Interactive visualisation in time and space of flights in the United States

Arthur Aubret* Université Lyon 1
Caroline Gaud† Université Lyon 1
Brice Letcher‡ Université Lyon 1

1 Introduction

Air traffic data is dense, multidimensional, and constantly evolving. It it hard to understand using only tables. Data visualisation techniques can play a key role in understanding the main trends in the data.

While we had a personal preference for analysing French air traffic data, and in particular departures/arrivals at Lyon Saint-Exupéry, open access data on flights is remarkably hard to obtain. We speculate this is for mainly commercial and security reasons. We thus concentrate on the best data we found, which is from the US.

In the United States (US), there are almost 5000 public airports. Over 9.5 million flights were recorded in 2015 carrying more than 900 million passengers [4]. The volume of traffic is therefore considerable.

Our main objective in this project is uncovering, through data visualisation, interactions between the spatial and temporal dimensions of air traffic in the US. We focus in particular on exploring the links between air traffic volume/direction and time/seasonality. We can then look to answer questions such as: are flights directed towards more sunny regions during summer and towards more mountainous regions in winter? When in the week, and in the calendar year, do flight numbers peak? From which airports, and by which airlines?

Those kinds of questions can be useful for airports to efficiently predict flights throughtout the calendar year. Their concern lies in satisfying as many customers as possible, taking into account main airlines and directions while limiting the number of low-fill flights, and thus reducing cost.

The importance of building effective visualisations is clear in this context, as these questions require representing complex data in intuitive ways to produce understanding. The visualisations will be built using the D3 JavaScript library and presented on our GitHub web page.

In summary, we will highlight the evolution of flights in space (according to countries, continents), in time (depending on the season or school holidays) and according to the company.

2 Related Work

This section will describe the state of the art of flight visualizations in space, time and companies.

2.1 Graphs: geography and time

Since our main visualisation will be with a geographical map including a graph, considering airports as nodes and flight routes as edges, we focused our attention on work related of graphs. One well-known issue of graphs is the large number of edges, creating visual clutter.

Some work already proposed a solution, it may be very general [5], or specific to graphs. Force-Directed Edge [2] provides, in bundling edges with the action of physical forces whithout moving nodes, a way to eliminate the clutter. This work has been extended [3] in order to integrate the direction of the edges. This allows for separating each bundle into two parts. An other approach has been proposed [6], which consists in bringing together adjacent nodes in a hierarchial manner. Each edge will have a direction towards the center of the destined cluster.

EdgeLens [7] considers an interactive method instead of an automatic method. The user can decide which area he wants to focus on.

There is also work on temporality [1], which considers dynamic graphs and shows the evolution of a variable on an edge. One of the techniques uses an efficient image-based bundling method to create smoothly changing bundles.

Finally, many visualisations [8] and tools about air-traffic exist, but a significant fraction of them have different objectives from ours. For example, they are focused on accurately analyzing the trajectories of aircrafts.

2.2 Calendars and streamgraphs

Given we will develop the visualisations in D3, we found particular appeal for two types of visualisations presented by D3 founder Mike Bostock: the 'Calendar View' and the 'Streamgraph'.

An example of a 'Calendar View', and one developed especially for flight data, is presented in Figure 1. It is developed by Stefan Walther [9] and inspires us to build a similar visualisation adapted to our purposes.

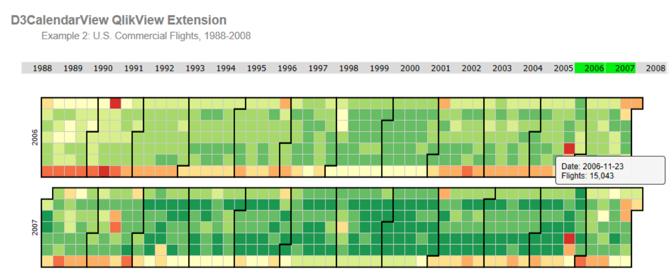

Figure 1: A calendar heatmap depicting flight traffic density on a day-to-day basis

The streamgraph visualisation is also interesting for us, the reasons for which are detailed in the following section.

3 Conception guidelines

3.1 Main visualisation

We will start with a geographical map showing the main American airports and the traffic between them, with a color or symbol encoding for amount of flights. We then wish to implement user interactivity allowing for some of the following:

1. Selecting a relevant timeframe, measured either in month of the year or day of the week, in which to restrict the data.

2. Superimposing traffic from different seasons: for instance, winter vs summer.

*e-mail: arthur.aubret@etu.univ-lyon1.fr
†e-mail: caroline.gaud@etu.univ-lyon1.fr
‡e-mail: brice.letcher@etu.univ-lyon1.fr

3. A 'summary wheel' giving the main directions of flights; this could help seeing what the differences are in time, and possibly why.

A search bar for airports and airline companies would be useful to implement for the user to get an understanding of the connectivity of the network. In addition, the possibility of selecting the top x most busy airports can give the main trends in the data.

We designed a hand-made sketch illustrating the main features we seek to represent. It is shown in Figure 2 below.

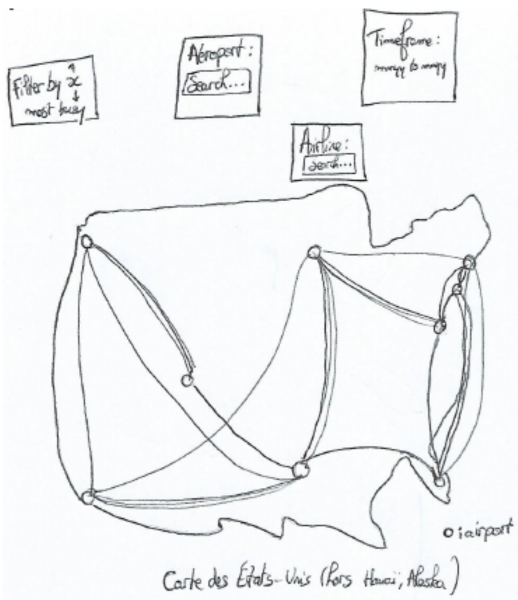

Figure 2: Sketch of a geographical map showing the main American airports and the traffic between them

The encoding of the amount of flights is crucial. We will probably experiment with at least:

- Symbol encoding: width of edges proportional to number of flights/passengers OR force-directed bundles of edges whose number will be proportional to traffic.

- Color encoding: because flight data is dense, the graph may be too cluttered, so that we could use a colour encoding on states or counties, portraying the volume of flights transiting through those regions.

a color or symbol encoding for amount of flights

We have two additional visual representations we are looking to implement. We consider building them because they are quite simple so not time-consuming, and because they could provide useful informational added-value.

3.2 Supporting visualisations

The first supporting visualisation is a calendar year heatmap. An example of one is provided in the Related works section in 2.2. It is an efficient representation of the main periods of air traffic intensity. It would also make it easy to contrast week-ends and weekdays, and also pinpoint holidays and how they affect air traffic.

The other supporting visualisation is the streamgraph. It will represent the volume of total air traffic contributed by the different airlines and/or airports. This information cannot easily be represented using a geographical map. Here as well, a temporal dimension can be added by selecting timeframes.

Naturally, realising these supporting visualisations is subject to time constraints and represent an initial work plan only. We prefer developing one precise and informative visualisation in the form of a geographical map as a priority.

REFERENCES

[1] S. I. F. T. R. K. A. C. T. Christophe Hurter, Ozan Ersoy. Bundled visualization of dynamic graph and trail data. *IEEE*, 2007.

[2] J. J. v. W. Danny Holten. Force-directed edge bundling for graph visualization. *Computer Graphics Forum*, 28:983–990, 2009.

[3] J. H. David Selassie, Brandon Heller. Divided edge bundling for directional network data. *IEEE*, 17:2354–2363, 2011.

[4] U. S. department of Transportation Office of Public Affairs. 2015 u.s.-based airline traffic data, 2015. https://www.rita.dot.gov/bts/press_releases/bts018_16.

[5] A. D. Geoffrey Ellis. A taxonomy of clutter reduction for information visualisation. *IEEE*, 13:1216–1223, 2007.

[6] D. Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE*, 12:741–748, 2006.

[7] S. G. Nelson Wong, Sheelag Carpendale. Edgelens:an interactive method for managing edge congestion in graphs. *IEEE*, pp. 51–58, 2003.

[8] M. D. Ronald Azuma, Howard Neely III and R. Geiss. Visualization tools for free flight air-traffic management. *IEEE*, 20:32–36, 2000.

[9] S. Walther. Calendar heatmap qlikview extension (d3calendarview), 2013. http://www.qlikblog.at/3070/calendar-heatmap-qlikview-extension-d3calendarview/.