Analyzing Top 6 US Airlines’ Route Networks

Samet Ayhan, sayhan@umd.edu, 11/08/2012

Objectives

The objectives of this analysis are as follows:

- Rank the airports in the US (based on direct connectivity with other airports and how central role they play in connecting other airports together)
- Identify airline dominance on airports (based on number of departure/arrivals per airline)
- Determine hub airports (based on number of departure/arrivals) and most popular pair of airports (based on number of flights taking place) per airline

Dataset

Aircraft Situation Display to Industry (ASDI) [1] data is used for this project. ASDI feed is a product of the Enhanced Traffic Management System (ETMS). It is provided by Federal Aviation Administration (FAA) to provide flight plan and track information in addition to arrival and departure information from the National Airspace System (NAS) to airlines and other organizations in real-time and near real-time. Average amount of daily ASDI messages is over 7 million. Since the project aimed analyzing route networks of major airlines in the US, only arrival and departure information of ASDI feed was used.

Data Preparation

6 major airlines in the US [2] with their arrival and departure information were processed for the month of July, 2011. Filtered data attributes are as follows:

- Flight number (identifying airline)
- Departure airport and
- Arrival airport

In order to produce the necessary data, a number of steps were taken including database querying using SQL, cleaning and transformation using various scripts in Perl. Datasets were generated for the following airlines:

- Delta Airlines (DAL)
- United Airlines (UAL)
- Southwest Airlines (SWA)
- American Airlines (AAL)
- US Airways (AWE) and
- JetBlue Airways (JBU)
Number of Excel rows (edges) for the combined datasets exceeded 10000. Also, to prevent cluttering on the graphs, 4-letter ICAO codes were used for airports, instead of their full names. Full list of airport names along with their 4-letter ICAO codes are available [3]. For ease of reading, some major airports with their ICAO codes are listed at the end of the report.

Headlines

**Headline -1: No wonder why KATL (Atlanta) is the world’s busiest airport**

For this visualization, duplicate edges per airline were removed outside NodeXL. Then, Eigenvector Centrality was mapped to vertex size, and Betweenness Centrality was mapped to vertex opacity. Also, Degree was mapped to Vertex X and Betweenness Centrality is mapped to Vertex Y.

As clearly seen in the following graph, KATL (Atlanta) has the largest degree as well as highest Betweenness Centrality among all US airports. The other 6 airports, forming the top 7 in no specific order is as follows: KDFW (Dallas), KORD (Chicago O’Hare), KMSP (Minneapolis), KMCO (Orlando), KLAS (Las Vegas), and KPHX (Phoenix)

Figure 1. KATL is the most central airport with highest number of connections to other airports
Headline-2: The world’s busiest airport deserves the world’s largest airline

In the following graph, Fruchterman-Reingold layout was used. The graph illustrates the entire air traffic network in the continental US. Using Clauset-Newman-Moore algorithm, 7 groups were created and each group was displayed in its own box. Also, Degree was mapped to vertex size, and each airline was represented by a distinct color. Per the following graph, it is obvious that as the world’s largest airline, Delta [4] dominates the world’s busiest airport KATL (Atlanta). Additionally, Delta dominates KMSP (Minneapolis), American Airlines dominates KDFW (Dallas), Southwest dominates KMDW (Chicago Midway), and United dominates KORD (Chicago O’Hare). Rest of the airports is shared among all airlines with the fact that Delta, United and Southwest Airlines’ dominance is obvious. Please note that edges are tightly bundled.

![Graph](image.png)

**Figure 2.** KATL is dominated by Delta Airlines
**Headline-3: Airlines seem to have exclusivity on major hub airports**

The following set of network graphs illustrates major hub airports per airline. They also present most popular routes per airline. The following common settings were applied to all network graphs from here on:

Harel-Koren Fast Multiscale layout was used. A number of groups were created per graph using Clauset-Newman-Moore algorithm and each group was displayed in its own box. Also, edge weight was generated by counting duplicate edges, as the data set included all flights between airport pairs. Edge weight was mapped to edge width and opacity, Degree was mapped to vertex size, and Between Centrality was mapped to vertex opacity. Dynamic filtering, and filter opacity allowed discovering the following insight, overall:

- Airlines don’t seem to share hub airports
- A pair of major hub airports does not necessarily form the most popular route

Figure 3a illustrates that KATL (Atlanta) is Delta Airline’s major hub airport. In addition, KATL (Atlanta) – KLGA (New York LaGuardia) is the most popular route Delta Airline flies.

![Figure 3a. Delta Airline’s Routes](image-url)
Figure 3b illustrates that KORD (Chicago O’Hare) and KDEN (Denver) are United Airlines’ two major hub airports. In addition, KORD – KSFO (San Francisco) and KSFO – KLAX (Los Angeles) are the most popular two routes United Airlines flies.

Figure 3b. United Airlines’ Routes
Figure 3c illustrates that KMDW (Chicago Midway) and KPHX (Phoenix) are Southwest Airlines’ two major hub airports. In addition, KLAS (Las Vegas) – KLAX (Los Angeles) is the most popular route Southwest Airline flies.

Figure 3c. Southwest Airlines Routes
Figure 3d illustrates that KDFW (Dallas) is American Airlines’ major hub airport. In addition, KDFW – KORD (Chicago O’Hare) is the most popular route American Airlines flies.

Figure 3d. American Airlines Routes
Figure 3e illustrates that KJFK (New York JFK) and KBOS (Boston) are JetBlue Airways’ two major hub airports. In addition, KJFK – KMCO (Orlando) is the most popular route JetBlue Airways flies.

Figure 3e. JetBlue Airways Routes
Figure 3f illustrates that KCLT (Charlotte) and KPHX (Phoenix) are US Airways’ two major hub airports. In addition, KLGA (New York LaGuardia) – KBOS (Boston), KBOS – KDCA (DC National), KLGA – KDCA, KBOS – KPHL (Philadelphia) are the most popular routes US Airways flies.

![US Airways' Routes](image)

**Figure 3f. US Airways’ Routes**

**NodeXL Critique**

NodeXL offers a rich set of features for network analysis including flexible import and export capabilities, direct connections to social networks, and various layout options using a number of algorithms. It supports dynamic filtering, task automation, and vertex grouping. It also allows graph metric calculation, and overall lets users make better sense of network data. Its open-source template is free and can be extended.

Although with these rich set of features NodeXL is a powerful tool, there is still room for improvement. Below is a list of capabilities that might provide additional power if added:

- The tool is missing “Undo” capability, which is a very critical feature.
• The current clustering algorithms provided by NodeXL are not customizable. Users may want to make changes to the parameters, the algorithms use and re-run to experiment and discover new patterns.

• Location sensitive map background would add a tremendous value. This way, vertices with location information (such as latitude and longitude) could automatically be placed on a map.

• Cross-platform installation & execution: The current version, 1.0.1.226 is good only on Windows platforms. It would reach out to many more users if it were available on other platforms such as Linux, and Mac OS.

• COMException and InvalidCastException errors were received a few times during the tool execution, as shown below.
Appendix

Some major airports with their ICAO codes

- KATL - Atlanta Hartsfield International Airport
- KBOS – Logan International Airport
- KCLT – Charlotte Douglas International Airport
- KDCA – Ronald Reagan National Airport
- KDFW - Dallas Forth/Worth International Airport
- KLAS - Las Vegas McCarran International Airport
- KLAX - Los Angeles International Airport
- KORD - Chicago O'Hare International Airport
- KPHL – Philadelphia International Airport
- KPHX - Phoenix Sky Harbor International Airport
- KSFO – San Francisco International Airport