Using ADOPT Algorithm and Operational Data to Discover Precursors to Aviation Adverse Events

Vijay Janakiraman, Bryan Matthews and Nikunj Oza
Data Sciences Group
NASA Ames Research Center, CA, USA

AIAA Science and Technology Forum and Exposition (SciTech 2018)
Session: IS-15, Learning, Reasoning, and Data-Driven Systems I
11th January 2018
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
Outline

• **Background**
  – Precursor discovery problem, uses, challenges

• **Methodology**
  – ADOPT algorithm

• **Case Studies**
  – Take-off Stall Hazard
  – STAR procedure adherence

• **Summary**
Precursor discovery
Precursor discovery

1000 ft, ~3 miles out

Speed Exceedance
Airspeed > threshold

Distance to touchdown

Runway

3mi
0mi
Precursor discovery

- Turn to final
  - Engine speed unusually high
  - Autopilot Mode mismatch

- Final Flaps not set
  - Tailwinds high

- Engine speed reduced
  - Flaps half way down
Precursor discovery

Data from one flight

~ 2500ft, 10 miles out
- Turn to final
  - Engine speed unusually high
  - Autopilot Mode mismatch

~ 2500ft, 5 miles out
- Engine speed reduced
- Flaps half way down

~ 2700ft, 15 miles out
All variables normal

~ 1500ft, 5 miles out
- Final Flaps not set
- Tailwinds high
Precursor discovery

- ~2500ft, 10 miles out
  - Turn to final
  - Engine speed unusually high
  - Autopilot Mode mismatch

- ~1500ft, 5 miles out
  - Final Flaps not set
  - Tailwinds high

- ~2700ft, 15 miles out
  - All variables normal

- ~2500ft, 5 miles out
  - Engine speed reduced
  - Flaps half way down

Runway 3mi 0mi

Speed Exceedance
Precursor discovery

Precursors

~ 2500ft, 10 miles out
  - Turn to final
  - Engine speed unusually high
  - Autopilot Mode mismatch

~ 2700ft, 15 miles out
  All variables normal

~ 2500ft, 5 miles out
  - Engine speed reduced
  - Flaps half way down

~ 1500ft, 5 miles out
  - Final Flaps not set
  - Tailwinds high

Speed Exceedance

Flight timeline
Precursor discovery

Precursors

~ 2500ft, 10 miles out
- Turn to final
- Engine speed unusually high
- Autopilot Mode mismatch

~ 2700ft, 15 miles out
All variables normal

Corrections

~ 2500ft, 5 miles out
- Engine speed reduced
- Flaps half way down

~ 1500ft, 5 miles out
- Final Flaps not set
- Tailwinds high

Speed Exceedance

Flight timeline

Probability of speed exceedance

Precursors

~ 2700ft, 15 miles out
All variables normal

Corrections

~ 2500ft, 5 miles out
- Engine speed reduced
- Flaps half way down

~ 1500ft, 5 miles out
- Final Flaps not set
- Tailwinds high

Speed Exceedance

Flight timeline

Probability of speed exceedance
Precursor discovery

Adverse event may be any event of interest

- **Single flight** safety events such as exceedances, go-around, stall,
- **Multi-flight** safety events such as loss of separation, TCAS events,
- **Airspace** or **NAS level** events such as GDP, congested sectors, delays,
- **Performance** events such as high throughput, mission success,
Why find precursors?

Forensic analysis of past events
- Accident investigation
- Hazard identification
- Operations

Precursors
- ~2500ft, 10 miles out
  - Turn to final
  - Engine speed unusually high
  - Autopilot Mode mismatch

~2700ft, 15 miles out
All variables normal

Precursors
- ~1500ft, 5 miles out
  - Final Flaps not set
  - Tailwinds high

Corrections
- ~2500ft, 5 miles out
  - Engine speed reduced
  - Flaps half way down

Speed Exceedance
Why find precursors?

Real-time decision support

- Crew alerting, Situational awareness, Action recommendation
Why find precursors?

• **Forecasting adverse events better and earlier**
  – Generate a knowledge base (precursors)

• **Develop decision support tools**
  – Alerting systems
  – Recommendation systems on corrective actions

• **Improve operator training**
  – Response and recovery from precursors

• **Predictive maintenance**
  – Precursors to component failures
Challenges in Precursor Discovery

• **Human expert analysis is not scalable**
  – Not easy to find patterns in 100s of time series.
  – Visualization is almost impossible.
  – Subjective variations among experts
  – Costly and slow

• **Data mining is not easy**
  – High dimensions (100s of variables)
  – High velocity of data (1000s of flights per day)
  – Data heterogeneity (continuous, categorical, text, voice, video)
  – Precursors are unlabeled.
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
Anatomy of a Safety Event

Scenario = (Context, Behavior, Outcome)
Precursor discovery using data

Data matrix $[X]_{d \times L}$

Time: $1 \ 2 \ 3 \ \ldots \ \ldots \ k \ \ldots \ L \ L + 1 \ \ldots \ T$

Event at time $k$

$\begin{bmatrix}
  x_1(k) \\
  x_2(k) \\
  \vdots \\
  x_d(k)
\end{bmatrix}$

Start of adverse event
Problem setup

• Data
  – Adverse time series data \( \overline{\mathcal{N}} = \{X_i\}, \ i = 1, 2, \ldots, \overline{N}; \)
  – Nominal time series data \( \mathcal{N} = \{X_i\}, \ i = 1, 2, \ldots, N; \)
  – Unsupervised

• Event
  – A time slice of data
    \[
    \begin{bmatrix}
    x_1(k) \\
    x_2(k) \\
    \vdots \\
    x_d(k)
    \end{bmatrix}
    \]
  – Data is a sequence of events \( X_i = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L_i)]_i \)
Precursor Definition

Given a sequence of events $X = [x(1), x(2), \ldots, x(L)]$, an action is any state transition $a_k : \mathbf{x}(k) \rightarrow \mathbf{x}(k + 1)$ where $1 \leq k \leq L$, then $a_k$ is a precursor to $E_A$ if

$$V(a_k) - V(a_k^*) > \delta,$$

where $\delta > 0$,

$a_k^*$ is the expert’s action at $k$,

$V(z)$ is the value function $\propto P(E_A|z)$
Related work

- Precursor discovery in multivariate time series is a new problem
  - No direct algorithm exists

**Challenges**
1. Unsupervised (no ground truth on precursors)
2. Temporal (long sequences make it hard)
3. High dimensionality

**Possible approaches**

- **Rule Mining**
  - Temporal rule mining [1,2] ① ②
  - Motif mining [7] ① ②
  - Clustering [10] ③

- **Model Based**
  - HMM [9] ① ②
  - Utility based rules [8] ① ②

- **Causality**
  - Causal Bayesian models, Granger causality [3,4,5,6] ① ②

**Issues/Drawbacks**
① Computationally expensive (scales combinatorial/exponential with number of items).
② Doesn’t handle continuous data (or needs discretization which grows combinatorial).
③ Similarity metric not easy to define for high dimensional data.
References

Background: Markov Model

An MDP is a tuple \((\mathcal{S}, \mathcal{A}, P_{s,a}, \gamma, R)\)

- \(\mathcal{S} = \mathbb{R}^d\) is a continuous state space with \(d\) state variables,

- \(\mathcal{A} = \mathbb{R}^l\) is an action space with \(l\) action variables,

- \(\{P_{s,s'}^a\}\) (or \(P_{ss'}\) if actions are unknown) are the state transition probabilities,

- \(\gamma \in [0, 1]\) is the discount factor,

- \(R : \mathcal{S} \rightarrow \mathbb{R}\) is the underlying reward function,

\[
\text{policy: } \pi(s, a) = p(a|s),
\]

\[
\text{optimal policy: } \pi_E(s) = a^*,
\]
Background: Value Function and Bellman’s Optimality

- The value of state $x_0$ under policy $\pi$ is
  \[
  V^\pi(x_0) = E[R(x_0) + \gamma R(x_1) + \ldots + \gamma^L R(x_L)|\pi]
  \]
  where the expectation is over the distribution of sequences starting from $x_0$.

- The expert’s policy $\pi_E$
  \[
  \pi_E(x) \geq \pi_i(x) \iff V^{\pi_E}(x) \geq V^{\pi_i}(x) \quad \forall \pi_i
  \]

- Bellman’s optimality
  \[
  \pi_E(x_k) = \arg \max_{\text{feasible } x_{k+1}} V^{\pi_E}(x_{k+1})
  \]
ADOPT Framework

1. Adverse time series, $\bar{N}$
   - Time series database
   - Nominal time series, $N$
   - Inverse Reinforcement Learning
     - Expert's reward model, $\hat{R}(s)$

2. Reinforcement Learning
   - Expert's value model, $\hat{V}_{\pi_E}(s)$

3. Precursor Discovery
   - Test trajectory $X_T \in \bar{N}$
   - Precursors $P_{E_A}$
Step 1: Expert’s Reward Model

- \( R(x; \alpha) = \alpha_1 \phi_1(x) + \alpha_2 \phi_2(x) + \ldots + \alpha_m \phi_m(x) \)
  
  - A general model of the expert’s reward
  
  - \( \alpha = [\alpha_1 \quad \alpha_2 \quad \ldots \quad \alpha_m]^T \) to be estimated
  
  - \( \phi_i(x); i = 1, 2, \ldots, m \) are some known basis functions (gaussian)

- \( \alpha^* = \arg \min_{\alpha} \{ E_{x_0} [V^{\pi_{adv}}(x_0; \alpha)] - E_{x_0} [V^{\pi_E}(x_0; \alpha)] \} \)
  
  - such that \( |\alpha_i| \leq 1, i = 1, 1, \ldots, m \)

\[
\hat{R}(x) = f_R(x; \alpha^*) = \sum_{i=1}^{m} \alpha_i^* \phi_i(x)
\]

ADOPT Framework

Time series database

Adverse time series, $\mathcal{N}$

Nominal time series, $\mathcal{N}'$

Inverse Reinforcement Learning

Test trajectory $X_T \in \mathcal{N}$

Expert’s reward model, $\tilde{R}(s)$

Precursor Discovery

Expert’s value model, $\hat{V}_{\pi E}(s)$

Precursors $P_{E_A}$
Step 2: Expert’s Value Model

- Value estimation by Monte Carlo method
  - Known reward (from previous step)
  - Time series data as Monte Carlo samples
  - Return $\text{Ret}(x) = \sum \gamma^k R(x_k)$ for each state $x$ as accumulated rewards

- For every labeled pair $(x_i, \text{Ret}(x_i))$, a regression model $\hat{V}^{\pi_E}(x; \theta)$ parameterized by $\theta$ can be built

$$\theta^* = \arg\min_{\theta} \frac{1}{N_s} \sum_{i=1}^{N_s} \| \text{Ret}(x_i) - \hat{V}^{\pi_E}(x_i; \theta) \|^2 + \frac{\mu}{2} \| \theta \|^2$$

$$\hat{V}^{\pi_E}(x) = f_V(x; \theta^*)$$
Step 3: Precursor Discovery

Given a sequence of events \( X = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L)] \), an action is any state transition \( a_k : \mathbf{x}(k) \rightarrow \mathbf{x}(k + 1) \) where \( 1 \leq k \leq L \), then \( a_k \) is a precursor to \( E_A \) if

\[
V(a_k^*) - V(a_k) > \delta.
\]

- requires finding the “optimal” decision
  - Bellman’s optimality
    \[
    \mathbf{x}_{k+1}^* = \arg \max_{\{\text{feasible} \ \mathbf{x}_{k+1}\}} V^{\pi_E}(\mathbf{x}_{k+1})
    \]

- requires scoring the suboptimal decisions
  - \( PI_k = V^{\pi_E}(\mathbf{x}_{k+1}^*) - V^{\pi_E}(\mathbf{x}_{k+1}) \)
  - A weighted contribution from reward may be added to tradeoff short term vs long term precursors
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
Take-off Stall Hazard

Adverse event: Drop in airspeed after take-off by at least a 20 knots

Goal: To find precursors using flight recorded data
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
Factors affecting drop in airspeed

- **Human Factors**
  - Errors in reference speed calculations, estimating AC weight, energy management.
  - Human-machine interactions, fatigue, aggressive flying, mode confusion.

- **Environmental**
  - Tail winds, wind shear, sensor failure

- **Procedural**
  - Avoiding terrain, flying over restricted area
ADOPT analysis

400 nominal flights
400 adverse flights
200 (100+100) holdout set

Time series database

Adverse time series, $\mathcal{N}$

Nominal time series, $\mathcal{N}'$

Inverse Reinforcement Learning

Test trajectory

$X_T \in \mathcal{N}$

Expert’s reward model, $\hat{R}(s)$

Precursor Discovery

Precursors

$P_{EA}$

Reinforcement Learning

Expert’s value model, $\hat{V}^{\pi_E}(s)$

SVM

# basis functions: 5000

Spread of Gaussian: 0.05
Flight analysis 1 – reference speed set incorrectly

<table>
<thead>
<tr>
<th>Time</th>
<th>4s</th>
<th>30s</th>
<th>35s</th>
<th>40s</th>
<th>45s</th>
<th>50s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailwind</td>
<td>Pitch Angle</td>
<td>PFD Spd</td>
<td>PFD Spd</td>
<td>Pitch Angle</td>
<td>Pitch Angle</td>
<td>PFD Spd</td>
</tr>
<tr>
<td>Pitch Angle</td>
<td>Altitude</td>
<td>Pitch Angle</td>
<td>Tailwind</td>
<td>PFD Spd</td>
<td>PFD Spd</td>
<td>Pitch Angle</td>
</tr>
<tr>
<td>Roll Angle</td>
<td>Roll Angle</td>
<td>Auto-throttle</td>
<td>Pitch Angle</td>
<td>Tailwind</td>
<td>Tailwind</td>
<td>Auto-throttle</td>
</tr>
<tr>
<td>Altitude</td>
<td>Tailwind</td>
<td>Roll Angle</td>
<td>Roll Angle</td>
<td>Auto-throttle</td>
<td>Auto-throttle</td>
<td>Auto-throttle</td>
</tr>
<tr>
<td>Auto-throttle</td>
<td>Auto-throttle</td>
<td>Altitude</td>
<td>Altitude</td>
<td>Autopilot</td>
<td>Autopilot</td>
<td>Autopilot</td>
</tr>
</tbody>
</table>
Flight analysis 2 – reference speed set incorrectly

<table>
<thead>
<tr>
<th>Time = 1s</th>
<th>4s</th>
<th>20s</th>
<th>25s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch Angle</td>
<td>Tailwind</td>
<td>Tailwind</td>
<td>Pitch Angle</td>
</tr>
<tr>
<td>Tailwind</td>
<td>Pitch Angle</td>
<td>Roll Angle</td>
<td>PFD Spd</td>
</tr>
<tr>
<td>Altitude</td>
<td>Roll Angle</td>
<td>Roll Angle</td>
<td>Tailwind</td>
</tr>
<tr>
<td>Roll Angle</td>
<td>Altitude</td>
<td>PFD Spd</td>
<td>Roll Angle</td>
</tr>
<tr>
<td>Auto-throttle</td>
<td>Auto-throttle</td>
<td>Altitude</td>
<td>Altitude</td>
</tr>
</tbody>
</table>
Flight analysis 1 – Nominal Flight
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
STAR procedure adherence

Adverse event: Drop in airspeed after take-off by at least a 20 knots

Goal: To find precursors using flight recorded data
Adverse event: Drop in airspeed after take-off by at least a 20 knots

Goal: To find precursors using flight recorded data
ADOPT analysis

400 nominal flights
400 adverse flights
200 (100+100) holdout set
Flight analysis 1 – reference speed set incorrectly

<table>
<thead>
<tr>
<th>Time</th>
<th>Tailwind</th>
<th>Pitch Angle</th>
<th>Roll Angle</th>
<th>Autopilot</th>
<th>Auto-throttle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>Pitch Angle</td>
<td>Roll Angle</td>
<td>Roll Angle</td>
<td>Autopilot</td>
<td>Auto-throttle</td>
</tr>
<tr>
<td>20s</td>
<td>Roll Angle</td>
<td>Tailwind</td>
<td>Auto-throttle</td>
<td>Roll Angle</td>
<td>Tailwind</td>
</tr>
<tr>
<td>25s</td>
<td>PFD Spd</td>
<td>Auto-throttle</td>
<td>Tailwind</td>
<td>Pitch Angle</td>
<td>Tailwind</td>
</tr>
<tr>
<td>30s</td>
<td>Tailwind</td>
<td>Auto-throttle</td>
<td>Tailwind</td>
<td>Pitch Angle</td>
<td>Pitch Angle</td>
</tr>
<tr>
<td>35s</td>
<td>PFD Spd</td>
<td>Auto-throttle</td>
<td>Tailwind</td>
<td>Pitch Angle</td>
<td>Pitch Angle</td>
</tr>
</tbody>
</table>
Flight analysis 2 – reference speed set incorrectly

<table>
<thead>
<tr>
<th>P</th>
<th>C</th>
<th>Je</th>
<th>D</th>
<th>Jo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>Slope-possible</td>
<td>Slope-possible</td>
<td>Slope-possible</td>
<td>Slope-possible</td>
</tr>
<tr>
<td>Ground Spd</td>
<td>Descent Rate</td>
<td>Ground Spd</td>
<td>Descent Rate</td>
<td>Altitude Skip</td>
</tr>
<tr>
<td>Descent Rate</td>
<td>Tailwind</td>
<td>Descent Rate</td>
<td>Tailwind</td>
<td>Descent Rate</td>
</tr>
<tr>
<td>Tailwind</td>
<td>Ground Spd</td>
<td>Tailwind</td>
<td>Ground Spd</td>
<td>Ground Spd</td>
</tr>
<tr>
<td>Altitude Skip</td>
<td>Altitude</td>
<td>Altitude Skip</td>
<td>Altitude Skip</td>
<td>Tailwind</td>
</tr>
<tr>
<td>Slope-possible</td>
<td>Altitude Skip</td>
<td>Altitude Skip</td>
<td>Altitude Skip</td>
<td>Altitude</td>
</tr>
</tbody>
</table>
Flight analysis 1 – Nominal Flight
Outline

• Background
  – Precursor discovery problem, uses, challenges

• Methodology
  – ADOPT algorithm

• Case Studies
  – Take-off Stall Hazard
  – STAR procedure adherence

• Summary
ADOPT’s features

• Data mining based precursor discovery algorithm

• Input
  – Feed in time series data with adverse event
  – Feed in nominal time series data
  – Data could be continuous, categorical, text, images

• Output
  – Precursor time instants
  – Precursor variables
  – Probability score

• Correlation and not Causation
ADOPT’s features

• Use any/all domain knowledge
  – Selecting variables
  – Scoping problems in space, time
  – Hand-engineering features

• Use any classifier of choice
  – SVM, decision tree, K-NN, logistic regression

• Extends to multiple adverse events
  – Holistic analysis, safety margins

• Parallelizable
  – Multiple CPUs
  – Analyze multiple airports, airspaces, aircrafts in parallel
Summary

• Precursor discovery is an important problem with uses in multiple applications in Aviation.

• ADOPT is an efficient data mining solution to find precursors.

• Two case studies are presented to show the setup, working and features of ADOPT.

• ADOPT will be open-sourced in the near future.
Thank You