Identification of Robust Routes using Convective Weather Forecasts

Diana Michalek
Operations Research Center, MIT

Prof. Hamsa Balakrishnan
Dept. of Aeronautics & Astronautics, MIT

8th ATM R&D Seminar
Napa, California, June 29-July 2 2009
Weather has a major impact on air traffic operations

- In 2007, demand for National Airspace System (NAS) resources resulted in increased delays
  - 76.9% of the total NAS delay (in minutes) was weather-related [IATA]
  - 24% of all flight delays were weather-related [BTS]
  - Delays were responsible for $41 billion in costs to the US economy, including $19 billion in costs to airlines and $12 billion in costs to passengers [Joint Economic Committee, May 2008]
  - 213,079 flights were delayed because of terminal-area weather (Ground Delay Programs) [FAA OPSNET data]
  - The total amount of delay was 15.2 million minutes, and the average delay per delayed flight was 71.3 minutes [FAA OPSNET data]
- Convective weather is responsible for a significant fraction of flight delays in the NAS
Air Traffic Flow Management

- Make strategic decisions a few hours ahead of operations to balance supply (capacity) and demand of NAS resources
  - Requires dealing with the impact of adverse weather conditions
  - Typically assumes as input:
    - Expected sector capacity
    - Sector capacity scenarios and probabilities
Prior work on weather-ATM integration

- **Air Traffic Flow Management algorithms**
  - Bertsimas and Stock; Nilim and El Ghaoui; Menon et al., Sun et al.
  - Assume presence of airport/airspace capacity forecasts
- **Algorithms to efficiently synthesize routes through airspace affected by weather (Prete and Mitchell)**
  - Static, fine-grained, time-varying weather data as input
- **Krozel et al. look at creating models of airspace capacity in the presence of weather**
  - Theoretical capacity using continuous minimum-cut theory
  - Uses static weather forecasts, does not incorporate uncertainty or validate forecast accuracy
- **Mitchell et al. incorporate a region of uncertainty around weather forecasts**
- **National Convective Weather Forecast 2 (NCWF-2, Seseske et al)**
- **Route Availability Planning Tool (DeLaura and Allan)**
  - Models jet route blockage deterministically
Robust scheduling in the face of weather

- Traditional focus of air traffic management (ATM) algorithms

- Our approach

- “The forecasts we have, rather than the forecasts we would like to have”
Each 1km x 1km pixel contains value of Vertically Integrated Liquid (VIL)
  - Translated into a “Level”; pilots typically avoid level 3+

At each time $T_0$, a deterministic forecast is available for times $T_0+5, T_0+10, \ldots, T_0+120$

Large data sets: Each day of data is $\sim 30$GB
Are pixel-based comparisons what we really need?

- Pixel-based comparisons tend to suggest that the weather forecasts are quite poor
- Forecasts may be “accurate” for TFM purposes, even though they have large error on individual pixels
  - The general weather trends may be predicted correctly
  - General storm type may be forecast correctly
Problem Statement: Given a weather forecast for some time in the future and a set of predetermined potential routes, which routes are likely to be open in the actual weather that materializes, and what is the uncertainty associated with this prediction?
Contributions of this paper

- **Robustness of routes**: If aircraft are routed along trajectories through the current forecast, what is the probability that these trajectories will actually be clear?
  - **Feature selection**: Can we identify features of paths through forecast weather that are highly-correlated with blockage?
  - **Classification**: Given a convective weather forecast and standard arrival (departure) routes, can we identify the most robust routes?

- **Optimization**: Given a route robustness model, can we use it to improve terminal area operations?

- **Data-driven approach**:
  1. Select sample routes through the forecast weather
  2. Validate these routes against the actual weather
  3. Correlate features of forecast paths with blockage to determine characteristics of robust paths
  4. Develop a classification algorithm to predict route blockage
Selection of routes through dynamic weather forecasts (for arrivals)

Forecast

t₀

Observed

t₀

100 km

10 km

All
Validate each route against the observed weather

- Route $P$ is defined to be **open** or **clear** in the observed weather if there exists a route that is not impacted by weather within a small neighborhood of the original route.
- To validate $P$, find route $P'$ in the observed weather grid that:
  - Does not pass through any node containing L3+ weather
  - Is within $B$ km of $P$, allowing for wiggle room (we set $B=8$)
- This is done by solving a shortest path problem (with turn penalties) as an integer program.
Arrival route in weather forecast (left) is open in the actual weather (right)
Departure route in weather forecast (left) is blocked in the actual weather (right)
Approximately 400 routes were sampled from 14 “worst weather” days at Hartsfield-Jackson Atlanta International Airport (ATL) during June and July 2007.

<table>
<thead>
<tr>
<th>t₀</th>
<th># Paths</th>
<th>Fx Open (%)</th>
<th>Act. Open (%)</th>
<th>% Act. Open</th>
<th>% Act. Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>408</td>
<td>51</td>
<td>78</td>
<td>99</td>
<td>57</td>
</tr>
<tr>
<td>30</td>
<td>408</td>
<td>54</td>
<td>78</td>
<td>96</td>
<td>58</td>
</tr>
<tr>
<td>60</td>
<td>384</td>
<td>55</td>
<td>78</td>
<td>93</td>
<td>60</td>
</tr>
<tr>
<td>90</td>
<td>392</td>
<td>63</td>
<td>78</td>
<td>88</td>
<td>62</td>
</tr>
<tr>
<td>100</td>
<td>408</td>
<td>65</td>
<td>78</td>
<td>87</td>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>t₀</th>
<th># Paths</th>
<th>Fx Open (%)</th>
<th>Act. Open (%)</th>
<th>% Act. Open</th>
<th>% Act. Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>408</td>
<td>55</td>
<td>79</td>
<td>99</td>
<td>55</td>
</tr>
<tr>
<td>30</td>
<td>408</td>
<td>53</td>
<td>79</td>
<td>95</td>
<td>61</td>
</tr>
<tr>
<td>60</td>
<td>384</td>
<td>55</td>
<td>79</td>
<td>94</td>
<td>62</td>
</tr>
<tr>
<td>90</td>
<td>392</td>
<td>61</td>
<td>79</td>
<td>91</td>
<td>60</td>
</tr>
<tr>
<td>100</td>
<td>408</td>
<td>66</td>
<td>79</td>
<td>90</td>
<td>58</td>
</tr>
</tbody>
</table>

Routes that are forecast open are overwhelmingly open in the actual weather, for all time horizons.

This raw data suggests it is reasonable to plan routings at a 2-hour forecast time horizon.
The following 11 characteristics of the forecast weather were chosen for their potential correlation with route blockage:

1. Mean VIL along path
2. Standard Deviation of VIL along path
3. Minimum distance to level 3+ weather along path
4. Mean distance to level 3+ weather along path
5. Maximum VIL in neighborhood of path
6. Theoretical capacity for weather scenario
7. Number of bottle necks in weather scenario
8. Length of bottleneck through which path passes
9. Minimal bottleneck
10. Maximum density of level 3+ weather along path
11. Maximum VIL density of level 3+ weather along path

Compute mutual information between each feature and blockage, to evaluate features for classification, and better understand their correlation.

\[
I[X;Y] = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)}
\]
Example: Mutual Information values for the features, compared across time horizon.
Visually, many features tend to correlate with route blockage

Feature 1 (average VIL along path) correlated with blockage
(60-min time horizon, arrival, B=8km)

- The empirical probability that a route is blocked increases with average VIL along the route

[Michalek and Balakrishnan, Annual Meeting of the American Meteorological Society 2009]
Classification training objectives

- Standard two-class confusion matrix:

<table>
<thead>
<tr>
<th></th>
<th>Predicted Open</th>
<th>Predicted Blocked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Open</td>
<td>True Positive (TP)</td>
<td>False Negative (FN)</td>
</tr>
<tr>
<td>Actual Blocked</td>
<td>False Positive (FP)</td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>

- For our application, we want to minimize scenarios in which we predict that the route is open, but in reality it is blocked (FP)
  - In other words, we seek to maximize the recall:

\[ \alpha^{-} = \frac{TN}{TN + FP} \]
Classifying robust routes

- Trained 2 classifiers
  - EnsSVM - ensemble of Support Vector Machines
  - WRF - weighted random forest (a weighted ensemble of decision trees)
- Trained on 70% of data instances, tested on remaining 30%, careful not to introduce bias in the split
- Created a separate classifier for each of 5 time horizons (10, 30, 60, 90, and 100-mins), and for both departures and arrivals

[Diagram Source: Liu et al, Boosting Predication Accuracy on Imbalanced Datasets with SVM Ensembles, 2006]
Results for Ensemble of SVMs

<table>
<thead>
<tr>
<th></th>
<th>10-min</th>
<th>30-min</th>
<th>60-min</th>
<th>90-min</th>
<th>100-min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EnSVM</td>
<td>Fx</td>
<td>EnSVM</td>
<td>Fx</td>
<td>EnSVM</td>
</tr>
<tr>
<td><strong>Arrivals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc</td>
<td>81.77</td>
<td>74.22</td>
<td>75.57</td>
<td>71.4</td>
<td>71.64</td>
</tr>
<tr>
<td>a⁻</td>
<td>89.95</td>
<td>96.97</td>
<td>86.17</td>
<td>90.6</td>
<td>71.28</td>
</tr>
<tr>
<td>a⁺</td>
<td>79.22</td>
<td>68.43</td>
<td>71.49</td>
<td>65.14</td>
<td>70.95</td>
</tr>
<tr>
<td>g-mean</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.77</td>
<td>0.71</td>
</tr>
<tr>
<td>% TP</td>
<td>64.15</td>
<td>55.32</td>
<td>55.99</td>
<td>50.95</td>
<td>57.43</td>
</tr>
<tr>
<td>% FP</td>
<td>1.72</td>
<td>0.45</td>
<td>3.16</td>
<td>2.29</td>
<td>5.18</td>
</tr>
<tr>
<td>% FN</td>
<td>16.5</td>
<td>25.33</td>
<td>21.27</td>
<td>26.30</td>
<td>23.19</td>
</tr>
<tr>
<td><strong>Departures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc</td>
<td>77.3</td>
<td>71.97</td>
<td>78.2</td>
<td>72.27</td>
<td>68.94</td>
</tr>
<tr>
<td>a⁻</td>
<td>98.31</td>
<td>99.31</td>
<td>88.55</td>
<td>90.52</td>
<td>82.35</td>
</tr>
<tr>
<td>a⁺</td>
<td>69.91</td>
<td>62.55</td>
<td>73.91</td>
<td>66.1</td>
<td>64.83</td>
</tr>
<tr>
<td>g-mean</td>
<td>0.83</td>
<td>0.79</td>
<td>0.8</td>
<td>0.77</td>
<td>0.73</td>
</tr>
<tr>
<td>% TP</td>
<td>53.05</td>
<td>47.45</td>
<td>57.77</td>
<td>51.51</td>
<td>51.92</td>
</tr>
<tr>
<td>% FP</td>
<td>0.44</td>
<td>0.17</td>
<td>2.32</td>
<td>1.99</td>
<td>3.42</td>
</tr>
<tr>
<td>% TN</td>
<td>24.25</td>
<td>24.52</td>
<td>20.43</td>
<td>20.76</td>
<td>17.01</td>
</tr>
<tr>
<td>% FN</td>
<td>22.26</td>
<td>27.86</td>
<td>19.48</td>
<td>25.74</td>
<td>27.64</td>
</tr>
</tbody>
</table>

- At the shortest time horizons, EnsSVM outperforms forecast in accuracy, but not in recall rate
- At the longer time-horizons (90- and 100-mins), the recall rates of the classifiers are higher (for arrivals, there is an 18% improvement in recall at 90-mins, and 37% at 100-mins) than those of the forecast
There is a tradeoff between False Positive Rate and accuracy

- The performance of the WRF classifier is similar to that of the EnsSVM
- The tradeoff between FP and accuracy rates is illustrated by varying the weight parameter (relative penalty on FN vs. FP) of the WRF

![Comparison of false positive and accuracy rates as a function of weight and time horizon](image)
The classifier gives a probability that a route is open in actual weather

Empirical validation of classifier’s probability of route being open

The classifier’s probability tends to give a conservative estimate of true probability of blockage, across time horizons and for both arrivals and departures
Creating capacity forecasts from predictions of route blockage

- For a terminal area with $m$ arrival gates, a stochastic arrival capacity can be estimated by:
  - Sample standard arrival route through each arrival gate
  - Run the route through the classifier, obtaining a probability that the route will be open in observed weather

- The capacity of the arrival airspace can be forecast as $\frac{C k}{m}$ with probability
  \[ \Pr(\text{exactly } k \text{ of the arrival routes are open}), \]
  where $C$ is the clear-weather capacity of the airspace.

ATL terminal area sectors and predefined fixes
Summary

- A data-driven approach is used to correlate features of a route with blockage.
- Using techniques from machine learning, we have developed classification algorithms to predict route blockage.
  - Can be extended to probabilistic forecasts of airspace capacity.
- This is promising for integration with ATM algorithms under uncertain weather.
- To our knowledge, the first attempt to develop ATM-specific validation and metrics for aviation weather forecasts.
  - Has led to the development of new scoring metrics for aviation weather forecast products [Matthews et al., 2009].
Thank you, any questions?