Identifying Temporally Persistent Flows in the Terminal Airspace via Spectral Clustering

Marco Enriquez
The MITRE Corporation
Center for Advanced Aviation System Development (CAASD)
McLean, Virginia, USA
menriquez@mitre.org

***Abstract*—Given a specified amount of flight trajectory data, data reduction and clustering methods (e.g., Principal Components Analysis and k-means) have become established tools for identifying flows (i.e., a group of similar flight trajectories). However, most flow identification algorithms in the literature rely solely on spatial clustering, without considering the temporal dimension. Temporal characterization of flows is important, as it: enables identification of salient air traffic features, provides a basis for scenario ("what-if") analyses, and allows for a more robust distillation of large and time-varying air traffic datasets.**

To address this shortcoming, this work proposes a methodology for identifying flows which persist over an arbitrary time span. This process leverages a generic Spectral Clustering framework, building upon the methodologies established by Enriquez and Kurcz in [4]. This algorithmic approach produces robust results, while remaining easy to implement and being computationally inexpensive. We present two examples to show the promise of this algorithm. First, the algorithm is used to automatically identify days in which irregular air traffic patterns occur in the Miami International Airport (MIA) terminal airspace. Second, we use this algorithm to help identify the minimum required number of new Performance Based Navigation (PBN) arrival and departure procedures in the National Airspace System (NAS), based on six months of historic data.

***Keywords*—Trajectory Clustering; Terminal Area Flow Identification; Spectral Clustering; Eigenvalue Decomposition; Graph Cut**

I. Introduction

Despite recent economic hardships, the National Airspace System (NAS) traffic growth is still projected to rise by more than 90 percent by 2032 – accommodating roughly 500 million more passengers [5]. In order to successfully plan and accommodate for the increased number of flights, we must understand dominant trajectory trends in the NAS. This paper focuses on the terminal area airspace, where identifying such patterns can reveal insights such as: how well procedures are being utilized, how aircraft have historically handled adverse weather conditions during arrivals, etc.

Advanced mathematical analyses can be leveraged to help with such discoveries, and has been discussed in the literature in the context of flow identification. (For the duration of this paper, we define the term "flow" to be a collection of flights that have similar spatial trajectories.) Eckstein coupled Principal Component Analysis (PCA) and the k-means clustering algorithm to realize a flight taxonomy in [3]. Gariel et al. [7] also used PCA in their work, but first augmented the dimensionality of the data (by adding heading, angular position, etc.) and used the DBSCAN clustering algorithm. The DBSCAN algorithm holds notable advantages over the k-means algorithm, as it does not require a-priori selection of cluster size and features outlier identification. Marzouli et al. [8] also leveraged PCA and DBSCAN to identify flows, from which a mathematical graph (network) was created. More recently, Enriquez and Kurcz used spectral clustering to identify flows in the terminal and en-route airspace [4]. The hierarchical clustering algorithm in Enriquez and Kurcz' algorithm only required positional data (as opposed to *operational* data, such as distance from corner post or procedure used) to yield robust results. Further, similar to DBSCAN, the algorithm in [4] did not require a-priori selection of cluster size.

To the author's knowledge, only the spatial dimensions (i.e., patterns observed laterally and perhaps vertically) are considered in most terminal flow analysis applications; methodological methods for identifying temporal relationships between flows are not discussed. This is unfortunate, as there is an emerging interest within the aviation community to characterize persistence and uncertainty in operations. This paper seeks to provide insight to the former, the temporal characteristic of flows.

This paper is written as a continuation and extension of the work and methodologies established in [4], and also employs a spectral clustering framework to identify temporally persistent flows. Enriquez and Kurcz created a methodology to distill flows into *nominal* lines, which can be regarded as a descriptor of a flow. Furthermore, the nominal lines have been shown to coincide with existing arrival and departure procedures in the terminal area airspace. This paper establishes the following fact: since nominal lines describe a flow for a given time duration, they too can be clustered across the temporal dimension to identify the temporal persistence of the flow.

An advantageous by-product of this research is that it also makes the spectral clustering process more computationally tractable for large flight datasets in the terminal area domain. It is well known that spectral clustering generally incurs an $O(n^3)$ computational cost[1] due to the eigenvalue decomposition, where n is the number of data elements being

[1]Sparsification techniques (e.g., using the $k-$nearest neighbor or an $\epsilon-$neighbor approach, as mentioned in [12]), sampling-based methods (e.g., the Nyström sampling method [1]) and iterative methods for computing eigenvalues (e.g., the power method, the Arnoldi method [11], [9]) may further reduce this cost, but may also alter the quality of the obtained clusters.

compared. This can be prohibitively expensive when we are considering one or multiple months' worth of airspace traffic. The methodology shown here will allow use of spectral clustering for large, time-varying air traffic datasets at a significantly reduced computational cost, since individual flight comparisons are being replaced with flow comparisons (i.e., a group of flights). This approach can be considered related to the following approaches in large-scale spectral clustering: the KASP algorithm [13] and the "Landmark-Based Spectral Clustering" (LBSC) algorithm [2].

This paper is organized as follows. Section II-A reviews the mathematical and algorithmic framework established in [4]. Section II-B then examines how to extend the clustering methodology in [4] to identify temporal trends in the trajectory data. We then use this temporal (or "4D") trajectory clustering algorithm on two applications: automatically determining anomalous arrival flights at Miami International Airport (MIA), and identifying the minimum number of PBN procedures required to sufficiently support airspace traffic at the top airports in the NAS. A detailed description of these applications, and how the 4D clustering algorithm aided such analyses, is discussed in Section III. The final section highlights future research directions and concludes.

II. Spectral Clustering Methodology

This section outlines the methodology to identify temporally persistent flows, by leveraging the generic spectral clustering framework in [4]. Hence, this section begins by reviewing the mathematical background and algorithmic approach behind spectral clustering for identifying terminal area flows. (Please consult [4] for further discussion of the mathematics and details behind the flow detection algorithm.) We then extend the above methodology to the temporal dimension, which will be discussed in the latter half of this section.

A. Trajectory Clustering Algorithm

The trajectory clustering algorithm begins by parametrizing track positions of terminal area flights as a function of time. This is done by linearly scaling all the aircrafts' position report times to the interval $[0, 1]$. Each aircraft's lateral position is then interpolated at m uniformly spaced locations in $[0, 1]$, which enables direct position comparisons between flights. Specifically, each flight trajectory can be written as a vector with m components:

$$\mathbf{f}_i = [(x_1, y_1)^{(i)}, (x_2, y_2)^{(i)}, ..., (x_m, y_m)^{(i)}],$$

where $(x_k, y_k)^{(i)}$ corresponds to the lateral position of aircraft i at parametrized time k. Our task is to partition a collection of n flight data elements $\{\mathbf{f}_i\}_{i=1}^n$ into similar groups, hence defining flows.

We then turn to a graph partitioning approach to accomplish the grouping of flights. We first define a network by constructing a similarity matrix, $W \in \mathbb{R}^{n\times n}$ whose entries are computed via the Gaussian kernel as $W_{i,j} = e^{\frac{-\|\mathbf{f}_i - \mathbf{f}_j\|_2^2}{2\sigma^2}}$ for some local scale parameter σ. The local scale parameter σ is important, as it dictates the "width" of each cluster. For a discussion of the effects of the scale parameter on clustering quality, as well as a strategy for choosing this parameter adaptively, see [14]. For the work considered here, it sufficed to set $\sigma = 1$. Our goal is to partition this network (represented by W) into groups such that the similarity between groups is minimized and similarity within a group is maximized.

Algorithm 1 Spectral Clustering Algorithm

1: def **spectralCluster**(W, ω_{min})
2: $\quad D = \sum_{j=1}^n W(i,j)$
3: $\quad L = D - W$
4: $\quad v$ = second smallest eigenvector of L
5: $\quad i_l = \{i | v_i \geq 0\}$ (indices of v with positive elements)
6: $\quad i_r = \{i | v_i < 0\}$
7: $\quad$ **if** stop(W_{i_l,i_l}) > ω_{min}
8: $\quad\quad$ **spectralCluster**(W_{i_l,i_l}, ω_{min})
9: $\quad$ **else**
10: $\quad\quad$ save i_l
11: $\quad$ **if** stop(W_{i_r,i_r}) > ω_{min}
12: $\quad\quad$ **spectralCluster**(W_{i_r,i_r}, ω_{min})
13: $\quad$ **else**
14: $\quad\quad$ save i_r

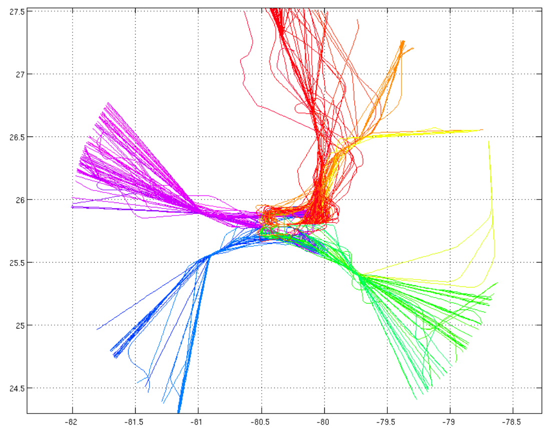
It has been proved that the second smallest eigenvalue of the graph Laplacian $L = D - W$, where D is a diagonal matrix given by $D_{ii} = \sum_{j=1}^n W_{ij}$, illuminates the semi-optimal data partition [10], [6]. Specifically for this work, we use the sign of the second eigenvalue's elements to determine the partition (i.e. $\{i|v_i \geq 0\}$ corresponds to one group and $\{i|v_i < 0\}$ corresponds to the other). This procedure can be applied recursively until a stopping criteria is met, as written in algorithm 1.

We note that in algorithm 1, the notation W_{i_l,i_l} (line 7) denotes the submatrix of W formed by taking the rows and columns of W corresponding to the index set i_l. Also, the function `stop` implements user-defined stopping criteria. For example, it could measure intra-cluster variance or as considered in the work here, the ratio of maximum distances:

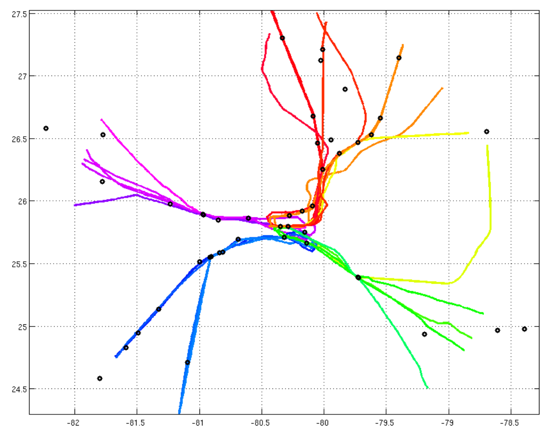
$$\frac{\max(W_{i.,i.})}{\max(W)},$$

which converges to zero as the hierarchical partitioning progresses. If the ratio above is less than the specified tolerance (lines 7 and 11), we save the index set and stop clustering that data grouping (lines 10). Otherwise, we make a recursive call to the clustering function (lines 8 and 12).

Figure 1 displays the output from algorithm 1, using one day of arrival traffic in MIA on March 1, 2011 as inputs. Figure 1a displays the flight clusters by color. Since each flight is parametrized, we can take the point-wise median of each flow, resulting in the nominal line. Figure 1b displays the nominal trajectories associated with each flow. In Figure 1b, circular markers denote fixes associated with Standard Terminal Arrivals (STARs). Note how the nominal lines align with the fixes.



(a) Flights colored by flow.



(b) Nominal trajectories associated with each flow.

Figure 1: Flows for arrival flights at MIA on March 1, 2011.

B. Identifying Temporally Persistent Flows via Spectral Clustering

The output algorithm 1 produces flights grouped by spatial trends, which we referred to as a "flow", during a set time period (e.g., a day). For most analyses involving the terminal airspace, examining one day is not enough – we must consider longer timespans. This, however, presents a dilemma: examining flows for a short period of time is computationally inexpensive, but insufficient for analysis. Increasing the timespan significantly, however, is prohibitive since the eigenvalue decomposition needed by spectral clustering generally incurs an $O(n^3)$ computational cost. Aside from computational feasibility concerns, we also wish to understand the temporal characteristics of the flows identified by algorithm 1. How often does a particular flow appear? How does it vary with time? Which flows can be classified as irregular?

We can address the above concerns and questions with the following insight: we can cluster the nominal lines (obtained from daily flows) across the temporal dimension. This idea is depicted in Figure 2. Since the nominal line is a reduced-dimension descriptor of a flow, this methodology will simultaneously reduce spectral clustering's computational cost while allowing us to gain insights about flows' temporal behavior. Specifically, if m denotes the number of nominal

Algorithm 2 4D Clustering Algorithm

foreach period **in** timespan
  pre-process track data
  form similarity matrix $W^{(period)}$
  $flows^{(period)}$ = **spectralCluster**($W^{(period)}$, ω_{min})
  foreach flow in $flows^{(period)}$
    $nominals_{(flow)}$ = point-wise median of flow
  end foreach
  $nominals^{(period)}$ = **union**($\{nominals_{(flow)}\}$)
end foreach
$nominals^{(all)}$ = **union**($\{nominals^{(period)}\}$)
create similiarity matrix $\bar{W}$ from $nominals^{(all)}$
$4Dflows$ = **spectralCluster**($\bar{W}$, ω_{4D})
post-process $4Dflows$

lines identified from the dataset – hence implying $m \ll n$ – the computational cost of this approach becomes $O(m^3)$. A further advantage of this approach is that it enables the re-use of the generic spectral clustering framework established in [4]. The pseudo-code implementing this idea can be seen in algorithm 2. In algorithm 2, the dataset's full timespan is equally partitioned into "periods" such as one day, one week, etc. Note that we use the `spectralCluster` function defined in algorithm 1, with two different tolerances ω_{min} and ω_{4D}. The choice of these parameters is application dependent. We advise to choose ω_{min} in such a way that the spatial flows produced capture sufficient resolution as dictated by the application. The parameter ω_{4D} should be chosen such that $\omega_{4D} \leq w_{min}$. The final step of algorithm 2 is intentionally vague, as there are many ways to post-process the "4D" flows. For example, one could examine 4D flows which contain a significant amount of flights and which also persist over a long time, as this would give the dominant terminal area flows. Consequently, we could also examine which of the 4D flows that do not regularly appear, implying possible operational changes. We note that this is a purely data-driven approach, as we do not use operational knowledge to obtain the daily or the 4D flows – only trajectory information. This is a further advantage of this approach as using operational knowledge may not always be feasible, or appropriate, for analysis.

To the author's knowledge, the closest related approaches in the large scale spectral clustering literature are the LBSC algorithm [2] and the KASP algorithm [13]. Chen and Cai [2] proposed the LBSC algorithm, which leverages fixed "landmark points" and Sparse Coding theory to reduce the computational complexity of the spectral clustering algorithm. In the context of the work presented here, "landmark points" are analogous to nominal lines. It is not clear how Chen and Cai's approach of approximating the similarity matrix via Sparse Coding theory would affect the quality of the identified flows, though they cite success with various datasets such as the MNIST dataset. The KASP algorithm [13] uses the $k-$means algorithm to find clusters in the dataset, and then runs a spectral clustering algorithm on the $k-$means clusters' centroids. Though the authors cite good results from KASP, we note that leveraging the $k-$means algorithm comes

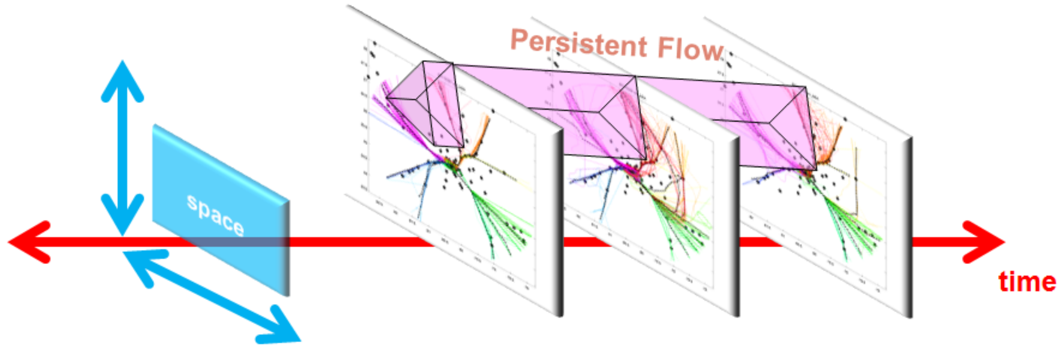


Figure 2: Picture depicting the idea behind the 4D clustering algorithm. By first clustering in the spatial domain, and then in the temporal domain, we are able to identify the persistence of a flow. We also reduce the computational effort required to perform spectral clustering on such datasets.

with certain drawbacks: it requires a-priori knowledge of the number of clusters expected out of the algorithm, and could suffer from reduced robustness as the $k-$means algorithm may not always converge to a favorable partitioning of the dataset. An advantage of the LBSC and KASP algorithms, however, is that they are general purpose algorithms that do not assume temporal structure in the dataset. Our approach, in contrast, exploits the knowledge of the trajectory datasets' temporal dimension in order to reduce the cost of spectral clustering.

III. Applications

We now present two applications that leverage the 4D trajectory clustering algorithm. First, we identify anomalous arrival flights at MIA for March 2011. Second, we use the 4D trajectory clustering algorithm to quantify the minimum number of RNAV procedures required to support airspace traffic at the busiest U.S. airports. Results from the second effort was included in the FAA response to Section 213 of the "FAA Modernization and Reform Act of 2012" (H.R. 658). We note that the trajectory data used for all the results in this section comes from Threaded Track, a MITRE data source of synthetic flight trajectories, which themselves are an amalgam of National Offload Program (NOP), Airport Surface Detection Equipment System (ASDE-X) and Enhanced Traffic Management System (ETMS) data.

A. Identifying Irregular Terminal Airspace Traffic at MIA

We first use the 4D clustering algorithm to distinguish between "regular traffic" and "irregular traffic" in the terminal area. Such knowledge, in turn, can be used to establish baselines at a given airport, or can be used to automatically detect large operational changes in the airspace traffic. The results in this section will use Algorithm 2 to highlight the latter, hence implying the former. We use March 2011 arrival traffic at MIA. We use $\omega_{min} = 0.035$ and $\omega_{4D} = 0.01$ for the parameters in Algorithm 2, and we define "irregular flow" to mean that the 4D flow consists of only one nominal line (i.e., the flow was not similar to any other flow in the timespan considered). Figure 3 shows the arrival traffic in MIA for four specific days, each with varying amounts of identified irregular flights. Notice that Figure 3c shows an unusually high number of holds on March 4 and Figure 3d shows a high number of anomalous flights on March 28. The number of irregular flights identified using the 4D clustering algorithm, separated by date, can be found in Figure 4. Some research on historical weather for the Miami region on March 28, 2011 shows that Miami incurred severe thunderstorms that day, with hail up to one inch in diameter and gusts up to 70 miles per hour.

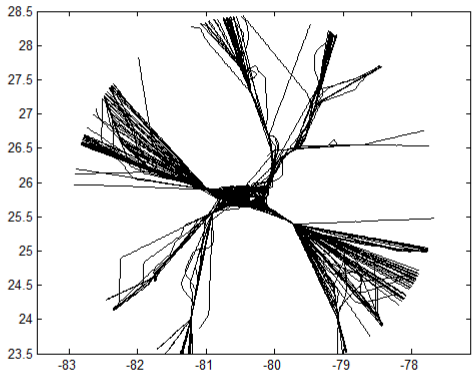
B. Identifying the Required Number of RNAV Procedures in the NAS

The MITRE Corporation recently conducted a "Top-Down" analysis in order to assess whether an existing or planned PBN procedure covers every arrival and departure flow seen at the busiest airports in the NAS currently. This analysis was included in the FAA's congressional response to Section 213 of the FAA re-authorization bill, known as "FAA Modernization and Reform Act of 2012" (H.R. 658). Section 213 of H.R. 658 specifies that the FAA shall provide plans and status updates of Area Navigation (RNAV) and Required Navigational Performance (RNP) procedure implementation at National Airspace System (NAS) airports looking forward three years.

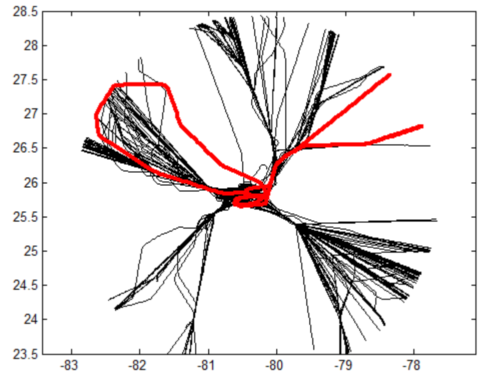
The analysis began by identifying airports with higher Area Navigation (RNAV) and Required Navigation Performance (RNP) equipage and traffic. This resulted in a list of 267 airports which had more than 13,000 operations each in CY 2011. Of those, 196 had greater than 70% RNAV-1[2] equipage and were considered candidate airports for new RNAV SIDs and STARs. The 196 airports we considered accommodated 10.8 million flights over the first six months in 2011.

We separated the above 10.8 million flights by each candidate airport, and then by departures and arrivals. The resulting tracks were used as the input to the 4D clustering algorithm 2, with parameters $\omega_{min} = 0.035$ and $\omega_{4D} = 0.025$. Since we are identifying the "procedure gap" at the candidate airports, we removed daily flows that corresponded to a published Standard Instrument Departure (SID) or STAR. This filtering

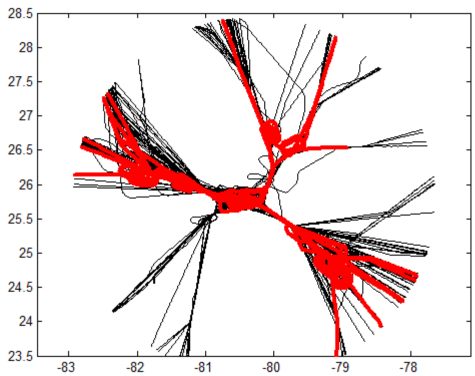
[2] Per the FAA Advisory Circular 90-101A, Section 2, Item j: "The RNP value designates the lateral performance requirement in NM increments associated with a procedure." Jeppesen expounds further: "Aircraft operating on RNAV-1 STARs and SIDs must maintain a total system error of not more than 1 nautical mile for 95% of the total flight time."



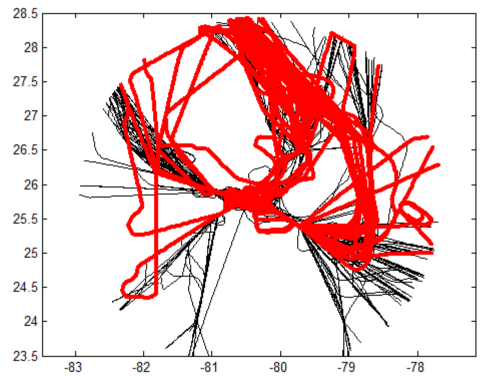
(a) March 3, 2011 Arrival Flights



(b) March 9, 2011 Arrival Flights



(c) March 4, 2011 Arrival Flights



(d) March 28, 2011 Arrival Flights

Figure 3: Four different days of arrival traffic at MIA. The lines in black denote flights deemed to be "normal" by the 4D clustering algorithm, while the red lines denote the identified anomalous flights. March 3 incurred no anomalous flights, and is shown here as a baseline. March 9, 4 and 28 incurred 4, 21 and 51 anomalous flights, respectively.

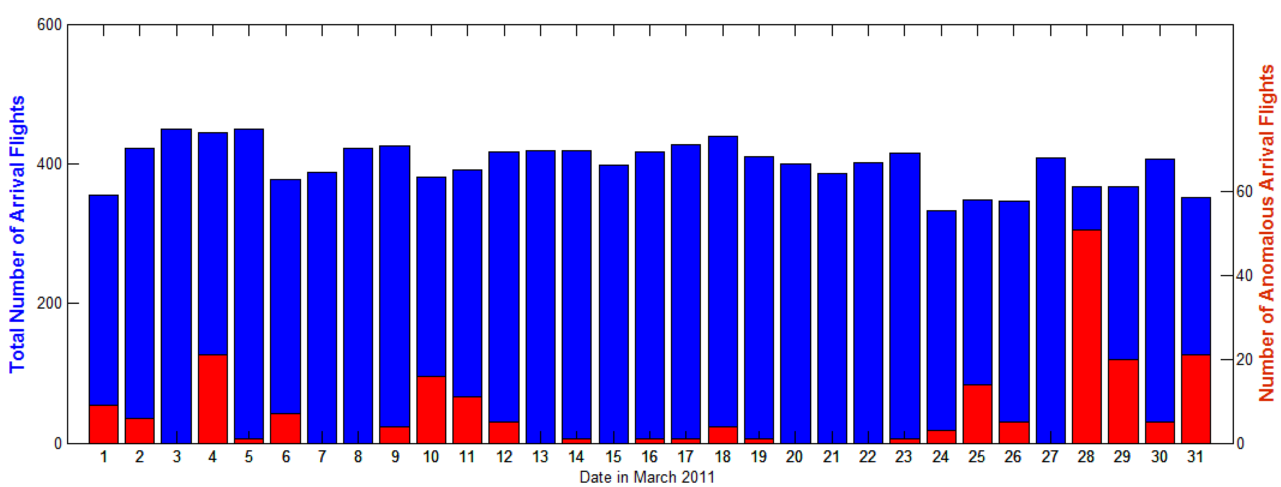


Figure 4: Plot of identified anomalous MIA arrival flights in March 2011 (red), and the total arrival flight count (blue). Please note that the red and blue bar plots have different y-axes.

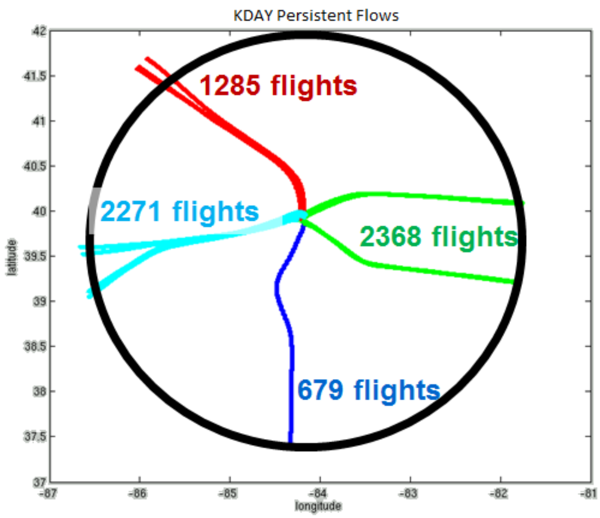


Figure 5: Persistent departure flows, grouped by their cardinal direction at Dayton International Airport (DAY). Each grouping of the persistent flows is colored uniquely. We comprehend this to mean that there is an opportunity to implement four new RNAV SIDs for DAY. We note that the flows seen above do not represent actual procedure designs, but can be used to inform such efforts.

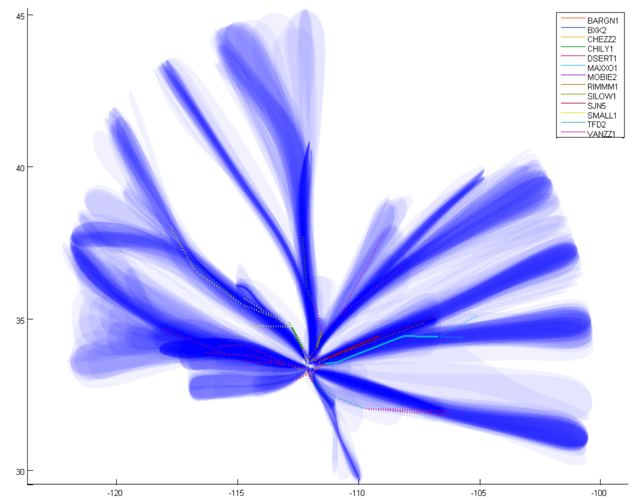


Figure 7: One month's worth of flow tubes for Phoenix Sky Harbor Airport (PHX) arrival traffic. Each flow tube was plotted as a transparent geometry and then overlayed on top of one another. Hence, darker regions indicate regions heavier traffic. We also plot PHX STARs here, for reference.

ATL - Hartsfield-Jackson Atlanta Intl	LGA - New York LaGuardia
BOS - Boston Logan Intl	MCO - Orlando Intl
BWI - Baltimore/Washington Intl	MDW - Chicago Midway
CLE - Cleveland Hopkins Intl	MEM - Memphis Intl
CLT - Charlotte Douglas Intl	MIA - Miami Intl
CVG - Cincinnati/Northern Kentucky Intl	MSP - Minneapolis/St. Paul Intl
DCA - Ronald Reagan Washington National	ORD - Chicago O'Hare Intl
DEN - Denver Intl	PDX - Portland Intl
DFW - Dallas/Fort Worth Intl	PHL - Philadelphia Intl
DTW - Detroit Metropolitan Wayne County	PHX - Phoenix Sky Harbor Intl
EWR - Newark Liberty Intl	PIT - Pittsburgh Intl
FLL - Fort Lauderdale/Hollywood Intl	SAN - San Diego Intl
HNL - Honolulu Intl	SEA - Seattle/Tacoma Intl
IAD - Washington Dulles Intl	SFO - San Francisco Intl
IAH - George Bush Houston Intercontinental	SLC - Salt Lake City Intl
JFK - New York John F. Kennedy Intl	STL - Lambert Saint Louis Intl
LAS - Las Vegas McCarran Intl	TPA - Tampa Intl
LAX - Los Angeles Intl	

Table I: A list of the airports considered for the Top-Down analysis.

step left approximately one million uncovered daily flows consisting of 5.2 million flights. In the post-processing step, we identified flows which appeared for at least 15 days as being "persistent". This resulted in the identification of 2601 persistent flows at the 196 airports that are not covered by an existing RNAV procedure. Finally, to compensate for the fact that multiple flows may be covered by the same procedure, we grouped the persistent flows by their cardinal direction. An example of this grouping logic can be seen in Figure 5.

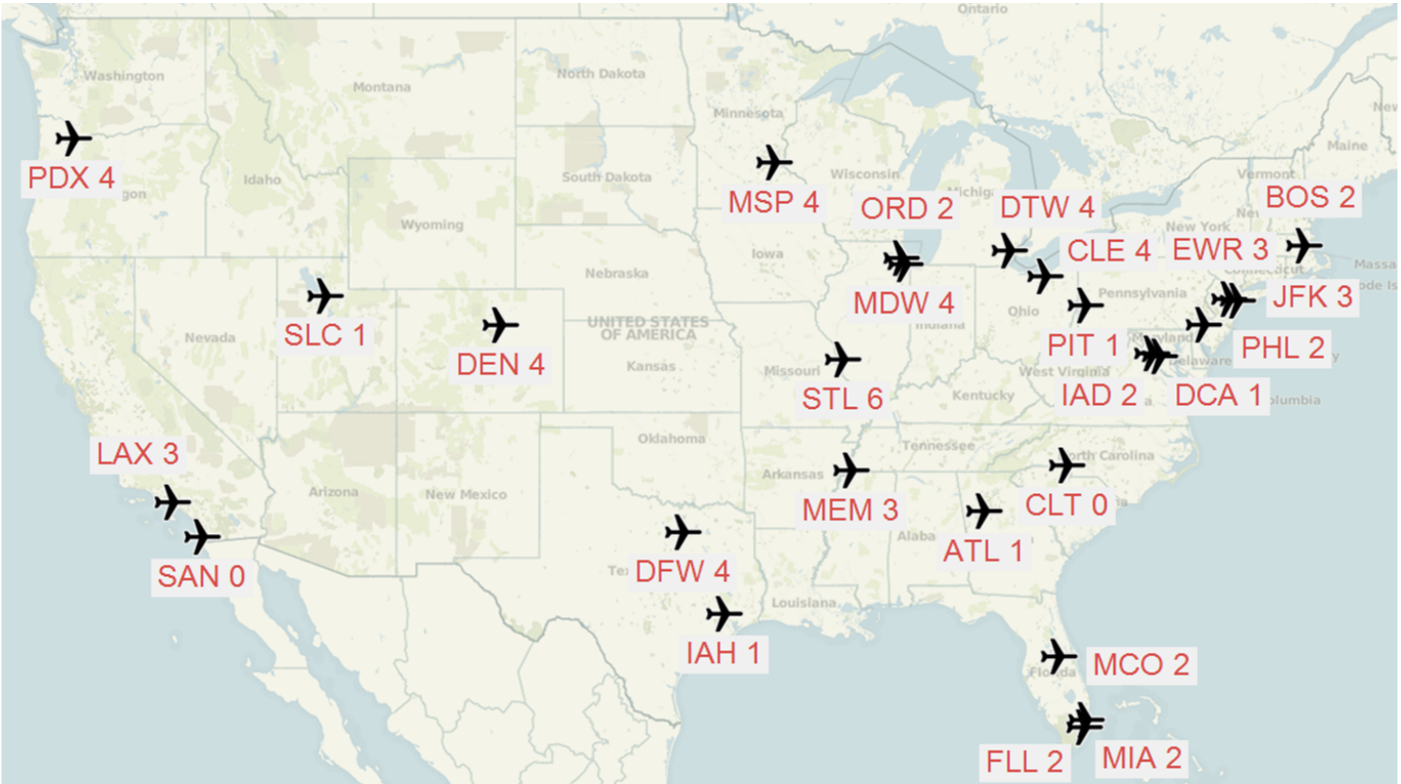
Results of the Top-Down analysis suggest a potential need for 87 new RNAV SIDs and 65 new RNAV STARs at the 35 busiest airports in the United States. These airports, formerly referred to as the OEP (Operational Evolution Partnership) 35, are listed in table I. The procedure gaps identified by the Top-Down analysis for each of the airports above can be seen in Figure 6. Compared to the June 2012 National Airspace and Procedures Team (NAPT) list of planned RNAV SID and STAR procedures, the Top-Down analysis shows an additional need for 47 SIDs and 38 STARs. For the remaining 161 airports considered in this analysis, 384 new RNAV SIDs and 330 new RNAV STARs are identified, beyond existing and planned procedures. We further note that local operational considerations such as Special Use Airspace (SUA), facility preferences, environmental restrictions, and other facility-specific constraints were not taken into account here, as the procedure suggestions which stem from such considerations may not be operationally feasible.

IV. Conclusion and Future Work

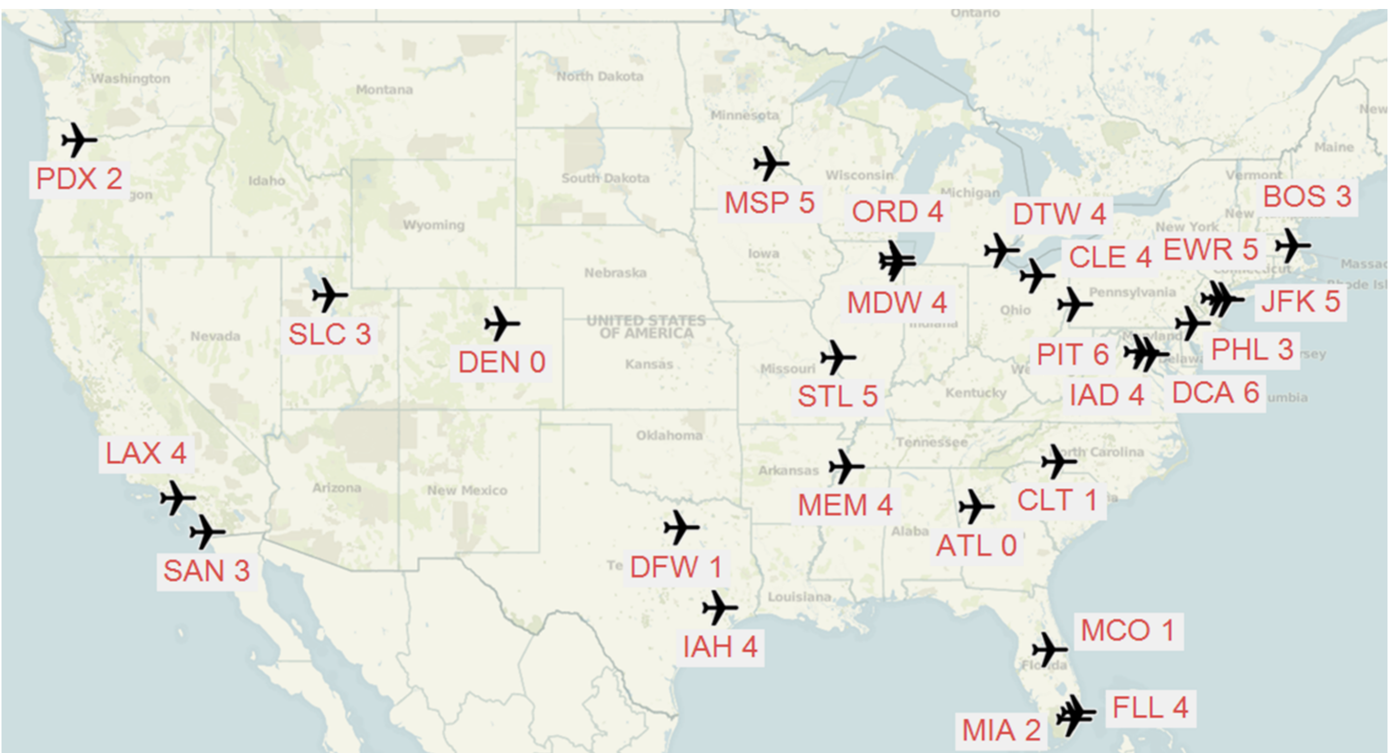
In this paper, we presented a purely data-driven methodology capable of identifying the temporal persistence of flows. Understanding this often overlooked feature is important as it allows us to characterize "normal" airspace traffic, automatically detect anomalous flights and also quantify how flows evolve over time. In turn, such insights will aid planners to accommodate the projected increase in NAS traffic.

The 4D trajectory clustering algorithm we presented here is a continuation of the spectral clustering flow algorithm established in [4]. The 4D trajectory clustering algorithm we present here has numerous advantages: it is simple to implement, it relies solely on spectral clustering (as opposed to the relying on the k-means algorithm), and it makes large-scale spectral clustering computationally tractable. We also applied the 4D trajectory clustering algorithm to two real-world examples, to show the promise of this methodology. We identified anomalous arrival flights at MIA, successfully illuminating the change in operations during a severe storm. We also identified the minimum number of PBN procedures that should be implemented at the U.S. airports which accommodate the highest number of RNAV and RNP-equipped aircraft.

Future directions for this research consists of including different characteristics of flows, aside from the nominal line, for 4D clustering. We can consider creating "flow tubes" from nominal lines, for example, by incorporating point-wise



(a) Procedure gap identified by the Top-Down analysis, for RNAV STARs.



(b) Procedure gap identified by the Top-Down analysis, for RNAV SIDs.

Figure 6: Procedure gap identified, per airport, by the “Top-Down” analysis. The procedure gap for RNAV STARs is shown in the top figure, while the procedure gap for the RNAV SIDs is shown in the bottom figure.

standard deviations. Flow tubes derived from the Phoenix Sky Harbor Airport (PHX) arrival traffic can be seen in Figure 7. It is then our intention to cluster the flow tubes across the temporal dimension as well, and compare and contrast the results with those obtained via Algorithm 2. Before proceeding with this effort, however, we must derive a robust metric to quantify the differences between two flow tubes. Presumably, the area of intersection between flow tubes should be a part of this metric. We would to also formally compare compute-time and cluster quality of our algorithm against various large-scale spectral clustering algorithms, some of which were mentioned here: the KASP algorithm, the landmark-based clustering algorithm, and Nyström sampling-based algorithms.

Acknowledgment

The author would like to thank Dr. Tom Becher and Mahesh Balakrishna at MITRE CAASD for their continued support for, and feedback on, this research. Further, the author thanks Stephen Levin for his contributions to the Top-Down analysis section, and for supplying the procedure gap figures used in this paper. Finally, the author would like to acknowledge the following, for their contributions towards the Top-Down analysis effort: Olga Gisin, Cameron Osborne, John Timberlake, Mushava Kodzwa and Faisal Arain.

References

[1] C.T.H. Baker. *The Numerical Treatment of Integral Equations*. Monographs on numerical analysis. Clarendon Press, 1977.

[2] Xinlei Chen and Deng Cai. Large scale spectral clustering with landmark-based representation. In *AAAI*, 2011.

[3] Adric Eckstein. Automated flight track taxonomy for measuring benefits from performance based navigation. In *Integrated Communications Navigation and Surveillance Conference*, 2009.

[4] Marco Enriquez and Christopher Kurcz. A simple and robust flow detection algorithm based on spectral clustering. In *ICRAT Conference*, 2012.

[5] FAA. FAA aerospace forecast: Fiscal years 2012-2032. Technical report, Federal Aviation Administration, 2012.

[6] Miroslav Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25, 1975.

[7] Maxime Gariel, Ashok N. Srivastava, and Eric Feron. Trajectory clustering and an application to airspace monitoring. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1511–1524, 2011.

[8] Aude Marzuoli, Vlad Popescu, and Eric Feron. Two perspectives on graph-based traffic flow management. In *First SESAR Innovation Days*, 2011.

[9] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition, 2003.

[10] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[11] Lloyd N. Trefethen and David Bau. *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 1997.

[12] Ulrike von Luxburg. A tutorial on spectral cluster. Technical report, Max Planck Institute for Biological Cybernetic, 2007.

[13] Donghui Yan, Ling Huang, and Michael Jordan. Fast approximate spectral clustering. Technical Report UCB/EECS-2009-45, EECS Department, University of California, Berkeley, Mar 2009.

[14] Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *NIPS'04*, 2004.

Author Biography

Marco U. Enriquez holds a Ph.D. and M.A. in computational and applied mathematics from Rice University in Houston, Texas and a B.S. in computer engineering and mathematics from Tufts University in Medford, Massachusetts.

He is currently a Senior Applied Mathematician at MITRE CAASD, located in McLean, Virginia. He works in the Airspace/Procedures Criteria & Standards department. His research interests include data reduction and optimization theory, and the numerical algorithms supporting such topics.