Probabilistic Conflict Detection for Air Traffic Control

Oliver Watkins

ojw26@cam.ac.uk

University of Cambridge, UK
Conflict Detection in ATC

- Background
- Markov Chains and Monte Carlo
- The real world
- Problems with the real world
- Results
Background
Air Traffic Control Architecture

- Rigidly Structured Airspace.
- ATC in Complete Control.
- Concentrates complex traffic patterns around gates.
- Simplifies problem for human operators.
ATC Structure Progression

Current system: ATC in absolute control
Near term: increased information flow between aircraft, improvements in situational awareness.
Mid term: ATC allows limited requests from aircraft for more direct routings.
Long term: ATC remains in command, but collaborative decision making between aircraft is implemented.
Free flight, the ultimate goal? ATC retains a strategic role, but mid term route management is distributed to aircraft.
Tools for ATC Enhancement

- Progression to more complex traffic flows.
- Requires skills beyond human ATC.
- Introduce tools to assist ATC.
- Particularly conflict probes.

Is this dangerous?
Conflict Detection

Various forms exist, but *Probabilistic Conflict Detection* appears to be the most effective.

We wish to evaluate:

\[ PC(t) = \mathcal{P}\{\exists t \in [t, t + T] | x_r(t) \in T\} \]

Where \( x_r(t) \) is the relative position of two aircraft and is a realization of a stochastic difference equation:

\[ \dot{x}_r(t) = f(x_r(t)) x_r(t) + g(x_r(t)) \eta(t) \]

We are therefore trying to solve a *stochastic reachability* problem.
Theoretical Approach
Methods of Solution

A variety of approaches exist to provide an approximate solution to the SR problem, for the particular application.

- Make a suitable approximation of the nature of $x_r(t)$ to give an analytical approximation.

- Extract a relevant measure of $PC(t)$, which can be used in its place.

These approaches have some drawbacks:

- Tend to be oversimplified and cannot capture all behaviour.
- The approximations break down under certain conditions.
- Undesirable in a safety critical application area.
Monte Carlo Approach

Simulate many flight paths; calculate the proportion which enter conflict.

- Can handle arbitrary levels of complexity.
- Minimal simplifying assumptions.
- MC is very simple in its most popular form:

\[
\mathbb{E} [\varphi(x)] = \frac{1}{M} \sum_{i=1}^{M} \varphi(x).
\]

- Can construct a rigorous formalism.
Monte Carlo Integration

We wish to evaluate:

\[ I = \mathbb{E}[g(x)] = \int g(x) \, dx \]

If \( x \) models a physical quantity in a real experiment, then:

\[ I = \mathbb{E}[g(x)] \approx \frac{1}{M} \sum_{i=1}^{M} g(x_i) \]

\[ \mathcal{P}(A) = \frac{1}{M} \sum_{i=1}^{M} 1_{x_i \in A} \]
Markov Processes

We have two processes:

- $Y_t$, the *observation process*; real-world observed trajectories.
- $X_t$, the *signal process*; simulated trajectories.
- Assume that $X$ and $Y$ match.

Both are *Markov Processes*:

$$ P(x_k | x_{k-1}, x_{k-2}, \ldots, x_0) \equiv P(x_k | x_{k-1}). $$
Transition Kernels

Define a *transition kernel* on $X$:\n
$$\Pr\{X_k \in A \mid X_{k-1} = x_{k-1}\} = \int_A K_k(dx_k, x_{k-1}).$$

Transition kernels are (for our purposes):

- Spatially homogeneous.
- Time inhomogeneous.

So a transition kernel maps probability distributions:

$$K(x_1, x_0) \rightarrow K(x_2, x_1)$$
Transition Kernels

Hence we may build up a joint distribution:

\[ \mu_0(x_0) \prod_{k=1}^{T} K(x_k, x_{k-1}) \]

Appropriate sampling from the joint distribution yields a realization of an SDE trajectory:

\[ X_{0:T}^{(i)} \sim \mu_0(x_0) \prod_{k=1}^{T} K(x_k, x_{k-1}) \]
MC for Conflict Detection

We choose a function $\varphi(x_{t:T+t})$:

$$
\varphi(x_{t:T+T}) = \max_{k \in t:T+t} \left( I_k = \begin{cases} 
1 & x_k \in \mathcal{T} \\
0 & x_k \notin \mathcal{T} 
\end{cases} \right)
$$

The MC integration can then be written in the form:

$$
PC(t) = \int \varphi(x_{t:T+t}) \delta_y(x_t) \prod_{k=t+1}^{T+t} K(x_k, x_{k-1}) dx_{t:t+T}.
$$

With:

$$
P^\ast C(t) = \frac{1}{M} \sum_{i=1}^{M} \varphi \left( X_{t:T+t}^{(i)} \right)
$$
Efficiency of MC

A natural question regarding MC in this form: is it not a bit inefficient?
Efficiency of MC

A natural question regarding MC in this form: is it not a bit inefficient?

\[ y_0 \quad y_t \]
A natural question regarding MC in this form: is it not a bit inefficient?
Efficiency of MC

A natural question regarding MC in this form: is it not a bit inefficient?

\[ y_0, y_t \]

\[ T \]
Sequential Monte Carlo

We wish to re-use old trajectories. How?
Sequential Monte Carlo

We wish to re-use old trajectories. How?

Observation

Estimate

Prior Estimate

Traditional ‘Kalman-esque’ Filtering - can be approached using particle methods
Sequential Monte Carlo

We wish to re-use old trajectories. How?

Our situation; exact or near exact observations preclude re-using particles
Sequential Monte Carlo

We wish to re-use old trajectories. How?

Look ahead one step, and the weighting becomes practical.
Optimal SMC

Lemma 1: (Optimal SMC Estimation) \( PC(t) \), given state observation \( y_t \) can be expressed as the following integral with respect to the joint distribution of signal trajectories 
\[
\delta_{y_0}(x_0) \prod_{k=1}^{T} K(x_k, x_{k-1}).
\]

\[
PC(t) = \int \varphi(y_t, x_{t+1:T+t}) \frac{K(x_{t+1}, y_t)}{\mu(x_{t+1})} \delta_{y_0}(x_0) \prod_{k=1}^{T+t} K(x_k, x_{k-1}) dx_{0:T+t}.
\]
Hence the probability of conflict is estimated as:

\[ PC(t) = \sum_{i=1}^{M} \varphi \left( y_t, X_{t+1:T+t}^{(i)} \right) \xi^{(i)}, \]

\[ \xi^{(i)} \propto \frac{K(X_{t+1}, y_t)}{\mu(X_{t+1}|y_0)}, \sum_{i=1}^{M} \xi^{(i)} = 1. \]

Subject to the estimate being sufficiently accurate.
Accuracy of SMC

The Chernoff Bound gives the accuracy of (unweighted) MC:

\[
P \left\{ \left| \hat{P}C(t) - PC(t) \right| > \epsilon \right\} \leq 2e^{-2M\epsilon^2}.
\]

The accuracy of SMC is given by a modified Chernoff Bound:

\[
P \left\{ \left| \hat{P}C(t) - PC(t) \right| > \epsilon \right\} \leq 2e^{-2\hat{M}\epsilon^2},
\]

where \( \hat{M} \in [1, M] \) is the Effective Sample Size:

\[
\hat{M} = \frac{1}{\sum_{i=1}^{M} (\xi(i))^2}
\]
SMC Algorithm

Initialisation

Sample trajectories from joint distribution

Repeat

Repeat

Receive state observation

Perform SMC estimation of $PC(t)$

Until $\hat{M} < M_{crit}$

Resample trajectories from last observation

Until $t > \tau$
The Real World
The Real World

To implement the CD algorithm, we need to characterize $K(\cdot, \cdot)$. Assume that:

- The aircraft dynamics are fairly straightforward.
- Disturbances are dominated by wind, acting additively on the velocity:
  \[
  \begin{align*}
  \dot{X} &= V \cos(\psi) + w_1, \\
  \dot{Y} &= V \sin(\psi) + w_2.
  \end{align*}
  \]
- Wind has some statistical structure: appears to be correlated in both time and space.
Wind

Work on this project has suggested a correlation structure for instantaneous wind innovations:

\[ \rho(t, t', x, x') = \sigma(z)^2 \exp(-\lambda |t - t'|) \exp(-\beta \|x - x'\|). \]

With \( \beta = 1.6 \times 10^{-6} \), \( \lambda = 6 \times 10^{-6} \) and \( \sigma(z) = 7.716 \).

How does this translate into wind innovations in discrete time?
Wind

Work on this project has suggested a correlation structure for instantaneous wind innovations:

\[ \rho(t, t', x, x') = \sigma(z)^2 \exp(-\lambda|t - t'|) \exp(-\beta \|x - x'\|). \]

With \( \beta = 1.6 \times 10^{-6} \), \( \lambda = 6 \times 10^{-6} \) and \( \sigma(z) = 7.716 \).

How does this translate into wind innovations in discrete time?

There exist matrices \( B_{i,j} \) such that with \( n_{u,i} \sim \mathcal{N}(0, 1) \):

\[ w_{u,k} = B_{k,k} n_{u,k} + B_{k,k-1} n_{u,k-1} + \cdots + B_{k,1} n_{u,1}, \]

with \( B \in \mathbb{R}^{2 \times 2} \) and \( w_{u,k} = [w_{u,k}^{(1)}, w_{u,k}^{(2)}]^T \).
The Markov Kernel

This seems to present a problem. Recall that for a Markov Process:

\[ \mathcal{P}(x_k| x_{k-1}, x_{k-2}, \ldots, x_0) \equiv \mathcal{P}(x_k| x_{k-1}). \]
The Markov Kernel

This seems to present a problem. Recall that for a Markov Process:

\[ P(x_k|x_{k-1}, x_{k-2}, \ldots, x_0) \equiv P(x_k|x_{k-1}). \]

But if

\[ w_{u,k} = B_{k,k}n_{u,k} + B_{k,k-1}n_{u,k-1} + \cdots + B_{k,1}n_{u,1} \]

Then the evolution at each state depends on what has happened at ALL previous time steps.
The Markov Kernel

This seems to present a problem. Recall that for a Markov Process:

\[ P(x_k| x_{k-1}, x_{k-2}, \ldots, x_0) \equiv P(x_k| x_{k-1}) \]

But if

\[ w_{u,k} = B_{k,k} n_{u,k} + B_{k,k-1} n_{u,k-1} + \cdots + B_{k,1} n_{u,1} \]

Then the evolution at each state depends on what has happened at ALL previous time steps.

Ways around this have been investigated...
The Markov Kernel

This seems to present a problem. Recall that for a Markov Process:

\[ \mathcal{P}(x_k|x_{k-1}, x_{k-2}, \ldots, x_0) \equiv \mathcal{P}(x_k|x_{k-1}). \]

But if

\[ w_{u,k} = B_{k,k} n_{u,k} + B_{k,k-1} n_{u,k-1} + \cdots + B_{k,1} n_{u,1} \]

Then the evolution at each state depends on what has happened at ALL previous time steps.

Ways around this have been investigated...

...and found to be unworkable.
Model Abstraction

- This seems to suggest that temporal correlation information cannot be exploited to improve CD (spatial is OK).
Model Abstraction

- This seems to suggest that temporal correlation information cannot be exploited to improve CD (spatial is OK).
- But temporal correlation is very strong; it’s really going to help detect conflicts.
Model Abstraction

- This seems to suggest that temporal correlation information cannot be exploited to improve CD (spatial is OK).

- But temporal correlation is very strong; it’s really going to help detect conflicts.

So compose another model, which preserves the instantaneous deviation statistics, but is generated by time independent innovations:
Model Abstraction

- This seems to suggest that temporal correlation information cannot be exploited to improve CD (spatial is OK).
- But temporal correlation is very strong; it’s really going to help detect conflicts.

So compose another model, which preserves the instantaneous deviation statistics, but is generated by time independent innovations:

\[ x_{k+1} = K_k x_k + \eta_k. \]

This then encodes valuable correlation information, but SMC can be applied.
Model Abstraction II

- Firstly track the relative deviation from the flight plan (along-track and cross-track) as a function of generic wind perturbations.
Model Abstraction II

- Firstly track the relative deviation from the flight plan (along-track and cross-track) as a function of generic wind perturbations.
- This establishes mean position $\mu(t)$, and covariance $V(t)$. 
Firstly track the relative deviation from the flight plan (along-track and cross-track) as a function of generic wind perturbations.

This establishes mean position $\mu(t)$, and covariance $V(t)$.

Then choose a model of the form:

$$x_{k+1} = K_k x_k + \eta_k.$$ 

Find $K_k$ such that the same $\mu(t)$ and $V(t)$ are generated.
• The abstract model respects the deviation statistics of the complex model.

• Some information on wind correlation should be preserved.

• Does it work?
Results
Alerting Logic

• What makes a ‘good’ conflict probe?
Alerting Logic

- What makes a ‘good’ conflict probe?
- In the current architecture - Issuing of effective *Alerts to ATC*
Alerting Logic

- What makes a ‘good’ conflict probe?
- In the current architecture - Issuing of effective *Alerts to ATC*
- False alerts (alert but no impending conflict) are wasteful.
Alerting Logic

- What makes a ‘good’ conflict probe?
- In the current architecture - Issuing of effective Alerts to ATC
- False alerts (alert but no impending conflict) are wasteful.
- Missed alerts (no alert, but subsequent conflict) are dangerous.
Alerting Logic

- What makes a ‘good’ conflict probe?

- In the current architecture - Issuing of effective Alerts to ATC

- False alerts (alert but no impending conflict) are wasteful.

- Missed alerts (no alert, but subsequent conflict) are dangerous.

- We have the freedom to choose the Alert Threshold

\[ \bar{C} \in [0, 1] \]
Alerting Logic

- What makes a ‘good’ conflict probe?
- In the current architecture - Issuing of effective *Alerts to ATC*
- False alerts (alert but no impending conflict) are wasteful.
- Missed alerts (no alert, but subsequent conflict) are dangerous.
- We have the freedom to choose the *Alert Threshold* 
  \[ \bar{C} \in [0, 1] \]
- Choose \( \bar{C} \) to optimize the trade-off between \( P(SA) \) and \( P(FA) \).
SOC Curves

Alerts are required some specified time before conflict occurs. The SOC curve plots $P(SA)$ against $P(FA)$ parametrized by alert threshold.

The point closest to $(0, 1)$ is the optimal operating point. Or:

$$
\tilde{C}_{opt} = \arg \min_{C \in [0,1]} \sqrt{P(FA)^2 + (1 - P(SA))^2}.
$$
Comparative Results

SOC curves for random geometries, required warning 1 min

SOC curves for random geometries, required warning 5 min
Comparative Results

SOC curves for random geometries, required warning 10 min

SOC curves for random geometries, required warning 15 min
Comparative Results

Consistently superior performance for all required warning times.
Conclusions

• MC methods are practical and appropriate for CD.

• SMC methods provide substantial improvements in efficiency.

• Temporal Correlation in disturbances presents some difficulties.

• Model abstraction enables us to exploit knowledge of correlation.

• The resulting methods give a 10% – 20% improvement in the performance measure.
Acknowledgements

This work has been funded by the Schiff Foundation, the European Commission under the project HYBRIDGE IST-2001-32460, and Sidney Sussex College.

Thanks to Arnaud Doucet and Maria Prandini for their assistance.