Network Characteristics of Air Traffic in the Continental United States

Banavar Sridhar*, Kapil Sheth**

* NASA Ames Research Center, Moffett Field, CA 94035
USA (Tel: 650-604-5450; e-mail: Banavar.Sridhar@nasa.gov).
** Tel: 650-604-5728; (e-mail:Kapil.Sheth@nasa.gov)

Abstract: Air traffic is undergoing big changes both in developed and developing countries. The demand for air traffic in the United States is expected to grow to 2 or 3 times the current levels of traffic in the next few decades. This paper models demand for air traffic as a network with several thousand nodes. Recently, methods have been developed to study the network characteristics of large complex systems in many natural and engineering fields. Several networks exhibit the "scale-free" property that as the network grows in size, a small number of components have a disproportionate influence on the successful operation of the network. This paper describes the network properties of the current air traffic system and uses future scenarios of air traffic growth to understand the performance of future air traffic networks. The air traffic network exhibits an exponentially truncated scale-free behavior. It is shown that a three-times growth in the overall traffic may result in a ten-times impact on the density of traffic in certain parts of the United States. Thus, in addition to bottlenecks at major airports, the risk of en route traffic saturation calls for route restructuring and the introduction of new operational concepts, such as automation to increase en route capacity.

1. INTRODUCTION

Air traffic in the United States continues to grow at a steady pace except for a dip immediately after the tragic events of September 11, 2001. There are different growth scenarios associated both with the magnitude and the composition of the future air traffic. Terminal Area Forecast (TAF), prepared every year by the Federal Aviation Administration (FAA), projects the growth of traffic in the United States [Anonymous 2005]. Both Boeing and Airbus publish market outlooks for air travel annually. Although predicting future growth of traffic is difficult, there are two significant trends in the growth of air traffic demand: (1) heavily congested major airports continue to see an increase in traffic, and (2) the emergence of Regional Jets and other smaller aircraft with fewer passengers operating directly between non-major airports. The interaction between air traffic demand and the ability of the system to provide the necessary airport and airspace resources can be modelled as a network. The size of the resulting network varies on the choice of the nodes of the network. It would be useful to understand the properties of this network to guide future design and development. Many questions such as the growth of delay with increasing traffic demand and impact of the en route weather on future air traffic require a systematic understanding of the properties of the air traffic network.

Recently, there has been a major advance in analyzing the behavior of biological and engineering networks with large number of components [Strogatz 2001]. Several networks exhibit a scale-free property in the sense that the probabilistic distribution of their nodes as a function of connections decreases slower than an exponential. These networks are characterized by the fact that a small number of components have a disproportionate influence on the performance of the network. Scale-free networks are tolerant to random failure of components; but are vulnerable to selective attack on components [Albert 2000, Newman 2003].

This paper examines two network representations for the current air traffic system. A network defined with the 40 major airports as nodes and with standard routes as links has a characteristic scale: all nodes have 60 or more links and no node has more than 460 links. A network defined with current aircraft routing structure exhibits an exponentially truncated scale-free behavior. Its degree ranges from 2 connections to 2900 connections, and 225 nodes have more than 250 connections. Furthermore, those high-degree nodes are homogeneously distributed in the airspace. A consequence of this scale-free behavior is that the random loss of a single node has little impact, but the loss of multiple high-degree nodes (such as occurs during major storms in busy airspace) can adversely impact the system. Two future scenarios are used to predict the growth of air traffic in the United States. It is shown that a three-times growth in the overall traffic may result in a ten-times impact on the density of traffic in certain parts of the United States.

The paper is organized as follows: Section on Complex Network Analysis provides a brief overview and terminology of the complex networks useful in the analysis of air traffic. This is followed by an application of the complex network methodology to the current air traffic system. Section on Future Air Traffic Scenarios considers different scenarios for the evolution of air traffic in the United States during the next 25 years and looks for changes in the behavior of the network. An analysis of the future air traffic scenarios shows that a three-times growth in overall traffic can result in a tenfold impact on the density of highly connected nodes in certain parts of the United States. Thus, in addition to bottlenecks at major airports, the risk of airspace congestion calls for route

restructuring and the introduction of new procedures and automation to increase airspace capacity. Concluding remarks and possible research in this area is discussed in the last Section.

2. COMPLEX NETWORK ANALYSIS

Complex systems have many agents or components interacting with each other, and their collective behavior is not a simple combination of the individual behavior. The pattern of interaction between agents can be studied as a network of connections between agents. Networks of many types are pervasive in modern society and scientists from many fields are trying to broaden their understanding of the structure of the networks. A network is made up of basic components called either vertices or nodes. Each node is connected to other nodes in the system. The line connecting two nodes is referred to as a link or an edge. An edge is directed if it runs only in one direction, and undirected if it runs in both directions. Degree is the number of edges connected to a node. Figure 1 shows a simple network with nodes A, B, C and D and edges connecting them. The degree of node A, B, C and D is respectively 3, 2, 2 and 1 in the figure.
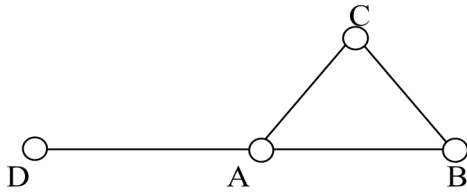

Figure 1. Nodes and edges in a network
The structure of the network has a strong influence on the functions performed by the network. It is possible to analyze networks of small sizes (less than 100) by drawing the picture of the network and analyzing its properties. The number of nodes in a complex engineering system, such as the worldwide web [Broder 2000], can easily be as large as a few million nodes. The behavior of random and regular networks with large number of components has been studied for a long time [Erdos 1960, Newman 2003]. Networks describing real systems are neither random nor regular. The ability to model real systems and capture the behavior of key variables is extremely difficult. Recent developments in complex networks examine the statistical properties of large networks and help answer questions about the dynamics and stability of such networks.

A major area of interest is the role played by the distribution of the degree of nodes in a network. Let p_k be the fraction of nodes in the network that has degree $\mathbf{k}$. The degree distribution for the network can be computed by generating a histogram of p_k. The degree distribution for large random graphs, where each edge is present or absent with equal probability, can be modeled as a Poisson distribution. Figure 2 shows the Poisson distribution, $e^{-\lambda}\lambda^k / k!$, for $\lambda = 4$. The distribution tends to peak for a small value of k and decays exponentially for large values of k. When using real data, to reduce the effect of noise in small datasets, the distribution of degree is expressed in terms of the complementary cumulative distribution function (ccdf), the probability of the number of nodes in the graph with degree greater than k, P_k. P_k is computed using the expression

$$P_k = \sum_{i=k}^{\infty} p_i$$

Figure 2 shows the ccdf for the Poisson distribution and Figure 3 shows it on a logarithmic scale.

It has been observed that the degree distribution for some real-world networks, such as the Internet and biological networks, are highly skewed with tails several times longer than the mean. These real networks have a small number of nodes, or hubs, with high degree of connectivity. Hubs play an important role in influencing the properties of the network. Real networks exhibiting a small number of hubs are referred to as "scale-free" networks [Barabasi 2003].
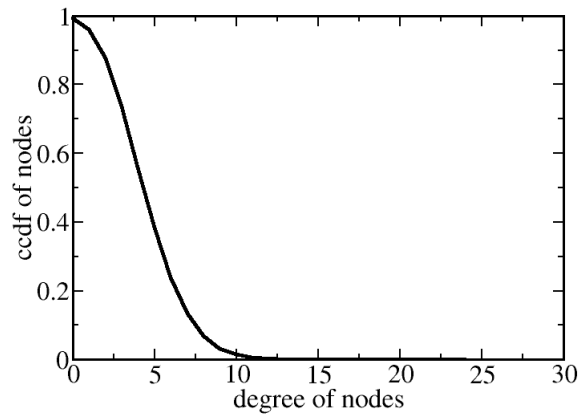

Figure 2. Poisson distribution
The distributions of degree for a number of networks, e.g. Internet, World Wide Web, collaboration network of mathematicians, etc., show a power law in their tails: $P_k \approx k^{-\alpha}$ for some constant value of α. Figure 4 shows the ccdf for α=2. Power law distributions appear linear in a logarithmic scale (Figure 5).
The appearance of hubs in scale-free networks is explained in terms of two behaviors both in people and real systems-growth and preferential attachment. There is a tendency for new growth to gravitate towards desirable locations. Generally, all the desirable locations already have some existing nodes. The preferential attachment mechanism leads to the creation of more powerful nodes or hubs. The various hubs compete to attract new nodes and absorb them, resulting in an increase of the number of links in the initial hubs.

Scale-free networks with power law distributions exhibit two important properties: (a) remarkable resistance to random failure of nodes and (b) extreme vulnerability to targeted attacks. The functionality and performance of the network depends on the existence of the edges between pairs of nodes. The distance between the remaining vertices gets longer as nodes are removed from a network and the removal of a certain number nodes may result in a collapse of the entire system.
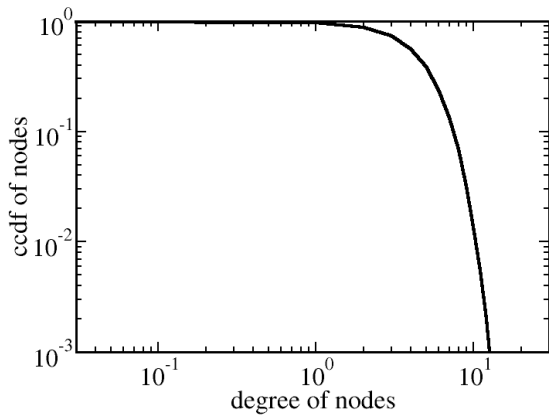
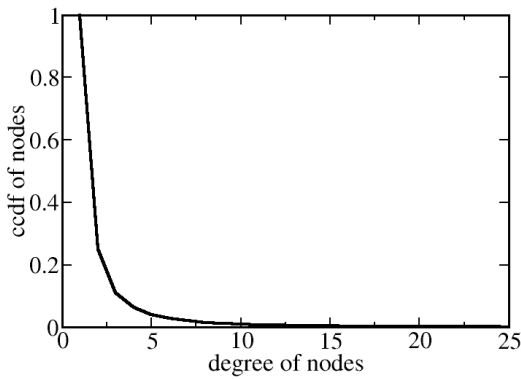Figure 3. Poisson distribution in logarithmic scale
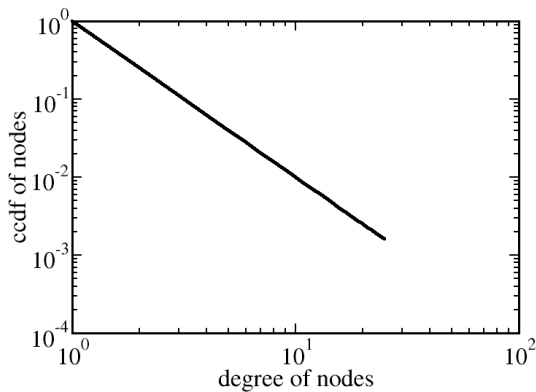

Figure 4. Power law distribution

Figure 5. Power law distribution in logarithmic scale

The tolerance of networks to failures or removal of nodes varies with their level of resilience. The nodes from a network could be removed randomly or in a coordinated approach. Several studies have shown that scale-free networks are generally robust to random removal of nodes. However, they are less tolerant to the selective removal of hubs.

The next section will examine the air traffic in the United States to see if it exhibits the properties of a scale-free network and draw analogies from the attack vulnerabilities of other complex networks. The methodology can also be used to understand the properties of the Next Generation Air Transportation System as it evolves over the next few decades [Pearce 2006].

3. CURRENT U.S. AIR TRAFFIC NETWORK

The National Airspace System (NAS) refers to the collection of hardware, software and people, including runways, radars, networks, FAA, airlines, etc., involved in air traffic management (ATM) in the U.S. It has been pointed from a modelling perspective that the NAS should be treated as a complex adaptive system [Donohue 2003]. NAS can be looked as a network at several different levels. The analysis of current traffic uses air traffic data from 2006 and constructs two different types of networks for a typical day of operations in the NAS. The analysis is based on actual traffic flow for a day in July 2006. The traffic flow data is processed using Future Air traffic management Concept Evaluation Tool (FACET) simulation software [Sridhar 2006] for conducting the network flow analysis.

Current air traffic is modelled in two different ways and the characteristics of the resulting air traffic networks are described in Table 1. The nodes in the Airport Network (AN) correspond to the 40 major airports in the U.S. Today, a set of predefined alternative routes is used for flying between city pairs. Each node in network AN is connected to another node through one or more routes. The Airport and Airspace Network (AAN) includes all airports in the U.S. and routes connecting these airports. A route between two airports is defined by a series of geographical positions or fixes in the airspace. Each fix together with the airports represent a node in the AAN network.

Table 1. Different types of ATM networks

Network	Nodes	Number of Nodes	Edges
Airport Network (AN)	40 major airports	40	Routes between 40 major airports
Airport and Airspace Network (AAN)	All airports and fixes along routes connecting all airports	8170	Routes between all airports

3.1 Behavior of the degree of nodes in ATM network

The ccdf for the nodes in network AN is shown in Figure 6. The figure shows the probability that an airport in the network has more than a certain number of routes originating or terminating at the airport. The degree of nodes at all airports exceeds 60 and Chicago has the maximum number (460) of connections. The distribution of the routes between airports does not show scale-free behavior. The airport connection distribution in the U.S is similar to the distribution of the connections between the network of world airports [Amaral 2000].

Figure 7 shows the ccdf of the nodes in network AAN. The ccdf of the nodes in network AAN shows a rapid decay of the percentage of nodes with the degree of the node. However, it has 225 nodes with more than 250 links, indicative of high volumes of traffic through these nodes.

Figure 8 shows the logarithmic behavior of the ccdf of the nodes in network AAN.

The nodes of network AAN initially follow the power law curve similar to several scale-free networks [Newman 2003]. However, towards the end of the tail, the distribution shows a deviation from the power law behavior indicating a limit on the number of nodes that have a high number of connections. The limit on the growth of the large hub nodes, in the distribution of the ATM network, is due to the constraints imposed on traffic demand by the location of the cities, economic development and government policies.

The existence of hub and secondary airports characterizes current air traffic operations [Bonnefoy 2004]. The network analysis provides an additional ability to study traffic behavior in the en route airspace. Figure 9 shows geographical distribution of the nodes with degree higher than 250 in network AAN. These nodes will be referred to as 250G nodes. 250G nodes represent approximately 2.7% of the total nodes in the network. Table 2 shows the number of 250G nodes in different traffic control regions, referred to as Centers. The Centers, Chicago (ZAU), Boston (ZBW), Atlanta (ZTL), Kansas City (ZKC), Washington, DC (ZDC), Indianapolis (ZID), Jacksonville (ZJX), Los Angeles (ZLA), Cleveland (ZOB) and New York (ZNY), each have more than ten of the 250G nodes. The Centers in the western part of the United States have fewer 250G nodes indicative of the lower traffic density in these Centers. It is informative to express the distribution of the 250G nodes in units of number of nodes per 10,000 square nautical miles (10Ksqnm). Using this measure of nodal traffic density, ZNY has the highest nodal density with six 250G nodes per 10Ksqnm.

The behavior of the degree of the nodes in an ATM network should not be surprising since the ATM network has evolved to serve population densities in the U.S. Network analysis helps to visualize and quantify the characteristics of the network.

3.2 Resilience of ATM networks

The tolerance of complex networks to random and targeted failures depends on their network structure. As observed earlier, a scale-free network is tolerant to random failure, since the hubs are few and the chance of a hub being selected randomly is low. However, the same network may be prone to targeted attacks on a small percentage of vital nodes. In ATM networks, weather can be regarded as an agent of attacks on the system.

Convective weather is a major source of uncertainty in ATM networks. One effect of severe weather is to make airspace unavailable for the flow of air traffic. The removal of airspace may result in the failure of hubs and increase the average path length in the network. The impact of weather on ATM system performance appears as delay [Sridhar 2005]. The tolerance of the ATM network to weather depends on the geographical distribution of the weather and the coverage of nodes with high degree.

4. FUTURE AIR TRAFFIC SCENARIOS

The demand for air transportation is projected to grow in proportion to the expansion of global trade. National and international projections of traffic growth indicate a tripling of passengers by 2025 [Pearce 2006].
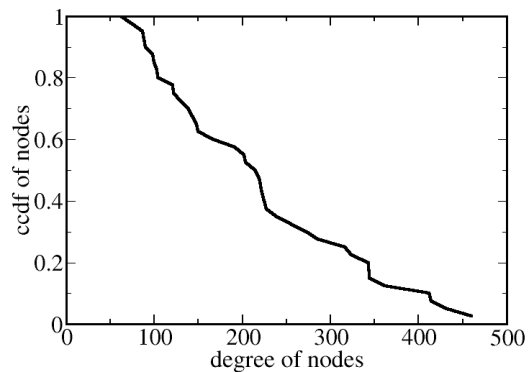

Figure 6. Cumulative distribution function for nodes (40 major airports) in network AN
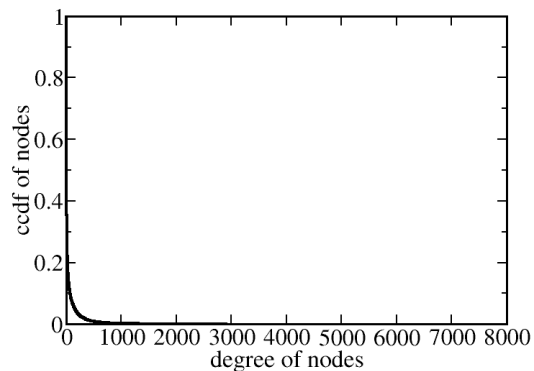

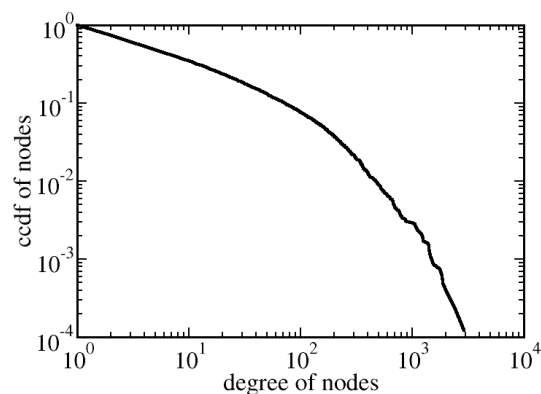Figure 7. Distribution of nodes in AAN with current traffic

Figure 8. Logarithmic behavior of nodes in network AAN

There may be increased traffic due to the growing presence of on-demand air taxis and unmanned air vehicles. It is estimated that 5000 micro jets may be operational by 2010 and 13500 by 2022. The TAF provides forecasts for airports in the NAS. The forecast for the 35 major airports in the United States receive more detailed modeling that takes into account local economic conditions and airline costs. TAF is the basis for most aviation demand forecasts. The TAF data published in

March 2006 provides traffic growth rates for the period 2005-2021. In order to achieve 3-times current day

Figure 9. Geographical distribution of 250G nodes in AAN under current traffic and severe weather polygons

traffic (3X) scenarios assumed in future planning, two different scenarios for post-2025 traffic growth are used in the traffic demand model [Huang 2006]. The compound extrapolation approach grows traffic until it reaches 3X by assuming the TAF airport growth rates for traffic beyond 2025. As an alternative, the homogeneous extrapolation approach assumes the same growth rates at all airports until the traffic level reaches 3X. These two scenarios will be referred to as 3X Compound (3XC) and 3X Homogeneous (3XH) scenarios in the rest of the paper. The scenarios are processed to compute the changes in the network properties of future ATM systems.
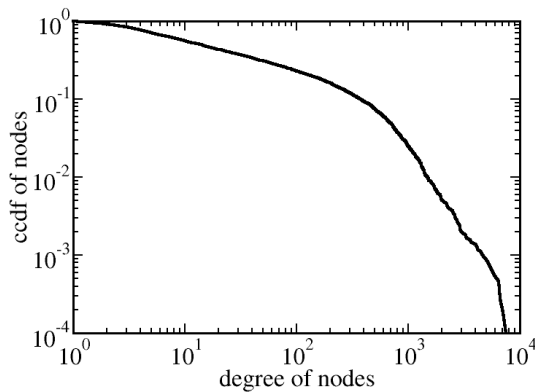

Figure 10. Logarithmic behavior of nodes in network AAN with 3X Compound traffic

4.1 Behavior of future ATM networks

The properties of future ATM networks can be derived similar to the properties of the current ATM networks. The computations can be used to compare the geographical distribution of the hubs compared to the distribution today. Figures 10 shows the distribution of the nodes in the AAN network for the 3XC traffic scenario. The result for the 3XH traffic scenario is similar.

The future ATM networks, under both scenarios, show exponentially truncated scale-free behavior similar to the current ATM network. The distribution of the hubs will have an impact on the growth of delay in future ATM networks subjected to severe weather. The vulnerability of the network to reduction in capacity caused by certain weather patterns

will be very significant and disruptive to the operation of the system. The impact of the weather on future ATM networks will be described subsequently in the paper.

Table 2. Distribution of 250G nodes by Centers under current and future traffic scenarios

Center	Current	3XH	3XC
ZAB	7	70	78
ZAU	15	67	70
ZBW	16	77	87
ZDC	22	134	143
ZDV	8	51	68
ZFW	3	59	68
ZHU	8	60	72
ZID	15	71	71
ZJX	14	71	81
ZKC	11	58	50
ZLA	16	102	119
ZLC	5	29	40
ZMA	4	44	54
ZME	10	38	42
ZMP	6	46	53
ZNY	14	80	88
ZOA	10	57	67
ZOB	13	90	89
ZSE	5	35	41
ZTL	15	59	62
Total	217	1298	1443

The geographical distribution of the 250G nodes under 3XC is shown in Figure 11. Table 2 shows the same distribution by Centers for current traffic and 3X scenarios. The 3X traffic demand creates close to a six-fold increase in the 250G traffic nodes compared to current traffic. Earlier, it was noted that ZNY has the highest 250G traffic nodal density. The hubs, 500G, 750G and 1000G, are defined similar to the 250G hub. Under the 3X demand, it can be shown that ZNY nodal density of 2006 is equaled or exceeded by a majority of the 20 Centers. Even more alarmingly, the nodal density of 250G nodes in ZNY is 6 times the current value, and Cleveland and Washington Centers have twice the nodal density of ZNY in 2006. As seen in Table 3, the 3X traffic also gives rise to nodes with even more connections.

Table 3. Nodes for different traffic scenarios

Types of nodes	Current	3XH	3XC
250G	225	1312	1468
500G	72	620	806
750G	33	304	452
1000G	22	165	262

4.2 Impact of Weather on ATM Networks

Another way to view the growth of traffic is to compare how similar weather patterns may affect current traffic and 3X traffic. The geographical distribution of the 250G nodes shown in Figure 11 is based on traffic during July 2, 2006. July 2, 2006 was a calm weather day with a total NAS aggregate delay of 11997 minutes [FAA 2004]. The traffic on

July 2, 2006 will be assumed as traffic unaffected by weather in this discussion [Sridhar 2005].

Whenever there is severe weather in the NAS, airspace capacity is reduced and traffic is rerouted or held on the ground causing delays in the system. Collaborative Convective Forecast Product (CCFP) is a model of severe weather activity and the areas marked blue in Figure 11 shows the CCFP for July 13, 2006. On July 13, 2006 the NAS experienced a significant total delay of 219350 minutes with ZNY, ZDC and ZTL Centers contributing a delay of 116654, 49400 and 23822 minutes respectively. Next, as shown in Figure 11, the same severe weather is overlaid in blue polygons on the geographical distribution of the 250G nodes under 3XC demand. Table 4 shows hub nodes affected by the weather today and in the future.
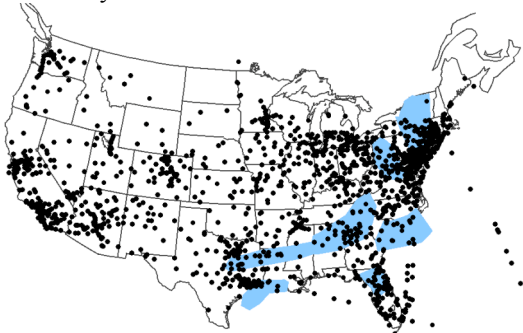

Figure 11. Geographical distribution of 250G nodes in network AAN under 3XC and severe weather polygons

Table 4. Nodes affected by severe weather

Types of Nodes	1X	3XH	3XC
250G	34	191	207
500G	9	96	118
750G	4	49	67
1000G	3	23	41

A greater number of hubs, between six to ten times, are affected by the same weather pattern under future traffic scenarios than today. If one considers the non-linear growth of delay, in regions such as ZNY, ZDC and ZTL where the demand on the airspace is close to capacity, the increased density of high traffic nodes in these regions will result in much larger delays compared to 2006.

5. CONCLUDING REMARKS

Air traffic in the United States can be modeled as a network to understand the impact of the predicted growth in the demand on the performance of the system. It is demonstrated that the air traffic network with current en route flight plan intersections as nodes and with the flight plans as links shows scale-free properties typical of several large engineering and biological networks. A consequence of this property is the non-linear growth of traffic in certain regions of the United States. A preliminary analysis indicates that a three-times growth in the overall traffic may result in a ten-times impact on the density of traffic in certain parts of the United States. The air traffic system currently experiences significant delay during periods of severe weather activity. The impact of weather of same severity will be magnified several times, especially in the northeastern parts of the United States, leading to lower system performance. Research must be conducted to determine whether this risk can be mitigated through restructuring routes or by introducing new operational concepts, such as automation assistance to controllers to increase airspace capacity. The network analysis described in the paper can be used to guide the development of various traffic flow management concepts to increase the efficiency of air traffic systems.

REFERENCES

[1] Anonymous, "Terminal Area Forecast Summary, Fiscal Years 2004-2020," Report No. FAA-APO-05-1, U.S. Department of Transportation, Federal Aviation Administration, Mar. 2005.

[2] Strogatz, S. H., "Exploring complex networks," Nature, 410, 268–276, 2001.

[3] Albert, R., Jeong, H. & Barabasi, A.L., "Error and attack tolerance of complex networks," *Nature* 406, 378–382, 2000.

[4] Newman, M. E. J., "The structure and function of complex networks," SIAM Review 45, 167-256, 2003.

[5] Broder, A. *et al.,* "Graph structure in the web," *Computer Networks*. **33,** 309–320, 2000.

[6] Erdos, P., and Renyi, A., "On the Evolution of Random Graphs," Publication of the Mathematical Institute, Hungarian Academy of Science, 5, 17-61, 1960.

[7] Barabasi, A.L., and Bonabeau, E., "Scale-Free Networks", Scientific American, 50-59, May 2003.

[8] Pearce, R.A., "The Next Generation Air Transportation System: Transformation Starts Now," Journal of Air Traffic Control, pages 7-10, Jan-Mar., 2006.

[9] Donohue, G., "Air Transportation is a complex Adaptive System: Not an Aircraft Design," AIAA/ICAS International Air and Space Symposium and Exposition, July 2003, Dayton, Ohio.

[10] Sridhar, B., Sheth, K., Smith, P., and Leber, W., "Migration of FACET from Simulation Environment to Dispatcher Decision Support System," *Proceedings of Digital Avionics Systems Conference*, Washington, D.C., Nov. 2005.

[11] Amaral, L.A.N., Scala, A., Barthelemy, M., and Stanley, H.E., "Classes of Small-world Networks," Proc. Natl. Acad. Sci., USA 97, 11149-1152, 2000.

[12] Bonnefoy, P.A. and Hansman, R.J., "Emergence and Impact of Secondary Airports in the United States," 4th AIAA Aviation Technology, Integration and Operations Conference, Sept. 2004, Chicago, IL.

[13] Sridhar, B., and Swei, S.M., "Relationship between Weather, Traffic and Delay Based on Empirical Methods," 6th AIAA Aviation Technology, Integration and Operations Conference, Sept. 2006, Wichita, KS.

[14] Huang, A.S., Schleicher, D. and Hunter, G., "Future Flight demand Generation Tool," 6th AIAA Aviation Technology, Integration and Operations Conference, Sept., 2006, Wichita, KS.