

Clusters and Communities in Air Traffic Delay Networks*

Karthik Gopalakrishnan¹, Hamsa Balakrishnan¹ and Richard Jordan²

Abstract—The air transportation system is a network of many interacting, capacity-constrained elements. When the demand for airport and airspace resources exceed the available capacities of these resources, delays occur. The state of the air transportation system at any time can be represented as a weighted directed graph in which the nodes correspond to airports, and the weight on each arc is the delay experienced by departures on that origin-destination pair. Over the course of any day, the state of the system progresses through a time-series, where the state at any time-step is the weighted directed graph described above.

This paper presents algorithms for the clustering of air traffic delay network data from the US National Airspace System, in order to identify characteristic delay states (i.e., weighted directed graphs) as well as characteristic types-of-days (i.e., sequences of such weighted directed graphs) that are experienced by the air transportation system. The similarity of delay states during clustering are evaluated on the basis of not only the in- and out-degrees of the nodes (the total inbound and outbound delays), but also network-theoretic properties such as the eigenvector centralities, and the hub and authority scores of different nodes. Finally, the paper looks at community detection, that is, the grouping of nodes (airports) based on their similarities within a system delay state. The type of day is found to have an impact on the observed community structures.

I. INTRODUCTION

Flights in the United States experienced nearly 16 million minutes of delay in 2014 [1]. Air traffic delays have been estimated to cost the US economy \$31-41 billion annually [2], [3]. When the traffic demand for an airport or airspace resource exceeds the available capacity of that resource (either because of deterioration in capacity or over-scheduling), congestion and delays occur. The networked nature of the air traffic system, and the flows of aircraft, crew and passengers, result in delays propagating from one part of the system to another.

Network models present a natural way to represent the air transportation system. The major portion of previous work in this area has considered models of traffic, either in terms of aircraft or passengers [4], [5], [6], [7], [8]. By contrast, this paper considers network models of air traffic delays, and how one might use operational data to build these models. A key difference between the problems is that network models

of operations or passengers tend to be undirected when aggregated over several hours or a day (in other words, traffic levels are similar in both directions between two nodes), while delay network models tend to be directed.

This paper models the US National Airspace System (NAS) at any time as a directed network, with airports as nodes and the current delays between origin and destination airports as weights on the edges. It then uses algorithms for graph clustering to identify characteristic delay states of the NAS. In contrast to previous work on delay propagation in air traffic networks [9], [10], our primary focus is *not* to cluster network nodes (airports) on the basis of their similarities, but instead, to cluster the networks themselves (i.e., the system state at different times) on the basis of their similarities [11], [12], [13]. In earlier work, the NAS delay state was characterized using the edge weights of the network; however, the connectivity and network structure of the system were not explicitly considered [14].

The main objectives of this paper are as follows: (1) Identify, cluster and characterize air traffic delay networks that are similar to each other; and (2) identify, cluster and characterize days (24-hour periods) in the NAS which are similar to each other. Objective (1) will yield the characteristic *delay state* of the NAS at any time, while objective (2) will help identify characteristic *type-of-days* that describe the evolution of NAS delays, accounting for the fact that every hour of the day has an air traffic delay network associated with it.

II. AIR TRAFFIC DELAY NETWORKS

Consider a graph $G = (V, E)$, where V is the set of *vertices* (nodes) and E is the set of edges between them. Let $n = |V|$ be the number of vertices. There is no directionality associated with the edges of an undirected graph, while in the case of a directed graph, each edge is represented as an ordered pair, (v_1, v_2) , where v_1 is called the *origin* (or tail) and v_2 is called the *destination* (or head) of the edge. A *weighted graph* has a weight $w : E \rightarrow \mathbb{R}$ associated with each edge. Let the weight on edge (i, j) be denoted by w_{ij} . Weighted graphs are also called networks.

The *adjacency matrix*, $A \in \mathbb{R}^{n \times n}$ of a weighted graph is given by

$$A_{ij} = \begin{cases} w_{ij}, & \text{if } (i, j) \in E, \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

From (1), one can infer that the adjacency matrix of an undirected graph is symmetric (and that it therefore has real eigenvalues), while the adjacency matrix of a directed graph is asymmetric in general.

*This work was partially supported by NASA under Air Force Contract No. FA8721-05-C-0002, and NSF under the CPS project FORCES, grant number 1239054. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

¹Karthik Gopalakrishnan and Hamsa Balakrishnan are with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. {karthikg, hamsa}@mit.edu

²Richard Jordan is with the MIT Lincoln Laboratory, Lexington, MA 02420, USA. richard.jordan@ll.mit.edu

A. Description of data

The air traffic delay networks considered in this paper are constructed using data from the Bureau of Transportation Statistics (BTS) [15]. This data set contains details of flights logged by US airlines that account for at least 1% of domestic scheduled passenger revenues. Data from 2011-2012 are used to develop a network that reflects the delay state of the NAS at any time, focusing only origin-destination (OD) pairs that serve at least 5 flights a day on average. This filtering yields an unweighted graph with 158 nodes (airports) and 1,107 edges (OD pairs).

The edge weight on each edge of the air traffic delay network corresponds to the delay level of that edge. For every hour in the two-year period, we use a moving median (low-pass) filter on every OD pair in order to determine its delay value at that time. Delay values for time-periods shorter than 6 hours when there were no flights on an edge are estimated using interpolation. Details on the data processing and network generation can be found in Hanley [16]. This procedure generates an air traffic delay network for every hour of each of the 731 days in the data set, that is, $24 \times 731 = 17,544$ networks.

III. FEATURES FOR COMPARING NETWORKS

In order to cluster “similar” networks, we need to identify properties or features that can be used to evaluate their similarity to one another. Many candidate scores can be used to compare two networks; their suitability can vary depending on the application being considered [17], [12]. For example, one can compare the edge weights of two graphs, that is, use the Euclidean distance between the vectors of edge weights as a measure of graph similarity [14]. However, such a representation requires $O(n^2)$ parameters to represent a graph, where n is the number of nodes. In this work, we consider promising features by which to compare networks that represent the delays in the air traffic system, and then use them to cluster similar networks, that is, identify characteristic delay states of the air traffic network. By considering features that represent the nodes and not the edges of the network, we obtain models with $O(n)$ parameters.

A. Degrees of nodes

The in- and out-degrees of a node in the air traffic network correspond to the total inbound and outbound delay (summed over all the incoming and outgoing edges from that node). In addition, one can also determine the total delay associated with a node as the sum of its inbound and outbound delays. For an airport i , these quantities can be determined from the adjacency matrix as:

$$d^{\text{out}}(i) = \sum_j A_{ij} \quad (2)$$

$$d^{\text{in}}(i) = \sum_j A_{ji} \quad (3)$$

$$d^{\text{tot}}(i) = d^{\text{out}}(i) + d^{\text{in}}(i) \quad (4)$$

In other words, the air traffic delay network at any time can be represented by the $n \times 1$ vectors, \vec{d}^{out} , \vec{d}^{in} and \vec{d}^{tot} .

B. Eigenvector centrality

A measure of the importance of a node in a network is given by the *eigenvector centrality* [18], [19]. The eigenvector centrality of a node increases if that node is strongly connected to other nodes with high eigenvector centrality. The eigenvector centrality vector, \vec{z} , can be calculated from the adjacency matrix as the eigenvector corresponding to the largest eigenvalue of A . However, the eigenvalues of A are guaranteed to be real only if it is symmetric, which leads to the eigenvector centrality being a useful metric for undirected graphs. By contrast, the air traffic delay networks are directed weighted graphs.

One approach to applying the concept of eigenvector centrality to directed graphs is by “symmetrizing” the adjacency matrix, for example, by considering the graph with adjacency matrix \tilde{A} :

$$\tilde{A} = (A + A^T)/2. \quad (5)$$

In the context of the air traffic delay network, the weight on each edge of the undirected graph is the average of the delays on the two corresponding edges of the directed graph.

C. Hub and authority scores

Hubs and authorities were introduced for directed networks in the context of the Hypertext Induced Topics Search (HITS) algorithm [20], [21]. The key idea is that there are two types of important nodes in directed networks: *hubs* and *authorities*. An important hub points strongly to many important authorities, while important authorities are nodes that are pointed to strongly by many important hubs. In the context of air traffic delay networks, a strong hub refers to an airport that has significant outbound delays to strong authorities (i.e., airports that have significant inbound delays from strong hubs). The hub scores (\vec{h}) and authority scores (\vec{a}) can be determined iteratively [20], or by considering the dominant eigenvectors of AA^T and A^TA respectively. Equivalently, they can be determined from the dominant eigenvector of \mathcal{A} [22], given by

$$\mathcal{A} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}. \quad (6)$$

Since \mathcal{A} is symmetric, the largest eigenvalue and corresponding eigenvector are guaranteed to be real. If A is an $n \times n$ matrix, then \mathcal{A} is a $2n \times 2n$ matrix. The first n terms of its dominant (nonnegative) eigenvector correspond to the hub scores of the nodes of the directed graph, while the second n terms correspond to its authority scores [22].

Example 1: Let us consider the example of the network shown in Fig. 1.

The vectors of incoming delays, outgoing delays and total delays at each airport are given by:

$$\vec{d}^{\text{in}} = [45 \ 5 \ 40 \ 30 \ 5 \ 15]^T \quad (7)$$

$$\vec{d}^{\text{out}} = [35 \ 10 \ 25 \ 50 \ 15 \ 5]^T \quad (8)$$

$$\vec{d}^{\text{tot}} = [80 \ 15 \ 65 \ 80 \ 20 \ 20]^T \quad (9)$$

While the quantities \vec{d}^{in} , \vec{d}^{out} and \vec{d}^{tot} reflect the delays associated with each node in the network, they do not reflect

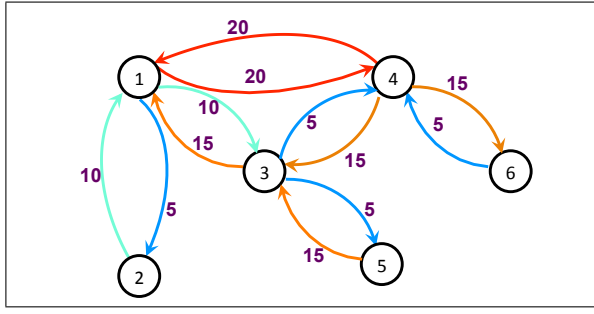


Fig. 1. Example of a weighted directed network, where the weights on the edges is the departure delay between the origin and the destination.

the edges, that is, the connectivity of those nodes. The eigenvector centrality can be determined from the undirected graph derived using (5), which is shown in Fig. 2.

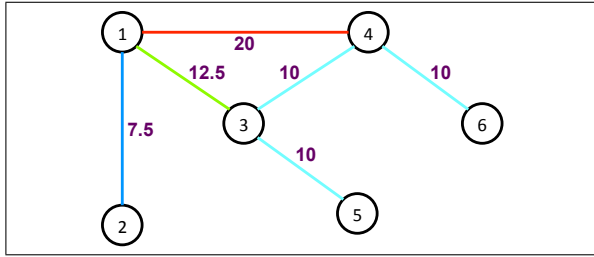


Fig. 2. Symmetrized version of the weighted directed network shown in Fig. 1.

The eigenvector centrality (i.e., the dominant eigenvector of \tilde{A}) in this case is given by

$$\vec{e} = [0.60 \ 0.14 \ 0.47 \ 0.59 \ 0.15 \ 0.19]^T \quad (10)$$

Comparing the values of \vec{e} and \vec{d}^{tot} in this example, we see that they are not identical. For example, the nodes 1 and 4 have the same total delay (80 units), but their eigenvector centralities differ, with node 1 being slightly more central. Similarly, nodes 5 and 6 have the same total delays, but different centralities. In both cases, we see that this is due to connectivity: Node 4 and node 1 are both similarly connected to each other and to node 3, but node 4 is more strongly connected to node 6 than node 1 is to node 2. Similarly, node 6 is more central than node 5 due to its connectivity to node 4 (which has high centrality) rather than to node 3 (which has lower centrality).

One can also compare the hub (\vec{h}) and authority scores (\vec{a}) of various nodes to their outbound (d^{out}) and inbound (d^{in}) delays, respectively. For this directed network,

$$\vec{a} = [0.70 \ 0.05 \ 0.56 \ 0.24 \ 0.05 \ 0.36]^T \quad (11)$$

$$\vec{h} = [0.32 \ 0.21 \ 0.35 \ 0.82 \ 0.25 \ 0.04]^T \quad (12)$$

Although node 4 has higher total inbound delay than node 6, it has a lower authority score since none of the three nodes it is connected to (1, 3 and 6) has a high hub score (and the weights on links (3,4) and (6,4) have low weights), while node 6 is strongly connected to node 4, which is a strong hub.

As the previous example suggests, the different features discussed above reflect different properties of the network. While the delay features (\vec{d}^{in} and \vec{d}^{out}) reflect the total inbound and outbound delays at a node, the hub and authority scores reflect the connectivity and the propensity for delay to propagate into and out of a node in a network.

IV. CLUSTERING OF NETWORKS

The features discussed in the previous section can be used to compare networks, and to cluster similar networks. In other words, two observed air traffic delay networks are judged to be similar to each other if these features are close to one another. The Euclidean distance between the feature vectors can be used to measure distance. Candidate feature vectors that were considered included:

- 1) Outbound and inbound delay vectors: $\begin{bmatrix} \vec{d}^{\text{out}} \\ \vec{d}^{\text{in}} \end{bmatrix}$.
- 2) Eigenvector centralities of the nodes of the symmetrized version of the air traffic delay network: \vec{e} .
- 3) Eigenvector centralities weighted by the total delay in the system: $(\sum_j \vec{d}^{\text{in}}(j)) \vec{e}$.
- 4) Hub and authority scores of the nodes of the air traffic delay network: $\begin{bmatrix} \vec{h} \\ \vec{a} \end{bmatrix}$.
- 5) Hub and authority scores weighted by the total delay in the system: $(\sum_j \vec{d}^{\text{in}}(j)) \begin{bmatrix} \vec{h} \\ \vec{a} \end{bmatrix}$.
- 6) Hub and authority scores along with the outbound and inbound delays (appropriately normalized to be in $[0, 1]$) at nodes: $\begin{bmatrix} \vec{h} \\ \vec{a} \\ \vec{d}^{\text{out}} \\ \vec{d}^{\text{in}} \end{bmatrix}$.

k-means or k-medoids algorithms were used to determine clusters of similar graphs. The Euclidean distance between the feature vectors was used as the objective function. The number of clusters is evaluated using various criteria, including the sum of distances to the nearest cluster centroids and silhouette plots [23]. The next section presents some illustrative results of clustering using different feature vectors.

A. Clustering by eigenvector centralities

Fig. 3 shows the 6 clusters (suggested by the silhouette plots [23]) that were identified by the eigenvector centralities, \vec{e} . These centroids can be qualitatively describes as follows: (1st row) High centrality in the West coast of the US (especially San Francisco (SFO) airport) exhibits strong centrality; (2nd row) Chicago O'Hare (ORD) has high centrality; (3rd row) Atlanta (ATL) airport has strong centrality; (4th row) a diffused state; (5th row) Elevated centralities at the East Coast airports, and (6th row) Dallas/Fort Worth (DFW) airport has strong centrality.

B. Clustering by Hub/Authority scores weighted by the total system delay

The feature vectors used in the previous section, eigenvector centralities, do not account for either the directed nature

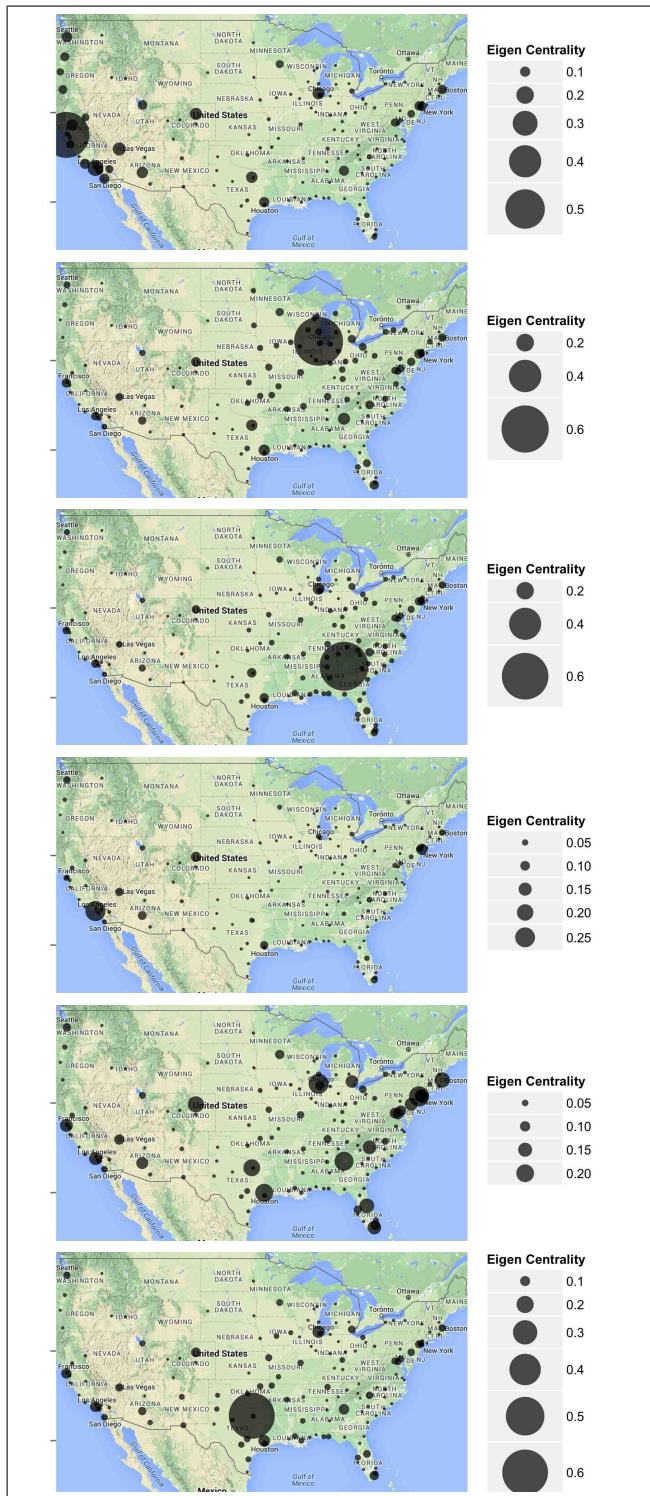


Fig. 3. Centroids of clusters identified by k-means clustering of eigenvector centralities. The radius of the shaded circle is proportional to its eigenvector centrality. We see that there is (top row) one state in which SFO exhibits strong centrality, (2nd row) one in which ORD airport does so, (3rd row) one in which ATL airport has strong centrality, (4th row) a diffused state, (5th row) a state with slightly higher centralities in the East Coast, and (6th row) a state in which Dallas Fort Worth (DFW) airport has strong centrality.

of the networks being clustered, or possible scalings of the network. In other words, the eigenvector centralities would

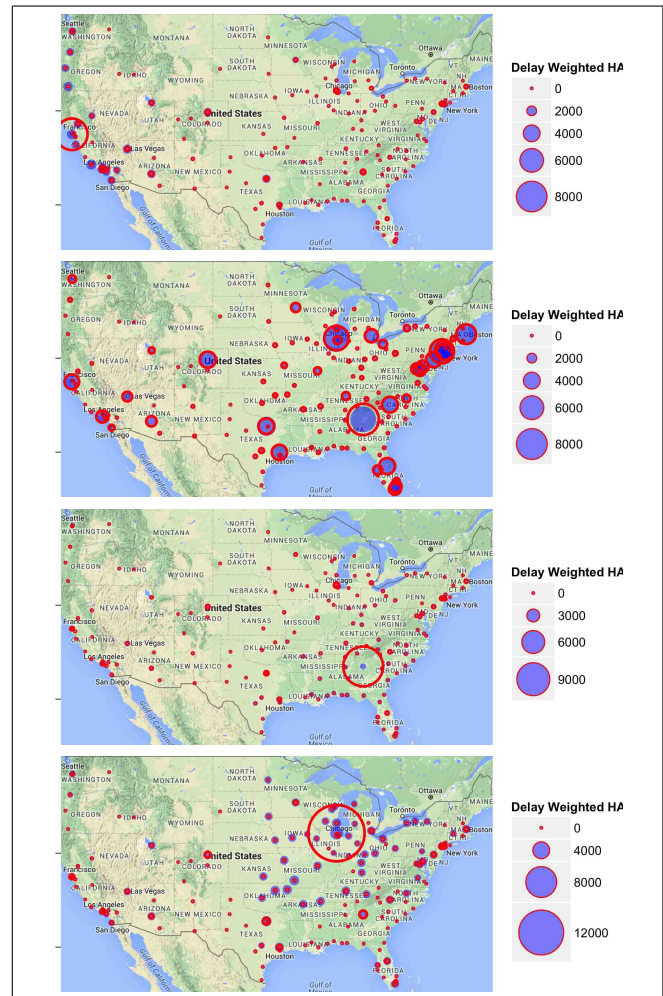


Fig. 4. Centroids of 4 of the 6 clusters identified by k-means clustering of the delay-weighted hub and authority scores. The radius of the blue shaded circle is proportional to the hub score of an airport weighted by the delay while that of the red circle is proportional to the delay-weighted authority score.

be the same for two networks that had adjacency matrices that were scaled versions of each other; however, the two networks would have very different system delays.

Fig. 4 shows 4 of the 6 clusters identified by the hub and authority scores, multiplied by the total system delay at that time. This feature vector accounts both for the directed nature of the network, as well as the total delay in the system corresponding to that network. The identified clusters can be qualitatively described as follows: (1st row) High delays at SFO which is also an authority; (2nd row) high delays and strong hub and authority scores in the Eastern US, including ATL, ORD and the Northeast airports; (3rd row) high delays and strong authority scores at ATL; and (4th row) very high delays and authority scores at ORD. The 5th and 6th states (not shown) have low and medium delays diffused across the system. It is interesting to note the differences between the eigenvector centrality clustering and clustering using hub/authority scores and delays. In particular, the latter does not identify a clear state in which DFW is dominant, instead opting for a medium NAS delays state in which

hub/authority scores are more diffused.

The NAS also exhibits temporal variations, and different delay states are more likely to occur at different times of the day. Fig. 5 shows the empirical frequencies of NAS delay occurrences by time of day. The low delay state is seen to occur consistently overnight, when traffic demand is low. The frequency of occurrence of this state is lowest during peak demand hours. The next most-frequently occurring cluster, which has medium delays distributed across major airports, tends to occur during the day when there are no major disruptions in the system. The “SFO high delay” state tends to occur with greatest frequency in the morning hours in the West coast of the US (i.e., afternoon in the East Coast), when the SFO area experiences fog in the summer months. The very high ORD delays cluster tends to peak around 6PM, when demand is at its highest. The high ATL and high East Coast delay states tend to occur in the late evening hours, when congestion peaks.

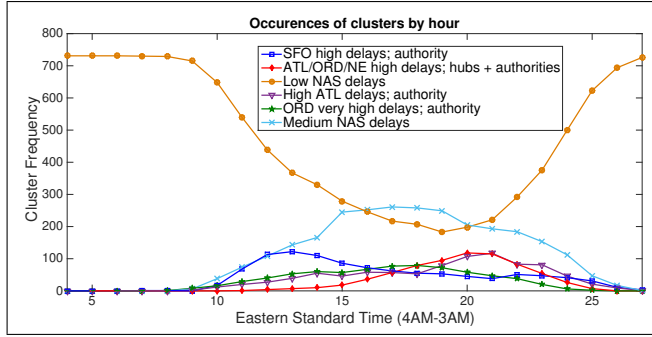


Fig. 5. Variation of occurrence of characteristic NAS delay state by time of day.

V. CLUSTERING OF SIMILAR TYPES OF DAY

While the state of the NAS at any time can be modeled as a network, a day is a time series of these states, and is therefore a time series of networks. Taking an approach similar to the one in the previous sections, characteristic days can be identified by clustering days that were most similar to each other. Each day of the 731 days in the data set consists of a 24 hour long time series of feature vectors, with each feature vector representing the air traffic delay network at that time. The candidate feature vectors at each time are still the ones discussed in Section IV.

The following sections present a few illustrative results from the clustering of these time series of air traffic delay networks.

A. Clustering time series of out- and inbound delays

Each day in the dataset corresponds to a time-series of length 24, with a air traffic delay network at each time. The delay network at time t on day n can be represented by the vector $\begin{bmatrix} \vec{d}_{t,n}^{\text{out}} \\ \vec{d}_{t,n}^{\text{in}} \end{bmatrix}$. Therefore, day n corresponds to the time series

$$D_n = \begin{bmatrix} \vec{d}_{1,n}^{\text{out}} \\ \vec{d}_{1,n}^{\text{in}} \end{bmatrix} \dots \begin{bmatrix} \vec{d}_{24,n}^{\text{out}} \\ \vec{d}_{24,n}^{\text{in}} \end{bmatrix} \quad (13)$$

Similar days can be determined using k-means clustering, by comparing the Euclidian distance between the time series D_n on different days.

Clustering using the outbound and inbound delays as feature vectors yields 6 characteristic types of day, as suggested by silhouette plots [24]. The centroids of these clusters, which are each a time series of 24 feature vectors, can be qualitatively described as follows: Days on which (1) Atlanta (ATL) very high delays, (2) West coast high delays, (3) ORD very high delays, (4) Northeast/ORD/ATL delays, (5) Medium NAS delays, and (6) Low NAS delays.

B. Clustering time series of hub and authority scores

Alternatively, we can use the feature vector (4) described in Section IV, namely the hub and authority scores, to represent the network at any time. In other words, day n is represented by the time series

$$D_n = \begin{bmatrix} \vec{h}_{1,n} \\ \vec{a}_{1,n} \end{bmatrix} \dots \begin{bmatrix} \vec{h}_{24,n} \\ \vec{a}_{24,n} \end{bmatrix}, \quad (14)$$

where $\vec{h}_{t,n}$ and $\vec{a}_{t,n}$ represent the hub and authority scores of the nodes of the network at time t on day n .

Silhouette plots for k-means clustering using time series of hub/authority scores for each day suggest a choice of 12 clusters. The centroids can be qualitatively described as follows: Days on which (1) Northeast has med HA scores, (2) Low/diffused HA scores, (3) ATL high H scores, (4) ATL high HA and ORD high H, (5) ORD med HA, (6) SFO high HA, (7) DFW high HA, (8) ATL med A and DEN, LAX med HA, (9) SFO high A, (10) IAH high HA, (11) ORD high H, and (12) ORD v. high A.

C. Clustering time series of inbound and outbound delays, and hub and authority scores

In considering the feature vectors in Sections V-A and V-B, one notes that the former only considers the edge weights (delays) but not the connectivities (the susceptibility of nodes to further inbound or outbound delays), while the latter only considers the connectivities but not the delays themselves. In reality, both kinds of feature vectors need to be considered. An analysis of the overlap between the two sets of clusters shows that there are relationships between them.

Fig. 6 shows the empirical probability distribution of which hub/authority-based cluster days from each of the five delay-based clusters would belong to. For example, a day classified as ATL high delays is most likely to also have high hub/authority scores at ATL, and there is a similar correlation between delays and hub/authority scores at the Northeast airports. High hub/authority scores at DFW are strongly correlated with medium NAS delays. Chicago delays (medium to very high) tend to be correlated with high ORD authority scores (and midwest hub scores). Finally, elevated hub/authority scores in the West coast tend to cause localized delays, and to not have a significant impact on systemwide delays. This statement is supported by the observation that most days with high hub and authority scores at SFO fall into the “West coast high delays” cluster,

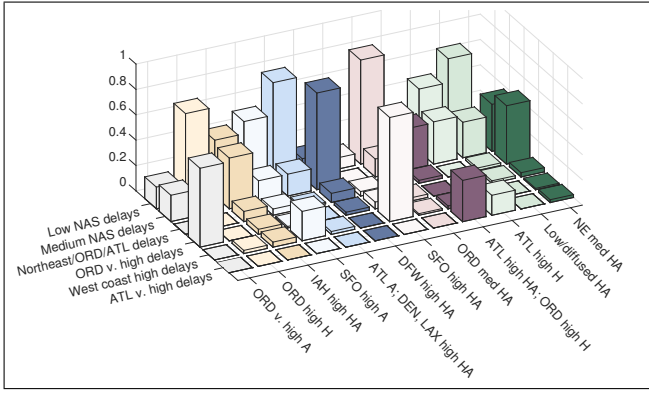


Fig. 6. Relationship between characteristic types of day identified using (1) outbound and inbound delays as feature vectors at each time, and (2) hub and authority scores as feature vectors at each time.

and most of the ones with high authority scores at SFO fall into either the low NAS delays cluster or the West coast high delays cluster. These days tend to have lower NAS-wide delays (not shown in figure). Fig. 6 also suggests that 10-12 clusters may be a reasonable number to describe the characteristic types of delay days seen in the NAS.

Another method for incorporating both inbound/outbound delays and hub/authority scores is to cluster days based on time series of feature vectors that include delays and hub/authority scores. Using the feature vector (6) described in Section IV, day n would be represented by the time series

$$D_n = \begin{bmatrix} \vec{h}_{1,n} \\ \vec{a}_{1,n} \\ \vec{d}_{1,n}^{\text{out}} \\ \vec{d}_{1,n}^{\text{in}} \end{bmatrix} \cdots \begin{bmatrix} \vec{h}_{24,n} \\ \vec{a}_{24,n} \\ \vec{d}_{24,n}^{\text{out}} \\ \vec{d}_{24,n}^{\text{in}} \end{bmatrix}, \quad (15)$$

where the delays at the nodes were appropriately normalized to be in $[0, 1]$. Clustering days using this approach yields 10 clusters, as was suggested by Fig. 6. These clusters can be qualitatively described as follows:

- 1) *Low NAS 1*: Low system delays; elevated delays and hub/authority scores at ORD, ATL and the Texas area.
- 2) *ATL*: Very high delays and hub/authority scores at ATL.
- 3) *ORD*: Very high delays and hub/authority scores at ORD.
- 4) *Low NAS 2*: Low system delays; elevated delays and hub/authority scores at ATL.
- 5) *Low NAS 3*: Low system delays; elevated delays and hub/authority scores at DEN.
- 6) *IAH*: Very high delays and hub/authority scores at IAH.
- 7) *NE/ORD/ATL*: High delays and hub/authority scores in the Northeast, ORD and ATL.
- 8) *SFO authority*: High delays and authority scores at SFO.
- 9) *SFO hub and authority*: High delays and hub/authority scores at SFO; elevated hub scores at LAX.
- 10) *DFW*: High delays and hub/authority scores at DFW.

VI. COMMUNITY DETECTION

It would be valuable to be able to identify airports that are most similar, that is, ones that tend to have similar delays (or propensity for delays). This problem is analogous to that of clustering similar nodes in networks [9], [10], [25]. As a first step, we consider the symmetrized version of the air traffic delay network, aggregated over a day. Community detection algorithms using the notion of *modularity*, as proposed by Girvan and Newman [26], determine groups of nodes (communities) such that there is stronger connectivity between nodes within a particular community than between nodes in different ones.

Given a partitioning of the network nodes into C communities, the modularity or quality of the partitioning is evaluated using the function, Q such that

$$Q = \sum_{i=1}^C (e_{ii} - a_i^2), \quad (16)$$

where e_{ij} is the fraction of total network edge weight that is on edges that connect nodes in community i to those in community j , and $a_i = \sum_j e_{ij}$. The partitioning is optimized over all values of C and possible community structures in order to maximize Q [26]. For example, community detection for the undirected graph shown in Fig. 2 results in two communities, one containing nodes 1, 2, 4 and 6, and the other containing nodes 3 and 5 (Fig. 7).

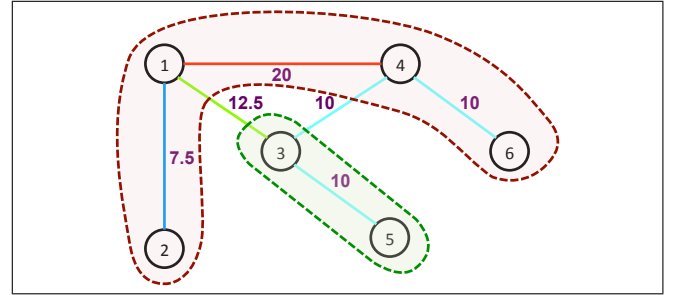


Fig. 7. Community detection for the undirected graph shown in Fig. 2.

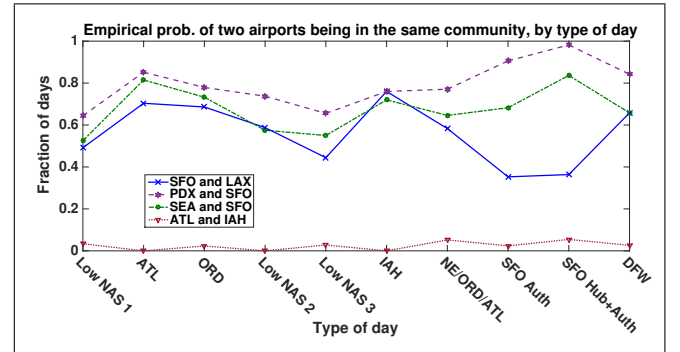


Fig. 8. Empirical probability of airports belonging to the same community on a particular type of day, for four pairs of airports: LAX and SFO, PDX and SFO, SEA and SFO, and ATL and IAH.

In air traffic delay networks, there is stronger connectivity amongst airports within a community (as indicated by the

delay weights on the edges) than to nodes in other communities. It is interesting to observe that the community structures vary depending on the type of day that is being experienced in the NAS. Fig. 8 shows the empirical probability that two airports belong to the same community, given that the day has been classified as being of a particular type. It shows that while Atlanta Hartsfield International (ATL) and Houston George Bush Intercontinental (IAH) airports are almost never in the same community, Portland International (PDX) and San Francisco (SFO) airports are in the same community on the majority of days. In addition, on days labeled as having high SFO delays, hub and authority scores, PDX and SFO are almost always in the same community. By contrast, the likelihood of SFO and Los Angeles (LAX) airports being in the same community can vary from nearly 80% on days when Houston (IAH) is impacted, to less than 40% on days on which SFO is impacted. Similarly, Seattle (SEA) and SFO are in the same community on 80% of days on which ATL is impacted or SFO is both a hub and an authority, but only 50% of days when delays are low, with slightly higher delays at Denver (DEN) airport.

VII. DISCUSSION AND NEXT STEPS

The results presented in this paper demonstrate the promise of clustering air traffic delay networks in order to identify characteristic delay states and characteristic types of day that take into account both spatial and temporal patterns and connectivity. The analyses help identify key airports and constraints that drive air traffic delays. The community detection analysis shows that groups of airports are strongly connected in terms of delays on certain types of days. Ongoing work on this topic includes comparing the characteristic types of days identified by clustering time series of networks (as was done in this paper) to types of days identified by clustering aggregate delays experienced at different airports. In addition, the evolution of delay states during the course of a day could be compared to the disruptions and control actions (i.e., Traffic Management Initiatives) that were seen on that day. Finally, the characteristic delay states, types of days, and significant features identified in this work can be used to develop predictive models of air traffic network delays, which may in turn enable improved decision support tools for stakeholders such as air traffic managers.

REFERENCES

- [1] Federal Aviation Administration (FAA), "OPSNET website," Accessed 2014, <http://aspm.faa.gov/>.
- [2] M. Ball, C. Barnhart, M. Dresner, M. Hansen, K. Neels, A. Odoni, E. Peterson, L. Sherry, A. Trani, and B. Zou, "Total delay impact study," 2010.
- [3] Joint Economic Committee, US Senate, "Your Flight has Been Delayed Again: Flight Delays Cost Passengers, Airlines, and the US Economy Billions," 2008.
- [4] G. Spiers, P. Wei, and D. Sun, "Algebraic connectivity maximization of the air transportation network," in *American Control Conference*, 2012.
- [5] P. Wei, G. Spiers, and D. Sun, "Algebraic connectivity maximization for air transportation networks," *IEEE Transactions on Intelligent Transportation Systems*, no. 2, pp. 685–698, 2014.
- [6] O. Lordan, J. M. Sallan, P. Simo, and D. Gonzalez-Prieto, "Robustness of the air transport network," *Transportation Research Part E*, pp. 155–163, 2014.
- [7] S. Bratu and C. Barnhart, "An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data," *Air Traffic Control Quarterly*, vol. 13, no. 1, pp. 1–27, 2005.
- [8] C. Nicolaides, L. Cueto-Felgueroso, M. C. Gonzalez, and R. Juanes, "A Metric of Influential Spreading during Contagion Dynamics through the Air Transportation Network," *PLoS One*, vol. 7, no. 7, 2012.
- [9] P. Fleurquin, J. Ramasco, and V. Eguiluz, "Systemic delay propagation in the US airport network," *Scientific Reports*, p. 1159, 2013.
- [10] —, "Data-driven modeling of systemic delay propagation under severe meteorological conditions," in *11th USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Lisbon, Portugal, June 2015.
- [11] S. E. Schaeffer, "Graph clustering," *Computer Science Review* 1, pp. 27–64, 2007.
- [12] B. Luo, R. Wilson, and E. Hancock, "Spectral clustering of graphs," in *10th International Conference on Computer Analysis of Images and Patterns, CAIP*, 2003.
- [13] P.-A. Champin and C. Solnon, "Measuring the similarity of labeled graphs," in *Proceedings of the 5th international conference on Case-based reasoning (ICCBR'03): Research and Development*, 2003, pp. 80–95.
- [14] J. J. Rebollo and H. Balakrishnan, "Characterization and prediction of air traffic delays," *Transportation Research Part C*, pp. 231–241, 2014.
- [15] Bureau of Transportation Statistics, "Airline On-Time Statistics and Delay Causes," 2015. [Online]. Available: <http://www.transtats.bts.gov/>
- [16] Z. J. Hanley, "Delay Characterization and Prediction in Major U.S. Airline Networks," Master's thesis, Massachusetts Institute of Technology, 2015.
- [17] L. A. Zager and G. C. Verghese, "Graph similarity scoring and matching," *Applied Mathematics Letters*, pp. 86–94, 2008.
- [18] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *Journal of Mathematical Sociology*, pp. 113–120, 2008.
- [19] S. P. Borgatti, "Centrality and network flow," *Social Networks*, pp. 55–71, 2005.
- [20] J. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM*, pp. 604–632, 1999.
- [21] T. Deguchi, K. Takahashi, H. Takayasu, and M. Takayasu, "Hubs and Authorities in the World Trade Network Using a Weighted HITS Algorithm," *PLoS One*, vol. 9, no. 7, 2014.
- [22] M. Benzi, E. Estrada, and C. Klymko, "Ranking hubs and authorities using matrix functions," *Linear Algebra and its Applications*, no. 5, pp. 2447–2474, 2013.
- [23] M. Charrad, N. Ghazzali, V. Boiteau, and A. Niknafs, "NbClust: An R package for determining the relevant number of clusters in a data set," *Journal of Statistical Software*, no. 6, 2014.
- [24] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, pp. 53–65, 1987.
- [25] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, no. 6, 2004.
- [26] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, pp. 7821–7826, 2002.