ADS-B Safety Data Mining and Exploratory Coverage Gap Analysis for Two Mission Segments

By:
Cliff Johnson, FAA Research Lead & Flight Test Engineer – Rotorcraft & Vertical Lift/eVTOL
Aviation Research Division
FAA William J. Hughes Technical Center,
Atlantic City, NJ
Feb. 24, 2021
ANOMALY DETECTION (KNOWLEDGE DISCOVERY) USING ADS-B DATA
Anomaly Detection in Flight Data Records

- Traditionally, anomaly detection in flight data records is based on **exceedance analysis**
  - Define parameter thresholds not to be exceeded
  - Combine parameters into safety events
  - Thresholds may depend on mission, helicopter make/mode, operator, event severity, etc.

- With large amounts of data, anomalies can be detected by observing patterns in the data

- Group data based on take-off and landing airports
Anomaly Detection in Flight Data Records

- **Goal**: identify clusters in set of flights from Airport A to Airport B

Flights are dissimilar in terms of distance in FPC space.

**ADS-B altitude data**

- **FPCA**
Rotorcraft Mission Segments

Air Tour

External Load

Airborne Law Enforcement

Aerial Firefighting

Search and Rescue

Helicopter Air Ambulance

Training

Offshore

Corporate/VIP Transport
ADS-B Data Analysis Process Flow

- **Raw Data Extraction**
  - Raw Data → Preprocessing Step → Trajectory Data → Time Series → FPCA → Clustering to find the grouping of a dataset
  - Raw Data → Preprocessing Step → Sequence Data → CV method: HOG → Clustering to find the grouping of a dataset

- **Trajectory data analysis**
  - Trajectory Data → Image Data → Feature extraction → Classification for feature comparison
    - Logistic regression
    - SVM
  - Trajectory Data → Image Data → Data labeling
    - Image resizing
    - Histogram of oriented gradient
  - Image Data → Feature extraction → Clustering to find the grouping of a dataset
Daily flight data recordings

If time gaps > 3 mins, create a new segment

Filtering the segments based on time gaps and flight duration

Filtering the segments based altitude threshold

Filtering the segments based velocity thresholds

Segmented flights

Retrieved the airfield data from NfdcFacilities.csv

Retrieved the airfield data from MaineHelipadInfo.csv

Combine the information retrieved from both airfield data sources

Segmented flights grouped by takeoff / landing sites
Data Quality Issues with ADS-B Data

Segment counts

- large time gap
- short duration
- short duration and large time gap
- very low altitude operation
- incomplete flight - starting high
- incomplete flight - ending high
- start or end speed too high
- valid flight

Useable Data
Fluctuation in Altitude Data
Fixing Altitude Data

- Multiple jumps in altitude data
- Define a cutoff threshold for filtering out the jumped data
- Use interpolation for filling the missing values
Fixing Altitude Data (Cont.)
# Duplicate Helipad Information

<table>
<thead>
<tr>
<th>Location 1</th>
<th>Location 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ‘ME02’ Eastern Maine medical center</td>
<td>‘HJOES’ St Joseph hospital</td>
</tr>
<tr>
<td>2 ‘ME95’ Central Maine medical center</td>
<td>‘HMARY’ St Mary hospital</td>
</tr>
<tr>
<td>3 ‘2B7’ Sebasticook hospital</td>
<td>‘HSBVH’ Sebasticook hospital</td>
</tr>
<tr>
<td>4 ‘KBXM’ Parkview Adventist medical center</td>
<td>‘HPARK’ Parkview hospital</td>
</tr>
</tbody>
</table>
Air Tour Dataset

- Operations grouped based on TO/LD sites
- Total number of flight: 365

Grouping data based on TO/LD sites

Sample set: from HNL to Kalaeloa

<table>
<thead>
<tr>
<th>FROM</th>
<th>TO</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>44</td>
<td>0</td>
<td>15</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>73</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>
Subgroups within the Air Tour Dataset

- Hilo to Hilo (73)
  - Subgroup 1 (34)
  - Subgroup 2 (8)
  - Subgroup 3 (31)

- Kohala to Waikoloa (63)

- Kalaeloa to HNL (52)

- Kahului to Kahului (44)

- HNL to Kalaeloa (25)

- Hana to Kahului (19)
Altitude Data
(Sample from the Air Tour Dataset)

• The dataset contains 24 altitude traces and the length (dimension) is 200
Kohala to Waikoloa (Position)
Air Medical Dataset

- Starting location for 555 flights
- Flight trajectories for 555 flights

- TO/LD sites are more distributed than concentrated
- Data jump observed after data cleaning
- Trajectories have radial pattern rather than circular

- N901CM – 460 flights/703.5hrs  
  - Apr. 2018-Dec. 2019
- N901EM – 406 flights/841.3hrs  
  - June 2018-Dec. 2019
- N710SB – 36 flights/67.8hrs  
  - Nov. 2019-Dec. 2019

- TOTAL: 900 flights/1612.6 flight hours
Air Medical Flight Visualization

[Graph showing altitude over time and a map indicating flight path with color-coded regions.]
Air Medical Flight Routes (Top 10)
Route: 68ME – ? = 68 Flights
Route: ME02 – ? = 35 Flights
Route: ? – ME02 = 29 Flights
Route: 68ME – ME02 = 14 Flights
Future Work

• ADS-B to Flight Hours Comparison
  – Compare ADS-B Flight Hours to Operator Recorded Flight Hours (i.e. Hobbs/Tach Time)
  – Perform statistical analysis of ADS-B tracks compared to onboard recorded data (i.e. Foreflight or other helicopter recorded data)
  – Identify additional gaps with ADS-B coverage at Low Altitudes

• ADS-B/Heliport Metrics
  – Calculate operations (i.e. counts) of takeoffs and landings at various locations (i.e. heliports)
  – Integrate ADS-B data into Heliport Risk Analysis and other metrics
Questions?