in Europe
- Information in flight data:
  - Rich about flight operations and risks
  - Underutilized by current practices
Objective

- Objective
  Identify emerging risks from routine flight data

- Issues: complexity of routine flight data
  - Large number of variables
  - Mix of relationships among variables
  - Variability among flights
    - Aircraft type
    - Procedures
    - Weather
    - Pilots
    - ...

3
FDR Data

- **Flight Data Recorder (FDR)**
  - Equips every commercial aircraft
  - Records 100+ to 1000+ flight parameters during each flight depending on aircraft and airline

- **Rich information about**
  - Aircraft
  - Environment
  - Crew operations

- **Challenge for analysis**
  - How to obtain useful information from massive data?

*Example of FDR data*

NTSB (2009), FDR Group Chairman's Factual Report, DCA09MA027
Current Data Analysis

- **Exceedance detection**
  - Exceedance of a value under certain conditions
  - List of events believed to be unsafe

- **Distribution analysis**
  - On specific queries, e.g. distribution of total energy at 900 ft AGL during approach, distribution of airspeed at takeoff, etc.

**Exceedance Event Examples**

<table>
<thead>
<tr>
<th>Event Code</th>
<th>Description</th>
<th>Measurement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>07A</td>
<td>Approach Speed Low Within 2 mins of T/D</td>
<td>nil</td>
<td>CASTDOWN</td>
</tr>
<tr>
<td>07B</td>
<td>Approach Speed Low Below 25ft Radio</td>
<td>CASATR30</td>
<td>CAS AT TOUCHDOWN</td>
</tr>
<tr>
<td>08A</td>
<td>Climb Out Speed High Below 400ft AAL</td>
<td>MNCLS3540</td>
<td>MIN CLIMB SPEED 35FT TO 400FT</td>
</tr>
<tr>
<td>08B</td>
<td>Climb Out Speed High 400’ to 1000’ AAL</td>
<td>MIN400150</td>
<td>MIN CLIMB SPEED 400FT TO 1500FT</td>
</tr>
<tr>
<td>08C</td>
<td>Climb Out Speed Low 35’ AGL to 400’ AAL</td>
<td>MNCLS3540</td>
<td>MIN CLIMB SPEED 35FT TO 400FT</td>
</tr>
</tbody>
</table>

[Larder, Brian, and N. Summerhayes. 2004. *Application of Smiths Aerospace Data Mining Algorithms to British Airways 777 and 747 FDM Data.*]

- **Limitations**
  - Only known safety issues are examined.
  - “You only get what you ask for.”
Proposed Approach

- Assumption
  - Majority of flights is safe

- Approach
  Develop a method to track detailed in-flight recorded data
  - Establish a *norm* of safe operations
  - Identify *anomalies*, or abnormal operations which indicate increased risks:
    - Vehicle impairment
    - External hazards
    - Inappropriate crew operations
    - ...
Data Analysis Method

- **Multivariate Cluster Analysis**
  - Use multiple variables over time
  - Cluster flights

- **Establishment of a norm**
  - *Norm*: flight contained in clusters
  - *Anomalies*: flights not belonging to a cluster

- Domain experts leverage on the results to identify emerging safety issues.

- **Advantages:**
  - Flights with unknown risks can be found
  - “You don’t need to specify what might be unsafe; but it tells”
Preliminary Study

Proof-of-concept demonstration on a limited FDR dataset

1. Pre-filter a relatively homogeneous dataset
   • All B777 arrivals at Abu Dhabi Int'l Airport (183 flights)
   • Focused on final approach phase

2. Transform multiple time series into one vector for each flight

3. Cluster the vectors to identify norm and anomalies
Subsets of flight parameters for Clustering

- Data limitation
  - 183 flights; 103 flight parameters
  - Too sparse to form clusters if all parameters are used for clustering

- In the preliminary study, clustering is based on subsets of flight parameters:

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>Longitude, latitude, height above threshold</td>
</tr>
<tr>
<td>Position (with respect to runway)</td>
<td>Distance to threshold, deviation to centerline, height above threshold</td>
</tr>
<tr>
<td>Position, heading, speed</td>
<td>Distance to threshold, deviation to centerline, height above threshold, heading relative to runway, speed measures</td>
</tr>
<tr>
<td>Engine</td>
<td>N1, fuel flow, EGT, thrust lever, EMS thrust, N3</td>
</tr>
<tr>
<td>Environment</td>
<td>Wind, temperature, pressure, air density</td>
</tr>
<tr>
<td>Motion</td>
<td>Speeds, accelerations, load factor, pitch change rate, roll change rate, yaw change rate</td>
</tr>
<tr>
<td>Control</td>
<td>Flap, slat, spoiler, elevator, stabilizer, trim, pitch, roll, yaw</td>
</tr>
<tr>
<td>Force</td>
<td>Drag, lift, gross weight, CG position, normal load factor</td>
</tr>
</tbody>
</table>
Cluster by Position

Parameters included: Longitude, Latitude, Height above touchdown
Cluster by Position

Cluster 1

Cluster 2

Outlier
Cluster by Position, Heading, & Speed

Height above touchdown (best estimate)

Calibrated airspeed
Cluster by Engine Parameters

Parameters included:
N1 for all engines, thrust, thrust lever, EGT for all engines, avg N3, avg fuel flow, etc.

Thrust lever angle (left inboard engine)

N1: average (all engines, percent of maximum)
### Summary of outliers identified by subset

<table>
<thead>
<tr>
<th>Position (Absolute)</th>
<th>Position (Relative)</th>
<th>Position, Heading, Speed</th>
<th>Engine</th>
<th>Weather</th>
<th>Motion</th>
<th>Control</th>
<th>Force</th>
</tr>
</thead>
<tbody>
<tr>
<td>373564</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>374577</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>377838</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>377844</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>368467</strong></td>
<td><strong>0</strong></td>
<td><strong>1</strong></td>
<td>1</td>
<td><strong>1</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>382554</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>369204</td>
<td>0</td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>1</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>370512</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td><strong>1</strong></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>375699</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>377288</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>385160</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>371927</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>379659</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>384089</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>382520</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>384512</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>370713</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>371044</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>371929</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>377860</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>379684</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>369202</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>1</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>383279</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>

14
Example Anomaly: High Approach
Easy to detect by current practices
Example Anomaly: Wind Gust
Difficult to detect by current practices
Summary & Future Work

• **Summary**
  - Proposed an approach to identify emerging risks from routine flight data
  - Developed a method to track detailed flight data and define norm and anomalies for flight operations
  - Performed preliminary analysis on a limited FDR dataset

• **Next Steps**
  - Extend the analysis to other phase of flight
  - Investigate parameters with no observable patterns over time
  - Apply method to full FDR dataset (Data Wanted)
Thank you!

Comments and questions?