Toward Optimal Transport Networks

Natalia Alexandrov∗, Rex K. Kincaid† and Erik P. Vargo‡

∗NASA Langley Research Center, Hampton VA 23681-2199, USA
†‡College of William & Mary, Williamsburg VA 23187-8795, USA

Strictly evolutionary approaches to improving the air transport system – a highly complex network of interacting systems – no longer suffice in the face of demand that is projected to double or triple in the near future. Thus evolutionary approaches should be augmented with active design methods. The ability to actively design, optimize and control a system presupposes the existence of predictive modeling and reasonably well-defined functional dependences between the controllable variables of the system and objective and constraint functions for optimization. Following recent advances in the studies of the effects of network topology structure on dynamics, we investigate the performance of dynamic processes on transport networks as a function of the first nontrivial eigenvalue of the network’s Laplacian, which, in turn, is a function of the network’s connectivity and modularity. The last two characteristics can be controlled and tuned via optimization. We consider design optimization problem formulations. We have developed a flexible simulation of network topology coupled with flows on the network for use as a platform for computational experiments.

Nomenclature

λ₁ The smallest eigenvalue of \( L \)
λ₂ The smallest nontrivial eigenvalue of the \( L \)
λₘₐₓ The largest eigenvalue of the \( L \)
\( A \) Adjacency matrix
\( a_{ij} \) Elements of \( A \); 1 if node \( i \) connected to node \( j \); 0 otherwise
\( c(G) \) Clustering coefficient
\( D \) Degree matrix
\( d_i \) Degree of node \( i \)
\( G \) Graph corresponding to a network
\( L \) Network Laplacian
\( N \) The number of nodes in \( G \)
\( s(G) \) A measure of the tendency of nodes of similar degree to be connected
\( s_{\text{max}} \) Maximum of \( s(G) \)
\( s_{\text{min}} \) Minimum of \( s(G) \)

I. Introduction

Large technological systems, such as the Internet and the air transportation system, have never been actively designed in their entirety; instead, they evolved gradually, in response to demands. Accommodating the projected near-future threefold increase in demand for air transportation, as well as the growing complexity of airspace, via strictly evolutionary changes is problematic. There appears to be a clear need for active and rigorous design methods.

∗Senior Research Scientist, MS 442, AIAA Senior Member.
†Professor, Department of Mathematics, P.O. Box 8795.
‡Graduate Student, Department of Mathematics, P.O. Box 8795.
Active rigorous design necessitates identification of design variables, as well as objectives and constraints whose functional dependence on the independent variables has been established analytically and experimentally. In other words, active design requires predictive models. Although design of systems such as air transportation may never be as amenable to active design as are complex physical artifacts (e.g., airplanes, engines), we conjecture that a significant portion of active design is possible with appropriately developed variables and metrics.

Air transportation is a highly complex network. The past several years have seen a surge in the study of both the structure (see, e.g., Newman\textsuperscript{1} for an overview) of complex networks and the dynamics (e.g., Strogatz\textsuperscript{2}) of flows on complex networks. The effect of network structure on the dynamics is the starting point in our investigation aimed at deriving quantifiable functional relations between global transport metrics (e.g., throughput, delays, capacity) and locally controllable structure (e.g., connectivity) of transport networks. Of particular interest is computational controllability of the local variables expressed as an optimization problem.

Recent groundbreaking work in a variety of domains (statistical physics, biology, Internet router design) derives models that link static topological network structure variables, such as degree distribution, clustering coefficient, and average shortest path distance, with the dynamic performance of flows over networks (e.g., synchronicity, throughput, congestion). In particular, Atay et al.\textsuperscript{3,4} Donetti et al.\textsuperscript{5} Motter et al.\textsuperscript{6,7} have begun to establish explicit functional dependences between topology and system dynamics on networks. Other measures, such as assortativity and \( s(G) \) (the tendency to connect with nodes of like degree) has been explored by other authors, e.g., Li et al.\textsuperscript{8,9}

While there is a superficial similarity between the systems under study in other domains and the air transport system, there are also substantial differences. For instance, the variation in the Internet packet size is not nearly as great as the variation in the plane capacity; the connections (routes) among nodes of the airport network are not as hard wired as for the Internet. Moreover, the assumptions made in the physics literature are not applicable to the transport network. For example, Atay et al.\textsuperscript{3,4} assume that all nodes are identical and conform to a generic discrete time equation. These assumptions are needed to simplify the problem in the initial stages of research, but a question remains whether the derived models would be sufficiently applicable to air transport.

Having conducted initial computational experiments, we believe that there is much to be learned from the research efforts in other domains and, although the models will not carry over one-to-one to air transport, the models from other domains can be developed further to accommodate transport assumptions and needs. In this paper, we summarize our results to date: a computational study of network topology metrics; construction of a test platform; and formulation of optimization problems. We will call our problem air transport network (ATN) optimization.

II. Spectral Graph Properties and Dynamic Processes

Several authors\textsuperscript{3,4,5,7,6} have recently examined network dynamics as a function of network topology and showed that different constrained topology optimization problems, such as maximizing synchronization or node proximity, lead to optimal topologies that, although not identical, share common features. Donetti et al.\textsuperscript{5} call these optimal networks entangled. The result appears relevant to optimization of transport networks. In this section, we review some of the basic concepts.

Let \( G \) be a graph (network) with \( N \) nodes. Associated with \( G \) is the adjacency matrix \( A = [a_{ij}] \), where \( a_{ij} = 1 \) if nodes \( i \) and \( j \) are connected and 0 otherwise. The degree matrix \( D = [d_i] \) is a diagonal matrix of node degrees. The Laplacian matrix \( L = D - A \) fully captures the connectivity structure of \( G \). The eigenvalues of \( L \) satisfy

\[
2d_{\text{max}} \geq \lambda_N \geq \ldots \geq \lambda_2 \geq \lambda_1 = 0,
\]

where \( d_{\text{max}} \) is the maximum degree of \( G \), and the eigenvector \((1, \ldots, 1)\) corresponds to the trivial eigenvalue \( \lambda_1 \).\textsuperscript{10} The first nontrivial eigenvalue \( \lambda_2 \) is known as the spectral gap or algebraic connectivity.

Donetti et al.\textsuperscript{5} give an illustrative example that foreshadows the topological significance of the spectral gap for networks dynamics. If a network consists of a number of disconnected subgraphs, its Laplacian is block-diagonal and the multiplicity of the trivial eigenvalue equals the number of disconnected subgraphs. Connecting the subgraphs weakly introduces small eigenvalues with nearly constant corresponding eigenvectors. This feature (small spectral gaps) provides criteria for graph partitioning in well-known algorithms.\textsuperscript{11}

Intuitively, small \( \lambda_2 \) values imply the existence of well-defined modules that can be disconnected by cutting
a small number of links, while large $\lambda_2$ values point to unstructured (entangled) graphs.

Several authors\textsuperscript{12,5,7,6} studied synchronizability of diffusive processes on networks, considering a general dynamical process

$$\dot{x}_i = F(x_i) - \sigma \sum_{j=1}^{N} L_{ij} H(x_j), i = 1, \ldots, N, \quad (1)$$

where $x_i$ are dynamical variables, $F$ is an evolution function, $H$ is a coupling function, and $\sigma$ is a coupling constant. Although diffusive processes are known to have synchronous states, the question is, under what conditions these states are stable. A linear stability analysis\textsuperscript{12} reveals that synchronized states are more stable for smaller $\frac{\lambda_2}{\lambda_N}$. Since the variability of the maximum eigenvalue is bounded\textsuperscript{13} for graphs under consideration, increasing stability of synchronized states amounts to maximizing the spectral gap $\lambda_2$. Other authors\textsuperscript{3,14} have used the spectral gap as an indicator of synchronization for discrete processes as well.

The normalized Laplacian, $L' = D^{-1} L$, and its eigenvalues $\{\lambda'_i\}$ also play an important role, especially in the study of random walks, a subject relevant to propagation of traffic through networks. Large spectral gaps increase the rate at which random walks move and disperse. A class of graphs with large spectral gaps, known as Ramanujan graphs, is described by Donetti et al.\textsuperscript{5} These graphs are regular, have a vanishing clustering coefficient, a small average shortest path distance and a large girth. Since air transport networks do not have this structure, general model (1) is not applicable. However, changing the coupling constant to $\frac{\sigma}{\sigma_t}$ and thus normalizing the effect of the neighboring nodes (in turn, increasing the relevance to traffic networks), results in an optimal topology when the normalized spectral gap is maximized.

In studying networks metrics, we also examine the function, $s(G)$, developed by Li et al.\textsuperscript{8,9} This metric distinguishes among network structures of a fixed degree distribution, based on the tendency of nodes of similar (or dissimilar) degrees to be connected. It serves as an indicator of performance in, e.g., router design. The metric is defined as

$$s(G) = \sum_{(i,j) \in \text{Links}} d_i \times d_j, \quad (2)$$

where $d_i$ is the degree of node $i$. The metric is often normalized as

$$S(G) = \frac{s(G) - s_{\min}}{s_{\max} - s_{\min}},$$

where $s_{\max}$ and $s_{\min}$ are the maximum and minimum values, respectively, of the sum taken over all possible connected graphs for a fixed degree distribution. Since it turns out that $s_{\min}$ is difficult to compute in practice, the metric is often approximated by

$$S(G) = \frac{s(G)}{s(G) - s_{\max}}.$$

Li et al.\textsuperscript{8} term graph realizations with large values of $s(G)$ scale-free and graph realizations with small values of $s(G)$ scale-rich. Consequently, in scale-free graphs high-degree nodes are more likely to be adjacent to other high-degree nodes, while in scale-rich graphs high-degree nodes are more likely to be adjacent to low-degree nodes.

### III. Transport Network Design Formulations

Entangled topologies optimal for such objectives as synchronization and maximum proximity of nodes bear some resemblance to topologies of transport networks derived by Guimerà et al.\textsuperscript{15,16} for transport with congested hubs. Intuitively, this makes sense: if hubs are congested, then the best way to disperse the load is by providing a massively connected network that is difficult to separate by cutting a few links. This supports the use of the spectral gap as one of the objectives for ATN design.

Our initial experiments with the spectral gap\textsuperscript{17} further support the notion of its usefulness as an indicator of performance. However, many questions remain open. As we mentioned, the assumptions underlying the construction of optimal entangled topologies are not, in general, applicable to realistic transport networks. For instance, transport networks are not regular; their degree distributions are not necessarily fixed. We need to introduce assumptions that will eventually allow us to model more practical conditions.
As a start, consider that the spectral gap of the entire network is too global a measure. We do not have the freedom of rewiring the network completely, especially considering the many players in the network (e.g., the air carriers). Moreover, even in the presence of congestion, doing away with hubs completely may not be an option because hub-like structures have a number of important benefits, such as the small world property (minimal hops) and robustness with respect to random disturbances. The optimal answer likely lies in a mix of modularity and entanglement. Is there a way to pose the graph optimization problem that preserves some degree of global modularity, while allowing for local entanglement within modules?

Let $G_0, G_1, \ldots, G_n$ be a decomposition of $G$, as follows. $G_0$ represents the subset of the network shared by all other subsets. This subset must maintain some degree of modularity. Let $V_i$ be the set of nodes (vertices) of graph $i$. $G_1, \ldots, G_n$ are subgraphs allowed to entangle locally. All subsets must maintain a fixed $s(G_i)$ and degree distribution

$$P_i(d) = \frac{1}{N_i} \sum_{j \in V_i | d_j \leq d} 1,$$

where $N_i$ is the number of nodes in subgraph $i$.

We will use problem formulations analogous to those used for multidisciplinary design optimization (MDO). However, due to a different nature of ATN (discrete and, eventually, mixed), the analogy to MDO formulations is purely structural. For now, the number and location of nodes will be fixed (as in the present airport network). Consider an example from a range of formulations being studied.

In this formulation, we employ bilevel optimization with no feedback. Let $G$ denote simple connected graphs. The outer problem will maximize the spectral gap of the global graph $G_0$, i.e., we solve approximately

$$\begin{align*}
\text{minimize}_{G_0} & \quad \lambda_2(G_0) \\
\text{subject to} & \quad s(G_0) = C_{G_0}^1 \\
& \quad P_0(d) = C_{G_0}^2 \\
& \quad G_0 \in \mathcal{G},
\end{align*}$$

(3)

where, $C_{G_0}^1$ and $C_{G_0}^2$ are constants. Then the subproblems will minimize their spectral gaps to increase connectivity, subject to maintaining the inter-subgraph connections fixed. That is, for each subproblem, we solve approximately

$$\begin{align*}
\text{maximize}_{G_i} & \quad \lambda_2(G_i) \\
\text{subject to} & \quad s(G_i) = C_{G_i}^1 \\
& \quad P_i(d) = C_{G_i}^2 \\
& \quad G_0 \in \mathcal{G} \text{ remains fixed.}
\end{align*}$$

(4)

Formulation (3–4) should maintain connections among the modules, while maximizing the fluency of the flow inside the modules. Other authors have considered other notions of partitioning the network into local and global subsets.

IV. Test Platform

Many sophisticated airspace simulation models have been developed in recent years. They include trajectory-based models, such as ACES and FACET; models based on network flow with physical analogies; empirical models; models based on Bayesian networks. See the last reference for a more complete list of models. The existing models operate under a variety of assumptions and predict various effects, such as congestion and delays. Despite the availability of a number of models, we have decided to develop our own research test platform. We do this for a number of reasons. To derive functional dependences in the system, we need a completely flexible definition of the system components. This implies control over the source code. Computational expense of some of the existing detailed airspace models is another consideration. Finally, we can use data generated by the existing models as a load on our system.

The Airport Network Simulation Program (ANSP) was designed as a first step in the development of an optimization framework able first to investigate and, we hope, eventually to model and design complex transport networks in the context of the total transportation system. Version 1 of ANSP is a discrete-event simulation, with events separated by discrete time intervals and processed chronologically. A global event list drives the simulation by coordinating interactions between the simulation’s two primary structures: nodes
(airports) and planes. While events are currently limited to arrivals, departures and vector updates, ANSP’s flexibility allows new events to be added easily. The target is an adaptive, gradually expanding simulation that will assist in numerical experimentation with transportation system concepts and design methods.

At completion, each simulation computes a number of metrics, currently including the total holding time for each flight, the average number of planes in queue for each airport; total number of hops, time in air, and distance traveled for each flight.

V. Computational Experiments

To-date, we have experimented extensively to observe the relations among various network metrics, such as the spectral gap, assortativity \((s(G))\), cluster coefficient, graph eccentricity and diameter, and others. The experiments confirmed that the spectral gap should prove useful, at some level of detail, for ATN design. They also raised questions about the correlation among the metrics. To answer these questions, we must transition from the purely static structure experiments to experiments that include traffic flow dynamics.

In this section, we give an example of static experiments with the spectral gap and report on the initial test of the ANSP platform that will allow us to introduce dynamics on the network. ANSP tests were conducted using actual 1990 U.S. Air Transport data.\(^{26}\)

V.A. Experiments with the spectral gap

We experimented with two types of abstract, 100-node, simple, undirected networks (no loops or multiple links): preferential attachment and geometric. Preferential attachment graphs are generated following the approach given by Barabási,\(^{27}\) where a network is grown by adding nodes and edges; for each node added, \(m\) edges are added preferentially, based on the current degree distribution. We chose \(m = 2\) to facilitate visualization.

Geometric graphs are generated by randomly selecting 100 points \((r, \theta)\) with values of \(r \in [0, 1]\) and values of \(\theta \in (0, 360]\). Edges exist between pairs of points if the Euclidean distance is less than a specified threshold (in our experiments thresholds between 0.17 and 0.25 were used). If the resulting graph is connected, it is kept; otherwise it is rejected and the process begins again.

The graphs plotted in Figures 1 and 2 were constructed by first generating a random instance of the particular graph class—geometric in Figure 1 or preferential attachment in Figure 2. Next, a simple tabu search heuristic was called to minimize or maximize the spectral gap, while keeping the degree distribution and \(s(G)\) fixed. Allowable re-wirings are pair-wise edge interchanges that preserve the degree distribution and \(s(G)\). Briefly, the tabu search checks to see if the move is acceptable, that is, if the move is improving and not tabu, or improving and tabu but leads to the best observed value of \(\lambda_2\) (aspiration criterion).

Figures 1 and 2 display networks with respect to the reciprocal of the eccentricity of each node \(u\). The eccentricity of \(u\) is its maximum (shortest path) distance. The graphs are generated by socnetv\(^a\). The goal of the plots is to uncover any qualitative differences between the graphs with small and large values of the spectral gap. Nodes with equal eccentricity values are plotted on the same (dashed line) circles. The circles with larger radii have larger eccentricity. Consequently, nodes near the center have shorter longest paths.

Qualitatively, when \(\lambda_2\) is small, the patterns are less organized, the eccentricity plots in Figures 1a and 2a are more dispersed and consist of many rings of constant eccentricity. The eccentricity plots with larger \(\lambda_2\) are more organized, with few rings of constant eccentricity. We refer the reader to an earlier paper\(^{17}\) for detailed description of the experiments and other metrics.

V.B. Experiments with U.S. 1990 air transport data

The functionality of ANSP was tested by comparing simulation statistics for the original network and several rewirings produced by tabu search.\(^{28}\) It is important to note that at this stage, ANSP alone cannot capture the true system complexity and dynamics; these computations are used purely to verify the performance of the platform, rather than to draw any definitive conclusions about the functional dependences.

U.S. 1990 airport activity statistics\(^{26}\) were used to generate an airport network with 336 nodes (airports). The network was then optimized, using four runs of tabu search. The four resulting networks were “optimal”

\(^{a}\)The source code and documentation can be found at \texttt{http://socnetv.sourceforge.net/}.
Figure 1. Geometric graphs: 100 nodes, $s(G) = 0.971$, fixed degree distribution

(a) $\lambda_2 = 0.009$, $c(G) = 0.426$
(b) $\lambda_2 = 0.314$, $c(G) = 0.297$

Figure 2. Preferential attachment: 100 nodes, $s(G) = 0.716$, fixed degree distribution

(a) $\lambda_2 = 0.006$, $c(G) = 0.210$
(b) $\lambda_2 = 0.365$, $c(G) = 0.101$
with respect to one of the four objectives ($s_{min}$, $s_{max}$, $\lambda_{2min}$, and $\lambda_{2max}$), subject to preserving the node degree sequence. Again, since tabu search is a heuristic, true optimality is not guaranteed.

We now had five networks: the original one and four networks optimal with respect to one of the four
metrics. The original airport network, $s(g) \approx 104,276,109$ and $\lambda_2 \approx 0.339$. The tabu search algorithms produced four networks with the following heuristically optimal metrics:

$$s_{\text{min}} \approx 90,988,359 \quad s_{\text{max}} \approx 109,205,715 \quad \lambda_{2\text{min}} \approx 0.132 \quad \lambda_{2\text{max}} \approx 0.382.$$

Now the five networks were imported into ANSP, which ran 100 independent trials, each generating 400 random origin-destination routes. We chose arbitrary large numbers of trials and routes to compute statistical averages of the total holding time (time in a holding pattern), total distance traveled, total time in air, time average number of planes in queue (arrival and departure) at each airport, and total hops for each flight were computed for each trial. The results were compared across all networks and are shown in the plots of Figures 3–5. (Time is in hours.)

In this experiment, we assign superior performance to lower numbers for holding time, distance traveled, time in air, and hops. A qualitative analysis shows that the original network and the $\lambda_{2\text{max}}$ rewiring exhibited the best overall performance. Both networks have relatively large $\lambda_2$ values, and hence high connectivity, relative to other networks with the same node degree sequence. The $\lambda_{2\text{min}}$ network was a low performer. We conjecture that low connectivity and the reliance on indirect shortest paths for flight routes may be a contributing factor. The $s_{\text{max}}$ network exhibits similar connectivity. However, the low performance of the $s_{\text{min}}$ network cannot be attributed directly to its spectral properties, as its spectral gap is large. Further analysis shows that the greatest performance detractor for the $s_{\text{min}}$ network is the reliance on major hubs. This may appear counterintuitive, since a smaller $s(G)$ typically corresponds to a less hub-like topology. However, the preponderance of east-west flights in the network appears to magnify the $s_{\text{min}}$ system’s use of hubs, specifically the Chicago O’Hare International Airport (ORD). After 100 simulations, approximately 58 flights per trial passed through ORD in the $s_{\text{max}}$ network and 95 for the $s_{\text{min}}$ network. ORD was consistently the most loaded airport in each network, which is a clear contributing factor to large average holding time and average time in air for the $s_{\text{min}}$ network. A tentative conclusion is that, while the spectral gap and $s(G)$ are informative for linking topological and dynamical properties, further assumptions and modeling elaborations are needed to account for realistic dynamic behavior.

A second test was conducted on a smaller network of 318 nodes (airports) that did not include Alaskan
airports. Some of the metrics are summarized in Tables 1–3. Table 1 contains metrics from the original network, Table 2 from the network rewired for $s_{\text{max}}$ and Table 3 from the network rewired for $s_{\text{min}}$. Note that the removal of Alaska increased the value of the spectral gap, compared to that of the 336 node network.
This stands to reason as removing a “module” lead to a more entangled network.

VI. Concluding Remarks

The present state of air transportation networks necessitates developing active and rigorous design methodologies to augment the traditional evolutionary approaches. We investigate the applicability of network models originating in statistical physics, Internet router design, and biology to air transport design. Although some of the modeling assumptions do not hold, the initial results are sufficiently interesting to warrant an effort in adapting the models for use in air transport analysis and optimization. We are cautiously optimistic that introducing increasingly realistic formulations will allow us to progress from abstract networks to models that approximate transport networks to a greater degree of fidelity.

References