Detection of individual anomalous arrival trajectories within the terminal airspace using persistent homology

Max Z. Li, Megan S. Ryerson, Ph.D.

University of Pennsylvania
Massachusetts Institute of Technology

June 28, 2018
1. Introduction and motivation
2. Method walk-through
   - Ideas from abstract math
   - Applied algebraic topology
3. Detection of anomalous trajectories within ORD data set
4. Preliminary results
5. Future work and conclusions
Jumping on the big data bandwagon

**LARGE and sparse data**
Preliminaries

Data, data, data

Jumping on the big data bandwagon

192003 × 58

And it gets worse from here ...
... much, much worse.

\[ \sim 4 \text{ gb} \text{ worth of trajectory data} \]
Jumping on the big data bandwagon

LARGE and sparse data

... and sometimes coordinate-free as well.
Shifting gears a little

topology
Topology – Esoteric and interesting. But useful?
Sampler: Low-dimensional topology (e.g. knot theory)
**Fact:** On the Earth’s equator, there *always* exists a pair of *opposite points* with the *same temperature* \(^1\).

\(^1\)Assuming that the surface temperature of the Earth can be thought of as a continuous function.
Fact: On the Earth’s equator, there always exists a pair of opposite points with the same temperature.

Theorem (Borsuk-Ulam)

Let $f$ be a continuous map from the $n$-sphere to the reals. Then there exists a point on the $n$-sphere such that $f(x) = f(-x)$. 
\[
\frac{d}{dx} f(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}
\]
$A \mathbf{v} = \lambda \mathbf{v}$
\[ F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x k} \, dx \]
From abstract to applied pt. 4 – algebra

\[ \mathbb{Z}/N\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z} \]
Topological data analysis and persistent homology (TDA/PH) uses results from algebraic topology to extract qualitative global information about the shape of your data set.

TDA/PH provides robust results since the overall shape of the data set should remain invariant even with injected noise (!)

(Ghrist, 2010; 2014), (de Silva & Ghrist, 2007)
Diversity of applications

Useful for problems involving large, high-dimensional data sets with high degrees of connectivity

A sampling of applications include:

- Networked neuroscience \textit{(Giusti et al., 2016)}
- **Higher-order graphs** \textit{(Huang & Ribeiro, 2015; 2016; 2017)}
- **Robotics** \textit{(Bhattacharya et al., 2013; 2014; 2015)}
- Fluid dynamics \textit{(Kasten et al., 2011)}
- Crystallography \textit{(Grottel et al., 2011)}
- LiDaR parsing \textit{(Keller et al., 2011)}
- Medical research \textit{(Szymczak, 2011)}
- **Aviation?**
Anomalous cyclic trajectories

Why do we care about airborne holds?

Why do we care about missed approaches/go-arounds?
Airborne holding (1/3)

- An important metric ...

  - Included as significant phase-of-flight for calculations of greenhouse gas emissions (*Evans & Schafer, 2011*)
  - Quantify improvements from new Multi-center Traffic Management Advisor (*Farley et al., 2005*)
  - Overall NAS performance under GDPs (*Ball et al., 2001*)
• An important **model specification/constraint** ...

- Model preference for assigning ground holding in lieu of airborne holding (*Ball et al.* 2003); (*Clark, 2009*)
- Upper bound constraint on number of airborne-holding aircraft (*Mukherjee & Hansen, 2007*)
- Inclusion of airborne holding cost in optimization model (*Andreatta et al., 2011*)
• An potential point of improvement ...

• Balancing airborne and ground delays $\rightarrow$ benefits on reducing total delay (Lulli & Odoni, 2007)
• New flight trajectories to increase benefits of airborne holding (Xu et al., 2017), (Xu & Prats, 2017)
1. A toy example ...
2. A dash of topology with a sprinkling of algebra
3. A most useful construction
A flight with a hold ...
... do you see the hole?
A fundamental question

It is **visually obvious** that a **hole** exists within the topology of this toy example ...

... but how do we formalize this observation?
A quick primer (1/4)

Definition (Homeomorphisms)

Map \( f : S \rightarrow U \) is a *homeomorphism* if \( f \) is a bijection and both \( f \) and \( f^{-1} \) are continuous mappings. Spaces \( S \) and \( U \) are *homeomorphic* if such a map \( f \) exists.

A *hole* is a non-trivial topology as it prevents *smooth deformations* from one space to another.
A **hole** is a non-trivial topology as it prevents *smooth deformations* from one space to another.

**Big idea:** *Homology* is a way of “keeping track” of holes between spaces (manifolds) using tools from algebra.

(Munkres, 1984; 2000), (Hatcher, 2002; 2005), (Ghrist, 2010)
Big idea: *Homology* is a way of “keeping track” of holes between spaces (manifolds) using tools from algebra.

Given some manifold $X$, with *cycle* $c$ and *boundary* $c'$ related to cycle $c$ via *boundary map* $c = \partial c'$, on the level of *algebra* ...

... we can construct a *chain complex* of abelian groups $C_n$ chained together by *boundary homomorphisms* $\partial_* : C_n \to C_{n-1}$.

In particular, we are interested in $\partial_{n+1} : C_{n+1} \to C_n$ and $\partial_n : C_n \to C_{n-1}$, where boundaries $\text{im} \; \partial_{n+1}$ form a subgroup under the cycles $\text{ker} \; \partial_n$. The $n^{th}$ *homology group* for $X$ is the equivalence class $H_n = \ker \partial_n \setminus \text{im} \; \partial_{n+1}$.

Theorem (Seifert-van Kampen)

Let $X = U \cup V$ be decomposable into open, path-connected sets $U, V$ and $U \cap V$ is also path-connected. Let $\iota : U \cap V \to U$ and $\kappa : U \cap V \to V$ be inclusion maps, and examine a point $x_0 \in U \cap V$. Suppose you know the fundamental groups of the decomposed spaces and their intersection, given by:

\[
\pi_1(U, x_0) = (a_1, \ldots, a_\alpha : r_1, \ldots, r_\rho)
\]
\[
\pi_1(V, x_0) = (b_1, \ldots, b_\beta : s_1, \ldots, s_\sigma)
\]
\[
\pi_1(U \cap V, x_0) = (c_1, \ldots, c_\chi : t_1, \ldots, t_\tau)
\]

Then the fundamental group of the original space $X$ can be written as:

\[
\pi_1(X, x_0) = (a_1, \ldots, a_\alpha, b_1, \ldots, b_\beta : r_1, \ldots, r_\rho, s_1, \ldots, s_\sigma, \\
\iota_* (c_1) = \kappa_* (c_1), \ldots, \iota_* (c_\chi) = \kappa_* (c_\chi))
\]
Theorem (Mayer-Vietoris)

Let \( X = U \cup V \) be decomposable into open, path-connected sets \( U, V \) and \( U \cap V \) is also path-connected. There exists a long exact sequence that relates the homology groups \( H_k(X) \) of the original space \( X \) to the homology groups of the decomposed spaces and their intersection.

\[
\cdots \xrightarrow{\partial} H_{k+1}(X) \xrightarrow{\partial} H_k(U \cap V) \rightarrow H_k(U) \oplus H_k(V) \rightarrow H_k(X) \xrightarrow{\partial} \cdots
\]

Mayer-Vietoris is analogous to Seifert-van Kampen, but for homology groups instead of fundamental groups (homology vs. homotopy).
Goal: Dynamically generate spaces $X$ out of our toy example with shifting topologies and look for \textit{persistent homological features}.

Idea:

1. Choose radius $r$
2. Construct $r$-neighborhood centered at lat-long $x = (\varphi, \lambda)$
3. Connect all pairwise $x$ that are \textit{no more than $r$ apart} – this is \textit{simplicial complex} construction
   - Points (0-simplex)
   - Edges (1-simplex)
   - Triangular-faces (2-simplex)
   - ...
4. Define the holes we’re looking for as \textit{loops consisting of only 1-simplexes}
5. Vary $r$ – \textit{what happens?}
**Goal:** Dynamically generate spaces $X$ out of our toy example with shifting topologies and look for *persistent homological features*.

- Some $r_i$ where significant hole is formed (*birth*)
- Some $r_f$ ($r_f > r_i$) where significant hole is covered up again (*death*)
Goal: Dynamically generate spaces $X$ out of our toy example with shifting topologies and look for *persistent homological features*.
Our path is clear(-ish)!

Compute *persistent homological features* for each flight’s trajectory data subset to look for anomalies such as cycles and holes.

We utilized the **TDA** package in **R** \((Fasy \ et \ al., \ 2014; \ 2017)\). Instead of the simpler *simplicial complex* construction, they used *Delaunay complex* construction, more specifically *Alpha complexes* (subcomplex of Delaunay complexes). More details in \((Edelsbrunner \ & \ Harer, \ 2010)\).

We will give a cursory overview for completeness.
**Alpha complex** $\mathbb{A}(X, r)$ construction (1/3)

**Input:** Latitude-longitude coordinates $(\varphi, \lambda) \in X \subset \mathbb{R}^2$

Create *Voronoi partition* via $\mathcal{V}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$

Create canonical $r$-neighborhood $\mathcal{B}((\varphi, \lambda), r)$

**Definition (Individual Alpha complex)**

An *individual Alpha complex* $\mathcal{R}((\varphi, \lambda), r)$ is the intersection of a $r$-neighborhood with the Voronoi partition for a specific point $(\varphi, \lambda)$.

$$\mathcal{R}((\varphi, \lambda), r) = \mathcal{B}((\varphi, \lambda), r) \cap \mathcal{V}((\varphi, \lambda)) \quad (1)$$
Definition (Alpha simplicial complex $\mathbb{A}(X, R)$)

Given trajectory data set $X$ composed of latitude-longitude coordinates $(\varphi, \lambda)$, the *Alpha simplicial complex* $\mathbb{A}(X, r)$ of circumradius $r$ is given by the collection of data points $\sigma \subseteq X$ satisfying Eq. 2:

$$
\mathbb{A}(X, r) = \left\{ \sigma \subseteq X \mid \bigcap_{(\varphi, \lambda) \in \sigma} R((\varphi, \lambda), r) \neq \emptyset \right\} \tag{2}
$$

*(Edelsbrunner & Harer, 2010), (Fasy et al., 2017)*
Alpha complex $\alpha(X, r)$ construction (3/3)

(Edelsbrunner & Harer, 2010)
Several trajectory data sets of ORD arrivals within terminal arrival airspace

Computation of Alpha complex $\mathbb{A}(\star, r)$

Persistent features via birth-death persistence diagrams

Presentation & discussion of preliminary results

What’s next?!
ORD terminal arrival airspace (1/2)

ORD terminal arrival airspace model

FYTTE FOUR

WYNDE EIGHT

WATSN THREE

ESSPO THREE

BENKY FOUR

VEECK THREE

TRTLL FOUR
ORD terminal arrival airspace (2/2)

ORD final approach corridor model

East Flow IAPs

West Flow IAPs

(not to scale)
Note the in-air holding in NE/SE gates as weather deteriorates and ORD’s configurations change.
Subset trajectory data set $X$ into $X_1, X_2, \ldots, X_N$ individual flight trajectory data subsets.

Form $A(X_i, r)$ for all $i = 1, 2, \ldots, N$, begin varying $r$ noting when features appear (are born) and vanish (die).

Record birth/death of features on birth-death persistence diagrams.
Given feature $\Delta_i$, it can be mapped to $(r_{\text{birth}}, r_{\text{death}})_{\Delta_i}$

Many ways to present $\{\Delta_i, (r_{\text{birth}}, r_{\text{death}})_{\Delta_i}\}$, one way is via **barcodes**

(Ghrist, 2008)
We use a **scatter plot representation**

This example is a **nominal arrival** within ORD data set with no in-air hold or missed approach/go around.
Recall the nominal case ...

Only $\Delta_0$ features, no non-trivial $\Delta_1$ loop features. Good!
Note the appearance of $\Delta_1$ loop features corresponding to cyclic in-air holding pattern at SW gate.
Holding prior to final approach (Example 1)

Note the appearance of $\Delta_1$ loop features corresponding to cyclic in-air holding pattern prior to ILS establishment.
Note the appearance of $\Delta_1$ loop features corresponding to cyclic in-air holding pattern prior to ILS establishment.
Conclusion

• Leveraged TDA/PH to detect in-air holding trajectories

• Preliminary ORD-specific observations:
  • In-air holding relatively weather invariant
  • Missed approach/go-around more common during bad weather
Adding variables in

We have only looked at the simplest, planar trajectory case

Think about including:

- Altitude
- Speed and climb/descent rates
- Heading

**Pro:** Deeper spatial relations that are impossible to determine visually

**Con:** Difficult to interpret + relate back to applications
Possible aviation applications (1/3)

- **Fuel flow rate** analysis, particularly using raw on-board FDR data

(Chati & Balakrishnan, 2013; 2017)
Future Applications of TDA/PH in Aviation

Possible aviation applications (2/3)

- Characterizing topology of holding stacks

LHR Easterlies, LHR Westerlies
Possible aviation applications (3/3)

• Linking trajectory generation via homotopy theory (Vidosavljevic et al., 2017) to TDA/PH using various lifting theorems and invariances

• TDA/PH in characterizing sensor networks and sensor coverage (Curry et al., 2012), (Dlotko et al., 2012)
  • ADS-B/MLAT
  • Internet-of-Things and UTM
Huang & Ribeiro (2017) examined weighted hypergraph of co-authorship networks within engineering journals (blue diamonds) and mathematics journals (red circles).

Analysis of these high-order co-authorship networks via TDA/PH yielded clear distinctions in co-authorship patterns between the two academic communities.

High-order networks in aviation?
Bhattacharya *et al.* (2015) tackles the problem of trajectory planning in environment with uncertain obstacles by examining the **most homologically persistent** trajectory classes that remain feasible across the largest range of obstacle probabilities.
Questions? Comments & suggestions?

Mug or donut?!
Homology groups helps in identification of **higher dimensional holes** by mapping \( n \)-dimensional circles into target space \( X \) that **cannot** be further extended to mappings from \( (n + 1) \)-dimensional objects into space \( X \).

Homology works in **all dimensions** and provides insight into **structure** of the space.
Homology group examples (adapted from slides by H. Gluck)

**Circle:** \( H_0(S^1) = \mathbb{Z}, \ H_1(S^1) = \mathbb{Z}, \ H_k(S^1) = 0, \forall k \geq 2 \)

**2-sphere:**
\[ H_0(S^2) = \mathbb{Z}, \ H_1(S^2) = 0, \ H_2(S^2) = \mathbb{Z}, \ H_k(S^2) = 0, \forall k \geq 3 \]

**Torus:**
\[ H_0(T^1) = \mathbb{Z}, \ H_1(T^1) = \mathbb{Z} \oplus \mathbb{Z}, \ H_2(T^1) = \mathbb{Z}, \ H_k(T^1) = 0, \forall k \geq 3 \]

**Klein bottle:** \( H_0(K) = \mathbb{Z}, \ H_1(K) = \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}, \ H_k(K) = 0, \forall k \geq 2 \)

**n-sphere:** \( H_0(S^n) = \mathbb{Z}, \ H_n(S^n) = \mathbb{Z}, \ H_k(S^n) = 0, \forall k \neq 0, n \)
Let $f, \tilde{f} : X \to Y$ be homotopic maps, i.e. there exists a map $F : X \times [0, 1] \to Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = \tilde{f}(x)$. Then on the level of homology the following homomorphisms $f_* = \tilde{f}_*$ map between the homology groups for $X$ and $Y$:

$$f_* = \tilde{f}_* : H_*(X) \to H_*(Y)$$ (3)