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

Max Z. Li
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, Pennsylvania, USA
lizhaoyu@seas.upenn.edu

Megan S. Ryerson, Ph.D.
Department of City and Regional Planning
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, Pennsylvania, USA
mryerson@design.upenn.edu

***Abstract*— In-air holding and missed approach/go-around procedures are highly undesirable within the terminal airspace. Such operations are fuel inefficient and add to controller workload, reducing airspace efficiency and predictability. However, the detection of isolated anomalous trajectories within a large aviation data set is difficult through classical methods. We propose using tools from topological data analysis to identify such anomalous trajectories. Our methodology – *cyclic trajectory detection via persistent homology* – identifies individual trajectories of aircraft that executed in-air holding and missed approach/go-arounds through the qualitative shape of the data. We preliminarily deployed this method on a data set of arrivals to the Chicago O'Hare International Airport and found that in-air holding incidences were relatively weather-invariant but missed approach/go-arounds were more common in inclement weather, particularly during runway configuration changes. We introduce topological detection of cyclic trajectories to address the difficulties in identifying, characterizing, and ultimately mitigating anomalous airspace events.**

Keywords: topological data analysis; persistent homology; anomalous trajectory identification; in-air holding; missed approach/go-around

I. Introduction and Motivation

In order to achieve operational efficiency and reduce delay and congestion, airport operators and controllers seek to increase *predictability* within the terminal airspace. Classical statistical methods and modeling tend to focus on general trends, even though characterizing the outliers may be important as well. When looking at trajectory data, outliers representing anomalous trajectories, particularly in-air holding and missed approach/go-around trajectories, are important to identify and characterize. These anomalous trajectories are fuel-inefficient and airspace resource intensive. Individual in-air holding and missed approach-go around cases are difficult to detect using classical clustering methods, particularly if their occurrence is uncommon compared to nominal operations. Topological data analysis and persistent homology offer a way to detect these individual anomalous trajectories, leveraging the trajectory's cyclical nature. A better detection and characterization of in-air holding/missed approach-go around trajectories will lead to a better understanding of what happened in each specific case, and how such anomalies could be prevented in the future. This improvement in predictability allows for more certainty in fuel uptake [1], potentially leading to economic and environmental benefits [2].

II. Literature Review

Topological data analysis and persistent homology use results from algebraic topology to extract qualitative global information such as connectivity, holes, and voids from high dimensional point clouds. Furthermore, the information teased out via topological methods are robust, since the overall shape of the data set should remain invariant under noise [3]. These are all desirable features when considering methods to examine and analyze trajectory data sets in aviation.

Topological data analysis and persistent homology have already found a diverse range of applications in engineering, particularly in problems involving large, high-dimensional, data sets with nontrivial degrees of connectivity and loops. These topological methods have opened new avenues in networked neuroscience [4], as well as in path-planning for autonomous robots [5, 6] and modeling coverage and exploration behaviors for teams of multiple robots [7]. Topology has also been used to identify persistent features within fluid flow [8], extract and track defects occurring within crystal structures [9], identify pertinent structures within LiDaR data sets [10], and characterize the complex connections and networks within the airways of the human lung [11]. Our preliminary work in this paper embraces the fact that aviation, like many other fields, is entering the era of big data. Such an abundance of high dimensional trajectory data lends itself well to topologically-qualitative methods from the emerging toolbox of topological data analysis. Our use of persistent homology to detect anomalous trajectories is a small survey of the possible applications of topological methods in aviation.

III. Overview of Methodology

In-air holding typically occurs during a period of time when the airport capacity drops dramatically and cannot accommodate any more aircraft entering from the en route and transition airspace. Compared to gate delays, airborne delays much more fuel inefficient. Furthermore, a new dimension of complexity is added to the terminal airspace due to the fact that controllers must now deal with re-injecting holding aircraft into the arrival stream. Although many Traffic Management Initiatives are enforced precisely to reduce the probability of in-air holding, they are still a common occurrence within busy terminal airspaces. Due to their relative rarity and significance in terms of airspace operations, it is important to identify individual in-air holding trajectories. Our approach leverages the topological differences between a normal trajectory and a trajectory with holding patterns using tools from persistent homology to identify significant *holes* in a data set.

We provide a brief overview of the fundamental ideas behind this identification methodology via the toy example in Figure 1. The toy example depicts a trajectory from an aircraft that executed an in-air holding pattern. As a result, there is a notable cycle within the trajectory where the aircraft loops around. Since the same type of trajectory features would appear in a flight that experienced a missed approach/go-around, this methodology can also be used in those cases.

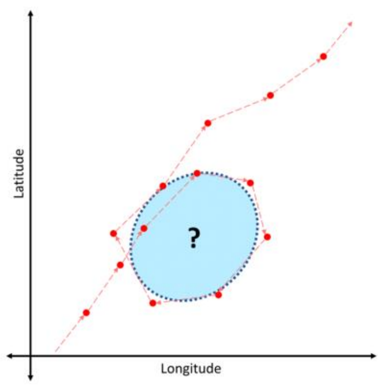

Figure 1. Toy example of anomalous trajectory with in-air holding

Visual inspection of the toy example in Figure 1 would indicate that there is a hole in the data set created by the cyclical trajectory. Informally speaking, a hole is considered to be a non-trivial object in topology, as it prevents *deformations* from one object to another. *Homology* is a way to represent topological spaces using concepts from algebra such as group, module, and ring theory. In this way, many topological features can be formulated algebraically and made more computationally tractable. It is through ideas and tools in homology that various computational software and toolboxes available to perform topological data analysis were created.

Using persistent homology, the visually-obvious hole in the toy example can be identified rigorously. Choose some radius r, and construct the r-neighborhood $\mathcal{D}(x,r)$ centered at a latitude-longitude data point $x \in \mathbb{R}^2$ with radius r. If we were to connect all pairwise points that are no more than r apart, the resulting structure – known as a *simplicial complex* – can be broken up into simpler geometric components. Loops consisting of only 1-simplices is the informal definition of a topological hole, and these are what we seek to detect. By varying r and observing the interactions of the r-neighborhoods, we notice that there is some r_i where the significant hole is formed, and another r_f where $r_f > r_i$ when the hole is covered again by the larger r-neighborhoods. We illustrate this with our toy example in Figure 2. We denote the radius r_i as the *birth* time of the significant hole, and r_f to be the *death* time. Insignificant holes that are due to sampling error or noise within the data will have a very short lifespan, since they will appear then are rapidly covered up again. Only significant holes, such as the one created by the holding pattern, will have a long, non-trivial lifespan. As the radius r varies from small to large, *persistence diagrams* can be used to represent the birth and death of topological features.

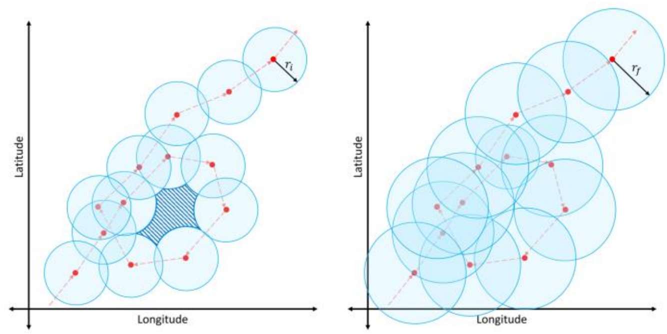

Figure 2. Simplicial complex at birth (*left*) and death (*right*) radii

To compute the complexes and form the persistence diagrams for individual aircraft trajectories in order to detect airborne holding patterns and missed approach/go-around instances, we utilized the TDA package in R [12] along with guidance from an introduction to the software package written by the authors of the package [13]. The algorithm computes *Alpha complex persistence diagrams*, which are analogous to the simplicial complexes described above except with some geometric differences in how the r-neighborhoods are formed. In lieu of simplicial complexes, Alpha complexes is a simplicial subcomplex of a *Delaunay complex*.

The formal construction of these Alpha complexes given by [14] begins by providing a set of data points $X \subset \mathbb{R}^d$. In our case, the set of data points $X = (\Phi, \Lambda)$ comes from the trajectory data set and are composed of latitude-longitude coordinates $(\varphi, \lambda) \in X \subset \mathbb{R}^2$. For every point $(\varphi, \lambda) \in X$, a Voronoi partition can be created consisting of all points in $\mathbb{R}^2$ that is closer to (φ, λ) than any other point in X. Denote this Voronoi partition by $\mathcal{V}((\varphi, \lambda))$, and construct the usual r-neighborhood $\mathcal{B}((\varphi, \lambda), r)$ centered at the point (φ, λ) with circumradius r. Whereas the simplicial complex consisted only of the r-neighborhoods, an Alpha complex consists of individual complexes defined by the intersection of a r-neighborhood of a point with the Voronoi partition associated with that point. Denote this individual complex as $\mathcal{R}((\varphi,\lambda),r)$, and it is explicitly defined by (1):

$$\mathcal{R}((\varphi,\lambda),r) = \mathcal{B}((\varphi,\lambda),r) \cap \mathcal{V}((\varphi,\lambda)) \tag{1}$$

The resultant Alpha complex of circumradius r for this trajectory data subset, denoted as $\mathbb{A}(X,r)$ is given by the collection of data points $\sigma \subseteq X$ satisfying the relation in (2):

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

In Figure 3, we compute the Alpha complex $\mathbb{A}(*,r)$ for a flight that did not experience any in-air holding or execute a missed approach/go-around. This flight proceeded smoothly from the arrival fix to the runway, as shown by the trajectory plot on the right-hand side of Figure 3. The Alpha complex did not detect any significant cycles, and no holes were born.

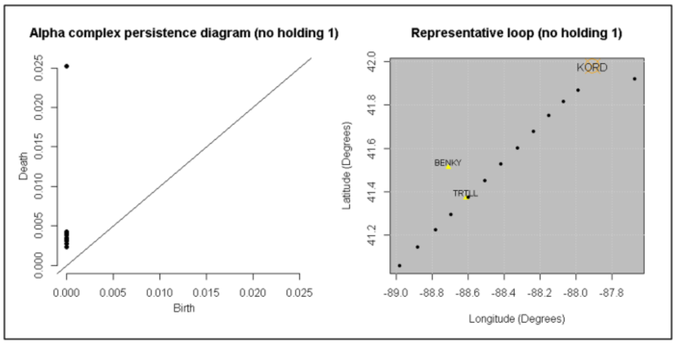

Figure 3. Alpha complex persistence diagram (*left*) for a normal arrival trajectory (*right*); black points are connected components, red triangles are loops detected within trajectory data subset

This stands in contrast to trajectories akin to the one presented in Figure 4 where the Alpha complex was computed for three anomalous trajectories. These trajectories all feature in-air holding, shown in the trajectory plots as loops in the data. Due to the existence of cycles in these trajectories, the Alpha complex was able to pick up the birth and death of several holes that appear. The most significant hole – the hole that lived the longest through larger and larger values of the circumradius r – is given by the highest red triangle in the Alpha complex persistence diagram. This long-lived topological feature corresponds to the red loop overlaid onto the trajectory plot and identifies these aircraft as having encountered in-air holding and/or executed a missed approach/go-around.

IV. Discussion of Early Results

We observed that the terminal arrival airspace for the Chicago O'Hare International Airport (ORD) tended to have in-air holding cases at an equal rate in both good and bad weather days. Identified missed approach/go-arounds were more common in the trajectory data subsets containing bad weather time periods. If crosswind components exceed allowable ranges, ORD will abandon its preferred east-west flow configuration and switch to the diagonal 22L/22R runways instead. Immediately before configuration changes is when missed approaches and go-arounds tend to happen the most. During this time period, it is obvious that conditions call for a configuration change, and any aircraft on final approach following Instrument Approach Procedures (IAPs) designated for the expired configuration are more likely to execute a missed approach and go-around to re-join the IAP for the new configuration instead.

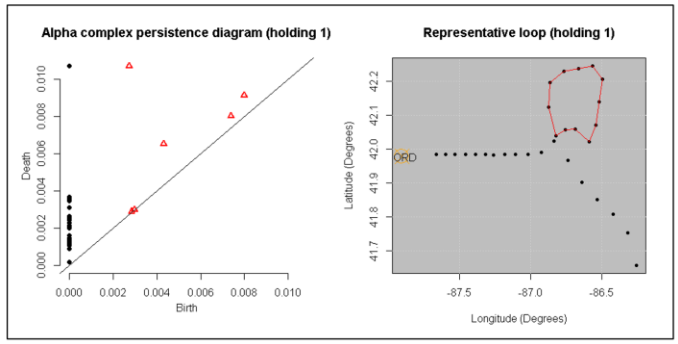

Figure 4. Alpha complex persistence diagram (*left*) for an anomalous arrival trajectory (*right*); see Fig. 3 for persistence diagram interpretation

V. Future Directions for Topological Data Analysis in Aviation

Another active research area in aviation that may benefit from an exploration using topological data analysis is taxi fuel flow rate analysis, particularly using raw data from onboard flight data recorders [15, 16]. An aircraft taxiing to and from the gate at a busy airport may encounter runway and ramp queues, causing a repeating pattern of stop-and-go thrust inputs as the aircraft taxis behind another aircraft within the queue itself [17]. This pattern of thrust inputs and corresponding fuel flow rate can be modeled periodically in time, and this periodicity can be mapped to cycles detectable through topological means.

Other areas in aviation that present opportunities for applications of ideas from algebraic topology and topological data analysis include origin-destination trajectory generation and sensor coverage. Homotopy theory can become very complex and rich when considering various exotic topologies, but for the purposes of aviation, the workspace is typically just the surface of the Earth, with various discontinuities that represent airspace features which will affect the resultant homotopies. [18] presents one possible application of homotopy theory in trajectory generation. Finally, [19] and [20] demonstrate the use of topological data analysis in sensor networks and sensor coverage. Applications in aviation research include characterizing the expansion of Automatic Dependent Surveillance-Broadcast (ADS-B) and Multilateration (MLAT) capabilities, particularly in regions of the world with poor primary radar coverage.

VI. Conclusion

We have demonstrated how persistent homology can be used to detect individual trajectories with cyclical patterns indicative of in-air holding and missed approach/go-around operations. With regards to our case study airport, we were able to successfully leverage tools from persistent homology to detect in-air holding trajectories, identifying flights that held at the entrance of the terminal airspace (around the arrival fixes) as well as the beginning of the final approach corridor. For the ORD terminal arrival airspace, we saw that occurrences of in-air holding tended to be relatively invariant under weather conditions but missed approach/go-around trajectories were more common in bad weather, particularly during runway configuration switches. We plan to continue refining the identification methodology and expand our scope to include other terminal arrival airspaces. This will allow us to observe how well our method works at detecting cyclic trajectories in other settings. We will also continue to explore other areas of aviation that may benefit from learning the underlying topology of the data using topological data analysis and persistent homology.

Acknowledgment

The student author would like to thank Dr. Gregory Henselman for encouraging the exploration of aviation data using topological methods, and Dr. Herman Gluck for his insightful explanations regarding topics in algebraic topology.

References

[1] Hao, L., Hansen, M., & Ryerson, M. S. (2016). Fueling for contingencies: The hidden cost of unpredictability in the air transportation system. *Transportation Research Part D*, *44*, 199-210. http://dx.doi.org/10.1016/j.trd.2016.02.016

[2] Ryerson, M. S., Hansen, M., Hao, L., & Seelhorst, M. (2015). Landing on empty: estimating the benefits from reducing fuel uplift in US Civil Aviation. *Environmental Research Letters*, *10*, 1-11. https://doi.org/10.1088/1748-9326/10/9/094002

[3] Ghrist, R. (2014). *Createspace: Elementary Applied Topology*. Retrieved from https://www.math.upenn.edu/~ghrist/notes.htmlJ. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] Giusti, C., Ghrist, R., & Bassett, D. S. (2016). Two's company, three (or more) is a simplex Algebraic-topological tools for understanding higher-order structure in neural data. *Journal of Computational Neuroscience*, *41*, 1-14. https://doi.org/10.1007/s10827-016-0608-6

[5] Bhattacharya, S., Lipsky, D., Ghrist, R., & Kumar, V. (2013). Invariants for homology classes with application to optimal search and planning problem in robotics. *Annals of Mathematics and Artificial Intelligence*, *67*, 251-281. https://doi.org/10.1007/s10472-013-9357-7

[6] Bhattacharya, S., Ghrist, R., & Kumar, V. (2015). Persistent Homology for Path Planning in Uncertain Environments. *IEEE TRANSACTIONS ON ROBOTICS*, *31*(3), 1-13. https://doi.org/10.1109/TRO.2015.2412051

[7] Bhattacharya, S., Ghrist, R., & Kumar, V. (2014). Multi-robot coverage and exploration on Riemannian manifolds with boundaries. *The International Journal of Robotics Research*, *33*(1), 113-137. https://doi.org/10.1177/0278364913507324

[8] Kasten J., Hotz I., Noack B.R., Hege HC. (2011) On the Extraction of Long-living Features in Unsteady Fluid Flows. In: Pascucci V., Tricoche X., Hagen H., Tierny J. (eds) Topological Methods in Data Analysis and Visualization. Mathematics and Visualization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15014-2_10

[9] Grottel S., Dietrich C.A., Comba J.L.D., Ertl T. (2011) Topological Extraction and Tracking of Defects in Crystal Structures. In: Pascucci V., Tricoche X., Hagen H., Tierny J. (eds) Topological Methods in Data Analysis and Visualization. Mathematics and Visualization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15014-2_14

[10] Keller P. et al. (2011) Extracting and Visualizing Structural Features in Environmental Point Cloud LiDaR Data Sets. In: Pascucci V., Tricoche X., Hagen H., Tierny J. (eds) Topological Methods in Data Analysis and Visualization. Mathematics and Visualization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15014-2_15

[11] Szymczak A. (2011) A Categorical Approach to Contour, Split and Join Trees with Application to Airway Segmentation. In: Pascucci V., Tricoche X., Hagen H., Tierny J. (eds) Topological Methods in Data Analysis and Visualization. Mathematics and Visualization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15014-2_17

[12] Fasy, B. T., Lecci, F., Maria, C., & Rouvreau, V. (2017). TDA: Statistical Tools for Topological Data Analysis (1st ed., Version 1.6) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/TDA/index.html

[13] Fasy, B. T., Kim, J., Lecci, F., Maria, C., Rouvreau, V., & CMU TopStat Group. (2014). *Introduction to the R package TDA*. Retrieved from https://cran.r-project.org/web/packages/TDA/vignettes/article.pdf

[14] Edelsbrunner, H., & Harer, J. (2010). *Computational Topology: An Introduction*. Retrieved from http://www.ee.oulu.fi/research/imag/courses/Vaccarino/Edels_Book.pdf

[15] Chati, Y., & Balakrishnan, H. (2013). Aircraft Engine Performance Study Using Flight Data Recorder Archives. *Proceedings of the 2013 Aviation Technology, Integration, and Operations Conference*, pp. 1-12. https://doi.org/10.2514/6.2013-4414

[16] Chati, Y. S., & Balakrishnan, H. (2017). A Gaussian Process Regression Approach to Model Aircraft Engine Fuel Flow Rate. *Proceedings of The 8th ACM/IEEE International Conference on Cyber-Physical Systems*, pp. 1-10. http://dx.doi.org/10.1145/3055004.3055025

[17] Levine, B. S., & Gao, H. O. (2007). Aircraft Taxi-Out Emissions at Congested Hub Airports and Implications for Aviation Emissions Reduction in the United States. *Transportation Research Board*, 1-15. Retrieved from https://trid.trb.org/view/802435

[18] Vidosavljevic, A., Delahaye, D., & Tosic, V. (2017). Homotopy Route Generation Model for Robust Trajectory Planning. *Air Traffic Management and Systems II*, *420*, 69-88. https://doi.org/10.1007/978-4-431-56423-2_4

[19] Curry, J., Ghrist, R., & Robinson, M. (2012). Euler Calculus with Applications to Signals and Sensing. *Proceedings of Symposia in Applied Mathematics*, pp. 1-71. Retrieved from https://www.math.upenn.edu/~ghrist/preprints/eulertome.pdf

[20] Dlotko, P., Ghrist, R., Juda, M., & Mrozek, M. (2012). Distributed computation of coverage in sensor networks by homological methods. *Applicable Algebra in Engineering, Communication and Computing*, *23*, 29-58. https://doi.org/10.1007/s00200-012-0167-7