Universitat
de les Illes Balears

Instituto de Física Interdisciplinar y Sistemas Complejos

DOCTORAL THESIS
2015

SYSTEMIC PROPAGATION OF DELAYS IN THE AIR-TRANSPORTATION NETWORK

Pablo Fleurquin

Universitat
de les Illes Balears

DOCTORAL THESIS
2015

Doctoral Program of Physics

SYSTEMIC PROPAGATION OF DELAYS IN THE AIR-TRANSPORTATION NETWORK

Pablo Fleurquin

Thesis Supervisor: José J. Ramasco Sukia
Thesis Co-Supervisor: Víctor Martinez Eguíluz
Thesis Rapporteur: Maxi San Miguel Ruibal

Doctor by the Universitat de les Illes Balears

Systemic propagation of delays in the air-transportation network

Pablo Fleurquin
Tesis presentada en el Departamento de Física de la Universitat de les Illes Balears

PhD Thesis

Directors: Dr. José Javier Ramasco Sukia, Dr. Víctor Martinez Eguíluz
Rapporteur: Maxi San Miguel Ruibal

Tesis doctoral presentada por Pablo Fleurquin para optar al título de Doctor, en el Programa de Física del Departamento de Física de la Universitat de les Illes Balears, realizada en el IFISC bajo la dirección de José Javier Ramasco Sukia, Víctor Martinez Eguíluz y como ponente Maxi San Miguel Ruibal.

Visto bueno
Directores de la tesis y ponente

Dr. José Javier Ramasco Sukia y
Dr. Víctor Martinez Eguíluz
Dr. Maxi San Miguel Ruibal.

Doctorando

Pablo Fleurquin

Palma, 19 de Diciembre de 2015

Whereof one cannot speak, thereof one must be silent.
Ludwig Wittgenstein

Resumen

La presente Tesis se encuentra enfocada en la descripción, análisis y modelaje de un Sistema Complejo pradigmático como lo es el sistema de transporte aéreo. La generación, propagación y eventual amplificación de los retrasos aéreos implican un gran número de mecanismos que interactúan. Tales mecanismos pueden ser clasificados como internos o externos al sistema de tráfico aéreo. Los principales mecanismos internos incluyen la rotaciones de aeronaves (los diferentes vuelos origen-destino que conforman un plan de vuelo diario), operaciones en aeropuertos, conexiones de los pasajeros y la rotación de la tripulación. Por otro lado factores externos tales como perturbaciones meteorológicas o amenazas a la seguridad, también modifican el normal funcionamiento del sistema y contribuyen a un alto nivel de congestión. En el presente sistema socio-técnico se suceden una multitud de decisiones humanas no siempre coordinadas con lo cual, dada la complejidad de las interacciones entre todos los agentes, hace necesario enfocar el problema desde la óptica de la teoría de los Sistemas Complejos. Especificamente por complejidad nos referimos a la aparición de un comportamiento colectivo como consecuencia de la interacción microscópica de los diferentes elementos que componen el sistema. Diversas herramientas han sido desarrolladas para hacer frente a sistemas de estas características. En este trabajo buscamos un enfoque a nivel de todo el sistema por medio de la teoría de redes complejas para lograr un entendimiento acabado de la propagación de retrasos aéreos.

Este trabajo se basa en la utilización de datos de tráfico reales, permitiéndonos un enfoque empírico, y de esta forma modelar, entender y comparar la dinámica de propagación de retrasos contra sucesos reales. A diferencia de otros procesos difusivos, tales como el modelado de enfermedas infecciosas, carecemos de las ecuaciones que rigen la dinámica del sistema. Por lo tanto, en esta Tesis se propone un modelo computacional basado en agentes que tiene valor no solo explicativo sino también predictivo. El enfoque seleccionado es interdiciplinar tomando elementos de diferentes campos como: Física, Ciencias de la Computación e Ingeniería de Transporte. La metodología seguida durante el doctorado es reflejada en la estructura de la Tesis: comienza con un análisis de los datos reales cuyo objetivo es explorar las principales características que hacen a la dinámica de los retrasos aéreos, a partir de las conclusiones de esta primera aproximación se desarrolla un modelo computacional para entender los mecanismos internos que generan dicha dinámica, dando paso finalmente al

estudio de la respuesta del sistema a perturbaciones externas y la cuantifación del grado de robustez que el mismo presenta.

La primera parte es una introducción de los principales conceptos teóricos y las herramientas utilizadas para caracterizar las redes espaciales con especial atención sobre los sistemas de transporte. Sobre el final se describen otros modelos de propagación de retraso aéreos.

La segunda parte comienza con una descripción de las principales características topológicas de la red mundial de aeropuertos (WAN), enfocando el análisis en la identificación de las comunidades topológicas en dicha red. El estudio de las comunidades es sumamente relevante para la comprensión de las propiedades estructurales de las redes y su posterior efecto sobre los procesos dinámicos que ocurren en ellas. Para ello realizamos un estudio comparativo de los resultados obtenidos al utilizar algoritmos de detección de naturaleza diversa. El principal aporte es el desarrollo de una metodología capaz de detectar las comunidades estadísticamente significativas y demostrar como las mismas tienen un mayor grado de similitud entre ellas en relación a comunidades no significativas. Los resultados obtenidos pueden servir como guia para la redefinición de las regiones supranacionales definidas por la IATA, pero en nuestro caso, de acuerdo a los patrones topológicos presentes en la WAN. Debido a la falta de datos que contengan la programación de los distintos vuelos comerciales a nivel mundial, concentramos nuestros esfuerzos en la red de aeropuertos de los Estados Unidos (USAN); en este caso los datos relativos a la programación y retrasos en los vuelos se encuentran disponible públicamente y sin costo. Es por esta razón que, disminuyendo la escala, pasamos a explorar las principales características topológicas agregadas de la USAN para luego centrarnos en las trayectorias de vuelo individuales y las características temporales del sistema de tráfico aéreo que se encuentran codificadas en la programación. El segundo capítulo de esta parte profundiza en el estudio de los retrasos aéreos definiendo métricas para evaluar el grado de congestión a nivel de todo el sistema. Por esta razón el foco de interés está puesto en la caracterización de los retrasos y cómo estos se transfieren y amplifican como resultado de las operaciones aereoportuarias, los llamados retrasos reaccionarios. Naturalmente los retrasos reaccionarios se propagan a través de la red, por lo que la comprensión de las características topológicas de la red de transporte aéreo, las propiedades estadísticas de los retrasos y las características de las rotaciones de las aeronaves son de gran importancia para el desarrollo del modelo.

La tercera parte está dedicada a la descripción del modelo basado en agentes para explorar con más detalle los aspectos dinámicos del sistema en estudio. Bajo el marco propuesto, los agentes son los aproximademente 5000 aviones de las principales aereolíneas comerciales, que componen el sistema de tráfico aéreo de los Estados Unidos cada día. Además, el modelo tiene valor predictivo ya que se encuentra basado en los datos reales de programación de las distintas aerolíneas. En esta parte, se describen los diferents objetos y subprocesos involucrados en la generación y eventual propagación de los retrasos aéreos. Como se mencionó

anteriormente estos son: la rotación de las aeronaves, la conectividad entre vuelos y la congestión aeropuertaria. Sobre el final se describen las diferentes maneras de inicializar el modelo.

La cuarta parte se centra en los mecanismos internos que impulsan las dinámicas de propagación de retrasos y muestra cómo el modelo es capaz de reproducir los patrones de propagación observado en los datos reales. Asimismo, puesto que el modelo es capaz de evaluar operaciones separadas, se identifica la conectivadad debido a los pasajeros en tránsito y a la rotación de la tripulación como el factor interno más relevante que contribuye con la propagación de los retrasos através del sistema. Además, los resultados indican que debido a las restricciones introducidas por la programación, existe un riesgo no despreciable de inestabilidad sistémica, incluso bajo condiciones de funcionamiento normales. En otras palabras, sin ningún tipo de perturbación externa importante que afecte al sistema, como por ejemplo condiciones climáticas extremas o huelgas de controladores aéreos, es posible alcanzar una congestión en todo el sistema debido a la complejidad de las interacciones entre los diferentes elementos que lo componen. Este análisis es capaz de proporcionar un marco de referencia para estudiar la estabilidad de los sistemas de transporte con horarios predefinidos por una programación. Su aplicación a otras redes de tráfico aéreo es, por supuesto, sencillo. Asimismo las métricas definidas para caracterizar la congestión son posibles de aplicar en otros ámbitos y, además, teniendo en cuenta las particularidades de otros sistemas es posible trasladarlo a otros medios de transporte con relativa facilidad.

Analizados las causas internas de la dinámica de retrasos, en la parte final exploramos los efectos de las perturbaciones externas al sistema. Comenzamos por cuantificar cómo las perturbaciones meteorológicas extremas generan los retrasos y motivan la intervención humana (mediante, por ejemplo, la cancelación de vuelos) para evitar el colapso del sistema. Para este análisis nos centramos en los hechos ocurridos el 27 de octubre del 2010 en los Estados Unidos. En esta fecha una tormenta extrema, que más tarde fue llamada “2010 Superstorm” asoló gran parte del centro y este de los Estados Unidos. Nuestro modelo es capaz de reproducir la evolución de la dinámica de los retrasos. A su vez, teniendo en cuenta las diferentes medidas de intervención, es posible mejorar los resultados del modelo para acercar la predicción cuantitativa a la realidad. Tomando como punto de partida este caso real en el siguiente capítulo analizamos la robustez del sistema en un marco general introduciendo métricas para cuantificar el impacto de las perturbaciones y la resiliencia del sistema de tráfico aéreo. Por último analizamos el grado de impacto de los aeropuertos en función de su tamaño e identificamos un patrón en la dinámica responsable de la propagación de los retrasos como un efecto dominó.

La Tesis termina con la conclusión y una serie de propuestas de investigación que continuan la linea de la misma.

Los objetivos de la presente Tesis se encuentran enmarcados por el alcance del programa de doctorando de ComplexWorld, un proyecto dentro del SESAR

WP-E cuyo objetivo es la promoción de la investigación en el área de tráfico aéreo e Ingeniería de Transporte. El doctorando ha publicado los siguentes artículos y una patente Española (asimismo una patente Europea esta pendiente de revisión), este material es utilizado en la presente Tesis:

- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2013). Systemic delay propagation in the US airport network. *Scientific reports*, 3.
- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M (2013). Data-driven modeling of systemic delay propagation under severe meteorological conditions. *In 10th FAA/Eurocontrol ATM Seminar, Chicago, USA*.
- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2014). Characterization of Delay Propagation in the US Air-Transportation Network. *Transportation Journal*, 53(3), 330-344.
- Fleurquin, P., Campanelli, B., Eguíluz, V. M., & Ramasco, J. J. (2014). Trees of Reactionary Delay: Addressing the Dynamical Robustness of the US Air Transportation Network. *transportation*, 11, 12.
- Campanelli, B., Fleurquin, P., Eguíluz, V. M., Ramasco, J. J., Arranz, A., Extebarria, I., & Ciruelos, C. (2014). Modeling Reactionary Delays in the European Air Transport Network. *In the Fourth SESAR Innovation Days, Madrid, Spain*.
- Campanelli, B., Lenormand, M., Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2014). Movilidad y transporte: un viaje a través del espacio, de la ciudad al mundo. *Revista Española de Física*, 28(3), 37-41.
- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. Spanish Patent ES2476566 (2015), Método para caracterizar la congestión aeroportuaria en una red de tráfico aéreo.

Abstract

The focus of this dissertation is to quantitative describe, analyze and model a paradigmatic socio-technical complex system such as the air-transportation system. The generation, propagation and eventual amplification of flight delays involve a large number of interacting mechanisms. Such mechanisms can be classified as internal or external to the air traffic system. The basic internal mechanisms include aircraft rotations (the different flight legs that comprise an aircraft itinerary), airport operations, passengers' connections and crew rotation. In addition, external factors, such as weather perturbations or security threats, disturb the system performance and contribute to a high level of system-wide congestion. Although this socio-technical system is driven by human decisions, the intricacy of the interactions between all these elements calls for an analysis of flight delays under the scope of Complex Systems theory. Complexity is concerned with the emergence of collective behavior from the microscopic interaction of the system elements. Several tools have been developed to tackle complexity. Here we use Complex Networks theory and take a system-wide perspective to broaden the understanding of delay propagation.

This work is driven by real traffic data, allowing us to take an empirical approach to the problem and model delay propagation dynamics against realistic situations. Unlike other spreading processes such as infectious disease modeling, we lack the mathematical equations that rule the dynamics. Therefore we propose an agent-based computational framework that has either explanatory and predictive value. The path chosen makes this thesis truly interdisciplinary tying fields as Physics, Computer Science and Transportation Engineering. The research methodology followed during this PhD is reflected in the thesis structure: it begins with a exploratory data analysis to identify the key drivers behind the dynamics of delay propagation, and then uses the output of this initial phase to develop a computational model that was later employed to understand the problem at hand focusing in the endogenous mechanisms that foster the spreading of flight delays. Finally, the last part explores the system response to weather perturbations and give insights regarding the system robustness to perturbations.

The first part provides an introduction of the main theoretical concepts and tools used to characterize spatial networks focusing on transportation systems, and describes other modeling approaches of flight delay propagation.

The second part starts with a description of the main topological characteristics of the World Airport Network with special attention to the identification of communities in this network. Community detection is of great relevance for the understanding of the structural properties of networks, and here we compare the output of widely used community detection algorithms and identify statistically significant and similar communities across methods. The results provide insights to define supranational regions according to topological patterns in the World Airport Network. Due to the lack of performance data for world-wide traffic we concentrate our efforts on the US Airport Network, for which data regarding flight delays is publicly available without cost. We therefore explore the main topological characteristics of the aggregated network and next we focus on the individual flight trajectories and temporal features of the system encoded in the schedule. The second chapter of this part deepens in the understanding of flight delays and defines metrics to asses the level of system-wide congestion. In the second part of the thesis the interest is put in characterizing delays and how they are transferred and amplified by subsequent operations, the so-called reactionary delays. Naturally reactionary delays spread across the network, so an understanding of the topological features of the air transportation network, the properties of aircraft rotations and the statistical features of flight delays is of great significance for subsequent modeling efforts.

The third part is devoted to the description of the data-driven agent-based model developed to further explore the dynamical aspects of the system under study. Under the proposed framework, the agents are the approximately 5 thousand aircrafts, from the main commercial airlines, that are part of the US air-traffic system every day. In addition it is data-driven because it uses the real airline schedules, which is why it has predictive value. In this part, we describe the object models and the subprocesses involved in the cascading mechanism of delay propagation. As mentioned before these are: aircraft rotation, flight connectivity and airport congestion. We also describe the different possible initial conditions for the simulations.

The fourth part focuses on the internal mechanisms that drive the flight delay dynamics and shows how the model reproduces the delay propagation patterns observed in the US performance data. Also since the model is able to evaluate separate operations, it identifies passenger and crew connectivity as the most relevant internal factor contributing to delay spreading. Furthermore, the results indicate that due to the constraints introduced by the existence of a schedule, there is a non-negligible risk of systemic instability even under normal operating conditions. In other words, without any major external perturbation affecting a large part of the system, such as extreme weather conditions or air controller strikes, the dynamics can move towards a network-wide congestion due to the intricacy of the interactions between the different elements of the system. Our analysis provides insights on how to study the performance and stability of networked transport systems with predefined schedules. Its translation to other airport networks is, of course, straightforward, and even though the modeling

of other transportation systems may require some particular details, the applicability of the metrics defined to measure network-wide congestion based on clustering is universal.

The final part explores instead the consequences of external perturbations to the system. We begin by quantifying how system-wide weather perturbations affect delay propagation and the effectiveness of managers and pilots interventions to prevent possible large-scale system failures. We focus on the events occurring on October 27 2010 in the United States. A major storm complex that was later called the 2010 Superstorm took place that day. Our model reproduces the evolution of the delay-spreading dynamics. By considering different intervention measures, we can even improve the model predictions getting closer to the real delay data. Next we generalize the problem of the system resilience to perturbations and introduce metrics to quantify the impact of perturbations and the robustness of the air-traffic system. We also provide insights depending on the airport size and identify the dynamical patterns that boost delay propagation.

This dissertation ends with the discussion and proposes future research lines.

The PhD objectives were formulated under the scope of the ComplexWorld PhD programme a SESAR WP-E project, which aims to promote research and Complex Science in aerospace and air-traffic management. During the PhD the candidate has published the following articles and also a Spanish patent (as well as a European patent with pending status), whose material is used in this dissertation:

- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2013). Systemic delay propagation in the US airport network. *Scientific reports*, 3.
- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M (2013). Data-driven modeling of systemic delay propagation under severe meteorological conditions. *In 10th FAA/Eurocontrol ATM Seminar, Chicago, USA*.
- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2014). Characterization of Delay Propagation in the US Air-Transportation Network. *Transportation Journal*, 53(3), 330-344.
- Fleurquin, P., Campanelli, B., Eguíluz, V. M., & Ramasco, J. J. (2014). Trees of Reactionary Delay: Addressing the Dynamical Robustness of the US Air Transportation Network. *transportation*, 11, 12.
- Campanelli, B., Fleurquin, P., Eguíluz, V. M., Ramasco, J. J., Arranz, A., Extebarria, I., & Ciruelos, C. (2014). Modeling Reactionary Delays in the European Air Transport Network. *In the Fourth SESAR Innovation Days, Madrid, Spain*.
- Campanelli, B., Lenormand, M., Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. (2014). Movilidad y transporte: un viaje a través del espacio, de la ciudad al mundo. *Revista Española de Física*, 28(3), 37-41.

- Fleurquin, P., Ramasco, J. J., & Eguíluz, V. M. Spanish Patent ES2476566 (2015), Método para caracterizar la congestión aeroportuaria en una red de tráfico aéreo.

Acknowledgements

En primer lugar quisiera agradecer a Maxi San Miguel por brindarme la oportunidad de estar escribiendo estas líneas. A José Ramasco y a Víctor M. Eguiluz por sus constantes enseñanzas.

A todos en el IFISC, en particular a mis compañeros del 07 por ser parte de este camino que juntos recorremos. A Ricardo que, aún siendo canario, me introdujo en las mejores costumbres mallorquinas. A Ángel por su apoyo a un foraster en su primer año y que pese a la distancia me hizo sentir como en casa. A Luis, Daniel y Xavi por como disfrutan los asados y así recordarme constantemente de mis raices rioplatenses. A mis compañeros de cubículo: a Adrián por nuestras enriquecedoras charlas sobre actualidad, a Víctor por su alegría caribeña y por que no a Enrico que siempre fue parte de nuestro cubículo aunque no fisicamente. Tampoco puedo dejar de agradecer al resto, en especial a unos 'recien' llegados como Julián y Peri, por crear un ambiente que dan ganas de ser parte. A Rubén y Antonia por hacernos la vida más fácil.

También quiero agradecer a mis padres, por su apoyo incondicional y ejemplo de vida que me ha guiado durante estos 30 años. A mis hermanos, por los momentos compartídos además de su cariño y sostén.

Por último, a mi mujer Josefina que sin ella todo esto no hubiera sido posible. Gracias por tu comprensión, confianza, por haber dejado todo y cruzar el charco. Gracias por cada día y por lo más importante a día de hoy: Alexa.

Contents

Part I

Introduction

CHAPTER 1

Theoretical framework: Complexity Science in transportation systems

The Science of Complexity has undergone a rapid expansion through a variety of interdisciplinary fields such as Computer Science, Technology, Biology and Socio-economical Systems. Complexity Science is a recent approach to comprehend a certain type of systems, the so-called Complex Systems, composed of a large number of interacting entities that produce emergent behaviors. Given this general framework, Complexity is not used to refer just to complicated phenomena within Science, it emphasizes the notion of collective behavior that surge from microscopic interaction scales.

Statistical Mechanics has provided a mathematical toolbox to analyze Complex Systems. Thus, modeling approaches from physics are combined with powerful computational resources to gain insight on fundamental questions poorly suited to traditional science. This is the case in systems, where a reductionist approach fails to explain the emergent behavior, thus following a Complex System's approach can shed light on it.

1.1 Tackling complexity: networks

One of the keystones of Complexity Science is Complex Networks. A network is a mathematical abstraction that represents systems of interacting entities. Naturally, entities are symbolized by nodes and the interactions by edges or connections between them. For instance, the World-Wide Web is a network of nodes representing web-pages and edges hyper-links between them. The use of network analysis to characterize Complex Systems has been generalized in the last decade. The potential of graphs describing social systems was pointed out almost a century ago (Moreno et al., 1934) and network theory was guided by the early work on random graphs by Erdós and Rényi (ERDdS and R&WI, 1959). These are networks with a majority of vertices that have roughly the same number of connections around the average value. However this modeling

paradigm switched at the end of the 90's after the seminal works by Watts and Strogatz (Watts and Strogatz, 1998) and by Albert and Barabasi (Barabási and Albert, 1999). In these works they found that real networks behave differently from the traditional approach, showing non-trivial topological characteristics. Ever since Complex Networks have been applied in a growing range of disciplines such as Technology (Huberman et al., 1998), Biology (Jeong et al., 2001), Social Systems (Castellano et al., 2009) or Economy (Snyder and Kick, 1979).

It should come as no surprise, though, that the topology of networks plays an important role on the dynamics taking place on them (Newman, 2010). Therefore the study of networks is of paramount importance to understand the behavior of, interconnected, real systems. A successful example is in the context of Social Science and the mechanisms behind social influence and information diffusion (Wasserman and Faust, 1994; Conte et al., 2012; Ugander et al., 2012; Christakis and Fowler, 2010). Other field where Network Science has been of fundamental importance is in the modeling of propagation of infectious diseases at a global scale that occurs when infected persons travel across the air-traffic network (Hufnagel et al., 2004; Colizza et al., 2006a; Balcan et al., 2009a,b; Tizzoni et al., 2012). The modeling concepts of epidemic spreading have a long tradition going back almost a century (M'Kendrick, 1925). One example is the SIR model, originally developed to understand epidemic spreading. In the SIR, a set of individuals are initially infected and for each time step they can randomly interact with each other so that a disease begins to propagate through the system. If the contacted neighbor is in the susceptible state (different from infected), then with probability λ will be infected. At the same time, the infection has a finite lifetime so there is a certain probability μ of an individual to recover and become immune.

In recent years these modeling efforts evolved to include real data. Cutting-edge technology has increased the availability of digital data concerning human activity patterns offering an opportunity to model complex socio-technical systems. Datasets capturing mobile phone calls, friends and followers within social networks, web-browsing, e-mail or transport ticketing have opened the path to understand systems at a realistic level. The amount of digital data created, replicated and consumed doubles every 2 years and it is estimated that for 2020 the size of the *digital universe* will be about 40 trillion of gigabytes. In other words, approximately 5 thousand gigabytes per person (Gantz and Reinsel, 2012). This gave rise to the term *Big Data*, a far-reaching concept that has an increasing influence in diverse areas such as new enterprise managerial systems (McAfee and Brynjolfsson, 2012), politics (Bond et al., 2012) or science (Provost and Fawcett, 2013). However data per se is pointless, one should look for novel ways to make use of it. In this sense, data-driven computational models have the potential to offer a predictive approach as well (Vespignani, 2012). As mentioned before, the modeling and forecasting of disease spreading patterns using air traffic data is a story of a notable success. One can, thus, wonder if this success can be extended to the propagation of other phenomena. In particular, we are interested in con-

sidering this conceptual framework as a stepping stone for the modeling of flight delay propagation and the way in which congestion can become a systemic risk.

1.1.1 The structure of real Complex Networks: definitions and topology

To understand the structure of networks and the dynamics happening on them, different concepts were developed that describe the main features of real complex networks. The set of metrics presented here is not exhaustive but rather the most relevant for the conceptual framework of this thesis; that is, spatial networks.

Adjacency matrix & degree distribution

Among all, the most salient is the degree distribution of a network. Besides lattices, nodes in a network may have different number of connections, the so-called degree of a node (k). By representing an undirected network with M edges and N nodes as a $N \times N$ adjacency matrix A:

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected;} \\ 0 & \text{otherwise.} \end{cases}$$

the degree of a node i is the number of neighbors and so it is defined as:

$$k_i = \sum_j A_{ij} \tag{1.1}$$

Therefore the degree distribution of a network $P(k)$, is a mathematical function accounting for the probability that a node selected randomly has exactly k edges. As pointed out before, in a Erdos-Renyi graph, most nodes will have a degree close to the average degree of the network ($\langle k \rangle$). In this case the degree distribution is a Binomial centered around the average degree, $\langle k \rangle$. In contrast, most real (complex) networks show heavy-tailed degree distribution, where some nodes, called hubs, have a number of edges orders of magnitude larger than the average network degree.

Assortativity measures

In the case of uncorrelated networks $P(k)$ fully determines their statistical properties, however this is not the case for most real complex networks as the degrees of two linked nodes are, in general, correlated. To account for this degree-degree dependency it is necessary to account for the conditional probability $P(k'|k)$ that an edge starting at a node with degree k points to a node of degree k'. Nevertheless the direct evaluation of this quantity is not easy, therefore one can overcome

this problem by defining the *average neighbors degree* of a node i as:

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \in N(i)} k_j \tag{1.2}$$

where $N(i)$ are the neighbors of node i and k_j is the degree of neighbor j. Also one can implicitly incorporate the dependency on k by calculating the average neighbors degree of nodes with degree k or *assortativity function* as:

$$k_{nn}(k) = \frac{1}{N(k)} \sum_{i/k_i=k} k_{nn}(i) \tag{1.3}$$

where the sum is over nodes i given that their degree is equal to k. The above definition can be expressed in terms of the conditional probability $P(k'|k)$ as in (Pastor-Satorras et al., 2001):

$$k_{nn}(k) = \sum_{k'} P(k'|k)k' \tag{1.4}$$

If $k_{nn}(k)$ is a decreasing function of k it implies that nodes with high degree values (hubs) tend to connect with nodes with relatively low degree (*disassortative* network). On the other hand, a positive correlation with k suggests that nodes have a tendency to connect with similar nodes regarding the degree (*assortative* network). Although the degree correlation present in the network is characterized by the behavior of $k_{nn}(k)$ it is sometimes useful, specially when the assortativity function gives noisy results, to compute the assortative mixing coefficient r of a network as defined in (Newman, 2002):

$$r = \frac{M^{-1} \sum_m j_m k_m - [M^{-1} \sum_m \frac{1}{2}(j_m + k_m)]^2}{M^{-1} \sum_m \frac{1}{2}(j_m^2 + k_m^2) - [M^{-1} \sum_m \frac{1}{2}(j_m + k_m)]^2} \tag{1.5}$$

where j_m, k_m are the degrees of the nodes at the ends of the mth edge, with $m = 1 \cdots M$. By these means, the r coefficient lies in the range $-1 \leq r \leq 1$ going from purely dissortative ($r = -1$) to assortative ($r = 1$) topological structure. Nevertheless, as we shall see in transportation networks the spatial constraints usually implies a flat assortativity function (or r close to 0).

It could happen though that degree-degree correlations may be masked by averaging over its neighbors. Another way to discern more subtle effects regarding these correlations is by means of the *rich-club coefficient*. The rich-club effect expresses the propensity between high degree nodes to be very well connected among each other. In social networks, this would suggest a kind of closed club behavior among highly connected individuals in contrast to a network structure comprised by groups of individuals with mixed connectivity features. A quantitative definition of the rich-club effect is given by the rich-club coefficient (Zhou and Mondragón, 2004):

$$\phi(k) = \frac{M_{i/k_i>k}}{\frac{1}{2}N_{i/k_i>k}(N_{i/k_i>k} - 1)} \tag{1.6}$$

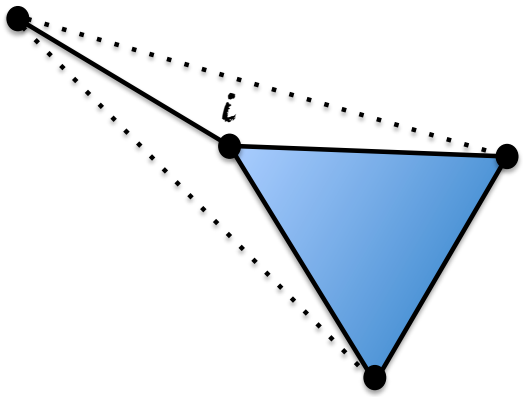


Figure 1.1: The clustering coefficient of node i is 1/3.

where $M_{i/k_i>k}$ is the number of edges among nodes i with degree higher than k ($N_{i/k_i>k}$) over the total possible number of edges within these nodes. As shown in (Colizza et al., 2006b) $\phi(k)$ uses a different projection of the conditional probability $P(k'|k)$ from the one used by r or $k_{nn}(k)$. Indeed, it is not trivially related to the above assortativity measures as it can be shown that dissortative networks can have rich-club effects.

Clustering coefficient

Another relevant parameter to characterize networks is the *clustering coefficient* C (Watts and Strogatz, 1998). It is a topological property that accounts for the degree of interconnectivity in the neighborhood of a node. In a social context, it would express how likely is that two individuals with a common friend know each other. For a node i of degree k_i the clustering coefficient is given by:

$$C(i) = \frac{M_i}{\frac{1}{2}k_i(k_i - 1)} \tag{1.7}$$

with M_i being the number of edges among the neighbors of i. It is useful to visualize it as the fraction of the actual number of triangles over the total possible number of triangles a node belongs to. In Figure 1.1 the clustering coefficient of node i is 1/3 since it belongs to only one triangle (solid one) of the three possible ones (dashed ones). Similarly, with the assortativity function it is more convenient to express C as a function of the node degree:

$$C(k) = \frac{1}{N(k)} \sum_{i/k_i=k} C(i) \tag{1.8}$$

Average shortest path

In particular for transportation systems it is useful to have a topological measure that can be related to the transportation efficiency of the network (see Section 1.2) and its characteristic size. One simple way to achieve this is by means of the average shortest path:

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} l(i,j) \tag{1.9}$$

being $l(i,j)$ the minimum number of edges between nodes i and j. For a small-word network the average shortest path scales as:

$$\langle l \rangle \sim \log N \tag{1.10}$$

On the other hand, d-dimensional regular lattices evince 'large-world' behavior scaling as:

$$\langle l \rangle \sim N^{1/d} \tag{1.11}$$

Weighted & directed networks

Importantly in many real complex networks not all links have the same weight. In fact, in some cases associating a value different from 1 may provide useful information concerning the structure and the dynamics of the system. In this case, the adjacency matrix of a weighted network is defined as:

$$A_{ij} = \begin{cases} w_{ij} & \text{if } i \text{ and } j \text{ are connected;} \\ 0 & \text{otherwise.} \end{cases}$$

where w_{ij} is a non-zero value that accounts for the weight of the link connecting nodes i and j. Weights are associated to a specific function of pair-wise connections and depend on a relevant property of the system under study. For instance, in (Barrat et al., 2004) weights refer to the number of available seats on a given connection for the World Airport Network (WAN) or the intensity of collaboration between authors in the Scientific Collaboration Network (SCN). On a biological context, weights can be associated to the contact energy between residues in amino acid networks (Jiao et al., 2007) or level of carbon flow between species in food webs (Luczkovich et al., 2003). Under this schema one can define the *strength* of a node in similar fashion to the degree of a node for unweighted networks:

$$s_i = \sum_j w_{ij} \tag{1.12}$$

and redefine the previous topological measures to take into consideration the available information concerning the role of edges beyond simple binary interaction.

In some systems, considering the link directionality may also provide useful information about the structure of the network that could be masked in the corresponding undirected graph representation (Bianconi et al., 2008). In the simple undirected case the adjacency matrix is symmetric therefore the degree is either defined as:

$$k_i = \sum_j A_{ij} \tag{1.13}$$

or:

$$k_i = \sum_j A_{ji} \tag{1.14}$$

whereas in its directed counterpart one has to distinguish between the in-degree k_i^{in} and the out-degree k_i^{out}:

$$k_i^{in} = \sum_j A_{ji} \quad k_i^{out} = \sum_j A_{ij} \tag{1.15}$$

In this thesis, if not otherwise stated, we will consider the unweighted air transportation network symmetric.

Community detection: exploring the mesoscale structure

Finally, an important topic in complex network studies is Community detection. The identification of communities is of great relevance for the understanding of the structural properties of networks and the emergence of macro-level collective behavior in complex systems (Fortunato, 2010). Despite the stochastic nature of their generation mechanisms, many real complex networks show a remarkable degree of organization at the mesoscopic level. Protein modules in biological networks (Hartwell et al., 1999; Jonsson et al., 2006), clusters based on geopolitical constrains in infrastructure networks (Guimera et al., 2005), scientific collaboration in coauthorship networks (Newman, 2004a) and group of friends or acquaintances in social networks (Ferrara, 2012; Grabowicz et al., 2012) are all examples of community structure in large-scale complex networks.

Intuitively a community is a tightly knit group of nodes with comparatively few edges joining nodes of different communities. Also this modular structure can be organized in different hierarchical levels as shown in Figure 1.2. The first methods to address the problem of community detection were based on graph partitioning and hierarchical clustering techniques (Kernighan and Lin, 1970; Scott and Carrington, 2011). In the last 10 years the field has been rejuvenated with several different methods. In a seminal paper published in 2002, Girvan and Newman (Girvan and Newman, 2002) fueled the field by proposing the concept of edge betweenness to detect and remove edges that connect nodes of different modules. A more recent approach relies on computing a global function, which gives an estimate of the quality of a network partition. The most popular algorithms of this kind are those employing modularity maximization

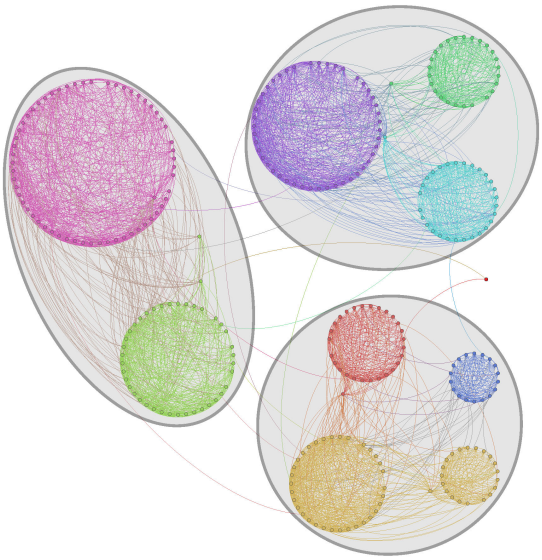


Figure 1.2: Example of modular structure with two hierarchical levels: in grey and a lower level differentiating smaller modules (color modules). From (Lancichinetti et al., 2011).

such as simulating annealing (Guimera et al., 2004), extremal optimization (Duch and Arenas, 2005) and greedy (Newman, 2004b) techniques. Information theory measures have also been used to explore community structure. A commonly used algorithm built upon this framework is INFOMAP (Rosvall and Bergstrom, 2007). Another class of methods are those employing clique-based techniques such as the Clique Percolation Method (Palla et al., 2005), which has the interesting feature of finding overlapping communities. A recent method also capable of finding overlapping communities is OSLOM (Lancichinetti et al., 2011). This method is based on the local optimization of a fitness function. Two unique characteristics of OSLOM are that it is capable of computing the statistical significance of each community and detect nodes that does not belong to any community (homeless nodes).

A recurring problem faced by researchers when comparing different community detection algorithms is how to measure their efficiency. Normalized mutual information (Strehl and Ghosh, 2003) since the work by (Danon et al., 2005) has been used by several authors to measure the performance of different algorithms in benchmarks (Fortunato, 2010; Lancichinetti et al., 2008, 2011; Aldecoa and Marín, 2013), where the underlying partitioning of the network is known. The mutual information $I(X, Y)$ measures the information shared between the original partition X and the one obtained by the algorithm, Y. Since this quantity has an upper bound given by the minimum entropy H between both partition, i.e., $I(X, Y) \leq \min(H(X), H(Y))$, a common normalization factor is the arithmetic average of the entropies:

$$NMI(X, Y) = \frac{2I(X, Y)}{H(X) + H(Y)}. \tag{1.16}$$

In a network with n nodes, let $C^{(a)}$ and $C^{(b)}$ be the set of clusters in partitions A and B, respectively. Define $n_i^{(a)}$, $n_i^{(b)}$ as the number of nodes in a specific cluster $i \in C^{(a)}$, $C^{(b)}$, and n_{ij} as the number of nodes from cluster $i \in C^{(a)}$ found in cluster $j \in C^{(b)}$. Finally, assuming that each partition has a total of $c^{(a)}$ and $c^{(b)}$ clusters, the *NMI* can be obtained as

$$NMI(A,B) = \frac{-2\sum_i^{c^{(a)}}\sum_j^{c^{(b)}} n_{ij}\log\left(\frac{n\cdot n_{ij}}{n_i^{(a)}\cdot n_j^{(b)}}\right)}{\sum_i^{c^{(a)}} n_i^{(a)}\log\left(\frac{n_i^{(a)}}{n}\right)+\sum_i^{c^{(b)}} n_i^{(b)}\log\left(\frac{n_i^{(b)}}{n}\right)}. \tag{1.17}$$

This quantity is 1 when A and B are identical and should tend to 0 the more dissimilar the two partitioning are.

Another way to measure the degree of similarity between groups of nodes is by means of a confusion matrix where we calculate the Jaccard index between every pair of communities between different algorithms. In this matrix, each row represents a community $C_i^{(a)}$ from algorithm A, while columns identify communities $C_j^{(b)}$ from method B. The entries are given by the corresponding Jaccard index, that is

$$M_{ij} = J\left(C_i^{(a)}, C_j^{(b)}\right). \tag{1.18}$$

By ordering the columns and rows of the matrix according to the Jaccard index, it is possible to visually identify how well represented the communities of one method are in the other. Identical partitioning would lead to the identity matrix, that is, with value one on the diagonal and zero elsewhere. Two methods with good agreement with each other should have a well defined, single peak of value close to 1 in each column and row of the confusion matrix, showing a strong one-to-one relationship between their communities. This matrix also have the advantage of offering the possibility to easily access the *recall score* (Hric et al., 2014) of the communities of one method with respect to the other, which is given by the maximum Jaccard index:

$$R(C_i^{(a)}) = \max_j\left(M_{ij}\right), \tag{1.19}$$

$$R(C_j^{(b)}) = \max_i\left(M_{ij}\right). \tag{1.20}$$

This quantity provides a measure of how well community i (j) from method A (B) is reproduced in method B (A). A perfect match will give a recall score of 1.

1.2

Transportation systems as Spatial Networks

Infrastructure systems such as transportation, Internet, mobile phone networks, power grids are all examples of networks which are characterized by a spatial embedding of its constitutive elements. The inclusion of spatial aspects

in these networks strongly affects the topological characteristics mentioned in Section 1.1.1 by associating a cost to the length of edges (Barthélemy, 2011). In particular, if there is an increasing cost related to distance, longer connections must be rewarded by some competitive advantage such as being linked to a network hub. For instance, this effect is clear in the air transportation network where is not common to have intercontinental flights between two regional airports, instead long haul flights connect large degree nodes.

This particular feature generates a trade-off between *cost* and *efficiency* driving decisions in the conception and development of infrastructure systems. Therefore systems of this kind evolve by balanced cost/efficiency optimization processes taken by numerous stake-holders in a short time scale (for instance, airlines constantly modify their schedule due to economical reasons). However, this does not mean, specially in the long term, a central planning process. On the contrary, optimization at the micro-level can produce suboptimal large scale systems and unexpected collective emergent behavior. However this optimization trait is crucial for understanding the topological structure of transportation systems (Banavar et al., 1999; Gastner and Newman, 2006; Helbing, 2001) as will be shown in Chapter 6 for the WAN.

In this regard, assuming that the cost is proportional to the length of the links in the system, it is therefore important to know which is the minimum-cost structure for a given set of nodes. For a set of N nodes the minimum number of edges required to connect the nodes is $E = N - 1$, thus producing a tree-like structure. Hence from all possible trees we should look for the minimum-cost structure by minimizing the total tree length:

$$l_T = \sum_{e \in M} d_M(e) \tag{1.21}$$

where $d_M(e)$ is the length of edge e. Following these arguments will lead us to the most economical structure given by the minimum spanning tree (MST). However, the MST is neither efficient in terms of transportation due to average longer paths nor resilient to failure because low connectivity. This kind of structures are very susceptible to errors/attacks as the probability of disconnecting a part of the network with the failure of a single edge is very high, and also it would produce large travel times between different locations. Nevertheless, it is interesting to evaluate the interplay between cost and efficiency by using the MST as a null model of the actual network. For instance, the relative cost of a network with respect to the minimum-cost structure is then given by:

$$Cost = \frac{l_T}{l_T^{MST}} \tag{1.22}$$

In the same way the transport efficiency or *transport performance* P can be defined as the ratio between the average shortest path of the network compared to its MST null model:

$$P = \frac{\langle l \rangle}{\langle l_{MST} \rangle} \tag{1.23}$$

One final distinction among spatial networks is the way space is embedded in the system. This is particular relevant in transportation networks, due to the fact that most networks can be considered up to a good extent as planar (although some intersections are unavoidable) except for the air transportation system. A planar network is defined as a graph that can be drawn in a 2-dimensional space in such a way that there is no edge intersection. The planar embedding produces some unique topological characteristics such as bounded average degree $< k > \leq 6$ and consequently sparse networks.

1.2.1 Modes of Transport

Land transportation

Road, rail and subway networks are all examples of transport by land systems which can be categorize to a good extent under planar graphs (Figure 1.3). The typical network representation of an urban road system consists of edges representing streets and end points and streets' intersections by nodes (Cardillo et al., 2006; Buhl et al., 2006; Batty, 2007; Lämmer et al., 2006). In the case of rail systems edges indicate rail-lines between nodes or train stations (Sen et al., 2003; Kurant and Thiran, 2006) and subway networks are represented in a similar way (Latora and Marchiori, 2002; Sienkiewicz and Hołyst, 2005; Von Ferber et al., 2009). These networks share some topological features that are very different to other transport networks, specially in the case of air transportation. As mentioned before the approximately planar embedding of such networks imposes constraints on the level of connectivity of the nodes and therefore on the degree distribution. Generally, transport by land systems show a distribution $P(k)$ peaked around the average degree value that it is always lower than 6. In most cases, the degree distribution it is shown to have an exponential form and a flat assortativity function $k_{nn}(k)$, implying a lack of degree-degree correlations. Furthermore, the average shortest path scales as $N^{1/2}$ a typical behavior of a 2-dimensional lattice.

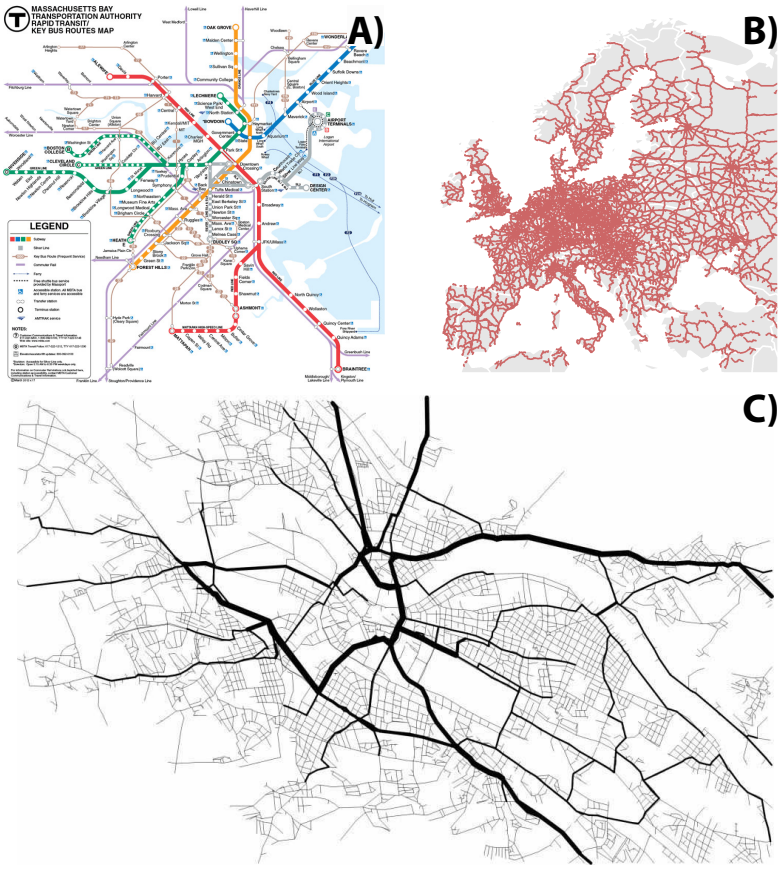


Figure 1.3: Examples of land transportation networks. A) Boston subway network (Latora and Marchiori, 2002), B) European railway network (Lenormand et al., 2014) and C) Dresden road network (Lämmer et al., 2006)

Maritime transportation

A critical infrastructure network for international trade is the Maritime Transportation Network (MTN). Shipping is the cornerstone of the global economy as it moves over 90% of the world trade. The world shipping fleet can be grouped depending on the type of cargo, each of which produces a different network structure where nodes are ports and edges direct routes between them (Kaluza et al., 2010). Broadly speaking, shipping is divided into containerized, bulk dry and crude/oil cargo. (Hu and Zhu, 2009) studied the structure of the MTN considering only containerized cargo. The network used is composed of 878 nodes (sea ports) and 7955 edges (direct routes) bearing a weight proportional to the number of container lines in a time period. The authors found that the network has a relatively small average shortest path of approximately 3.6 and a clustering coefficient of 0.40 following a truncated power-law behavior for k^{in} and k^{out}.

As noted in (Kaluza et al., 2010) one of the peculiarities of the MTN network within transportation systems is the heavy asymmetry of its links due to the the abundance of circular path. This behavior differs from the WAN network where back-and-forth trips are more statistically significant. With a larger dataset than in (Hu and Zhu, 2009) they claim that $P(k)$ is a heavy-tailed distribution but does not follow a power-law behavior (Figure 1.4). In addition, the weight $P(w)$ and the node strength $P(s)$ distributions are found to be power-laws with exponent $\mu = 1.71 \pm 0.14$ and $\mu = 1.02 \pm 0.17$ respectively. Finally, by exploring the relationship between node strength and degree the authors show that network hubs have, on average, heavy weighted links.

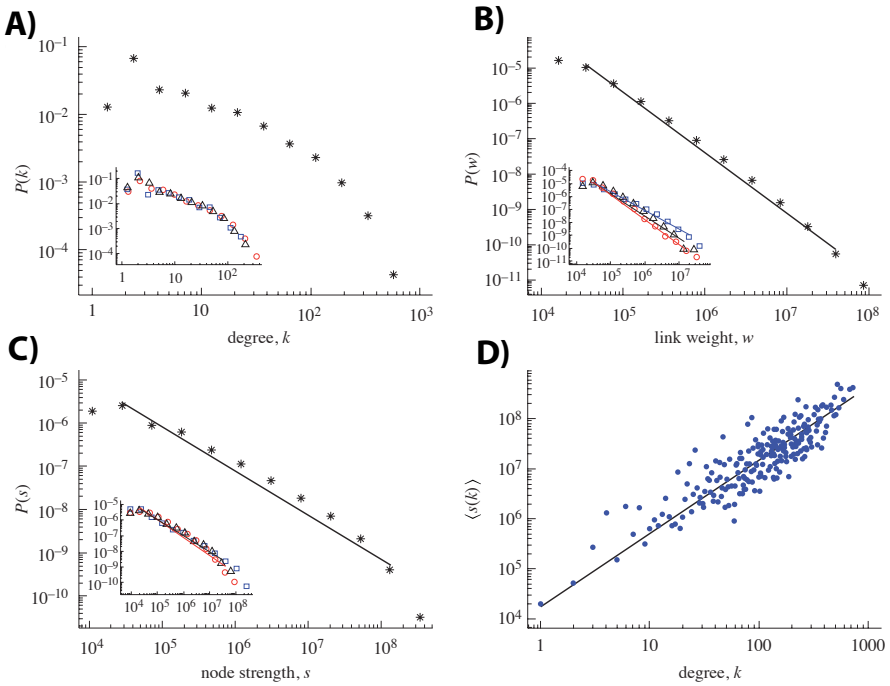


Figure 1.4: Statistical features of the MTN. A) Degree, B) Weight and C) Node strength distributions. D) Average node strength as a function of node degree. From (Kaluza et al., 2010).

Air transportation

Among all the different means of transport, the air transportation system is the one that has experienced the fastest growth (Heppenheimer and Heppenheimer, 1995). In 2013 domestic and international air passengers summed up 3023 millions[1] and is expected to increase by 6% this year[2]. However, the rapid increase in demand comes at a high price causing the system to become saturated. It is therefore of great importance to understand the interplay between the various components of the system, whereas demand and capacity are two sides of the same coin.

For this system the first results were published in 2004 and 2005. Air transportation systems have been traditionally described as static graphs over a certain time period with vertices representing airports and edges direct flights (Barrat et al., 2004; Guimera et al., 2005). Each edge also bears a weight corresponding to the number of seats available in the connection. Clearly, although spatially embedded the network is not planar as land transportation networks. In addition, unlike the MTN the unweighted version can be considered almost bidirectional. The initial works (Guimera et al., 2005) were done using the OAG MAX[3]

[1]The World Bank database: http://www.worldbank.org

[2]International Civil Aviation organization (ICAO), forecast of scheduled passenger traffic: http://www.icao.int

[3]OAG worldwide: http://oagdata.com/solutions/max.aspx

database for the period Nov 2000-Oct 2001 that consists of $N = 3883$ airports and $E = 27051$ unique airport pairs having non-stop connections. (Guimera et al., 2005) focuses on the network description with an analysis of the degree distribution and community structure. In particular the authors found that the tail of the degree distribution follows a truncated power law:

$$P(k) = k^{-\gamma} f(k/k_\star) \tag{1.24}$$

where $\gamma \simeq 2.0$, $f(k/k_\star)$ is a truncation function and $k_\star$ depends on the network size.

In addition the authors explore the community structure of the network using simulating annealing to find the network partition that maximizes the modularity. Other works have focused as well on the meso-scale structure of the WAN (Sales-Pardo et al., 2007), and in this thesis we draw some observations regarding the communities in it. Figure 1.5 depicts the partition found with each color representing a community. As the authors noted, using modularity optimization, the communities cannot be explained solely by geographical considerations. Distinguishing between edges that connect nodes of different communities (intracommunity) and edges linking nodes within a community (intercommunity) each node is classify according to its role in the network (Figure 1.6 A) and compared with the roles in a randomized version of the WAN (Figure 1.6 B):

- *Nonhub node categories*
  - Ultraperipheral nodes (black): nodes with all edges in their community.
  - Peripheral nodes (red): nodes with most edges within their community.
  - Nonhub connector nodes (green): nodes with many edges to other communities.
  - Nonhub kinless nodes (blue): nodes with homogeneously distributed edges among communities.
- *Hub node categories*
  - Provincial hubs (yellow): hubs with the majority of edges within their community.
  - Connector hubs (brown): hubs with many edges to most of the rest of the communities.
  - Kinless hubs (gray): hubs with homogeneously distributed edges among communities.

The classification is done through the *within-community degree z-score* z_i and the *participation coefficient* P_i. The former measures the level of connectivity of

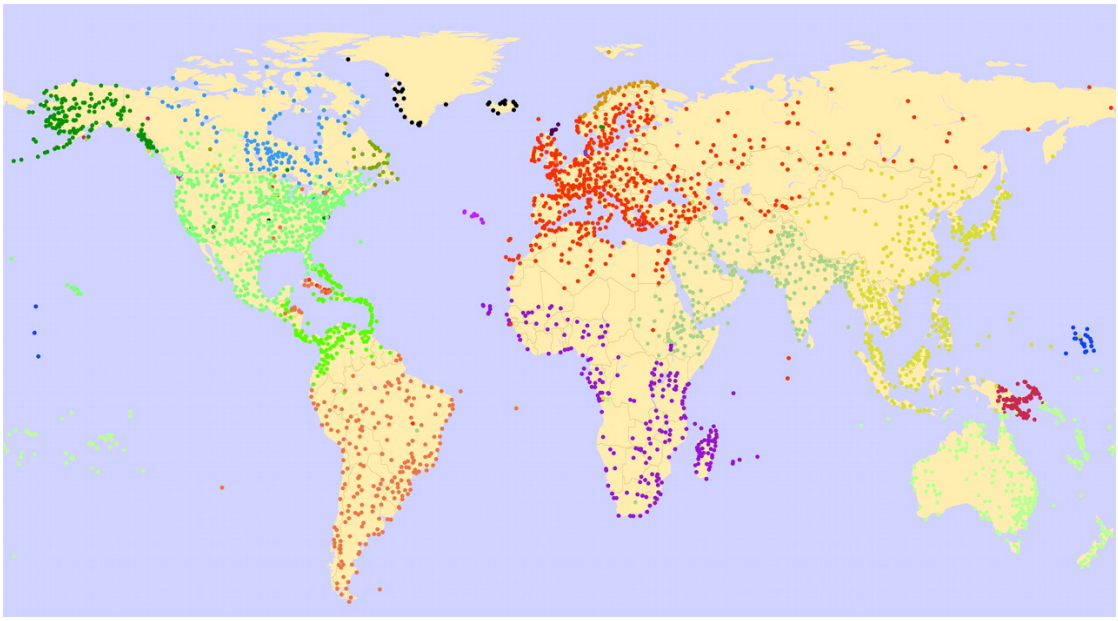


Figure 1.5: Communities in the WAN using modularity optimization. Each node corresponds to an airport and each color to a community. From (Guimera et al., 2005).

node i to other nodes in its community and the latter the level of connectivity to nodes in other N_M communities:

$$z_i = \frac{k_i - \bar{k}_{s_i}}{\sigma_{\bar{k}_{s_i}}} \qquad P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{k_{is}}{k_i} \right)^2 \tag{1.25}$$

where k_i is the number of edges of node i to other nodes in its community s_i, $\bar{k}_{s_i}$ is the average degree over all nodes in s_i, $\sigma_{k_{s_i}}$ is the standard deviation of degree in the community and k_{is} is the number of edges of node i to other community s.

A second work (Barrat et al., 2004) includes an analysis on the correlations between network topology and fluxes of passengers finding a non-linear relation between them:

$$w_{ij} = (k_i k_j)^{\theta} \tag{1.26}$$

where w_{ij} is the number of seats available in the connection between airports i and j, while k_i is the number of connections with other airports of airport i, and θ is a parameter whose value was estimated to be approximately 0.5. This non linear relation between the topology and traffic has been used later for modeling (Colizza and Vespignani, 2007). The airport network has also been studied at different geographical resolution scales, restricted, for instance, to a single country (usually the U.S. (USAN) (Opsahl et al., 2008; Gautreau et al., 2009; Wuellner et al., 2010; Lancichinetti et al., 2011) but also China (Li and Cai, 2004) or Europe (Burghouwt and de Wit, 2005)). A different formalism was used

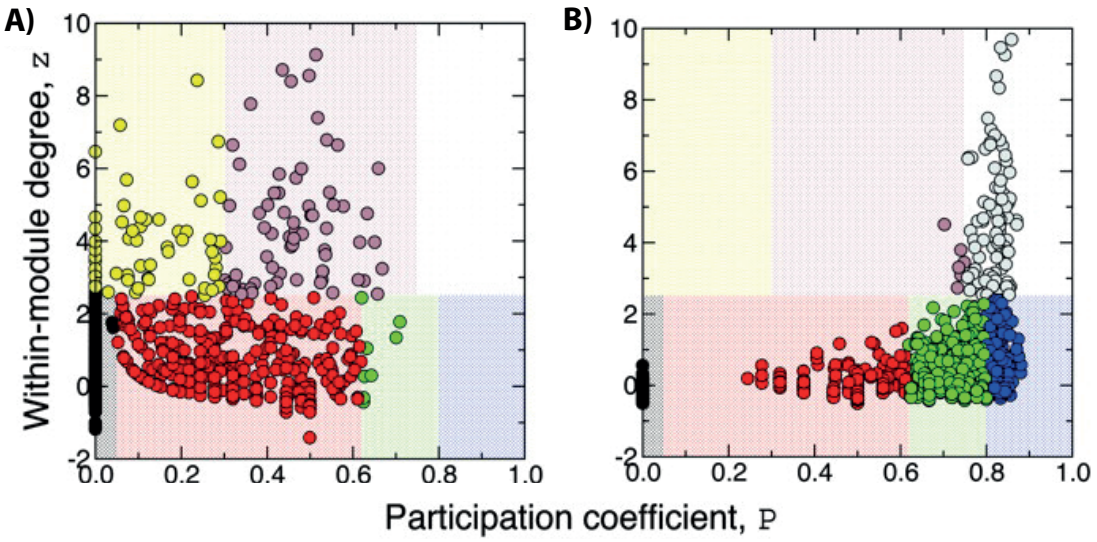


Figure 1.6: Each node corresponds to an airport and the colors correspond to different node classification depending on the role. A) Actual WAN network and B) randomized version of the network From (Guimera et al., 2005).

in (Cardillo et al., 2013) where the authors considered the European network composed of different layers each corresponding to an airline network.

Most studies, as the ones previously mentioned, consider the air transportation network as a static graph. Indeed, this is a simplifying approach, that works depending on the focus of the study and therefore on the time-scale of the relevant parameters compared to the time-scale of the network evolution. However, the structure of the air transportation network is far from been static. Airlines are constantly rescheduling flights based on unforeseen perturbations (canceled and diverted flights) and, in the long run, adapting its routes to the varying demand of passengers. Due to data availability there is a growing body of research in this topic. Specifically, the dynamics of the connections and the traffic levels have been analyzed for the USAN (Gautreau et al., 2009). In addition, for the European air-traffic system (Burghouwt et al., 2003) studied the effects of deregulation on the network evolution, (Zhang et al., 2010) explored the impact of China's economic reform on the geographic patterns of China air transportation system and (da Rocha, 2009) did a similar analysis regarding the Brazilian network.

1.3

Previous studies concerning flight delays

According to the 2008 Report of the Congress Joint Economic Committee, the monetary cost of domestic flight delays drained from the U.S. economy 40.7 bil-

lion dollars in 2007[4]. Likewise, in Europe, similar results are expected, squandering billions in system inefficiencies[5,6]. The situation can turn even grimmer in the next decade since the air traffic is envisaged to increase[7,8,9] (Jetzki, 2009). Delays damage companies' balances due to enhanced operation costs contributing to deteriorate their image with costumers (Folkes et al., 1987). Passengers suffer a loss of time, even more acute in case of missing connections, that translates into decreased productivity, missed business opportunities or leisure activities. Additionally, the impact of flight delays is not only an economic problem; it also raises environmental concerns given the additional climate-disrupting CO_2 emissions (Zou et al., 2013).

As a consequence of this challenging situation, a considerable effort has been invested in the area of Air Traffic Management to characterize the sources of initial (primary) delays (Rupp, 2007; AhmadBeygi et al., 2008) and the way in which they may be transferred and amplified by consequent operations, the so-called reactionary delays (Jetzki, 2009; Beatty et al., 1999; Mayer and Sinai, 2003; Bonnefoy and Hansman, 2007; Cook, 2007; Belobaba et al., 2009). The concept of delay itself implies a time difference with respect to the baseline provided by a predefined schedule (Rupp, 2007; Mayer and Sinai, 2003). The propagation of delays thus corresponds to the spreading of a malfunction across the system. The mechanisms responsible for it reflect the complexity of air traffic operations. Apart from the airport network structure and dynamics, other factors contributing to the delay propagation are airport congestion (Bonnefoy and Hansman, 2007), plane rotation or crew and passenger connection disruptions (Beatty et al., 1999; Cook, 2007; Belobaba et al., 2009). Airline schedules typically include a buffer time to deal with all these issues. However, when this time is not enough, the departure of the next flight gets delayed and can affect further operations in a cascade-like effect (Beatty et al., 1999). Additionally, other external factors can affect flight performance as, for example, weather, labor regulations and strikes or security threats. The intricacy and interaction between all these elements clearly qualifies delay propagation phenomena under the scope of Complexity Science.

Because of the inherent complexity of the mechanisms that produce and boost delay spreading, different modeling techniques were proposed. A first line of research focused on simulating the air traffic system as a network of queues without considering information on aircraft schedules (Schaefer and Millner,

[4]Joint Economic Committee of US Congress, Your flight has been delayed again: Flight delays cost passengers, airlines and the U.S. economy billions (May 22, 2008): http://www.jec.senate.gov

[5]ICCSAI Fact Book on Air Transport in Europe (2007-2011):http://www.iccsai.eu

[6]Eurocontrol Annual report (2008-2011): http://www.eurocontrol.int

[7]ICCSAI Fact Book on Air Transport in Europe (2007-2011):http://www.iccsai.eu

[8]Eurocontrol Annual report (2008-2011): http://www.eurocontrol.int

[9]Joint Economic Committee of US Congress, Your flight has been delayed again: Flight delays cost passengers, airlines and the U.S. economy billions (May 22, 2008): http://www.jec.senate.gov

2001). A second line of research was devoted to analytical approximations for modeling the airport runway operations as a dynamic queuing system with varying demand and service rate (Malone, 1995). Another analytical queuing model were used in (Pyrgiotis et al., 2013) and (Lacasa et al., 2009), the latter combining a diffusive process to study the onset of phase transitions in the system. In this work, airports were modeled as dynamic queues and implemented in a network. The authors ran the model in a network of 34 airports with a specific algorithm that accounts for downstream propagation of delays. An additional body of work uses statistical tools to predict the delay patterns observed in the data. Such techniques could be classified into traditional linear regression models (Churchill et al., 2007), artificial neural networks (Sridhar et al., 2009) and Bayesian networks (Xu et al., 2005). Other modeling attempts were devoted to understand the impact of external perturbation in the system (Rosenberger et al., 2002; Wang et al., 2003; Janić, 2005).

Some of the mentioned models have been limited to single airport or just a few major airports analysis with different level of detail, while others performed an aggregated analysis of the whole system. By considering an agent-based framework (Bazzan and Klugl, 2009; Shah et al., 2005) we can give insights, in a cost-effective way, of how micro-level interactions give place to emergent behavior from a network-wide perspective. To do so, we define metrics able to quantify the level of spread of the delays in the network. We then apply these metrics to a database with information on the operations in the U.S during 2010, and introduce a model that reproduces the delay propagation patterns observed in the data. The model shows also a notable capacity to evaluate the risk of development of system-wide congestion and to assess the resilience of daily schedules to service disruptions.

Part II

Data analysis

CHAPTER 2

Air Transportation: a topological perspective

The present chapter is the stepping-stone of this Thesis where the subsequent efforts aimed at modeling delay spreading in the air transport network are built on. The results obtained here together with the ones from the next chapter are relevant in order to better characterize flight delays from a statistical perspective and have to be taken into account for the model development and validation. We begin the analysis by exploring the data sets available to understand the main characteristics of the air transportation system. In this sense, we start by addressing some topological characteristics of the World Airport Network (WAN) using data from the International Air Transportation Association (IATA)[1]. Some of the results are well known for this network but still are relevant for the dynamics of delay propagation. We then continue our top-down analysis and focus on the United States Airport Network (USAN) a subset of the WAN. The reason for doing so is that, unlike the WAN system, there is performance data concerning flight delays. In this regard, the data sources used for the analysis can be categorized into three distinct groups: annual fraction of connecting passengers per airport, time zone conversion data and flight schedules and USAN network construction.

2.1 Large scale structure of World Airport Network

The evolution of the World Airport Network (WAN) is the result of several social and economic factors. Intuitively, geographical and political constrains are some of the organizing principles that drive its structural patterns (Guimera and Amaral, 2004). In addition, in a shorter time scale, its structure is influenced by the actions of airlines trying to maximize their economic benefit (Cardillo et al., 2013). Clearly the outcome of all these factors must be encoded in the

[1] International Air Transportation Association (IATA): http://www.iata.org

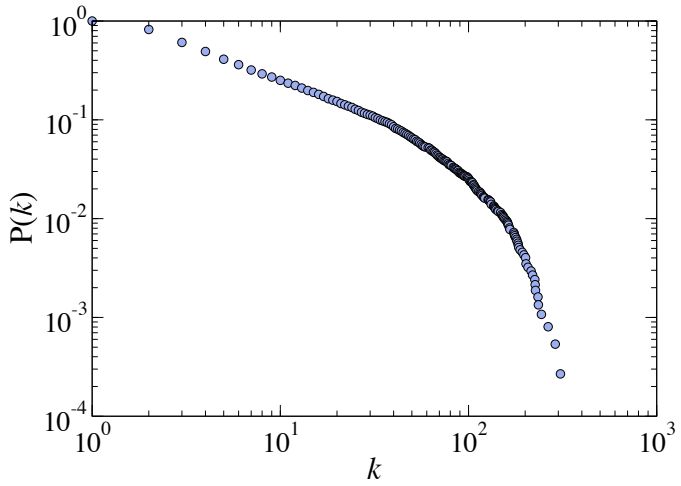


Figure 2.1: Cumulative degree distribution for the WAN.

network topology and these cannot be fully understood without a system-wide perspective.

Our WAN database comprises a network with 3728 airports and 24208 edges or direct flights between each node. A large percentage of connections between airports are bi-directional so in the following we will work with the undirected WAN. On average each airport is connected with 13 other airports, while the maximum degree found is 305. It is a sparse network with an average link density (ρ_G), defined as $\frac{L}{N(N-1)/2}$, of 1.7×10^{-3} (a network with L links and N nodes). Sparsity could be the result of balanced cost/efficiency optimization processes taken by individual airlines trying to maximize efficiency while diminishing the cost of having a densely connected network. In addition the assortativity mixing coefficient r of the network has a value of 0.011. As mentioned in Chapter 1 purely dissortative and assortative networks have, respectively, a value close (or equal) to -1 and 1. Thus, in the WAN case the topological structure shows, on average, no assortativity between nodes of similar degree. However, the average can hide some preference between highly connected nodes and this will become clear once we compute the rich-club ratio. Another relevant topological measure is the average clustering coefficient $\bar{C}$ and for the WAN network this has a value of 0.51. This parameter varies from 0 to 1, therefore the WAN network, has a higher number of triangles between neighboring nodes when compared to to the average clustering coefficient of its randomized version ($\bar{C} = 0.049$) (Guimera et al., 2005). In other words, we find airports linked to a third airport which are also linked between themselves. In addition the network can be categorized as a "small-world" network in which airports can be reached from every other by a relatively small number of links. This observation is confirmed by the average shortest path $\langle l \rangle$ between nodes of network which yields 4.4 (Guimera et al., 2005).

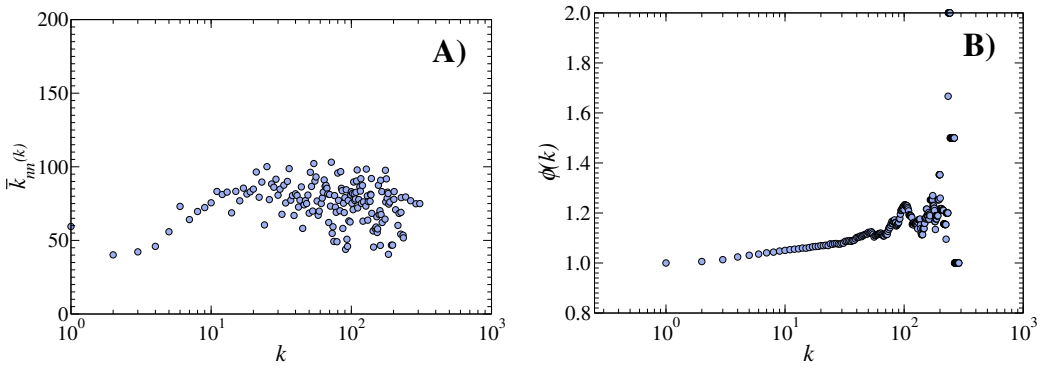


Figure 2.2: Degree-degree correlations. A) Average network degree and B) Rich club ratio for the WAN.

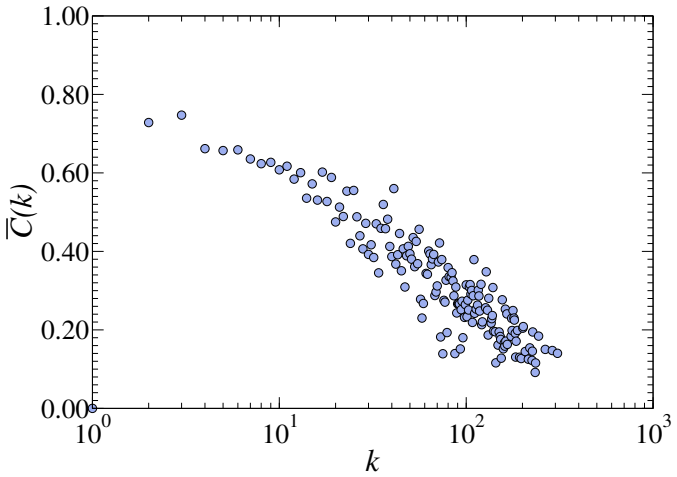


Figure 2.3: Average clustering coefficient as a function of the nodes degree for the WAN.

We have analyzed the main topological characteristics from an aggregate level making no distinction whatsoever with the airport size. In the following we will explore the topological characteristics as a function of the node degree. Therefore we begin by computing the cumulative degree distribution of the World Airport Network in Figure 2.1. In agreement with previous studies of the WAN, $P(k)$ shows a heavy-tailed behavior and as some studies have shown it can be fitted to a truncated power law:

$$P(k) = k^{-\gamma} f(k/k_{\star}) \tag{2.1}$$

Not surprisingly the truncated behavior is due to the fact that each airport can only withstand a finite number of connections.

Figure 2.2 shows the results for the neighbor average degree (k_{nn}) and the rich-club ratio (ϕ) as a function of the node degree. On the one hand, k_{nn} is an increasing function of the node degree for low values and then remains almost

flat (Figure 2.2). In B the rich-club ratio signals that it is proportional to the degree, meaning that nodes with high degree tend to be connected among them. This result could be related with the observation that long haul flights connect mainly continental and, therefore, network hubs resembling the idea of "superhighways" as described by Z. Wu *et al.* in (Wu et al., 2006).

Finally, in Figure 2.3, we compute the average clustering coefficient as a function of the node degree. The decreasing pattern is expected due to the fact that the clustering coefficient varies as k^{-1}. However we cannot discard the effect of the typical air transportation network hub and spoke structure where hubs connect smaller airports that are not connected between them, thus for largely connected airports, the network shows mix of star-like structures.

2.2

Community detection algorithms & the meso-scale structure of the World Airport Network

We continue our analysis by focusing on the meso-scale structure of World Airport Network. A classical line of action to tackle this subject is through community detection algorithms. There is a large and diverse family of algorithms. For this reason, it is important to be able to compare the relative performance of each method. One way of evaluating this is by determining how well they behave against artificial benchmarks with known community structure. The first tests were carried out using the benchmark by Girvan and Newman (Girvan and Newman, 2002). This benchmark consists of a synthetic network of 128 nodes split in four groups of 32 nodes each. The probability of edges belonging to the same community and between them are chosen to keep the average degree of a node equal to 16. By varying the average number of inter-community edges per node one can check the efficacy of different methods by comparing the fraction of correctly classified nodes. Lancichinetti, Fortunato and Radicchi (Lancichinetti et al., 2008) developed a more complete type of benchmark that accounts for heterogeneous node degree and community sizes distributions and can include overlapping nodes as well. The LFR (Lancichinetti, Fortunato and Radicchi benchmark) poses a much more challenging test for most algorithms than standard benchmarks as the Girvan and Newman. However very few validations have been done in real complex networks due to the lack of ground truth (Radicchi, 2014; Hric et al., 2014).

We apply two modularity optimization algorithms the Fast Greedy (Clauset et al., 2004) and Louvain (Blondel et al., 2008) methods; plus algorithms of a very different nature such as INFOMAP (Rosvall and Bergstrom, 2007) and OSLOM (Lancichinetti et al., 2011) to check for their consistency in a real complex network as the WAN. We show that the partitions found by each method are remarkably dissimilar according to community features such as number, size,

quality and visual inspection. In view of these results, we asses the quality of the modules against the basic community definitions and found that an important fraction of the obtained communities are statistically non-significant when compared to a randomized ensemble of the WAN.

A topological community can be thought as a cohesive group of nodes with more internal edges than edges linking community nodes with the rest of the network. This basic notion is at the core of all community detection algorithms. The methods used here are among the most popular algorithms and were also selected because they rely on different methodologies.

Fast Greedy & Louvain. These algorithms are among many methods that look for a partition of the network into communities that maximizes the modularity measure (Newman, 2006b). Modularity quantifies the quality of a partition by comparing the number of edges that belongs to a given module with the number we would expect from a random graph with the same degree distribution. According to these techniques a good partition would be one that maximizes the value of modularity. Finding the optimal partition is computationally expensive (Brandes et al., 2008) because of the large number of ways to divide a network into communities. For this reason several methods have been developed to at least find fairly good partitions, in our case we use the Fast Greedy for assessing community fundamentals and the Louvain algorithm, because of its stochastic nature, when assessing the community fundamentals in the ensemble of randomized versions of the WAN.

INFOMAP. Another way of exploring community structure in complex networks is with the INFOMAP algorithm, which aims to concisely describe the diffusion of information over the network. This process is characterized by the probability flow of random walks in the network. One way of describing the diffusion of random walkers in the network is by Huffman coding (Huffman et al., 1952) in which frequent events or objects retain short names and rare events long names. The algorithm applies a two-level description to separate relevant structural objects of the network. The code retains unique names for coarse-grain objects and recycle the individual nodes names within each structure. This description allows to highlight network structures where the walker will spend much of its time and will use fewer bits for describing the information flow than a one-level description. Therefore community detection, under this framework, is reduced to finding the network partition that minimize the description length of an infinite random walk process. This optimization process is done by integrating greedy search with simulated annealing.

OSLOM. Whereas the previous methods focus on finding the best partition given a global measure, OSLOM's approach is to locally optimize the statistical significance of each cluster. This difference can be very beneficial when exploring networks that are partially modular. The core idea of the cluster significance is to quantify how likely is to find that cluster as a subgraph of a random graph with the same degree distribution as the original network.The lower this likelihood is, the stronger the indication that the cluster is a real community and not just

a product of random fluctuations. The algorithm starts from a single node and adds an arbitrary number of the most significant neighbors. It adjust this group of possible community nodes by adding/removing neighboring nodes in order to increase the cluster significance. Because of its stochastic nature, several realization have to be performed. The method output is a cover of the network with the possibility of finding nodes in more than one community (overlapping nodes) and nodes in no community at all (homeless).

2.2.1 Communities in the World Airport Network

When a researcher aims to unveil the possible modular structure of a real network the first question that arises is; given the limited size of the WAN which of the various methods should be used? We are not interested in computational time efficiency but rather on the quality of the resulting partition. If a modular structure is present in the network, at least we shall expect the methods' results to be fairly similar signaling a unique partition of the network into communities. This structural characteristic must be unique because it is the product of the particular mechanisms that have generated the network in the first place. Furthermore, if this is not the case, community finding in real networks would be a poorly defined problem. Uniqueness is the fundamental idea behind the comparative analysis done with respect to the ground truth communities generated in synthetic benchmarks, so the same concept must be applied to real networks. We explore the consistency of the different approaches but with respect to the WAN, a extensively used network in the community detection field.

Let us start the analysis by visually inspecting the communities found by the three methods in the WAN (weighted and unweighted) spatial network. As shown in Figure 2.4 the results look strikingly different to one another. The Fast Greedy algorithm produces a more coarse-grained partition with the largest communities spanning different continents, as the one formed by Asia, most of Oceania and Middle East in the unweighted case. Other large communities are the ones constituted by Europe and Eastern Russia or the United States plus Central America for both the weighted and unweighted WAN. On the other hand, the cover produced by OSLOM is fine-grained but, therefore, more fuzzy. For instance, the United States, in the unweighted case, is divided into 4 communities besides homeless and overlapping nodes and Europe is divided in several communities matching no geopolitical subdivision of the continent. In the weighted case, the United States is divided in 2 communities (mainly East and West) with many homeless nodes. The INFOMAP partition seems like a trade-off between the results of the previous methods with communities at the level of countries in most cases. It is also important to highlight that while this representation of the community structure gives the impression that each of the methods Fast Greedy, OSLOM and INFOMAP obtain a relatively small number of clusters, in fact they return 48, 83 and 160 partitions, respectively.

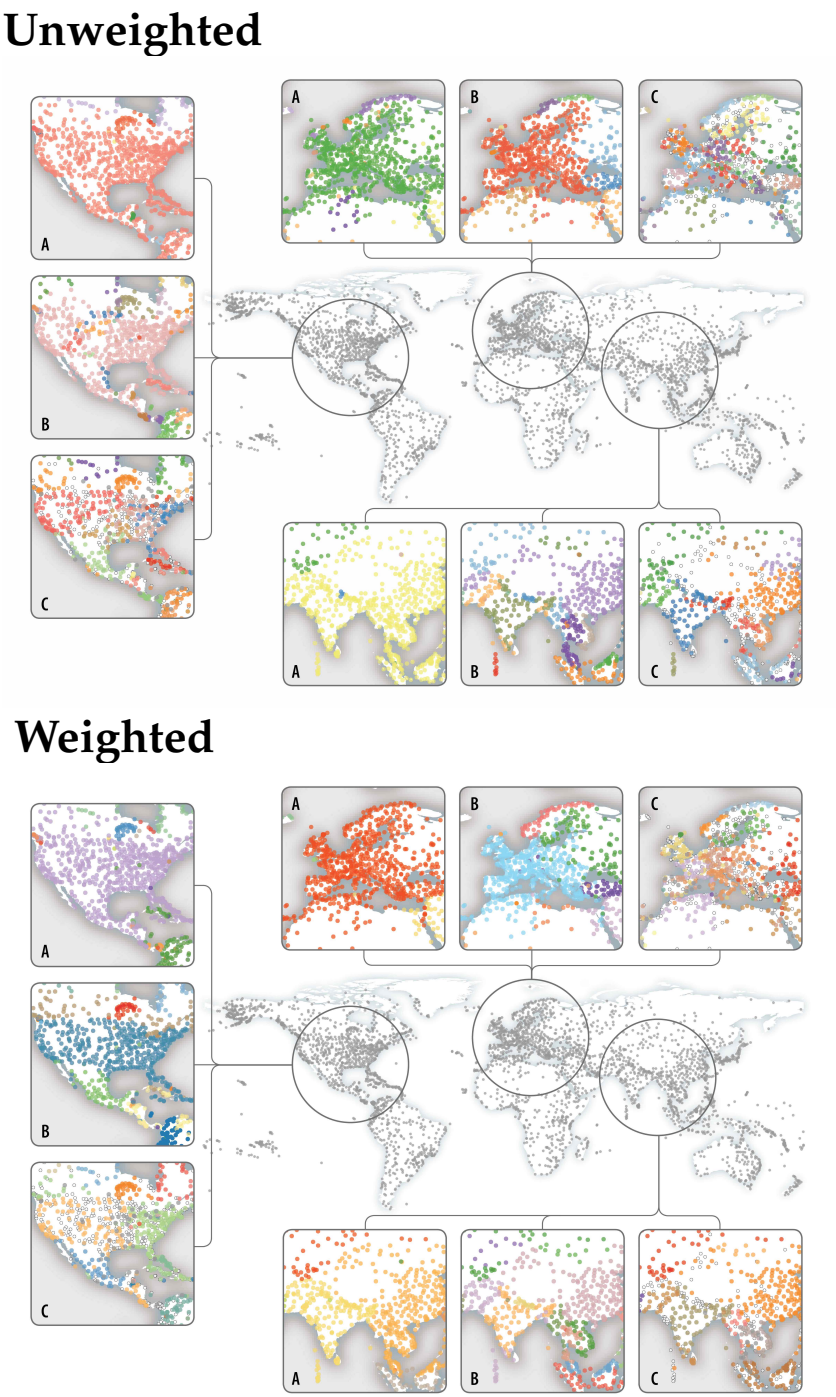


Figure 2.4: Community maps in the unweighted and weighted WAN network using the Fast Greedy (A), INFOMAP (B) and OSLOM (C) algorithms. White colored airports in the OSLOM community map represent homeless nodes (nodes that do not belong to any community), while grey ones are overlapping nodes (nodes in more than one community).

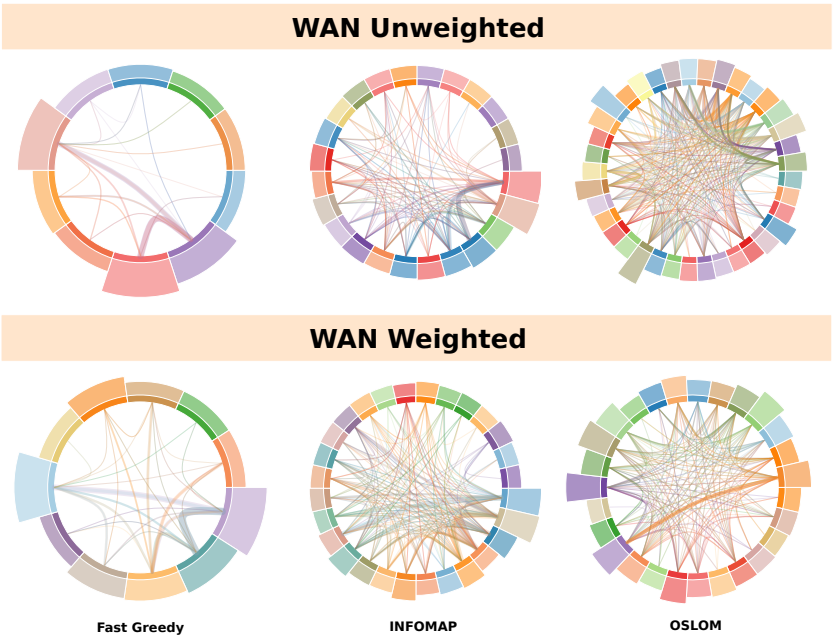


Figure 2.5: Network of communities for the unweighted (upper panel) and weighted (lower panel) WAN network using the Fast Greedy, INFOMAP and OSLOM algorithms. Only communities larger than 30 nodes are represented and sized according to the number of nodes that belongs to each community.

Furthermore, Figure 2.5 displays a network of the largest communities for the unweighted (upper panel) and weighted (lower panel) WAN, that is, only those that have more than 30 nodes ($\sim$ 1% of the nodes in the network). One immediately realizes the difference in the number of communities that share this property in each method. For the unweighted WAN, the edges between clusters are weighted according to the number of links connecting the communities in the original network. In the weighted case, according to the number of available flight seats between clusters. In both cases, the connectivity patterns differ with more homogeneously distributed weights for the cover found by OSLOM than for the INFOMAP and Fast Greedy. Moreover, Fast Greedy shows a great disparity between the link weights and INFOMAP's largest clusters are heavily connected within them than with respect to the rest of the communities. These results would lead to different conclusions if, for instance, the research focus was to explore the dynamics of information flow.

In the same line, the conclusions regarding the cities global roles in the study by Guimera et al. (2005) (see Section 1.2.1 of the Introduction) would had been very different if instead of using a simulating annealing algorithm for modularity optimization, the authors might have used the INFOMAP algorithm. In Figure 2.6 we reproduced Figure 1.6 of the Introduction but in this case we used the INFOMAP algorithm. The differences in node classification are quite clear,

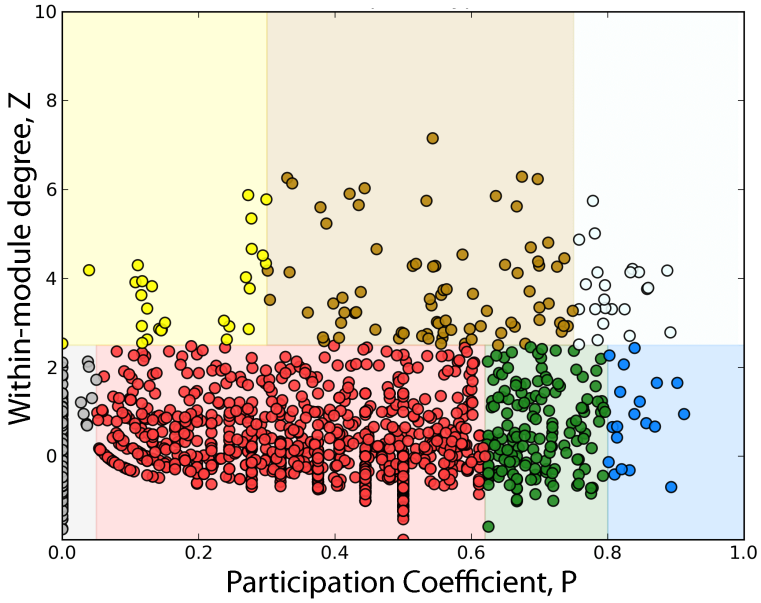


Figure 2.6: Each node corresponds to an airport and the colors correspond to different node classification depending on the role. The algorithm used is INFOMAP.

for instance the abundance of nonhub connectors (green) and kinless nodes (blue) and the existence of kinless hubs (grey) is something that cannot be reproduce by the modularity optimization algorithm used in Guimera et al. (2005). These results would yield different conclusions regarding the identification of most influential cities in dynamical processes taking place in the World Airport Network.

According to the community size probability density function (see Figure 2.7) we can conclude that Fast Greedy usually finds communities of large size while OSLOM's modules are of intermediate size. This is one of the reasons why the OSLOM cover is visually less clear. Similar conclusions are derived from the weighted version (see Appendix A.1). Importantly, the large size of the communities found by the Fast Greedy, is something well-known in the literature. It has been discussed in previous papers that modularity optimization techniques have a resolution limit which makes them unable to find communities smaller than a certain threshold (Fortunato and Barthélemy, 2007). The results reinforce the observation that the methods, not only qualitative but also quantitative, give very diverse partitions/covers. This is very puzzling, how can the algorithms detect partitions be so dissimilar?

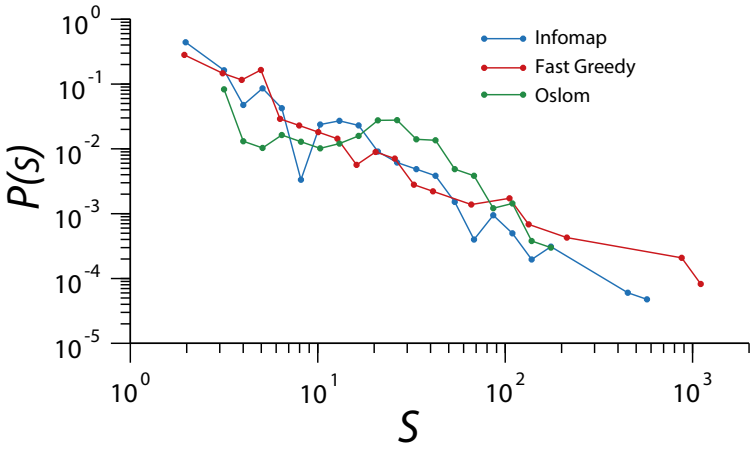


Figure 2.7: Community size probability density function.

2.2.2 Assessing community fundamentals

A starting point to understand why the methods analyzed here give such dissimilar results is by testing them against the basic definition of what topological communities are. However this is not an easy and trivial task to start with, due to the fact that there is no single definition of community universally accepted. The concept of topological community has much evolved in the last decade. The first definitions where based in the idea that there should be more edges linking nodes of the same community (intra-cluster links) than edges connecting nodes of the community with the rest of the network (inter-cluster edges) (Radicchi et al., 2004). On the other hand, the prevailing framework defines a community as a group of nodes more tightly interconnected among themselves than with the rest of the network, yet in this context, tightly means coherently connected not with more links. A possible approach to explore the "coherence" of a tightly group of nodes is by comparing it to what we would find, for the same algorithm, if we apply it to a null model or randomized version of the WAN network. The reason to do so is that in an ensemble of null models of the network we should not expect to find any community structure. Defined the scheme we will follow, we still lack a measure to explore the differences (or similarities) between the communities found by the methods used. In this regard we will use the link ratio l_r (for other measures see Appendix A.1) between the number of intra-cluster (l_i) and inter-cluster links (l_o):

$$l_r = \frac{l_i}{l_o} \,. \tag{2.2}$$

In addition we can simply extend this definition to account for link weights:

$$w_r = \frac{w_i}{w_o} \,. \tag{2.3}$$

Besides its simplicity and clear intuitive meaning these measures are the building blocks of the most challenging and used benchmark for community detection: the LFR benchmark. The mixing topological parameter μ of the benchmark, defined as the fraction of links shared by the community with other nodes of the network, can be rewritten as a function of the link ratio:

$$\mu = \frac{1}{1 + l_r} \,. \tag{2.4}$$

With the other measures explored in Appendix A.1 the general picture remains the same.

We begin by inspecting the identified modules based on the link ratio measure. In Figure 2.8 we plot for the unweighted network the logarithm of the ratio between the number of intra-cluster and inter-cluster links for the communities found by each algorithm. A value of zero means that a module has the same number of internal and external edges. We have also included the results for a synthetic network constructed using the LFR benchmark (B1). This artificial network is built from a power law distribution with exponent 2.5 for the degree sequence and exponent 1 for community size. The mixing topological parameter μ used is 0.2, which means that the synthetic modules have twice the number of links connecting nodes of the same module than links to other modules. The Figure confirms the high variability among the distributions of link ratio for the different methods spanning through a wide range of values for each method. In great contrast, the interquartile range is almost null for the synthetic network, evincing that they have detected the ground truth communities.

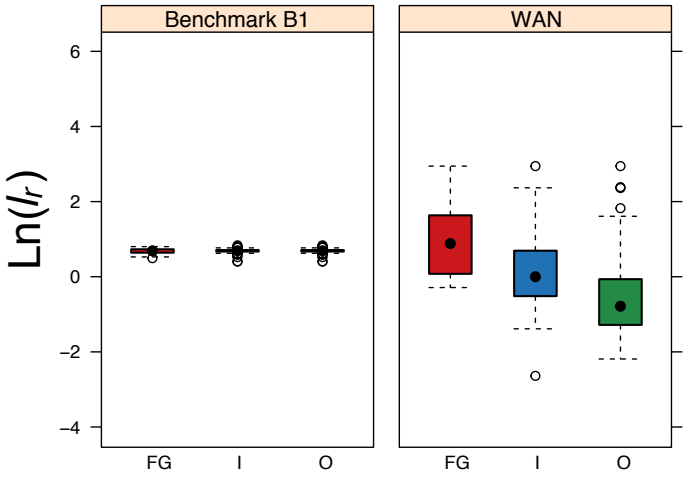


Figure 2.8: Distribution of the logarithm of the link ratio definition for the communities found by each method: Fast Greedy (FG), INFOMAP (I) and OSLOM (O).

Testing the communities with the weight ratio measure shows similar results (Figure 2.9). It is important to stress that in the OSLOM case its ability to

detect overlapping nodes can increase the occurrence of inter-cluster edges (Yang and Leskovec, 2013). Therefore with highly overlapping communities these measures can be overshadowed.

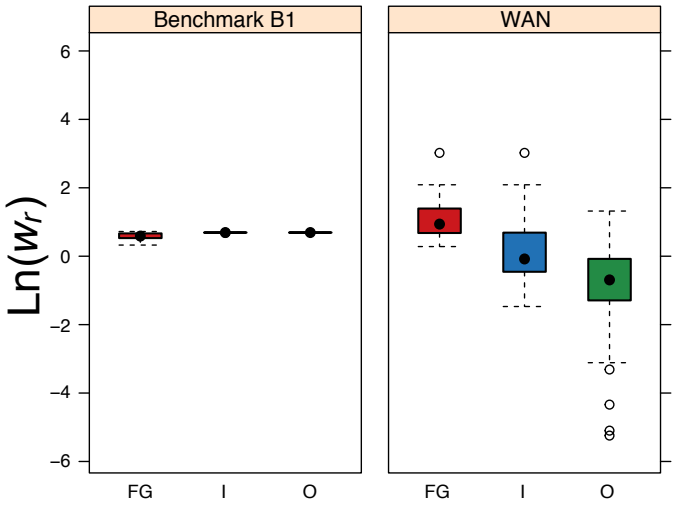


Figure 2.9: Distribution of the logarithm of the weight ratio definition for the communities found by each method: Fast Greedy (FG), INFOMAP (I) and OSLOM (O).

In the Appendix A.1 we further tested the link and degree ratio against artificial networks that by construction do not have modular structure. All artificial networks have 5000 nodes. We used the configurational model to construct two networks: one with a random degree sequence with mean 20 and standard deviation 1 (Random) and another with a degree sequence following a power law of exponent 2.5 (Scale Free). The third artificial network was generated with the LFR benchmark but with a mixing parameter of 0.65 (Benckmark B2). This corresponds to a k_r of 0.35 and a l_r of 0.27, thus the structures are not communities but *anti-communities* (Newman, 2006a) (communities with a number of intra-cluster links smaller than expected by chance and with a number inter-cluster links correspondingly bigger). Important to note is that the resulting OSLOM's cover for the Scale Free network has 88.6% of homeless nodes. For INFOMAP it found 2 big communities of 4090 and 515 nodes and the remaining divided in very small communities. This signals that the algorithms have detected the random nature of these two artificial networks. Worth noting are the results for the Benchmark B2 where INFOMAP and OSLOM were capable of finding the anti-communities, while the Fast Greedy algorithm manage to find other communities different from the ground truth ones. An interesting result though, that shows how INFOMAP and OSLOM are well trained to find anomalies in the topology even when these anomalies are not considered communities on the context the LFR benchmark. As suggested by Newman (Newman, 2006a) if instead of maximizing the modularity the algorithm minimize it, most likely the Fast

Greedy algorithm would have been capable of finding these anti-communities. The problem with testing the different algorithms against synthetic networks is that the methods are very well trained to spot anomalies as it the case with communities having all approximately the same link ratio. This regularity is not seen in real complex networks, and this is the main reason why algorithms results vary widely.

2.2.3 Comparison between real networks and their randomized counterpart

Despite of the disparity of the results, the topological structures can still be statistically significant when compared to a randomized version of the real network, thus coherently connected. To do so, we rewire the unweighted real networks by link swapping avoiding the creation of parallel edges and self loops. The Fast Greedy algorithm is deterministic so in order to work with an ensemble of random network versions of the WAN we have used a stochastic modularity optimization algorithm; the Louvain algorithm. Figure 2.10 depict the link ratio and community size distributions for the original and randomized WAN. Importantly, it is not possible to carry out the analysis using the OSLOM method. The reason is that OSLOM is constructed upon the idea of the b-score of a community (Lancichinetti et al., 2010). The b-score quantifies how significant a community is when compared to the configuration model (random null model). This is a major advantage that neither INFOMAP or the Louvain algorithm possess, which results in not finding any community when OSLOM is run over the randomized network. That being said the communities found by OSLOM, in the original network, whose score is lower than the p-value are significant per se, from the algorithm point of view.

With the proposed methodology we are able to identify a threshold under which it is not possible to assert if the community found is statistically significant or the result of the system internal noise given that the corresponding link ratio has also been found in the ensemble of WAN null model networks. By these means we consider a community to be statistically significant if the probability of finding a link ratio higher than a certain value in the random ensemble is less than 5%. This threshold is depicted for each method as a dashed vertical line (Figure 2.10); all communities above this threshold are significant regarding the algorithm used. Table 2.1 shows that for Louvain and OSLOM methods the communities found are all significant. In the INFOMAP case 44.4% of communities are labeled as significant and almost 72% of the networks nodes belong to significant communities.

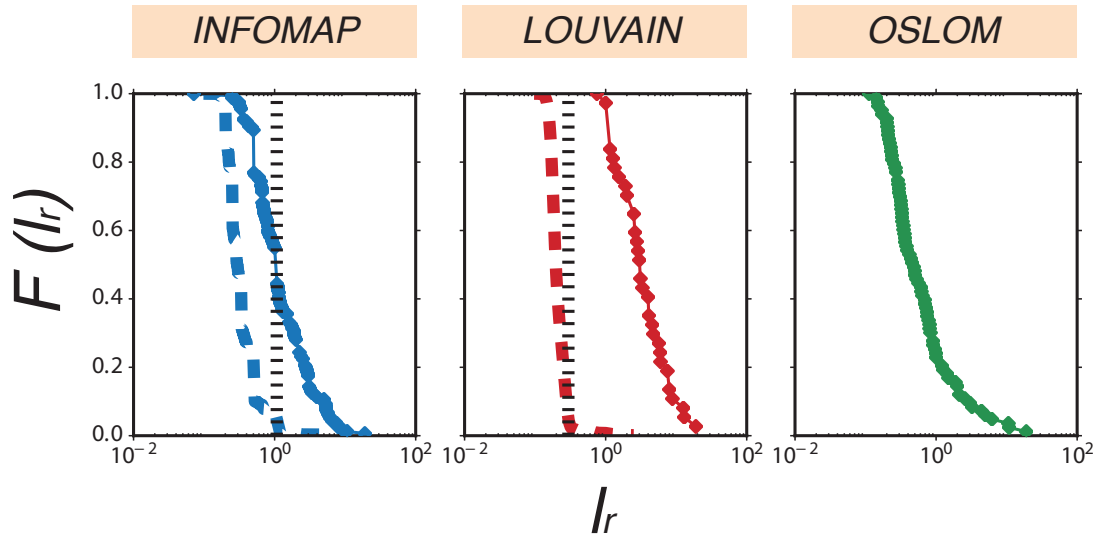


Figure 2.10: Link ratio cumulative probability. Comparison between the communities found by INFOMAP (blue), Louvain (red) & OSLOM (green) algorithms for the WAN network (dots) and their randomized version (dashed line). In the OSLOM case the method is able to recognize the randomized version, finding no communities whatsoever.

Method	Link ratio threshold	Significant Communities (%)	Significant Nodes (%)
INFOMAP	1.073	44.4	71.8
Louvain	0.296	100.0	100.0
OSLOM	-	100.0	100.0

Table 2.1: Link ratio threshold, % of significant communities and % of significant nodes for each algorithm.

2.2.4 Comparison within methods of the most significant communities

In Appendix B.1 we discuss the Normalized Mutual Information (Danon et al., 2005), a commonly used tool on network detection literature to asses the global quality of a community detection algorithm with respect to ground truth and to compare the performance of alternative methods. We analytically demonstrate that the Normalized Mutual Information has low resolution and even with a 50% of random spurious communities the score is relatively high. In the present work, the lack of ground truth for the real networks analyzed and the striking differences between algorithms call for a measure that allows quantifying not only a global similarity between partitions, but a local one as well. That is, a measure that enables the comparison of individual communities between methods. As shown in the previous section, when comparing the results obtained

by each method on the original network and its randomized counterpart, we have found a threshold under which it is not possible to assert that a community found is statistically meaningful or just the result of the system internal noise. One possible way to achieve is by means of the recall measure defined in the introduction.

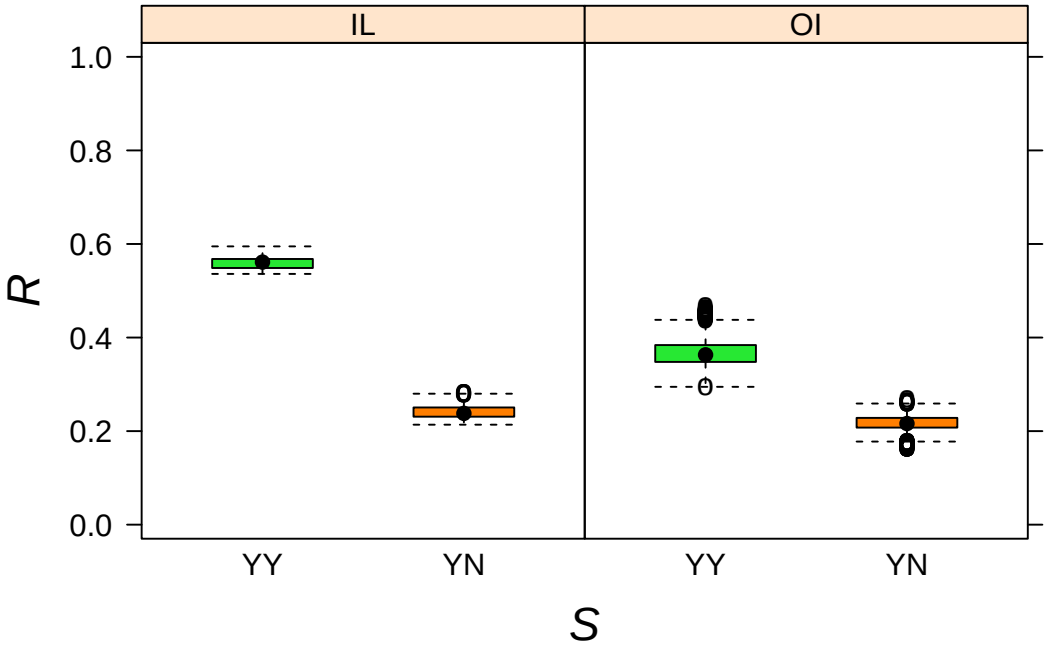


Figure 2.11: Recall score between sets of communities of different methods for the WAN network: INFOMAP-Louvain (IL), OSLOM-INFOMAP (OI) and OSLOM-Louvain (OL). X-labels are: significant/significant (YY) or significant/non-significant (YN)

With this in mind it is possible to asses how well significant communities across methods compare to each other. Are significant communities found by the different methods more similar than non-significant ones? In order to answer this, in Figure 2.11 we plot the recall score between the sets of significant communities between two methods (YY) and compare it against the set of non-significant communities between the same two methods (NN). In case a method has all of its communities significant with respect to the link ratio the comparison is done between significant/significant (YY) and significant/non-significant (YN). All communities found by Louvain and OSLOM are significant in this sense, so the methods cannot be compared for the difference with non-significant communities, this is the reason why we compared INFOMAP-Louvain and OSLOM-Louvain.

Importantly, from Figure 2.11 we can conclude that significant communities are more similar between them than non-significant and this stands for every

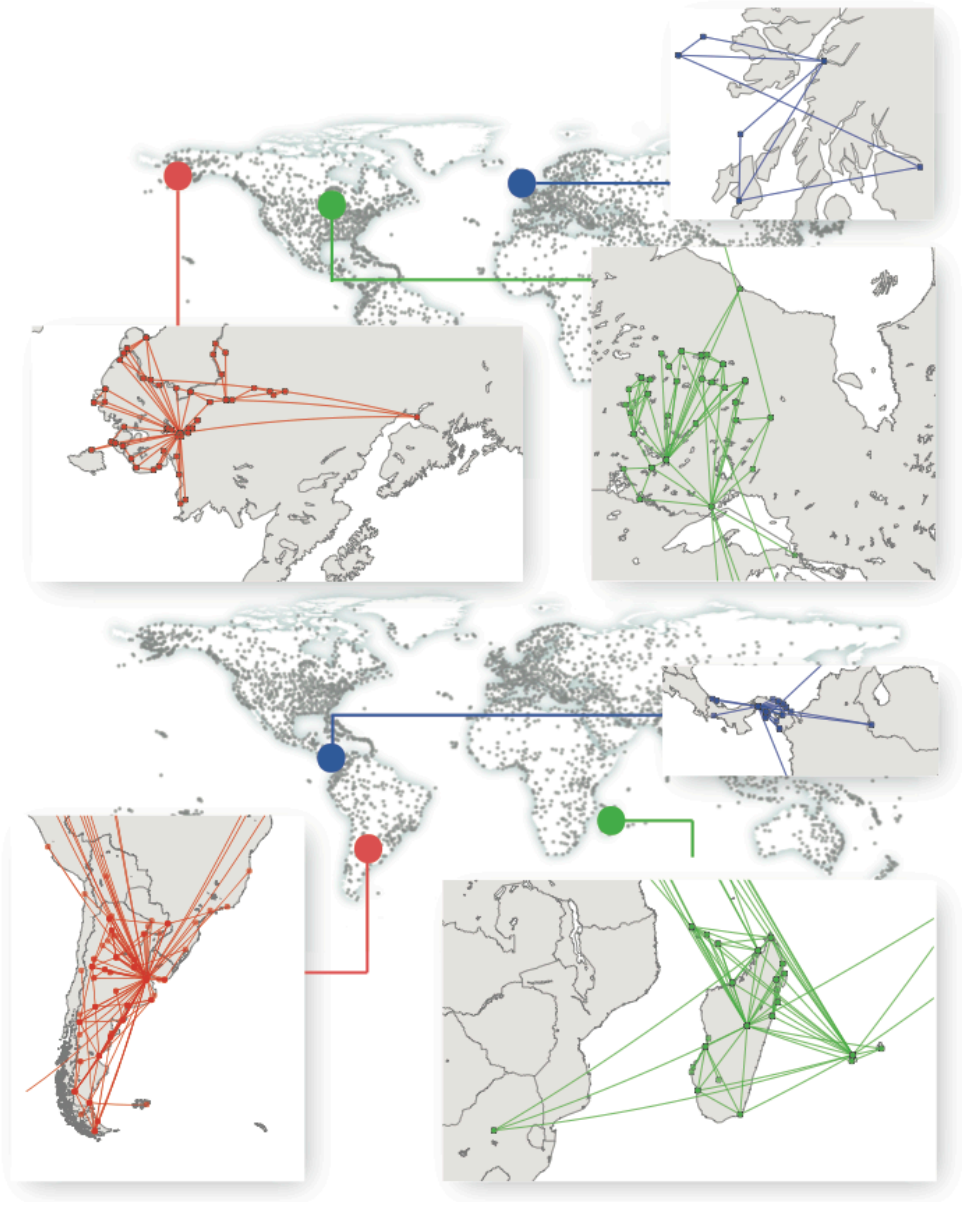


Figure 2.12: Significant and identical communities between: OSLOM & INFOMAP (RED), OSLOM & Louvain (BLUE), INFOMAP & Louvain (GREEN).

pair of methods used. This result alleviates the differences in the partitions found by each method and, therefore, one can pose the following question: how to determine which method is better given a particular real network? A possible way is to use the present framework and analyze the resulting partition against an ensemble of null models to separate significant structures from noisy ones. From the set of significant communities, Figure 2.12 shows some of the communities found with more than one algorithm. In this case we have selected an example of the set of identical communities, that is, communities A and B

(of different methods) that when compared the Jaccard index ($J(A,B)$) is equal to 1. Same communities found between OSLOM and INFOMAP, OSLOM and Louvain, and INFOMAP and Louvain are, respectively, in red, blue and green. As can be seen, the communities are of very different size ranging from 42 airports in Alaska to 6 in Scotland. Other examples of identical communities, can be found in Appendix A.1. The same exploratory analysis can be done with communities for which $J(A,B) < 1$ and analysis the results depending on the level of similarity, e.g: set of communities with $J(A,B)$ approximately 0.9, 0.8, 0.7, etc.

2.3

Characterizing the United States Airport Network

The US airport network is constructed using the information available at the Bureau of Transport Statistics[2]. We restrict our analysis to domestic flights conducted in the year 2010. In particular, we use the Airline On-Time Performance Data, which is built with flight data provided by air carriers that exceed one percent of the total domestic scheduled service passenger revenue. Added together this represents 18 carriers that combined, sum up $6,450,129$ scheduled domestic flights from 305 commercial airports. Considering all flights from 2010 and not only those that report On-Time Performance Data, the number of scheduled domestic flights totalizes $8,687,800$[3]. Consequently the used dataset represents 74% of all scheduled domestic flights of 2010. The database comprises relevant flight information that enables us to represent the US airport network and furthermore replicate the scheduled flights for every day of 2010. It is important to note that this schedule is based on real events, which in some occasions may differ from the original planned schedule of the companies. If a flight gets canceled or diverted the airline may introduce changes in the original schedule that are not possible to trace back. However, given that these flights represent, respectively, the 0.20% and 1.75% of all flights in the database, one can expect these changes not to be of large magnitude.

Among the available data fields, we consider the following: *Tail number, airline ID, airports of origin and destination, date of the flight, scheduled departure and arrival times, real departure and arrival times, and whether the flight was canceled or diverted*. The tail number is an alphanumeric code that identifies the aircraft and allows to track it along the daily plane rotation. Arrival and departure times (real or scheduled) refer to the event when the flight actually reaches or departs from the gate. Taxing or take-off and landing times are considered to be part of the departure (arrival) times. We will exclude of the

[2]Bureau of Transport Statistics of the US Government, RITA database. Available online at: http://www.bts.gov

[3]BTS press release of March 22, 2011, Available online at: http://www.bts.gov/press_releases/2011/bts017_11/html/bts017_11.html

coming analysis diverted and canceled flights since are a small fraction of the total operations.

2.3.1 Connecting passengers for each US commercial airport

Another key input for modeling the delay propagation over the network is the connection between flights. The previous database has no information regarding flight connectivity, neither for the crews nor the passengers. To approximate the heterogeneity of the airports we used the T100 Domestic Market (US carriers) and the DB1B Ticket information downloaded from the BTS page. These documents allow us to obtain an approximation of the annual fraction of connecting passengers per airport. The information of T100 corresponds to the total number of passengers who have a flight departing from an airport regardless of their real point of origin (P_{T100}). On the other hand, the database DB1B contains a 10% sample of the number of passengers whose itinerary originated in each given airport (P_{DB1B}). So for each airport we can get an approximation of the annual fraction of connecting passengers as:

$$C = \frac{P_{\text{T100}} - 10 \cdot P_{\text{DB1B}}}{P_{\text{T100}}} \tag{2.5}$$

Although our model is based on flight not passenger connectivity, these ratios are related given that a flight might wait for passengers if the economical impact of leaving them behind is higher than delaying the flight. Examples of these factors for the network hubs are given in Table 2.2.

Airport code	Fraction of connecting passengers
ATL	0.81
ORD	0.72
DFW	0.75
DTW	0.69
DEN	0.71
MSP	0.68
IAH	0.75
SLC	0.73
MEM	0.81
MCO	0.69

Table 2.2: Fraction of connecting passengers for the top ten airports in degree.

2.3.2 Time zone conversion

The United States spans through several time zones. In order to unify criteria and simplify the analysis, we transform local time to the East Coast local time. The selected timezone is the natural choice considering the daylight time flow in the United States. Olson or tz database[4] is used to ensure an accurately timezones conversion from the respective local times in the database to the East Coast time (EST in winter and EDT in summer time).

2.3.3 Flight schedules and network construction

Based on the data a network between airports can be built, where airports are the vertices and edges represent direct flights from one airport to another. Even though, the network is not completely bidirectional, i.e., if there is a flight from A to B there is always a flight from B to A, on a daily basis almost 98% of the edges are on average bidirectional. The day with the lowest percentage in 2010 has 92% of its links bidirectional. Small airports cause these minor anomalies. To simplify the analysis we symmetrized the network. As a result, on an annual basis, the resulting US air-transportation network comprises 305 commercial airports and 2,318 connections. A graphical representation of the network can be seen in Figure 2.13. Nodes are sized according to their degree.

The largest and most active airports for 2010, in the sense of the number of connections and flights, are represented in Table 2.3. The maximum degree corresponds to Atlanta International Airport (ATL) with 159 different connections and the average degree of the whole network is 15.2.

Airport code	# edges	# flights
ATL	159	809,869
ORD	147	608,981
DFW	140	524,206
DTW	128	314,369
DEN	125	470,592
MSP	116	246,245
IAH	107	362,562
SLC	94	246,245
MEM	86	152,730
MCO	83	241,851

Table 2.3: Major airports according to their degree (number of different destinations).

[4]pytz - World Timezone Definitions for Python: http://pytz.sourceforge.net

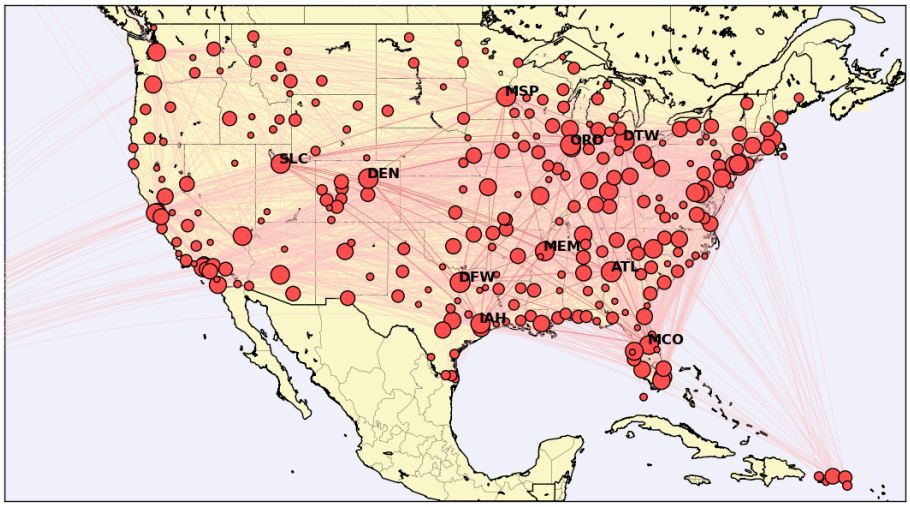


Figure 2.13: US Airport Network considering all flights from 2010. Nodes are sized according to their degree and the top 10 airports according to their degree are labeled.

Figure 2.14 depicts the complementary cumulative distribution of the number of flights and different connections for all the airports of the network in 2010. Both distributions are wide and confirm the presence of high heterogeneities in the airport network. Some few airports are large hubs with many different connections and flights while most of the airports have low traffic. These topological characteristics are well known for this network but still are relevant for the dynamics of delay propagation.

For cluster and individual airport dynamics, we will consider networks aggregated in a 24 hour period. The reason for this is that most airports are closed or have low operation activity in the early morning, causing a disruption in the flight dynamics. Figure 2.15 shows the departure probability as a function of the scheduled departure time. We can distinguish a operating zone with few or no activity that goes from 00:00 am to 04:00 am (local time).

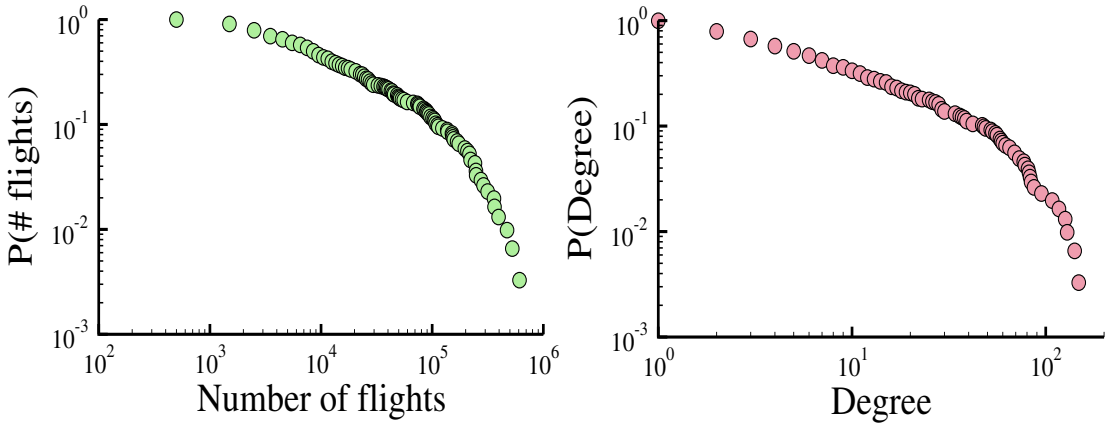


Figure 2.14: Complementary cumulative distribution of the number of flights and number of different connections (degree) for the airports in 2010.

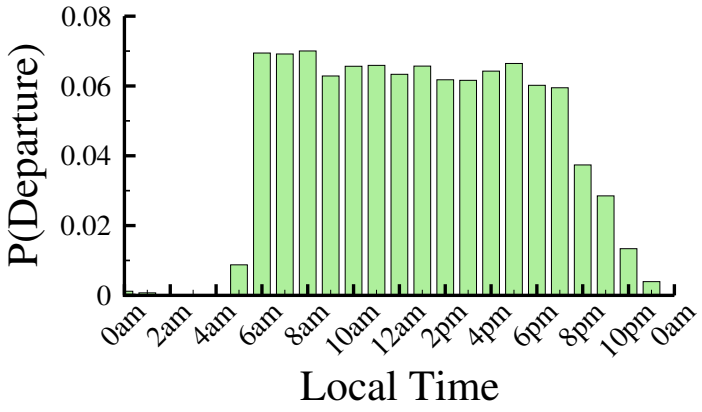


Figure 2.15: Probability of flight departure as a function of the scheduled departure hour.

2.3.4 Flight trajectories

Because the airport network can be thought as an ensemble of individual flight trajectories, understanding the airplane rotation is an important ingredient to characterize the propagation of reactionary delays. Using the airplane's tail number we trace back the airplanes movements. It can be seen that 80% of trajectories are composed of a number of leaps between 2 and 7 (Figure 2.16). Very few do longer rotations due to time constraints for the duration of the flights.

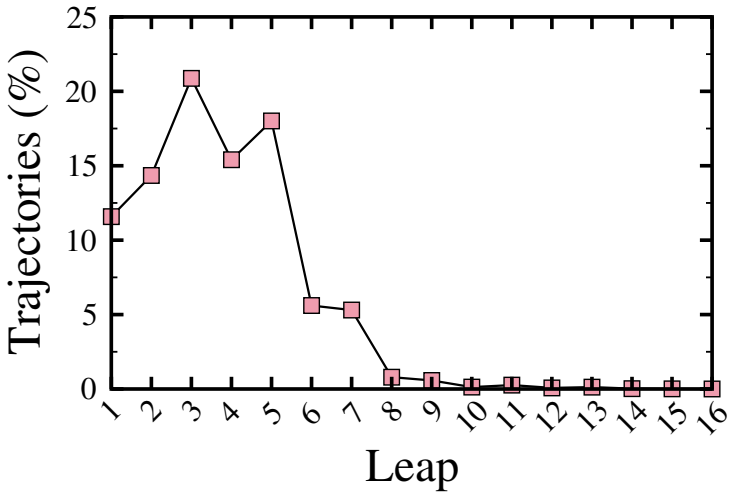


Figure 2.16: Percentage of trajectories as a function of the number of leaps.

A low percentage of aircraft trajectories follow a circular path, i.e., an airplane starts and finishes the day in the same airport (Figure 2.17). This finding does not mean that the trajectories cannot close taking into account longer periods of time (weeks, months or years).

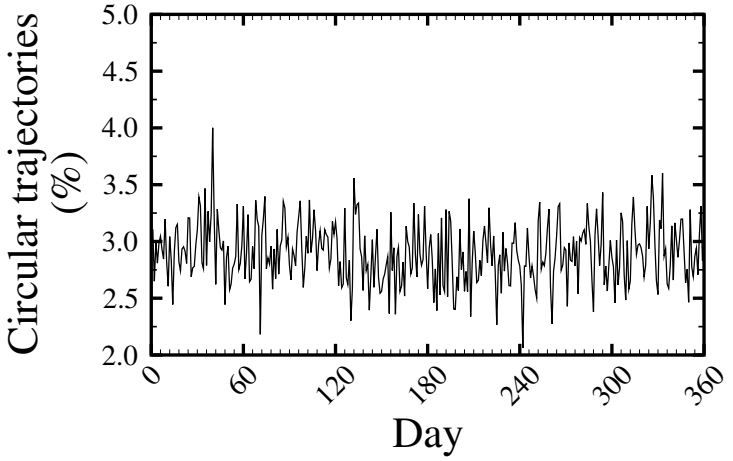


Figure 2.17: Percentage of daily trajectories that start and end in the same airport.

The airports can be classified according to the fraction of trajectories starting in them that are circular. Not necessarily, these airports are the ones with highest degree (see Figure 2.18). This clearly indicates that the network hubs (nodes with highest degree) do not always coincide with the airlines hubs. We are

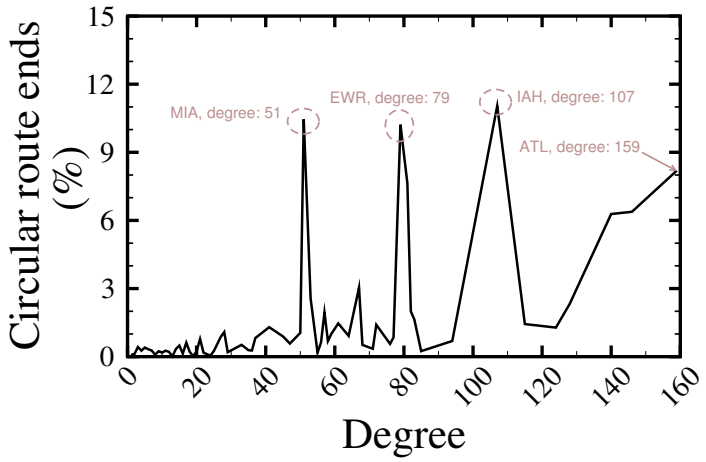


Figure 2.18: Percentage of trajectories ending at an airport as a function of the airport degree. IATA codes stands for: MIA (Miami), EWR (Newark), IAH (Houston) and ATL (Atlanta).

assuming here that the airline hubs are those airports with a larger percentage of closed rotations.

CHAPTER 3

Flight delay characterization

We have described the topology of the network and the rotation of the flights. The next step is to focus on the data regarding flight delays. There are several definitions of flight delays. For instance, according to the FAA a flight can be considered as delayed if the operation takes place 15 minutes after scheduled[1]. In our work, we follow the definition given in (Beatty et al., 1999; Rupp, 2007) and define delay as the time difference between real and scheduled operations (arrival or departure at the gate). This definition is more flexible and does not filter out small delays that can form part of a general state of congestion.

3.1 Flight delay distribution

We plot in Figure 3.1 A the complementary cumulative distribution of departure and arrival delays for 2010. Firstly, just like the degree and flight distribution, the delay distribution is broad showing a long tail. Secondly, the shape of these distributions is quite for arrival or departure. Another factor that does not modify the shape of the delay distribution is seasonality (Figure 3.1 B). As can be seen, there is no noticeable difference between the curve apart from that raising from smaller statistics.

In Figure 3.2 we depict the cumulative distribution for several airports. A peripheral airport like Honolulu International Airport (HNL) and two continental airports are displayed in the Figure: Atlanta Hartsfield-Jackson International Airport (ATL) and John F. Kennedy Airport (JFK). We can see that ATL and JFK show slight differences at the tail of the distribution. On the other hand, Honolulu displays a broader distribution. This is probably due to the longer duration of the flights with destination or origin in HNL that allow for an easier absorption of short delays. The delays in the islands can be, therefore, much larger than those in the continent and as a consequence the distribution be-

[1]FAA definition of variables: http://aspmhelp.faa.gov/index.php/APM:_Definitions_of_Variables

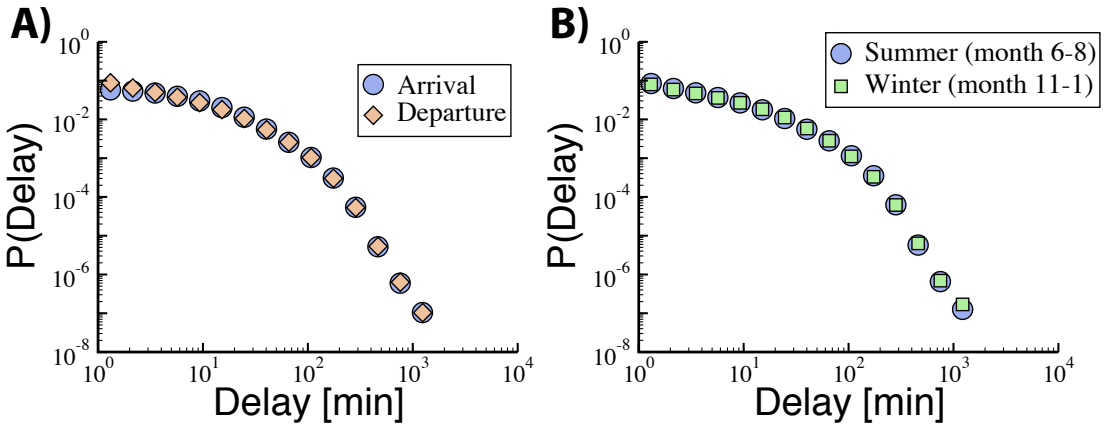


Figure 3.1: (A) Complementary cumulative distribution of delays per flight for arrival and departure. (B) Distribution of departure delays separating the flights according to the season.

comes more skewed. The heavy-tailed characteristic of these distributions is an indication of the complex nature of the delay spreading mechanisms.

In order to understand the tail of the distribution, we extract the flights with departure delay above 12 hours and compare them with all the flights of 2010. Plotting the departure delay as a function of the scheduled departure time we can distinguish how flights with delay higher than 12 hours are more abundant than the baseline at the beginning and at the end of the day (see Figure 3.3). Conversely the distribution taking into consideration all flights of the year is almost flat.

Another feature of the flights with departure delay longer than 12 hours is the relevance of the destination airport. In Table 3.1, we compare the data for long delayed flights with two sets of randomly selected flights: one among all the flights and the other with flights selected among those with any positive delay. From the data 51 airports (16.00%) are the destination of 414 delayed flights. If the 414 flights are randomly chosen, the number of destination airport increases up to 120 (more than double the results from the real data) regardless of the way we choose the flights. This shows a bias towards a smaller set of destination airports. Note that the same phenomenon is not observed for the departure airports that are in the same range both in the data and in the randomly selected flights.

Other variables as days, tail-number or air carries remain the same. This significance of the destination airports could be related to GDP or Ground Delay Program from the FAA Chang et al. (2001). This program is implemented to control air traffic volume to airports where the estimated demand is expected to surpass the Airport Arrival Rate. When a GDP is issued flights destined to the affected airport are not permitted to depart until their Controlled Departure Time. In Figure 3.4, we plotted the number of flights with long delays versus

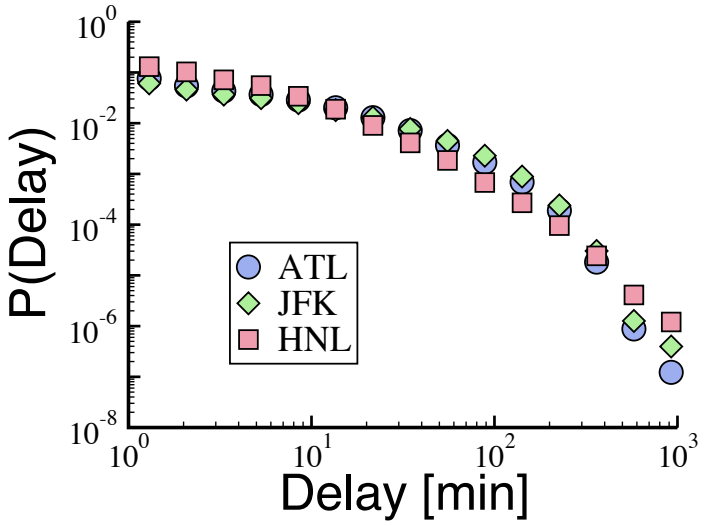


Figure 3.2: Complementary cumulative delay distribution for flights departing from Atlanta Hartsfield-Jackson (ATL), New York John F. Kennedy (JFK) and Honolulu (HNL) airports.

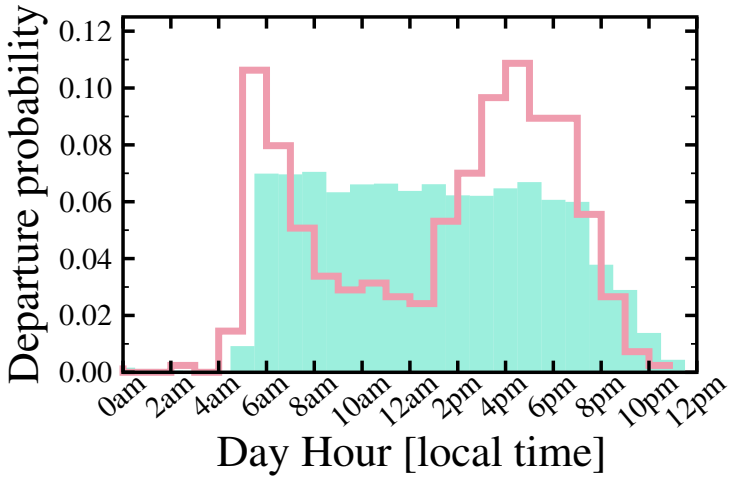


Figure 3.3: Probability of flight departure as a function of the scheduled departure hour. Blue bars represent the probability taking into account all flights of 2010. In red is the probability of flight departure only for those flights whose departure delay is 12 hours or more.

FLIGHTS WITH DEPARTURE DELAY ≥ 12 HOURS						
	Flights	Origin	**Destination**	Days	Planes	Airlines
With Problems	414	118	51	226	314	14
Total	6,341,340	305	305	365	5,081	18
Percentage	0.01%	38.0%	16.0%	62.0%	7.0%	77.0%
RANDOMLY CHOSEN 414 FLIGHTS						
	Flights	Origin	**Destination**	Days	Planes	Airlines
With Problems	414	114	120	248	392	18
Total	6,341,340	305	305	365	5,081	18
Percentage	0.01%	38.0%	39.0%	67.0%	8.0%	100.0%
RANDOMLY CHOSEN 414 FLIGHTS DELAYED						
	Flights	Origin	**Destination**	Days	Planes	Airlines
With Problems	414	112	120	246	383	18
Total	6,341,340	305	305	365	5,081	18
Percentage	0.01%	36.0%	39.0%	67.0%	7.0%	100.0%

Table 3.1: Statistical analysis of flights with departure delay higher than 12 hours.

the ranking of destination airport with respect to the number of long delayed flights. The data correspond to the blue bars while the randomly selected set of flights are the red curve. In the data, the first 8 airports are destination of 75% of the long delayed flights, while in the randomly selected set the first 8 airports totalize only 52%.

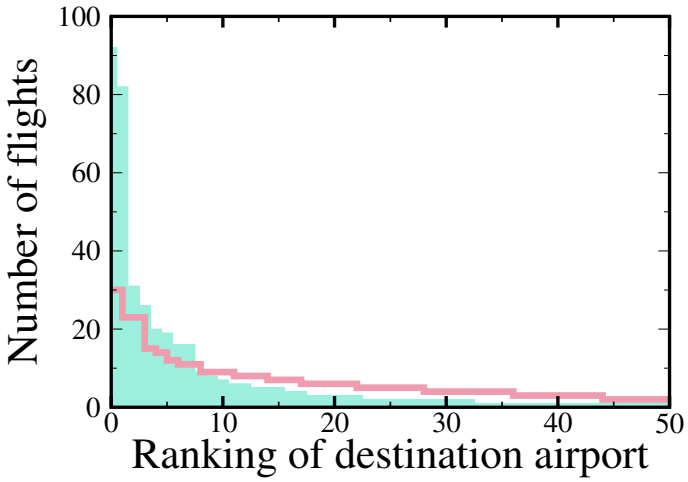


Figure 3.4: Ranking of the number of flights delayed 12 hours or more for the 51 destination airports for the data (blue bars) and the randomly selected airports (red line). For the sake of clarity, from the 120 destination airports from the random case we only plot the first 51 airports.

3.2

Average delay per delayed flight and Turn Around Time

The average delay per delayed flights (those with positive delays) for 2010 is 29 minutes. Furthermore, the annual average delay for the years between 2006 and 2009 is, as well, approximately 29 minutes. This value is used to define when an airport is considered congested an used for days of operational problems (unsatisfactory) or not (satisfactory). Those are respectively days whose average delay per delayed flight is over or below 29 minutes, respectively. Table 3.2 shows the ranking of the 20 best and worst days of the year according to their average delay for flights with positive delay.

Another informative measure is the Turn Around Time (TAT), which stands for the time spent by an aircraft on ground from arrival to departure from the gate. We refer as ΔTAT to the difference in TAT between scheduled and actual (real) times:

$$\Delta TAT = TAT_{\text{sch}} - TAT_{\text{act}} \tag{3.1}$$

This difference between the planned time an aircraft remain parked at the gate and the real time is related with the aircraft ground handling efficiency and accuracy. Hence, Turn Around Time is associated with schedule punctuality, airport

Unsatisfactory days		Satisfactory days	
DATE	Average delay (mins.)	DATE	Average delay (mins.)
Oct, 27	54.3	Apr, 19	16.9
Mar, 12	53.0	Oct, 09	17.2
Dec, 12	51.9	Nov, 11	17.3
Jan, 24	49.8	Apr, 14	17.6
Feb, 24	49.1	Oct, 08	18.0
May, 31	46.8	Set,11	18.4
May, 21	45.5	Apr, 15	18.4
May, 14	44.6	Oct, 13	18.5
Jun, 23	44.6	Apr, 17	18.5
Jul, 13	44.3	Nov, 10	18.8
Jun, 24	42.8	Nov, 09	18.9
Jul, 12	42.7	Mar, 06	19.1
Jan, 21	41.5	Oct, 12	19.2
Jul, 29	41.4	Mar, 17	19.3
Jun, 15	41.2	Feb, 28	19.5
Jun, 27	40.5	Oct, 16	19.5
Mar, 20	40.5	Apr, 13	19.5
Mar, 11	39.9	Nov, 26	19.5
Aug, 22	39.7	Set, 09	19.6
Jan, 25	39.5	Set, 20	19.7

Table 3.2: Ranking of the 20 worst/best days of the year 2010 according to their daily average delay for flights with positive delay.

operational efficiency and is a key factor in maintaining flight connectivity and aircraft rotational sequence stability (Wu and Caves, 2000).

To study how congestion impacts airport operations, in Figure 3.5, we compare two opposite scenarios: March 12, one of the days with the largest average delay and April 19, the lowest average of the year (see Table 3.2). It is clear how a congested day is characterized by higher fluctuations. The intra-day evolution of ΔTAT for Atlanta shows the capability of airports to recover, positive values of ΔTAT, even in one of the worst days of the year (Figure 3.5 A).

The distribution of ΔTAT also shows long tails both in the positive and negative values (Figure 3.6), another indication of the complex nature of this phenomenon.

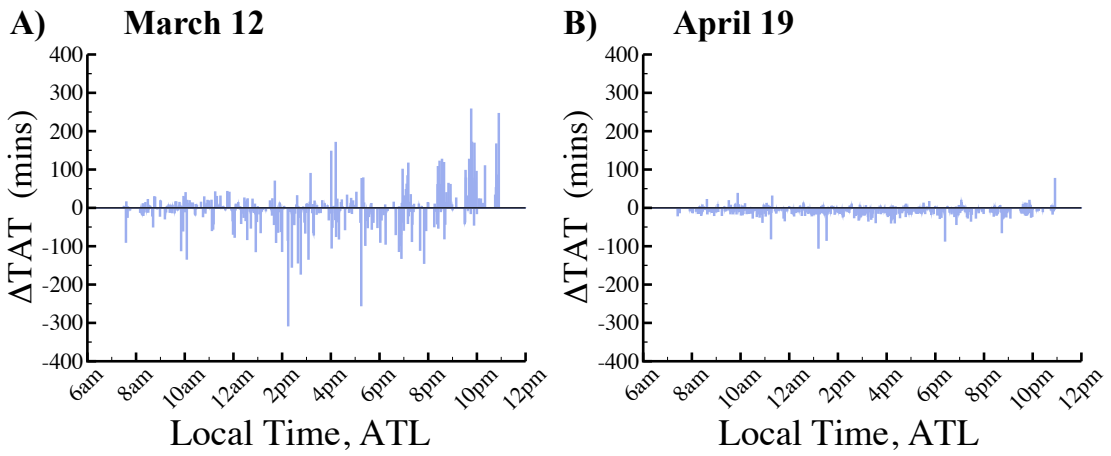


Figure 3.5: Difference between the scheduled and real Turn Around Time (ΔTAT) for operations in ATL on March 12 (A) and April 19 (B).

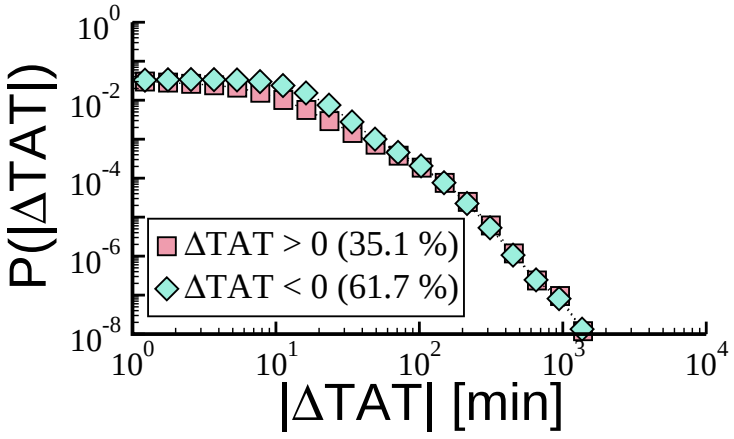


Figure 3.6: Distribution of the absolute value ΔTAT per flight, separating positive and negative contributions.

3.3

Cluster and airport dynamics

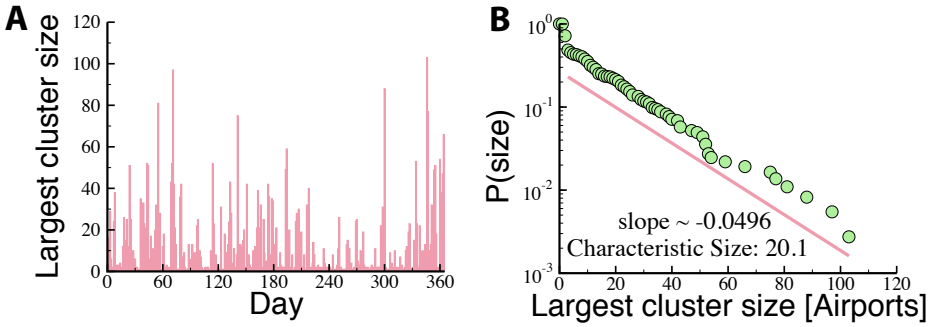


Figure 3.7: (A) Daily size of the largest cluster. (B) Complementary cumulative distribution of the size of the largest cluster (log-normal scale).

The focus so far has been on individual flight delays. We define now a metric of congestion for the full network. As mentioned in the previous section we consider an airport as congested during a given time period when the average delay of all its departing flights in that period exceeds the 29 minutes threshold. Because of the low operating activity in the early morning that disrupts the delay propagation dynamics from day to day (Chapter 2), a daily airport network is built using to assess whether congested airports form connected clusters. Note that being in the same cluster is a measure of spatio-temporal correlation of congestion but not necessarily a sign of a cause-e ffect relation. Maps with the congested airports and the connections between them are shown for di fferent days in Figures 3.8 A-C. We analyzed days with di fferent level of congestion given by the daily average departure delay: March 12 high congestion, April 19 low congestion (see Table 3.2) and March 9 in order to explore what happen at an intermediate level of congestion we selected. The scenario dramatically changes from day to day: in some days a large cluster surges covering 1 / 3 of all airports (high congestion), while in others only one or two airports cluster together (low congestion). At an intermediate level some airports rise as congested but they are not able to merge into a cluster. These behaviors indicates that connectivity

is thus an important factor to produce high congestion and consequently delays are propagating through connected airports in an intra-day time period.

Taking into account all days of 2010 the largest connected cluster size is explored for each day (Figure 3.7 A). A strong variability is thus the main characteristic of the delay dynamics, with congested days scattered among low congested ones. The cumulative distribution of the cluster size is displayed (Figure 3.7 B) is compatible with an exponential decay.

In-depth analysis considering an intra-day basis shows the characteristics of delay propagation dynamics (Figure 3.9). Not only the cluster size changes significantly from day to day, but in a congested one (Figure 3.9 A) the cluster emerges and decays, from hour to hour, in a recognizable pattern that involves days marked with a network-wide spreading of the delays (further examples in the following Chapter). In this case we can differentiate a growing and a decaying phase, that do not occur in a low congested day (Figure 3.9 B).

This feature combined with the variation in the number of clusters throughout the day says much about the importance of connectivity of airports in the network. We distinguish again two different behaviors for congested and uncongested days. The former, exhibits a fluctuating number of clusters throughout the day, implying that no stable merging of clusters takes place (Figure 3.10).

The latter shows some clusters surging within the first hours of the morning and from then on diminishing almost monotonously. In fact, the first hours of this decaying process takes place simultaneously with the growing phase concerning the cluster size which implies the development of merging events. This high-level interaction dynamics between clusters appears to be crucial in the evolution of an unsatisfactory day, where high-degree nodes play an important role to make this merging event come about.

In addition, events involving individual nodes occur and varies dramatically from time to time. Figure 3.11 displays how nodes that belong to the largest cluster of the day vary their condition rapidly. One hour they are above the 29 minutes threshold and next they recover and vice versa. Most nodes switch from one state to the other very quickly, although some few nodes repeat their condition at least two time steps (red series).

An open question is whether the congested airports are recurrent. In Figure 3.12, we calculate the Jaccard index to compare the sets of airports in the largest cluster in consecutive days or for the top 20 worst and best days. This index is 1 if the clusters are equal and 0 if they are strictly different. Interestingly, the index is relatively low for days with large clusters, which implies that the largest congested cluster is not persistent in time. However, some airports are consistently part of the largest cluster (Table 3.3). This entails that the largest cluster is composed of a hard core regularly present in it and a group of airports that alternate in time. Although some airports appear in both lists, the order changes and both sets are not exactly equal. In these lists, there is a strong component of airports located in the West Coast. We think that this is due to the

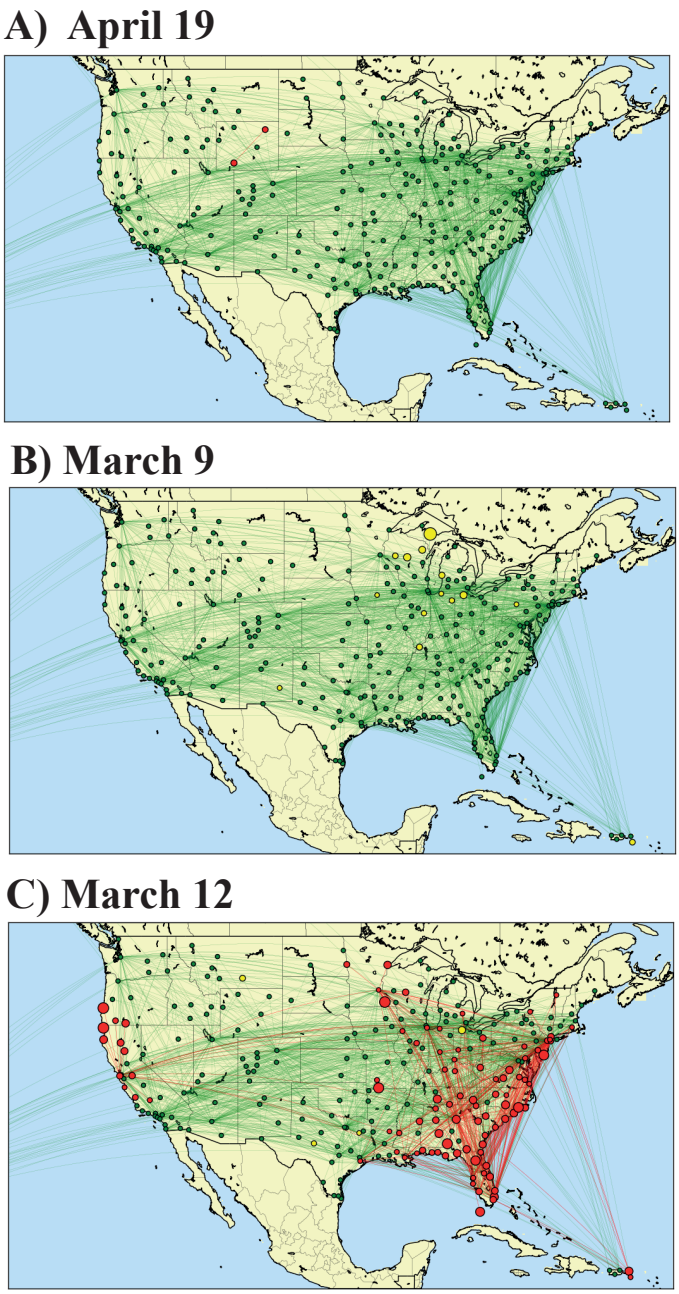


Figure 3.8: Maps of the congested airports showing also connections between them: (A) low, (B) intermediate, and (C) high level of congestion according to the daily average departure delay. The airport color codes are: red, congested airport belonging to the largest cluster; orange,congested airport not belonging to the largest cluster; green, airport not congested. Links connecting airports in the largest cluster are in red.

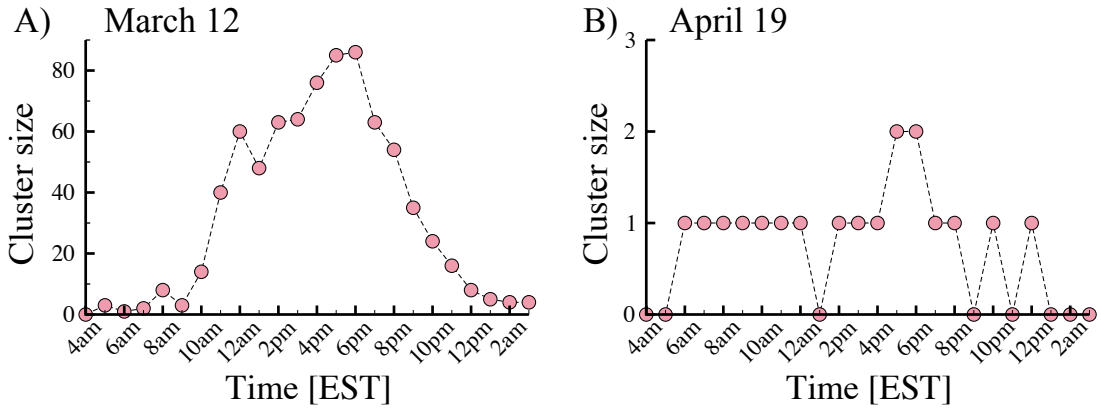


Figure 3.9: (Evolution of the largest cluster size per hour. The selected days are A) the second day with the largest delay (March 12) and B) the one with the lowest delay (April 19)

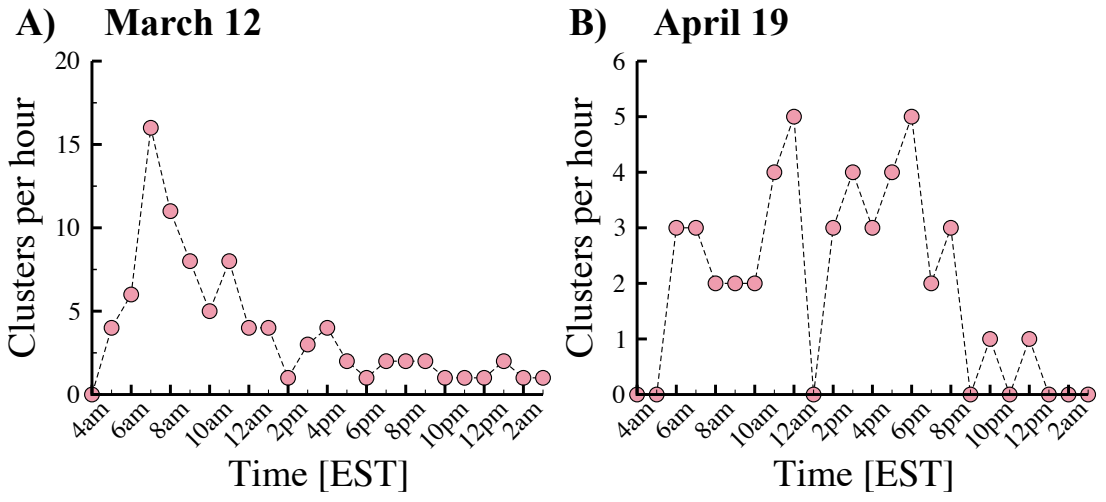


Figure 3.10: Evolution of the number of clusters for: A) March 12 and B) April 19.

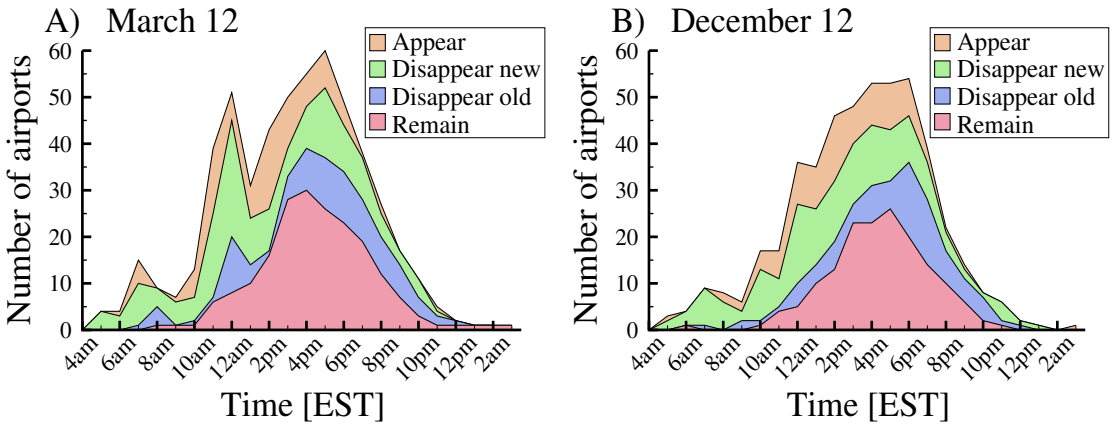


Figure 3.11: Number of airports that belongs to the largest cluster of the day for A) March 12 and B) December 12. Red color indicates the number of "old" airports (that their average departure delay per flight has been > 29 minutes at least in the previous hour as well), while new airports that match this condition are shown in orange. The nodes which their average departure delay per flight will drop below 29 minutes in the next hour are shown in green (if they have been in problem for only one hour) and blue (if they have been in problem at least for two hours).

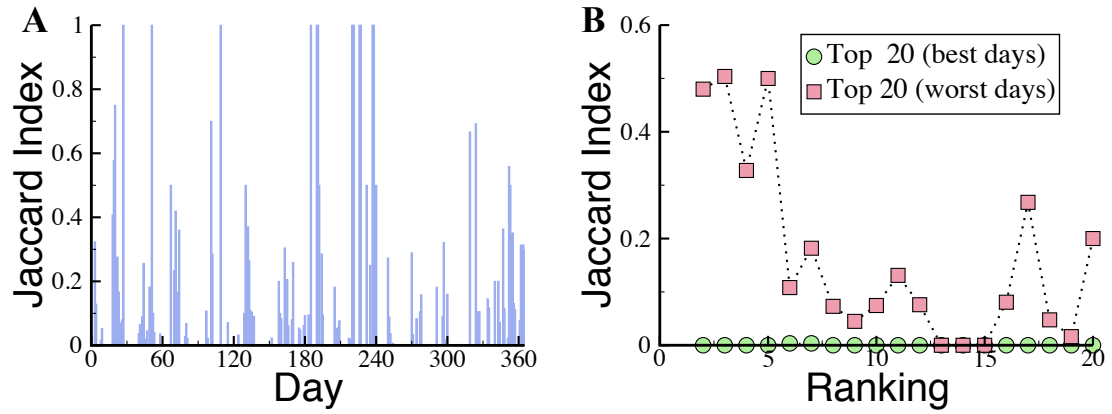


Figure 3.12: (A) Jaccard index comparing airports belonging to the largest cluster in consecutive days, or (B) ranking positions according to the top 20 days with largest or lowest average delay.

Airport code	days in largest cluster	Airport code	days with problems
ACV	100	OTH	167
CEC	80	CEC	138
SFO	54	ACV	136
OTH	52	LMT	111
MOD	49	MOD	90
EWR	45	CIC	86
CIC	45	MFR	70
LMT	44	BRW	62
MFR	43	CRW	60
CRW	41	MLB	60

Table 3.3: Top 10 raking of airports in number of days belonging to the largest congested cluster or in number of days with problems.

time difference between East and West Coasts. Flight operations initiate before in the East Coast and so the delays can propagate Westwards toward the end of the day.

The results of this Chapter evince a very different behavior between non congested and congested days. Furthermore, the broad delay distributions signals the heterogeneity present in the system. Therefore, several questions regarding the nature of delays dynamics still have to be answered and these will be addressed in the upcoming chapters.

APPENDIX A

A.1

Community detection in Real Complex Networks

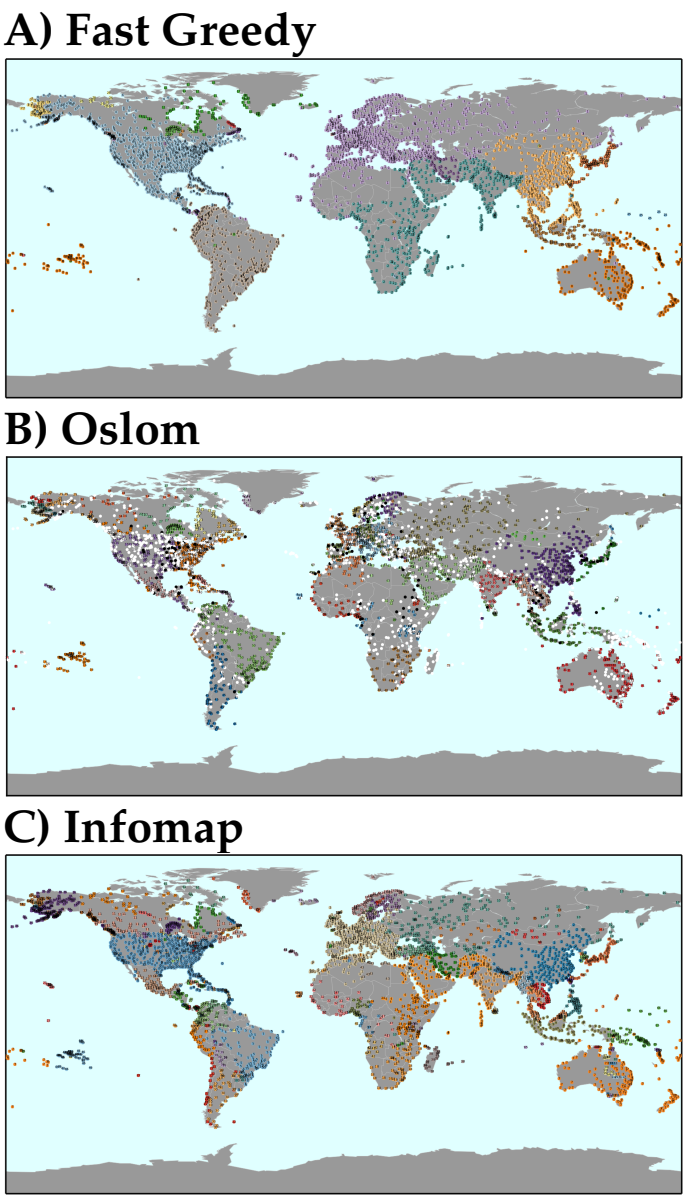


Figure A.1: Community maps of the weighted WAN network using the Fast Greedy (A), Oslom (B) and Infomap (C) algorithms. White colored airports in the Oslom community map represent homeless nodes and black ones overlapping nodes.

A.2

Other measures for assessing the quality of a community

Besides the link ratio one can also compute the ratio between the sum of the internal degree of the vertices (k_i) and the total degree of the community (k_t, sum

of degrees for all nodes belonging to a community) Radicchi et al. (2004):

$$k_r = \frac{k_i}{k_t} \,. \tag{A.1}$$

This definition can be quite lenient, since in an undirected graph internal degrees are summed twice. On the other hand, the degree ratio has the convenience of being equal to $1 - \mu$ (mixing parameter) in the LFR benchmark. In the case of weighted networks these definitions can be easily extended accounting for the sum of the link weights instead of the number of links and the node strength instead of the degree.

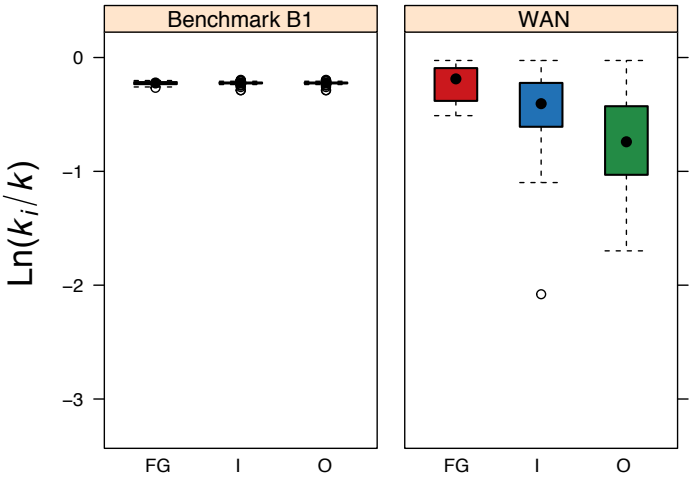


Figure A.2: Distribution of the logarithm of the degree ratio definition for the communities found by each method: Fast Greedy (FG), Infomap (I) and Oslom (O).

From Figure A.4 we can conclude that both INFOMAP and OSLOM algorithms perform really well on the synthetic network B2. They are able to find exactly the ground truth anti-communities. On the other hand, the Fast Greedy algorithm is not capable of finding these ground truth anti-communities.

APPENDIX A.

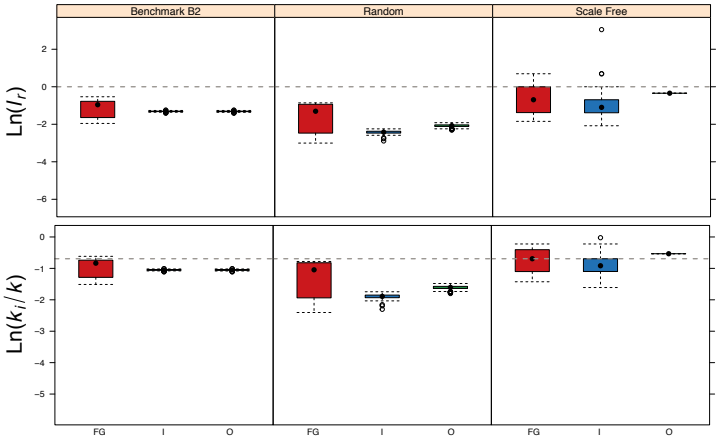


Figure A.3: Upper panel: Distribution of the logarithm of the link ratio for the artificial networks. Lower panel: Distribution of the logarithm of the degree ratio.

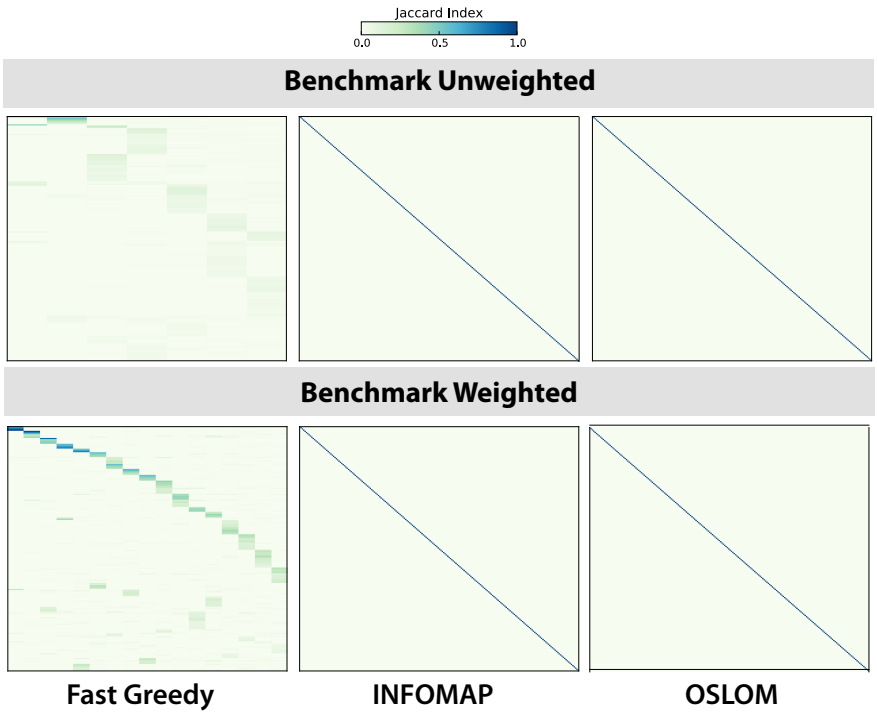


Figure A.4: Jaccard matrices for the Benchmark B2 unweighted (upper panel) and weighted (lower panel) synthetic network for each method. Each matrix row correspond to a ground truth anti-community of the benchmark and each column to a anti-community found by the algorithm.

A.3

Significant & identical communities found between the different algorithms

Figure A.5: Significant and identical communities between: OSLOM & INFOMAP (RED), OSLOM & Louvain (BLUE), INFOMAP & Louvain (GREEN).

APPENDIX B

B.1

Normalized Mutual Information

A recurrent problem faced by researchers when comparing different community detection algorithms is how to measure their efficiency. Normalized mutual information (Strehl and Ghosh, 2003) has been used by several authors to measure the performance of different algorithms in benchmarks (Fortunato, 2010; Lancichinetti et al., 2008, 2011; Danon et al., 2005; Aldecoa and Marín, 2013), where the underlying partitioning of the network is known. The mutual information $I(X, Y)$ measures the information shared between the original partition X and the one obtained by the algorithm, Y. Since this quantity has an upper bound given by the minimum entropy H between both partition, i.e., $I(X, Y) \leq \min(H(X), H(Y))$, a common normalization factor is the arithmetic average of the entropies:

$$NMI(X, Y) = \frac{2I(X, Y)}{H(X) + H(Y)}. \tag{B.1}$$

In a network with n nodes, let $C^{(a)}$ and $C^{(b)}$ be the set of clusters in partitions A and B, respectively. Define $n_i^{(a)}$, $n_j^{(b)}$ as the number of nodes in a specific cluster $i \in C^{(a)}$, $i \in C^{(b)}$, and n_{ij} as the number of nodes from cluster $i \in C^{(a)}$ found in cluster $j \in C^{(b)}$. Finally, assuming that each partition has a total of $c^{(a)}$ and $c^{(b)}$ clusters, the NMI can be obtained as

$$NMI(A, B) = \frac{-2\sum_i^{c^{(a)}}\sum_j^{c^{(b)}} n_{ij}\log\left(\frac{n \cdot n_{ij}}{n_i^{(a)} \cdot n_j^{(b)}}\right)}{\sum_i^{c^{(a)}} n_i^{(a)}\log\left(\frac{n_i^{(a)}}{n}\right) + \sum_i^{c^{(b)}} n_i^{(b)}\log\left(\frac{n_i^{(b)}}{n}\right)}. \tag{B.2}$$

This quantity is 1 when A and B are identical and tend to 0 the more dissimilar the two partitioning are.

Unfortunately, this measure can be very insensitive to errors, as can be demonstrated by the following construction. Suppose one knows the *ground truth*

communities of a given network. This will be our set $C^{(a)}$. Now let us divide each of the $c^{(a)}$ clusters in two, by taking a fraction x of the $n_i^{(a)}$ nodes and forming a new cluster with them. The set $C^{(b)}$ will be this new partitioning, with $c^{(b)} = 2c^{(a)}$ clusters. Therefore, for every original community i in $C^{(a)}$ we have two communities in $C^{(b)}$ with common nodes. For simplicity, lets order the indexes of communities in $C^{(b)}$ so that those formed by the $(1-x)n_i^{(a)}$ nodes from cluster i in $C^{(a)}$ have the same index in $C^{(b)}$. The complementary partition will be denoted by index i_s. Also, lets define Λ the set of community indexes in $C^{(a)}$ and Λ_s the set of indexes i_s. Therefore, the number of nodes in each cluster j in the set $C^{(b)}$ is

$$n_j^{(b)} = \begin{cases} (1-x)n_i^{(a)} & \text{if } j = i \in \Lambda \\ xn_i^{(a)} & \text{if } j = i_s \in \Lambda_s. \end{cases} \tag{B.3}$$

Similarly, the number of common nodes between communities in each partitioning follows the same relation:

$$n_{ij} = \begin{cases} (1-x)n_i^{(a)} & \text{if } j = i \in \Lambda \\ xn_i^{(a)} & \text{if } j = i_s \in \Lambda_s \\ 0 & \text{o/w.} \end{cases} \tag{B.4}$$

In this scenario, the calculation of the *NMI* is straightforward since the sum over the communities in $C^{(b)}$ can be split in those two groups. Making use of this property, that is,

$$\sum_j^{c^{(b)}} = \sum_{j\in\Lambda} + \sum_{j\in\Lambda_s}, \tag{B.5}$$

we obtain that the *NMI* can be written in closed form as a function of x and $H(A)$. The mutual information $I(A,B)$ gives us

$$\begin{aligned} I(A,B) &= \sum_i^{c^{(a)}} \sum_j^{c^{(b)}} \frac{n_{ij}}{n} \log\left(\frac{n \cdot n_{ij}}{n_i^{(a)} \cdot n_j^{(b)}}\right) \\ &= \sum_i^{c^{(a)}} \frac{(1-x)n_i^{(a)}}{n} \log\left(\frac{n}{n_i^{(a)}}\right) + \sum_i^{c^{(a)}} \frac{xn_i^{(a)}}{n} \log\left(\frac{n}{n_i^{(a)}}\right) \\ &= H(A), \end{aligned} \tag{B.6}$$

while the entropy of partition B is

$$\begin{aligned} H(B) &= -\sum_i^{c^{(b)}} \frac{n_i^{(b)}}{n} \log\left(\frac{n_i^{(b)}}{n}\right) \\ &= -\sum_i^{c^{(a)}} \frac{(1-x)n_i^{(a)}}{n}\left[\log(1-x) + \log\left(\frac{n_i^{(a)}}{n}\right)\right] \\ &\quad -\sum_i^{c^{(a)}} \frac{x n_i^{(a)}}{n}\left[\log x + \log\left(\frac{n_i^{(a)}}{n}\right)\right] \\ &= -\log\left(x^x\,(1-x)^{1-x}\right) + H(A). \end{aligned} \tag{B.7}$$

Therefore, the Normalized Mutual Information is simply

$$NMI(A,B) = \frac{2H(A)}{2H(A) - \log\left(x^x\,(1-x)^{1-x}\right)}, \tag{B.8}$$

where $H(A)$, the entropy associated to the original partitioning, is given by

$$H(A) = -\sum_i^{c^{(a)}} \frac{n_i^{(a)}}{n} \log\left(\frac{n_i^{(a)}}{n}\right). \tag{B.9}$$

It is immediate to notice that the minimum value for the NMI in this situation is when $x = 0.5$, as should be expected since this construction is symmetric. Therefore, the minimum is simply

$$\min_x \{NMI(A,B)\} = \frac{H(A)}{H(A) - 0.5\log 0.5}. \tag{B.10}$$

Equation B.8 allows us to construct a heat map in the phase space $(H(A), x)$ to analyze how sensitive is the Normalized Mutual Information when faced with this construction (Fig. B.1). With that equation we can see that even for small values of entropy ($H \sim 1$) we already have a minimum value of $NMI > 0.85$.

Of course, there is still the question of what are the typical values for the entropy. As Eq. B.9 shows, this quantity will depend on the specific distribution of community sizes. To give an illustration of possible values, we present three scenarios from typical benchmark structures. In the case of homogeneous community size, that is, when all communities have the same size $n_i^{(a)} = n/c^{(a)}$, the entropy H_h is simply

$$H_h = \log c^{(a)}. \tag{B.11}$$

Therefore, with more then 10 communities we are already in the area of $H > 1$ (Fig. B.2), which places us in the region of NMI closer to 1 even for $x \sim 0.5$. For instance, the Girvan-Newman benchmark (Girvan and Newman, 2002) uses 4 communities of homogeneous size, which would already give a minimum of 0.8 with this construction.

APPENDIX B.

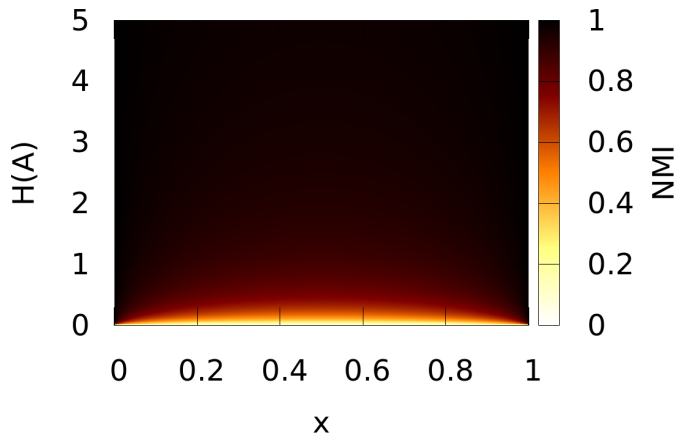


Figure B.1: Normalized Mutual Information variation with respect to the network community structure entropy H and the fraction of nodes used to form sub-communities from the original ones. Each original community in the partitioning A is divided in two, with a fraction of x and $1 - x$ of its original nodes in each.

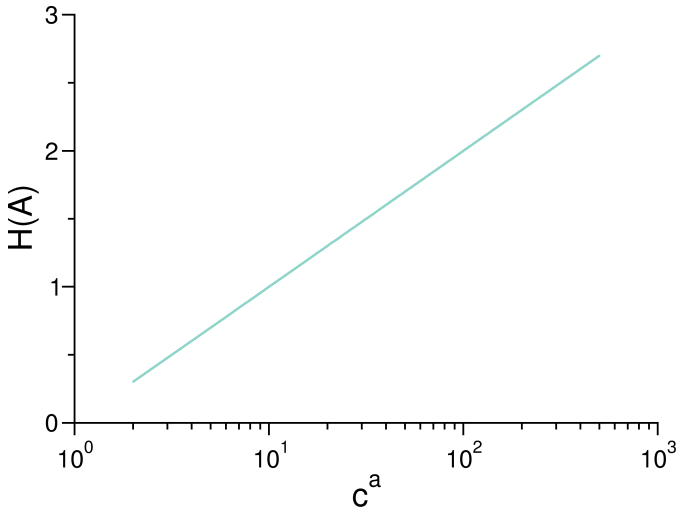


Figure B.2: Entropy $H(A)$ as a function of number of communities $c^{(a)}$ in the case of homogeneous community size distribution.

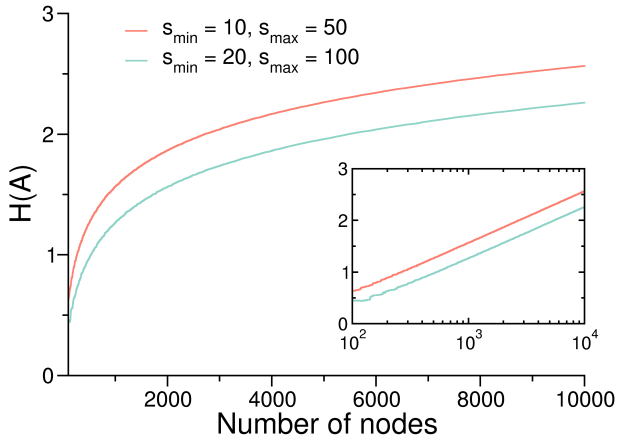


Figure B.3: Entropy $H(A)$ as a function of number of nodes in the network in the case of community size distribution $\sim s^{-1}$ with two different specified boundaries for minimum and maximum size. The *small* scenario has $s \in [10, 50]$ (red line) while the *big* one has $s \in [20, 100]$ (blue line). The inset presents the same plot in semi-log scale.

While the homogeneous size distribution might illustrate the situation for the mentioned benchmark, for large scale real world networks one expect to find a heterogeneous distribution, say a power-law, for the communities size. With that in mind, the LFR benchmark Lancichinetti et al. (2008, 2011) has been developed for constructing communities with such distribution. In their work, they compare different community detection algorithms using two scenarios of their benchmark, both with size distribution with an exponent of -1 and two different boundaries. In the first case, called *small*, the communities have a fixed minimum size $s_{min} = 10$ and maximum $s_{max} = 50$, and *big* scenario where $s_{min} = 20$ and $s_{max} = 100$. By inspecting the curves of entropy obtained in those two scenarios for different network sizes(Figure B.3), we can see that for networks with more than 1000 nodes we reach an entropy larger than 1 even in the best case scenario of *big* communities. This leads to the result that even for small networks ($n \sim 100$) the NMI barely lowers from 1 when the communities are disrupted (see Fig. B.4).

These results should not be interpreted as diminishing the utility of the benchmarks mentioned, but rather as shedding light on the lack of sensitivity of the NMI when used to evaluate the success of community detection algorithms on those benchmarks.

So far, the construction proposed divides the communities in two but preserves the subtracted nodes together. That is, nodes that are originally from different communities are not mixed on the new partitioning. Such construction could be deemed unfair to NMI, since it preserves a higher overlap between the

APPENDIX B.

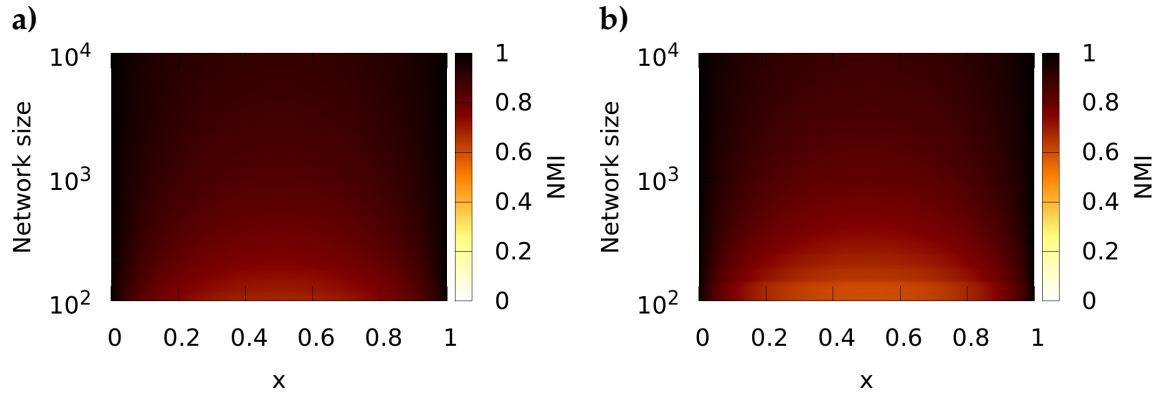


Figure B.4: Normalized Mutual Information variation with respect to LFR networks of size n and the fraction of nodes used to form sub-communities from the original ones. Each original community in the partitioning A is divided in two, with a fraction of x and $1-x$ of its original nodes in each. The original community size distribution is a power-law with exponent -1 and lower(upper) boundary $s_{min}(s_{max})$ fixed. a) $s_{min} = 10$, $s_{max} = 50$. b) $s_{min} = 20$, $s_{max} = 100$.

new artificial communities and the original ones, resulting in higher mutual information. The other extreme would be to mix the subtracted nodes from all communities and draw at random the new ones, keeping the sizes constrained at xn_i, which would give the lowest mutual information possible with such constraint. On the other hand, if a detection algorithm is good, we expect that when it fails to detect a community in its entirety, it should give subgroups within that set of nodes, with a low occurrence of mixture. Therefore, we expect the mutual information to lay between the preservation of subgroups (upper bound) and the random structure (lower bound).

To compute the NMI for the randomized case, again we have to determine the overlap between the original communities and the ones in the new partition. As in the previous case, from each original community i in the set $C^{(a)}$ we will remove $xn_i^{(a)}$ nodes and preserve the remainder. As before, for simplicity we will preserve their indexes and make use of the label sets Λ and Λ_s The new communities will be formed by $(1-x)n_i^{(a)}$ and $xn_i^{(a)}$ nodes. The first ones will preserve its $(1-x)n_i^{(a)}$ original nodes, while the former will be populated by nodes chosen at random from the pool of *orphaned* nodes. This pool is formed from the collection of the $xn_i^{(a)}$ subtracted from all original communities $i \in C^{(a)}$. Since we assume that every node belongs to a single community, this pool will have $\sum_i xn_i^{(a)} = xn$ nodes. Therefore, the probability of picking a node that was originally from community i is simply $n_i^{(a)}/n$. This allows us to calculate the expected number of nodes from cluster $i \in C^{(a)}$ found in cluster $j \in C^{(b)}$, which is

given by

$$n_{ij} = \begin{cases} (1-x)n_i^{(a)} & \text{if } j = i \in \Lambda \\ xn_j^{(a)} n_i^{(a)}/n & \text{if } j = i_s \in \Lambda_s \\ 0 & \text{o/w.} \end{cases} \tag{B.12}$$

Since the entropy is a function of the community size distribution alone, this construction does not alter $H(B)$. The only quantity affected in Eq. B.2 is the mutual information itself, $I(A,B)$. Once more, we will make use of the splitting of the sum over communities in $C^{(b)}$, so that

$$\begin{aligned} I(A,B) &= \sum_i^{c^{(a)}} \sum_j^{c^{(b)}} \frac{n_{ij}}{n} \log\left(\frac{n \cdot n_{ij}}{n_i^{(a)} \cdot n_j^{(b)}}\right) \\ &= \sum_i^{c^{(a)}} \frac{(1-x)n_i^{(a)}}{n} \log\left(\frac{n}{n_i^{(a)}}\right) + \sum_i^{c^{(a)}} \sum_j^{c^{(a)}} \frac{xn_j^{(a)}}{n} \frac{n_i^{(a)}}{n} \log\left(\frac{n \cdot xn_j^{(a)} n_i^{(a)}/n}{n_i^{(a)} \cdot xn_j^{(a)}}\right) \\ &= (1-x)H(A). \end{aligned} \tag{B.13}$$

Since both $H(A)$ and $H(B)$ are the same as in the previous structure, by comparing Eq. B.13 to Eq. B.6 we see that the NMI will have the same expression as in Eq. B.8, multiplied by $1-x$.

$$NMI(A,B) = (1-x)\frac{2H(A)}{2H(A) - \log\left(x^x (1-x)^{1-x}\right)}. \tag{B.14}$$

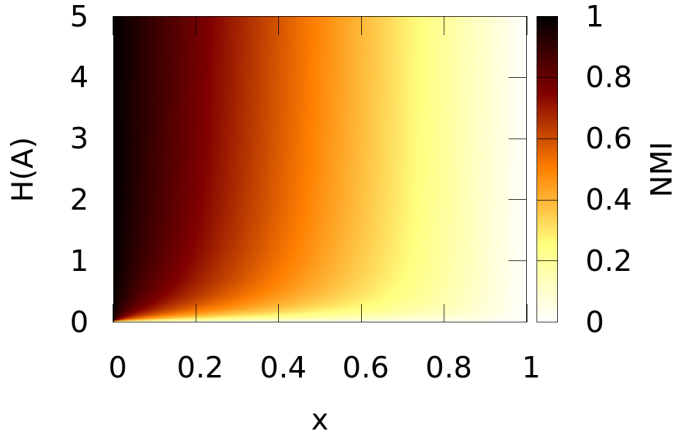


Figure B.5: Normalized Mutual Information variation with respect to the network community structure entropy H and the fraction of nodes used to form sub-communities from the original ones. Each original community i in the partitioning A has x of its n_i nodes removed. For each one of those communities, a new one is formed by taking at random xn_i nodes from the pool of all orphaned nodes.

From Figs. B.5- B.6 we can see that, although this scenario presents a better picture for the NMI measure, we still have to remove more than 20% of the nodes from each community to be able to detect a drop below 0.8 in almost all situations. One would need a relatively small network ($n \sim 100$) and *big* clusters in order to have a noticeable decrease in the NMI for a removal of 10% of the nodes in each community, as shown in Fig. B.6b. But it is important to notice that in this situation, since the lower boundary for the community size is already 20 nodes, we would have a maximum of ~ 5 communities, each with $\sim 20\%$ of the nodes in the network.

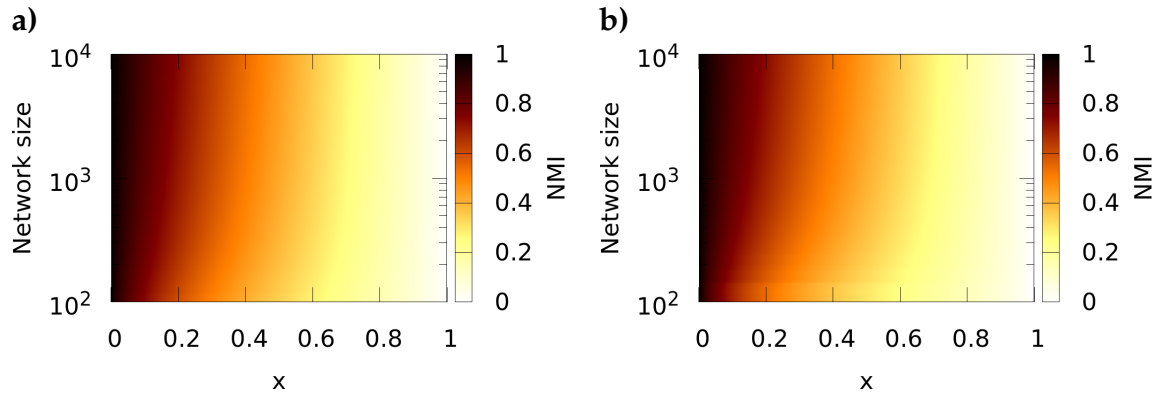


Figure B.6: Normalized Mutual Information variation with respect to LFR networks of size n and the fraction of nodes used to form sub-communities from the original ones. Each original community i in the partitioning A has x of its n_i nodes removed. For each one of those communities, a new one is formed by taking at random xn_i nodes from the pool of all orphaned nodes. The original community size distribution is a power-law with exponent -1 and lower(upper) boundary $s_{min}(s_{max})$ fixed. a) $s_{min} = 10$, $s_{max} = 50$. b) $s_{min} = 20$, $s_{max} = 100$.

Part III

Agent-based modeling

CHAPTER 4

NewCat: Network-wide Congestion Assesment Tool

We propose an agent-based model at the level of aircrafts and data-driven in the sense that the daily schedules and the primary delays are obtained directly from real records in the database. The model also combines within the same framework queuing behavior and a schedule based approach dynamics. This level of realism is necessary to confront the model predictions with the real unfolding of the delay events during each day. Importantly, the schedule is not the a priori airline's flight plans, meaning that the data that we have is modified by the events occurred during a day. However, for simulation purposes we will initially use days with low or almost no heavy external perturbations, and therefore days with low levels of canceled and diverted flights are used.

The purposes of this model is to understand how delays propagate and magnify considering internal operational factors and schedule. As it will be explained further below, "extrinsic" or primary delay is given at the initial conditions of the simulation to the first flight of the day for some aircraft rotations, and then let this perturbation evolve multiplying or diminishing the delay according to the particular structure of the system. Concretely, the model dynamics will be based on three subprocesses which are: (i) aircraft rotation, (ii) flight connectivity and (iii) airport congestion. The last two are independent from each other, and can be turned on/off to explore the relevance of each subprocess in the delay propagation dynamics. Aircraft rotation, on the other hand, is intrinsic to the schedule and so we do not switch it off.

We use one-minute intervals as the basic time step unit in the model and proceed in each simulation until the schedule of a selected day is completed (all flights had completed their itinerary). In most cases, this means slightly more than $1,440$ minutes. This time interval allows the simulation to execute actions at a realistic concurrent time-scale and is the finest level available in the data. As shown in the previous Chapter most airports have a operating zone with few activity between from 00:00 to 04:00 am local time. It should also be noted that the United States has six different time zones: four of them corresponding to the continental United States (eastern, central, mountain, pacific), and the other two corresponding to Alaska time zone (one hour behind pacific time) and Hawaii

time zone (one hour behind Alaska for most part of the year). We should decide which time zone is best appropriated and the time to start the simulations in the model. Given that most US traffic is initially located in the East and low operating zone between 00:00 and 04:00 (local time) we select 4am East Coast local time as the starting point for airport operations and to begin the aircraft rotational sequences. By selecting 4 am Eastern Time (EST) as the starting point we ensure that most aircraft rotational sequences are sorted correctly and is the natural choice considering the daylight time flow in the United States. So as to arrange the schedule in a sequential order we converted time data from Local to Eastern Time.

4.1

Hierarchy of Objects

4.1.1 Aircraft (tail-number)

The airplane is the primary fundamental agent of the simulation. The number of airplanes that participate in the simulation varies with the day considered and is approximately $4,000$. Each aircraft is unique and comes identified by their tail number. This code allows us to reconstruct the rotational sequence of the plane during the day. This sequence can be subdivided in individual flight legs or point-to-point flights.

4.1.2 Point-to-point flight

This is the basic schedule unit. It is the minimum package of information used as an input to relocate an aircraft from an origin to a destination airport, meeting the planned schedule. During their itinerary an aircraft can be in one of two flight phases: block-to-block or turn-around phase. The former is the time elapsed from the airport origin gate to the airport destination gate. The latter is defined as the time the aircraft remains parked at the airport gate (Figure 4.1).

Flights are characterized by a tail-number, origin airport, destination airport, schedule departure time ($T_{\text{sch.d}}$) and schedule arrival time ($T_{\text{sch.a}}$). Block-to-block time (T_{b}^{ij}) for an aircraft between two airports i (origin) and j (destination) is calculated as:

$$T_{\text{b}}^{ij} = T_{\text{sch.a}}^{j} - T_{\text{sch.d}}^{i} \tag{4.1}$$

Another issue worth noting in our model is that, in the block-to-block phase we do not allow for delay absorption or reduction. This can only be achieved in the turn-around phase by means of the difference between the actual arrival time of the previous flight leg and the scheduled departure time of the next flight leg.

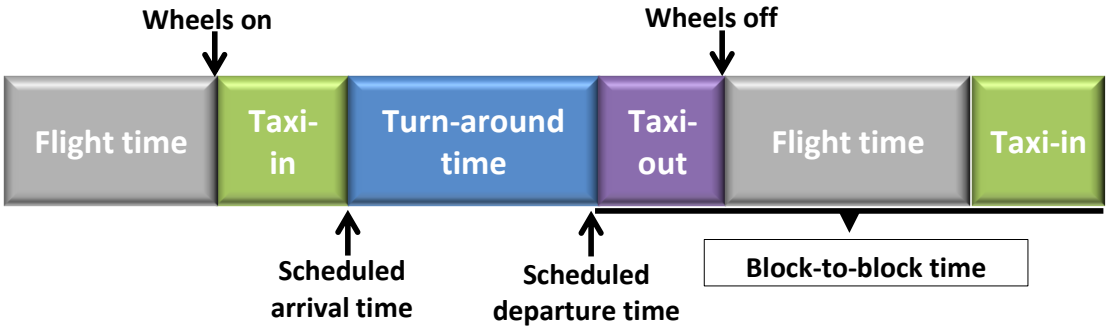


Figure 4.1: Turn-around and block-to-block time/phase definition.

4.1.3 Air carrier (airline id)

Air carriers are the second level in the model. Each aircraft has an airline associated via the airline code id. Only aircrafts having the same airline id are allowed to interact during the process of flight connectivity (see section 4.2.2 for further details).

4.1.4 Airport

The airport is an intermediate-level entity located in space coordinates, where interactions among aircrafts take place. This interaction occurs indirectly through the schedule, flight connections or airport queues (see section 4.2.3 for further details). Each airport is different from the others because of their planned capacity and the local aggregation of the schedule. Airports play the role of nodes in the transport network.

4.1.5 Clusters of congested airports

This is a high-level entity that represents the interactions between airports. The clusters are formed by airports whose average (departure) delay per flight is higher or equal to 29 minutes and are linked by a direct connection (see C.1.1 for further details). In most cases, we are interested in the largest cluster of the full day (or by hour in some cases). The size of a cluster is measured by the number of airports that belong to it.

Figure 4.2 shows a representation of two clusters (Cluster A and B) constituted by airports whose average departure delay per flight in a certain time period is equal or larger than 29 minutes (red dots). Apart from this condition airports within these clusters are linked by a connection. In this case, cluster A correspond to the largest cluster in a certain time period according to the number of airports that form this cluster.

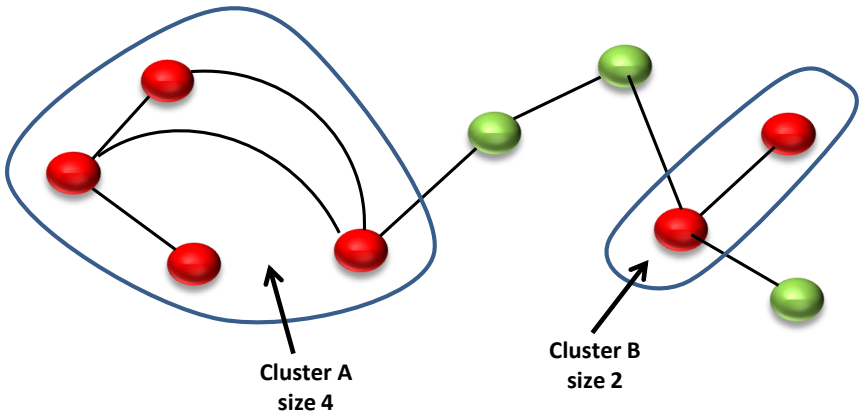


Figure 4.2: Red dots correspond to airports whose average departure delay per flight in a certain time period is > than 29 minutes. Green dots correspond to airports whose average departure delay per flight in a certain time period is ≤ than 29 minutes.

4.2

Subprocesses

4.2.1 Aircraft rotation

During a day each aircraft has an itinerary to accomplish, which for approximately 88% of the trajectories consists of two or more flight legs. Naturally, to complete a flight leg, the previous ones have to be fulfilled, e.g., it is not possible to depart from San Francisco to Honolulu if the airplane has not completed the previous leg from Atlanta to San Francisco. Besides this evident situation, if an aircraft arrives late (inbound delay) and the delay cannot be absorbed by the turn-around time it will depart late in the next flight leg (Figure 4.3). A buffer time is included in the turn-around phase to absorb this type of delay but this is already incorporated in the schedule obtained from the data.

Another feature of this subprocess, is that in the turn-around phase each aircraft, when arrived, has to comply with a minimum service time T_s, set as 20 minutes. This service time includes operations such as refueling, passenger unboarding/boarding, luggage handling, safety inspection, etc. The equation that govern the rotation subprocess is given by:

$$T^j_{\text{act.d}} = \max[T^j_{\text{sch.d}}; T^j_{\text{act.a}} + T_s] \tag{4.2}$$

where j corresponds to destination airport and i to the origin one. The subindexes act.d,act.a and sch.d correspond respectively to Actual Departure, Actual Arrival and Schedule Departure.

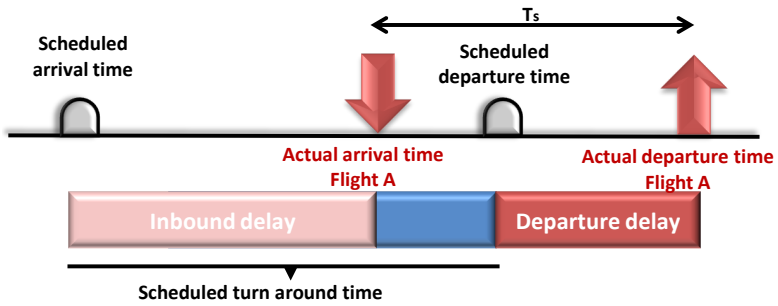


Figure 4.3: Aircraft rotation description.

4.2.2 Flight connectivity

In addition to rotational reactionary delay, the need to wait for load, connecting passengers and/or crew from another delayed airplane from the same fleet may cause, as well, reactionary delay.

To simulate this, for each flight at a particular airport, connections from that airport are randomly chosen as follows. Firstly, we take a ΔT window prior to the scheduled departure time of the flight. Secondly, we distinguish possible connections of the same airline from other flights, that have a scheduled arrival time within the ΔT window (Flights B and D in the example of Figure 4.4). Finally, from these possible connections we select those with probability α multiplied by the flight connectivity factor. The flight connectivity factor defined in the previous Chapter is the annual fraction of connecting passengers for each US commercial airport and α is an effective parameter of control that allows to modify the strength of this effect in the model. For instance, $\alpha = 0$ means that there is no connection between flights with different tail number, while $\alpha = 1$ makes the fraction of connecting flights of the same airline equal to the fraction of connecting passengers in the given airport. The parameter α is varied according to the case under study and ΔT is always taken to be 180 minutes (3 hours). We have tried different values of ΔT and those below 180 make the model very difficult to reproduce the real congestion because connections are scarce and above 180 the additional increase in connections has negligible effect because the buffer time is larger. Let us suppose that from the previous example

Figure 4.4: Possible connections within flights of the same airline.

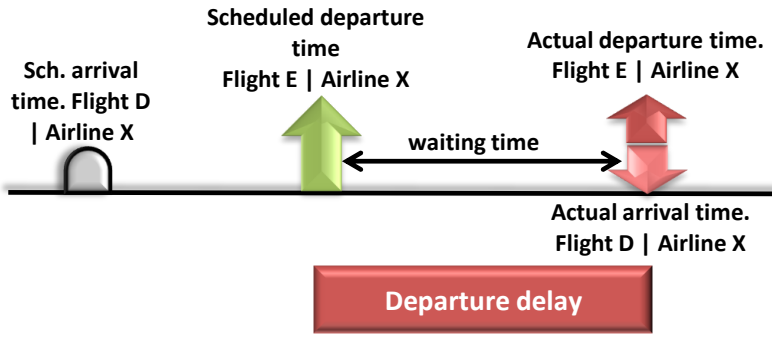


Figure 4.5: Flight connectivity description.

Flight D was randomly selected. By this subprocess an airplane is able to fly if and only if their connections have already arrived to the airport, if not it has to wait until this condition is satisfied (Figure 4.5). It is important to note that flight connectivity is the only source of stochasticity in the model due to a lack of knowledge about the real flight connections within the daily schedule. In this case the Actual Departure time of the next flight leg is given by:

$$T^{j}_{\text{act.d}} = \max[T^{j}_{\text{sch.d}}; T^{j}_{\text{act.a}} + T_{\text{s}}; \max[T^{i'j}_{\text{act.a}}]] \ , \ \forall i' \neq i \tag{4.3}$$

The index i' corresponds to the connections that the flight has to wait for in order to depart (Flight D). Notice that the rotation subprocess is included because it is intrinsic to the schedule. It is the baseline subprocess.

4.2.3 Airport congestion

Since airports are entities with a finite capacity, the possibility of congestion has to be introduced in the model. By considering this effect there exists a possibility of spreading the delays between flights of different airlines. This occurs indirectly through an airport's queue. That is to say that delays from aircrafts of different airlines can delay others because they congest the airport. In this case the propagation is not one-to-one as in the previous cases, it requires a cumulative effect of several delayed aircrafts to perturb the airport efficiency and once this condition is met the delay spread to other aircrafts and affect other airlines. We assume a "First in-First Served" queuing protocol, the most widely used queue operation in the US and simple to introduce in the model. In the simulations each airport will have a capacity that varies throughout the day according to the Scheduled Airport Arrival Rate (SAAR). This means that for every airport the nominal capacity for each hour of the day is the scheduled flights that arrive per hour (Figure 4.6). Due to reactionary delays aircrafts may not arrive as planned and the Real Airport Arrival Rate (RAAR) will vary. Whenever RAAR > SAAR, a queue begins to form with the arriving aircrafts. Naturally, airplanes that are not in queue are being served and this service time

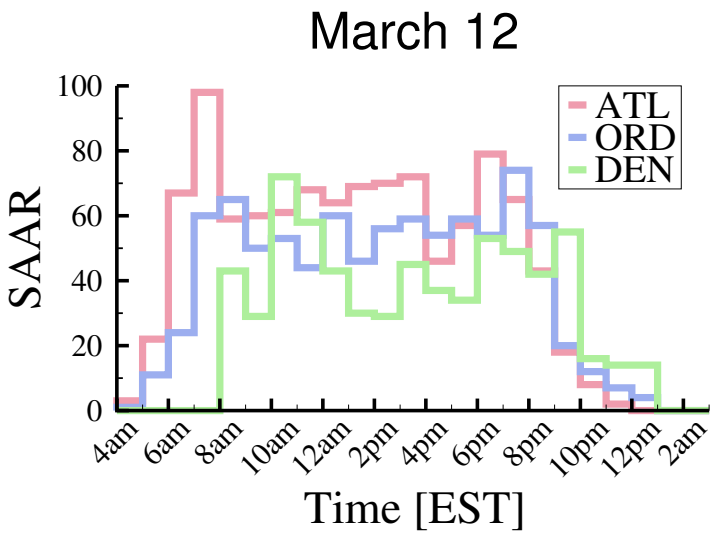


Figure 4.6: Example of SAAR for three major airports: Atlanta International Airport (ATL), O'Hare International Airport (ORD) and Denver International Airport (DEN).

takes T_s (see section 4.2.1 for further details). It should be noticed that once an aircraft starts to be served this process cannot be interrupted no matter how SAAR varies. We define another effective control parameter β in order to modify the nominal capacity of the airports. This parameter multiplies the SAAR and in the simulations presented here affects all the airports in the same way. For instance, if we want to introduce a buffer capacity of 20%, β is set to 1.2.

When aircraft rotation and airport congestion is present the equation is ruled by:

$$T^j_{\text{act.d}} = \max[T^j_{\text{sch.d}}; T^j_{\text{q}} + T^j_{\text{act.a}} + T_s] \tag{4.4}$$

where q means the time spent by the aircraft in the queue waiting to be served. Finally, the full model dynamics is govern by a combination of the three subprocesses:

$$T^j_{\text{act.d}} = \max[T^j_{\text{sch.d}}; T^j_{\text{q}} + T^j_{\text{act.a}} + T_s; \max[T^{i'j}_{\text{act.a}}]] \ , \ \forall i' \neq i \tag{4.5}$$

4.3 Initial conditions

Initial condition refers to the situation of the first flight of an aircraft sequence, meaning when, where and the departure delay of this flight. As will be later shown, variations on this situation can have a great impact on the delay prop-

agation. In other words, as we will see in the next Chapter the dynamics of delays over the network is highly sensitive to the initial conditions.

We took as initial conditions the average delay per flight for the first flights of all the aircraft sequences and by the fraction of rotations whose first flight is delayed fraction of airplanes. Comparing the ranking of the 20 worst and best days of 2010 (Figure 4.7) we can observe that it is most likely that if a day started with unfavorable initial conditions it will likely produce large congested clusters. We initialize the simulations by two different ways depending on the

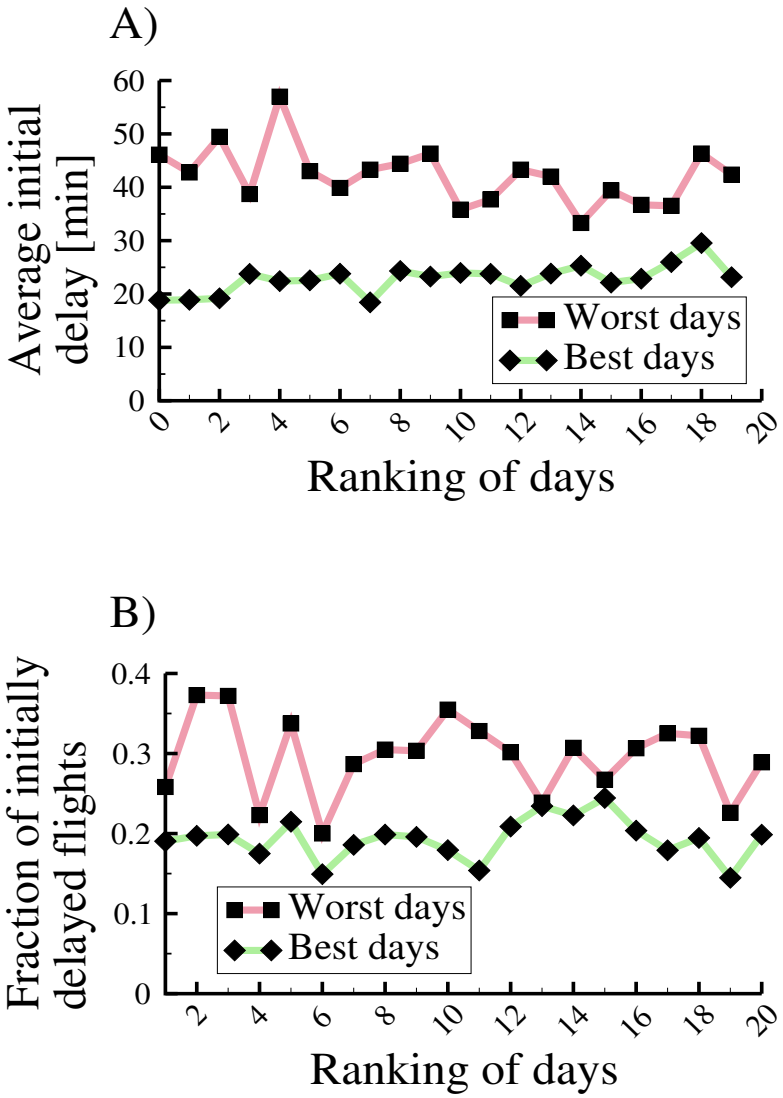


Figure 4.7: Initial conditions of the 20 worst days (red) and of the 20 best days (green) of 2010.

case under study: from data or random initial conditions.

4.3.1 From the data

Initializing the model *from the data* means to replicate exactly the situation of the first flights of all the aircrafts sequences for a particular day.

4.3.2 Random initial conditions

When random initial conditions are set, initial delays are reshuffled among all possible aircrafts, so when and where may vary. Two inputs are needed: initial delay and fraction of flights initially delayed. For instance,

Initial delay: 20 minutes

Percentage of airplanes initially delayed: 10%

Suppose that the number of aircrafts for one day simulation is $4,000$. In this example, 400 aircrafts will have their first flight departing with an initial delay of 20 minutes.

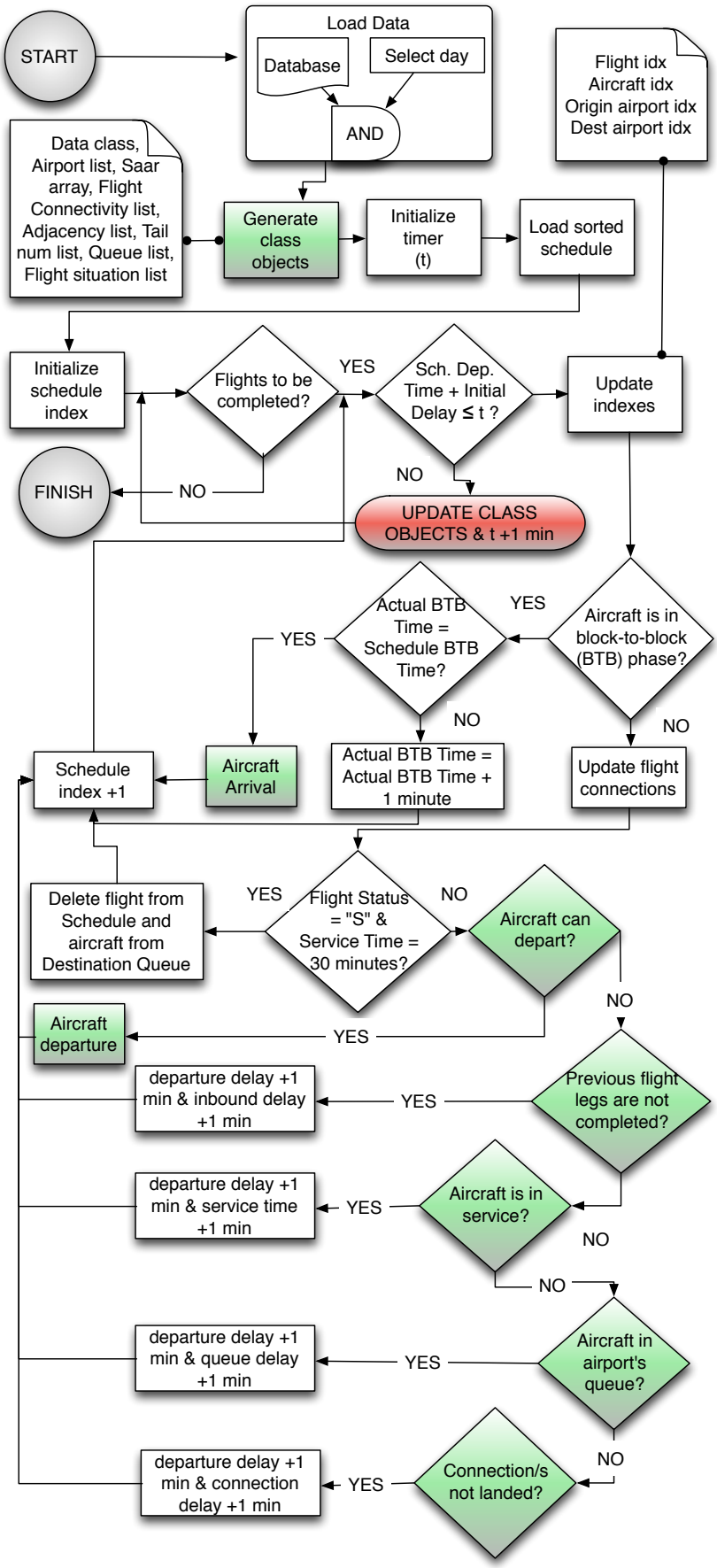
APPENDIX C

C.1

Decision Tree & clustering

Model flowchart summary including all the subprocesses. Flowchart objects in green and red will be explained separately.

APPENDIX C.


START
Load Data
Database
Select day
AND
Flight idx
Aircraft idx
Origin airport idx
Dest airport idx
Data class, Airport list, Saar array, Flight Connectivity list, Adjacency list, Tail num list, Queue list, Flight situation list
Generate class objects
Initialize timer (t)
Load sorted schedule
Initialize schedule index
Flights to be completed?
YES
Sch. Dep. Time + Initial Delay ≤ t ?
Update indexes
FINISH
NO
NO
UPDATE CLASS OBJECTS & t +1 min
Actual BTB Time = Schedule BTB Time?
YES
YES
Aircraft is in block-to-block (BTB) phase?
NO
NO
Schedule index +1
Aircraft Arrival
Actual BTB Time = Actual BTB Time + 1 minute
Update flight connections
Delete flight from Schedule and aircraft from Destination Queue
YES
Flight Status = "S" & Service Time = 30 minutes?
NO
Aircraft can depart?
NO
Aircraft departure
YES
departure delay +1 min & inbound delay +1 min
YES
Previous flight legs are not completed?
departure delay +1 min & service time +1 min
YES
Aircraft is in service?
NO
NO
departure delay +1 min & queue delay +1 min
YES
Aircraft in airport's queue?
departure delay +1 min & connection delay +1 min
YES
Connection/s not landed?
NO

- *Generate class objects*:
  Once the data is loaded for a particular day into the data class object, the remaining class objects are created using this data structure. These objects are:
  - Airport list: Indexation of all the airports that operated that day.
  - SAAR matrix: Includes the hourly capacity (schedule airport arrival rate) for every airport in the list.
  - Airport Flight Connectivity Factor List.
  - Adjacency list: Contains the network structure for that day.
  - Tail number: Indexation of all the aircrafts that operated that day.
  - Schedule: For each flight the schedule object contains the information described in section 4.1.2, initial delay (see 4.3), flight index, flight status (on land "L", flying "F" and in service or in queue "S"), inbound index (previous flight leg index) and connections (see 4.2.2). All flights are initialized with flight status "L".
  - Tail number situation: For each aircraft contains the origin airport, the destination airport, the scheduled and actual block-to-block time and the departure delay (initial, inbound, queue and due to connections).
  - Airports tail number queue: For each airport contains the aircrafts ordered as First in - First served.
  - Airports flight queue: The same as the previous one but indexed with flight number.

- *Update class objects & t +1 min*:
  Objects as Schedule and Airport tail number and flight queues are synchronous updated for each time step.

- *Aircraft arrival*:
  The flight status is changed from "F" to "S" and the airport's tail number and flight queues are updated.

- *Aircraft can depart?*:
  The aircraft can depart if the service time (20 minutes) is complete and the are no flight connections to wait for. Initial flight legs of an itinerary are considered as already served.

- *Aircraft departure*:
  Tail number situation and origin airport queues are updated. The actual block-to-block time is reset. Flight status is changed from "L" to "F".

- *Previous flight legs are not completed?*:
  Check if the inbound index is among the flight connections and the flight status is "L".

- *Aircraft is in service?*:
  Inspect if the flight status for the aircraft is "S" and the service time is different from zero or the aircraft position at the airport queue is less than airport capacity.
- *Aircraft is in airport's queue?*:
  Check if the flight status is "S" and the service time is zero.
- *Connection/s not landed?*:
  Verifies if the number of connections in the schedule for the flight is zero and the flight status is "L".

C.1.1 Clustering

1. Create a cluster list with all airports labeled as −1 (unexplored).
2. Create an empty list (active list) to include the airports to inspect while traversing the adjacency list (network).
3. While unexplored airports continue to exist in the cluster list:
   - For each airport in the cluster list:
     - Check if the airport is unexplored and the average delay per flight for the airport is greater than 29 minutes.
     - If it is so, label the airport with its index and insert the airport index in the active list.
     - Else, label the airport as −2 (not delayed).
     - While the active list continue to have airports to explore:
       - For each airport in the active list:
         - Explore its neighbors in the adjacency list.
         - Check if they are labeled as unexplored and their average delay per flight is greater than 29 minutes.
         - If it is so, label them with the same index as before and insert the airport index in the active list.
         - Else, label the airport as "not delayed".
       - Remove from the active list the airports that their neighbors had been explored.

Part IV

Endogenous mechanisms of delay

CHAPTER 5

Delay propagation dynamics

In this chapter we will analyze the performance of the US Air-transportation system in terms of the internal mechanisms defined in the previous chapter, these are: aircraft rotation, flight connectivity and airport congestion. We also explore the model performance and show that is able to reproduce the delay propagation patterns observed in the U.S. performance data. Our results indicate that there is a non-negligible risk of systemic instability even under normal operating conditions.

We begin with an overview of the model parameters used for the individual day simulations of the most and less congested days of 2010. In section, we explore the model accuracy and sensitivity to α for all days of this same year. In section 5.1 and 5.2 we asses some results concerning the variability of the cluster dynamics regarding its time evolution throughout the day. In section 5.3, we explore the relative importance of each internal mechanism to propagate the delays and finally in section 5.4, we give some insights concerning the model sensitivity to initial conditions.

Parameter	Oct 27	Mar 12	Dec 12	Jul 13	Oct 9	Apr 19
T_s [min]	30					
ΔT [min]	180					
α	0.263	0.190	0.265	0.075	0.020	0.020
β	1.0					
I. Cond.	"From the data"					

Table 5.1: Overview of default values of the model's parameters. The values of α correspond to the best fit for the day.

5.1

Model validation and sensitivity to α

In order to compare empirical results and model predictions regarding the evolution of the cluster of congested airports, we run the model fixing the airport capacity parameter $\beta = 1$ and fitting the flight connectivity factor α to obtain a maximum cluster size as the one observed in the data. By fixing β to 1, we are assuming the same airport capacity as in the data. The results for the evolution of the congested cluster size hour by hour can be seen in Figure 5.1 for December 12 (A), July 13 (B) and October 9 (C). Note that the fit of α is essential to get the maximum of these curves, however the cluster size evolution predicted by the model follows that of the real data.

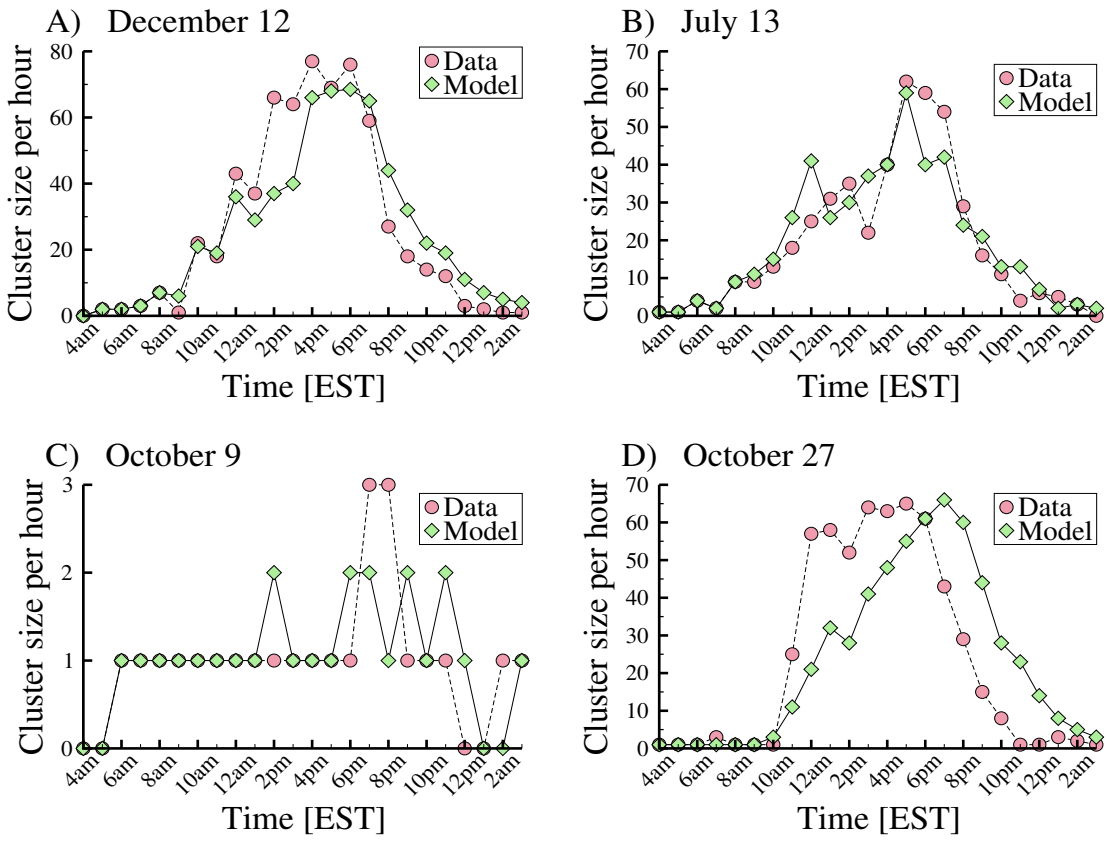


Figure 5.1: Evolution of the largest cluster size for A) December 12 (α = 0.265), B) July 13 (α = 0.075), C) October 9 (α = 0.002) and D) October 27 (α = 0.263)

In the case of October 27 (Figure 5.1 D), the size of the cluster evolved much faster than the model prediction. Analyzing the possible explanation to this difference, we found that severe weather conditions occurred that day across

an important part of the country[1] affecting flights in airports such as Hartsfield-Jackson (Atlanta), John F. Kennedy (New York), La Guardia (New York), St. Paul (Minneapolis), O'Hare (Chicago), Philadelphia and Newark. External perturbations were not explicitly introduced in the model so we cannot expect to be able to reproduce well delay dynamics of days with large external perturbations. In the next Chapter, we deal with this case and adapt the model to take into account external perturbations.

Besides the evolution of the size of the largest cluster per hour, we also checked the results concerning the evolution of the number of clusters during the day (see Figure 5.2). Results for March 12 (A), December 12 (B), April 19 (C) and October 9 (D) confirm that the model is in good agreement. In the previous sections

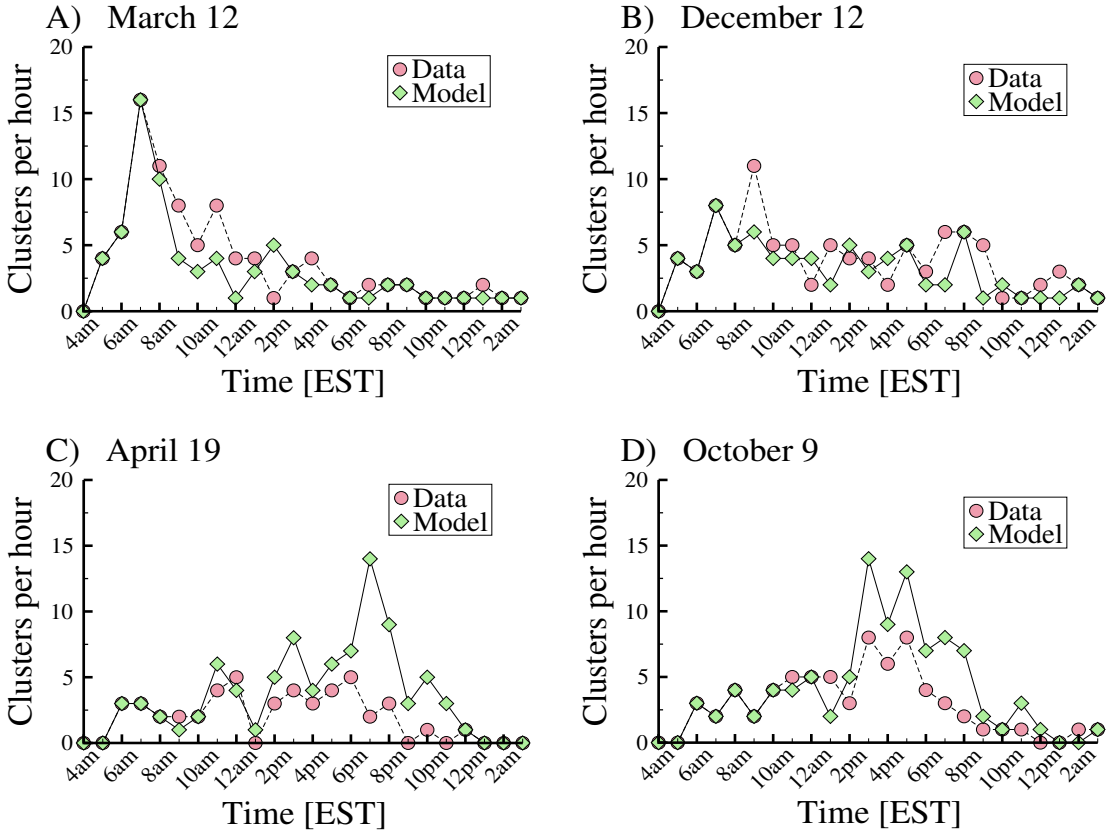


Figure 5.2: Evolution of the number of clusters. Comparison between data and model results for: A) March 12, B) December 12, C) April 19 and D) October 9 of 2010.

we have defined days/airports with problems as those whose average delay per delayed flight was over 29 minutes. Another way, of classifying the days is by means of the largest cluster size of the day. To do so, we set a cluster size that

[1]Severe weather occured on October 27. See NOAA: http://www.weather.gov/arx/oct2610 and CNN: http://edition.cnn.com/2010/US/10/27/us.weather/index.html?_s=PM%253AUS

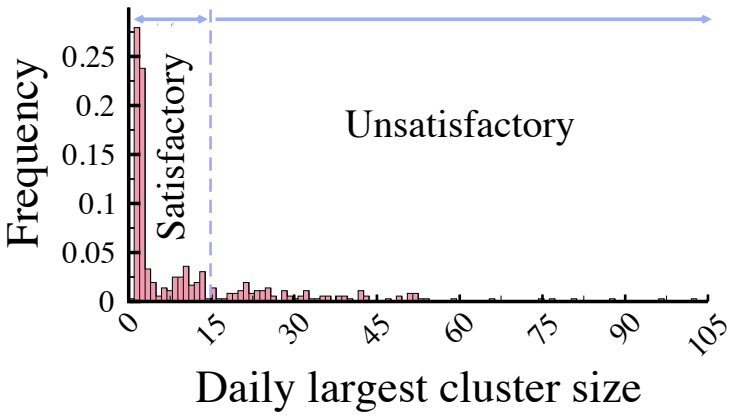


Figure 5.3: Frequency of the largest cluster size for all days of 2010.

corresponds to 15 airports so that if the largest cluster size in a day is higher than this threshold the day is labeled as problematic or unsatisfactory. On the other hand, if it is less than 15 airports the day is labeled as satisfactory. This threshold was selected because in the distribution of largest cluster size there exists a small depression at this value (Figure 5.3). This particular value for the threshold is arbitrary. Still, we have repeated the analysis with some other thresholds and checked that the main conclusions are maintained. The introduction of the

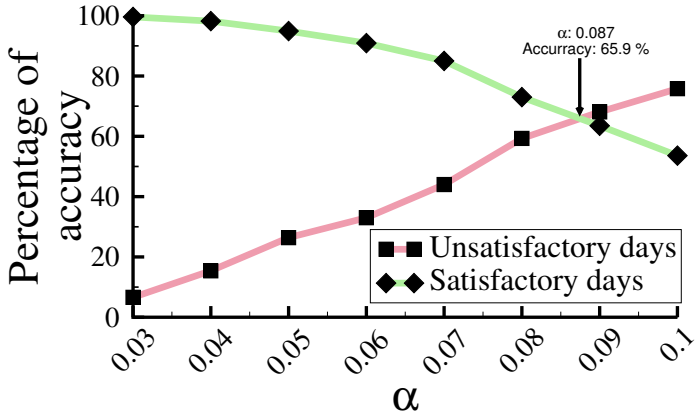


Figure 5.4: Model forecast accuracy as a function of the α parameter for the two classes (unsatisfactory/satisfactory days). All days of 2010 are taken into account.

threshold allows us to define a binary variable associated to the performance of the network each day. Since the model requires a fit in α to reproduce the precise dynamics of the congested clusters, the aim of this exercise is to set a generic

value of α and study how many of the satisfactory/unsatisfactory days are actually predicted. According to our definition, during 2010, 75% of the days get a satisfactory performance. In order to assess the model correspondence with reality, we have to take into account that satisfactory days outweigh unsatisfactory ones. Naturally, with a high α the model simulations predict unsatisfactory days with high accuracy but provide many false positives for satisfactory days. On the other hand, with a low α, most of days with small clusters are successfully predicted but not those with large congested clusters. Bearing this in mind, we defined the percentage of accuracy as a tradeoff between the percentage of accuracy for satisfactory and unsatisfactory days. Figure 5.4 show the fraction

Unsatisfactory days		Satisfactory days	
DATE	Accurate Prediction	DATE	Accurate Prediction
Oct, 27	No	Apr, 19	Yes
Mar, 12	Yes	Oct, 09	Yes
Dec, 12	Yes	Nov, 11	Yes
Jan, 24	No	Apr, 14	Yes
Feb, 24	Yes	Oct, 08	No
May, 31	No	Set, 11	Yes
May, 21	Yes	Apr, 15	Yes
May, 14	No	Oct, 13	Yes
Jun, 23	Yes	Apr, 17	Yes
Jul, 13	Yes	Nov, 10	Yes
Jun, 24	No	Nov, 09	Yes
Jul, 12	Yes	Mar, 06	Yes
Jan, 21	Yes	Oct, 12	Yes
Jul, 29	Yes	Mar, 17	No
Jun, 15	Yes	Feb, 28	Yes
Jun, 27	No	Oct, 16	Yes
Mar, 20	Yes	Apr, 13	Yes
Mar, 11	Yes	Nov, 26	Yes
Aug, 22	Yes	Set, 09	No
Jan, 25	Yes	Set, 20	Yes

Table 5.2: Ranking for the top 20 days by the average delay for flights with positive delay. Model accuracy according to the classification of each day in satisfactory or unsatisfactory. The model is able to predict unsatisfactory days with an accuracy of 70% and satisfactory ones with an 85%.

of correct predictions both for satisfactory and unsatisfactory days. Both curves cross at a value of $\alpha = 0.087$ and at an accuracy rate of 65.9%. Obviously, this is a simplistic technique to measure performance. A more elaborate technique should include appropriate economic considerations to take into account that the cost related to false positives, claiming that a day is going to have a large congested cluster without actually occurring, and false negatives, not being able

to predict a major collapse, are different. Even so, this simple method provides us with a quantitative framework to validate the model and to assess the importance of including further mechanisms in the simulation. Another accuracy test

March 12		
Airport Code	Percentage Realizations	Accurate Prediction
ATL	100.0	Yes
CWA	100.0	Yes
DFW	100.0	No
DLH	100.0	Yes
EAU	100.0	No
EYW	100.0	Yes
FLL	100.0	Yes
GGG	100.0	No
MGM	100.0	Yes
MIA	100.0	Yes
ORD	100.0	No
SJT	100.0	No
STT	100.0	Yes
TOL	100.0	No
BHM	99.8	Yes
CAK	99.5	Yes
CHA	98.6	Yes
FAY	98.4	Yes
MEM	98.3	Yes
HSV	97.9	No

Table 5.3: Top 20 ranking of airports that appear more frequently in the largest cluster for the model results compared to what actually occurred on March 12.

was done to check if the model is able to predict not only the size but the airports that comprises the largest cluster of the day. We selected March 12 whose largest cluster is formed by 97 airports. The model is stochastic, so we run it for 1500 realizations. Comparing the data with the model results for the top 97 airports most frequently appearing in the largest cluster, the model accurately identify 57.8% of them. Table 5.3 displays the Top 20 airports which are more prevalent in the simulations showing if they appeared in the real data as part of the largest cluster for March 12. This is a first comparison, since the real cluster is coming from a single realization in a particular day it cannot be taken as a definitive validation of the model. However, an accuracy of 57.8% with such a simple framework is already encouraging.

5.2

Further results on cluster dynamics

Because of the stochasticity included in the flight connectivity each realization has a slightly different outcome. Figure 5.5 displays the variability between model realizations of the results for March 12 considering a confidence interval of 95% with a fitted α value of 0.190. Simulations in this case were done using initial conditions "from the data"; this means that the stochasticity is caused only by flight connectivity. No matter which set of flight connections are randomly selected, March 12 will continue to display a large cluster. In Figure 5.5 we can

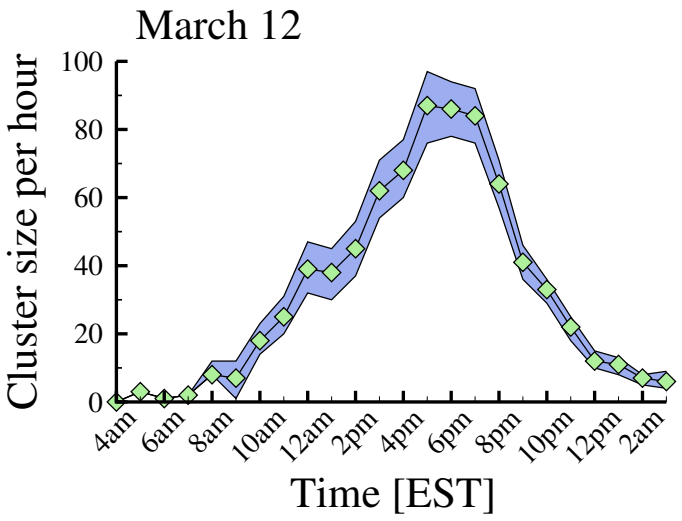


Figure 5.5: Variability of the model results for March 12. Blue margin represents the 95% confidence interval of the size of the largest cluster throughout the day.

differentiate a growing phase that goes from 4am to 5pm and a declining phase from 5pm onwards. As already said, merging is critical for the size evolution of the clusters. Because in the first hours of an unsatisfactory day there are several clusters, thus more possible combinations of merging events, the growing phase is characterized by a stronger variability than the declining phase. The latter, depicts a low variability and as Figure 5.2 A shows the number of clusters do not increase during this phase. All in all, this indicates that no atomization into smaller clusters is produce when the size diminishes. The cluster size dissolves continuously.

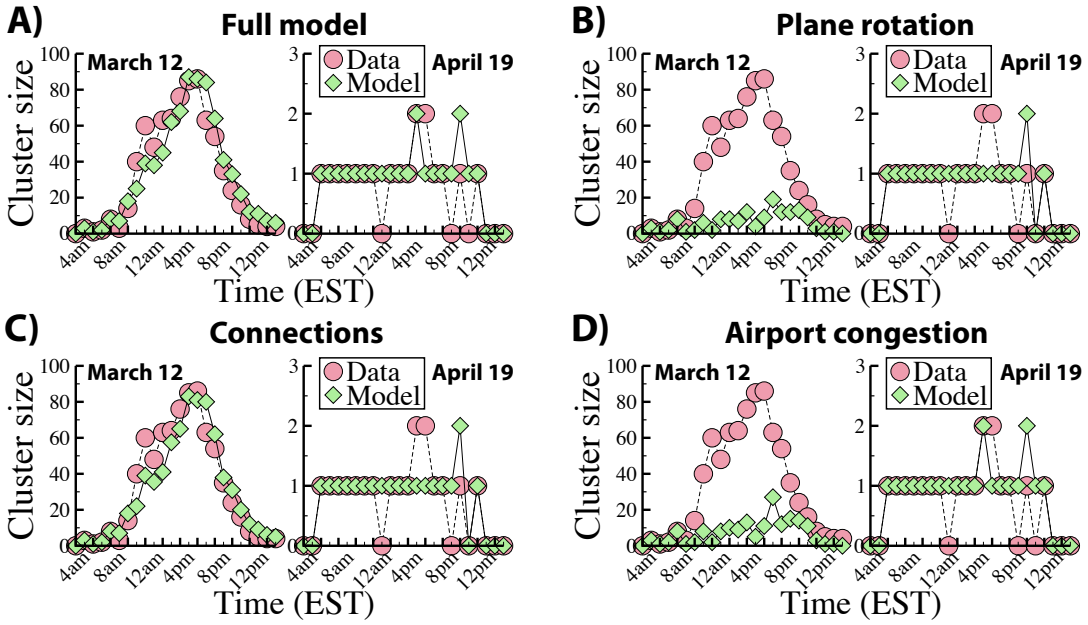


Figure 5.6: Evolution of the largest cluster per hour. A) for the full model, B) the model only with plane rotations, C) only with plane rotations and passenger connections and D) only with plane rotations and airport congestion. The selected days are the ones with the lowest delay (April 19) and the second day with the largest delay (March 12).

5.3

Relative importance of each subprocess

The model allows us also to explore which are the contributions of the main three ingredients (plane rotation, flight connectivity or airport congestion) to propagate delays. From Figures 5.6 B-C, we can conclude that flight connectivity is the most important factor. Only when the flight connectivity subprocess is turned on we can reproduce the cluster dynamics observed in the real perfomance data (Figure 5.6 C). One may still wonder if the picture changes when the capacity of the airports is modified. Actually, the model exhibits weak sensitivity to variations on the β coefficient as shown in Table 5.4. For April 19 Table 5.4 shows that only by diminishing the scheduled capacity by half, the day will start to have some problems according to the size of the largest cluster of the day. In the case of March 12 the scheduled airport capacity was increased by 50% and the results indicate that this increment did not change the overall picture. Furthermore, Figure 5.7 B shows that increasing airports' capacity will not ease off the propagation of delays. The reason for this is that the main cause of delay spreading, flight connections within the schedule, is independent of the airport capacity. Conversely, reducing β by 50% could worsen the situation (Figure 5.7 A). In this case, a decrease on the airports' capacity could act as a

April, 19		March, 12	
β	$< S >$ [num. airports]	β	$< S >$ [num. airports]
1.0	1.0	1.0	92.0
0.9	1.0	1.1	92.3
0.8	1.0	1.2	89.9
0.7	1.0	1.3	88.7
0.6	1.3	1.4	86.4
0.5	16.8	1.5	86.4

Table 5.4: Average largest cluster size $< S >$. β variation for April 19 with an α fixed at 0.02: airports should work at half the scheduled capacity to transform this day into an unsatisfactory one. β variation for March 12 with an α fixed at 0.19: an increase of 50% in β decreases the size of the cluster by only a 7%.

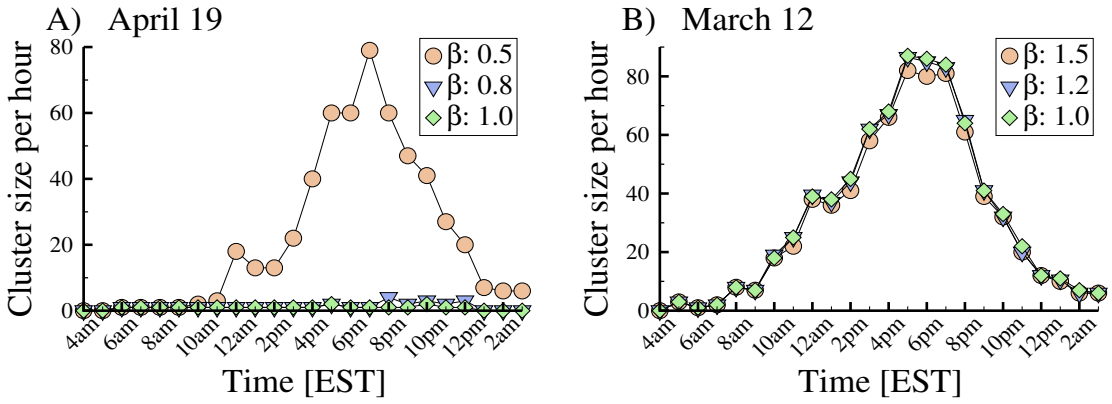


Figure 5.7: Dependence of the hourly largest cluster with a variation in β. A) April 19 and B) March 12.

trigger to new primary delays (different from the initial ones) that later on will spread in a cascading effect due to the flight connectivity. Although a decrease on the scheduled capacity of 50% for every airport in the network is not likely to occur in practice, a much realistic situation could be an airport or group of airports operating under-capacity when severe weather conditions are met.

In any case, airport congestion could be the source for primary delays but it is not able to propagate them throughout the network as it is the case of flight connectivity. These results suggests that models failing to incorporate flight connectivity as a source for the propagation dynamics may not fully capture the system-wide effect of this phenomenon.

5.4

Sensitivity to the initial conditions

The initial delays affect the outcome of the model. In the results of Figure 5.1 and 5.2, we take the primary delays for each aircraft from the data as initial conditions for the model. Introducing different initial conditions, we can assess the resilience of a day schedule to an increase of unexpected incidences. This question is explored in Figure 5.8 where a fraction of randomly selected flights are delayed. The size of the largest cluster is the fraction of delayed flights and of the intensity of the initial delays. For the sake of simplicity, we set all the initial delays in the simulation equal to a fixed value (delay intensity in Figure 5.8). The results are displayed for the schedules of two days: April 19 and March 12, which respectively show a very small and very large cluster in the real data. In particular, the average flight delay on March 12 was the second largest in 2010. The congestion on the worst day of the year, October 27, can be explained due to extreme meteorological conditions, while on March 12 no major external event was reported. Therefore, the network-wide propagation of delays in that day was likely caused and driven by internal mechanisms of the system. Comparing in Figure 5.8 the curves for March 12 and April 9, one notices that the surface representing the largest cluster size for March 12 is displaced toward smaller values of the initial delay intensity or fraction of flights with primary delay. This shows a higher susceptibility of the schedule of this day to disruptive perturbations. Another interesting feature of the curves of Figure 5.8 is that, given enough primary delays, they show a non-negligible risk of systemic failure regardless of the schedule. The curves in Figure 5.8 for different values of α also confirm the relevance of connections and crew rotations for the spreading of delays.

The primary flight delays in a day of real operations do not necessarily localize randomly in the network. If the causes are bad weather, technical or labor issues delays are more prone to concentrate in a few airports. In Figure 5.9, this issue is explored by comparing the intra-day evolution of the cumulative size of the largest congested cluster when the initial delays are introduced in the model in two different ways. The first one is by using the primary delays given in the database. The second procedure is by randomly shuffling the flights affected by the primary delays. The values of the real delays in the database are maintained but they are assigned to flights selected at random. The comparison of the curves for the two cases with the real data shows that random perturbations are way more efficient to collapse the system. While airports in general have some capacity to recover delays, the random selection of delayed flights affect a larger number of them and besides concentrate a heavier burden on smaller airports which have less capacity to react. This result evinces that the method followed for schedule evaluation in Figure 5.8 is conservative in the sense that it considers the schedule under a non favorable scenario for the distribution of primary delays.

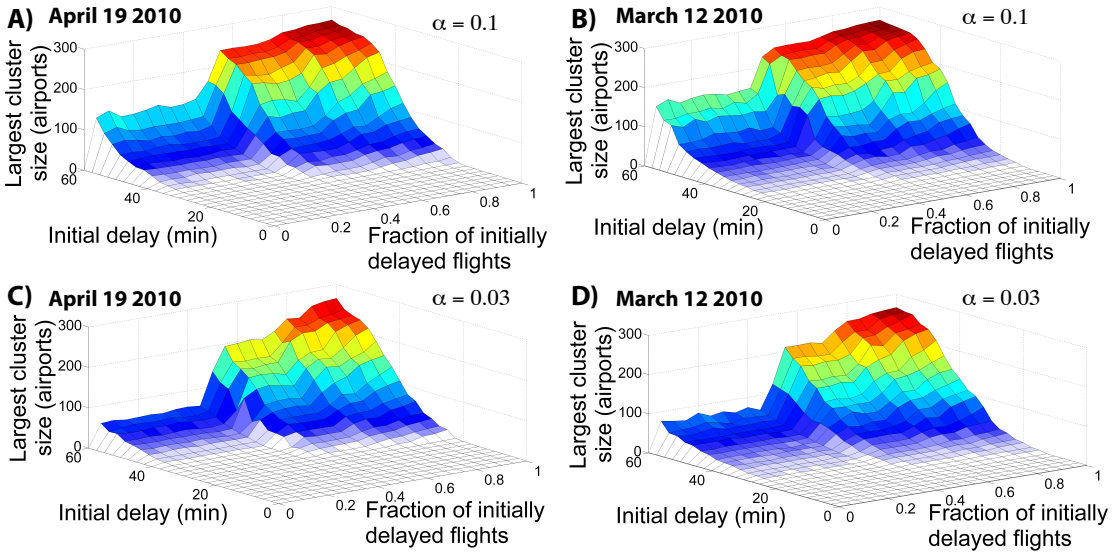


Figure 5.8: Assessment of the schedule resilience to develop large clusters. In the plots, the size of the largest congested cluster is displayed as a function of the fraction of initial delayed flights and of the intensity of the initial delays for a congested March 12 in B) and D), and for an uncongested day on April 19, A) and C), for two values of the flight connectivity factor α. An initial fixed delay is assigned to randomly chosen flights.

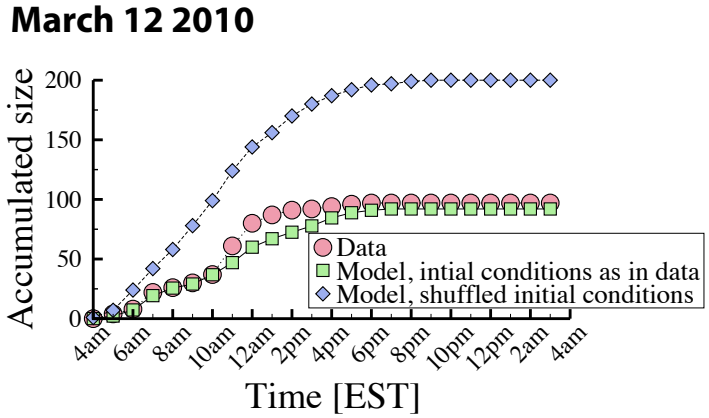


Figure 5.9: Time evolution of the cumulative size of the largest congested cluster for different initial delays of the flights: assigned as found in data or to randomly selected flights but keeping the same values as in the data.

Part V

Impact of Perturbations

CHAPTER 6

Case study: weather perturbations

The upsetting consequences of weather conditions are well known to any person involved in air transportation. Among the sources of primary delay, some of the most devastating are related to weather perturbations (Allan et al., 2001; Klein et al., 2007, 2009). Still the quantification of how these disturbances affect delay propagation and the effectiveness of managers and pilots interventions to prevent possible large-scale system failures needs further attention. In this chapter, we take a coarse-grained approach to understand the system response to the introduction of large-scale disruptions that will be followed by a more detailed view, in the next Chapter, regarding the system's dynamics due to perturbations.

External disruptions are commonly caused by adverse weather, ranging from reduced ceiling and visibility to convective weather. Therefore, we selected for simulation and comparison purposes October 27^{th} 2010, a day for which we count with real performance data and that turned to be the worst day in 2010 according to the average flight delay. The origin of this high congestion levels was a severe weather phenomena distributed across the country[1]. Indeed, the meteorological disturbances of this day were later known in the media as part of the 2010 Superstorm.

In the previous chapters, we have analyzed the system response to primary delays in the first flight leg of an aircraft itinerary. By going one step further, we want to understand the system behavior to disrupting events that may compromise the system stability. Previous attempts to understand the stability of the air traffic network were based on queuing (Janić, 2005) and percolation (Woolley-Meza et al., 2011) theory. Here we simulate the system performance under weather-disrupting inputs modeled as a shortfall on terminal capacity. Our viewpoint falls in line with earlier research on weather-related delays associated to capacity constraints (Allan et al., 2001; Klein et al., 2007).

By considering different intervention measures, we can improve the model predictions getting closer to the real delay data. Our model can thus be of

[1]Severe weather occured on October 27. See NOAA: http://www.weather.gov/arx/oct2610 and CNN: http://edition.cnn.com/2010/US/10/27/us.weather/index.html?_s=PM%253AUS

help to managers as a tool to assess different intervention measures in order to diminish the impact of disruptive conditions in the air transport system.

6.1

October 27 of 2010 Superstorm

October 27 2010 embodies the concept of external disturbances affecting a large part of the air transport network. The disturbances were the result of severe weather conditions generated by a low-pressure system that started during the early morning hours of the 26^{th} in the Southern Plains and moved North, producing a significant pressure gradient that caused strong wind gusts[2]. The massive storm complex continued throughout the 27^{th} and dissipated on the 28^{th} impacting some airports of the National Aviation System. According to the news reports, at least Hartsfield-Jackson airport in Atlanta (ATL), and the three main airports of the New York-New Jersey area, John F. Kennedy airport (JFK), La Guardia airport (LGA) and Newark airport (EWR), experienced large delays because of inclement weather[3]. Focusing on the On-Time Performance data, this day presented the largest average flight departure and arrival delay of the whole year. These values are, respectively, 54 and 53 minutes. Moreover, the largest congested cluster of the day amounts to 88 airports, the third worst performance of 2010 after March 12 (97 airports) and December 12 (103 airports). A map of the congested cluster of airports for October the 27^{th} is shown on Figure 6.1 A. The congested airports are plotted in yellow if they do not belong to the largest cluster and in red if they do. The size of the symbols of the airports in the largest congested cluster is proportional to the average delay. In the map, we can observe how the network congestion affected a vast area that spread from Central to Eastern U.S matching the area where the windstorm developed. In addition, the average departure delay for the first legs of the day for the flight rotations equals 46 minutes, which ranks among the top 10 days with worst initial conditions. Remember that these primary delays are the initial conditions introduced as external inputs for the simulations of our model. The map in Figure 6.1 A shows how weather perturbations can produce system-wide effects in the air transportation network. The managers and pilots of the different airlines and airports, of course, reacted to the problems generated by the weather disturbances. Typically, these reactions included flight delay but also cancellations and flight diversion to airports different from those of destination. These two latter factors introduce changes in the planned schedule that it is important to analyze. We depict in Figure 6.1 B the number of cancelled flights per hour for October 27 and October 20. October 20 was a low congested day with only 2 airports being part of the daily largest congested cluster and

[2]NOAA October 27 wind event: http://www.weather.gov/lot/2010oct26

[3]Delays in JFK, LGA and EWR: http://www.myfoxny.com/story/17450431/airport-delays-at-jfk-lga-newark?clienttype=printable

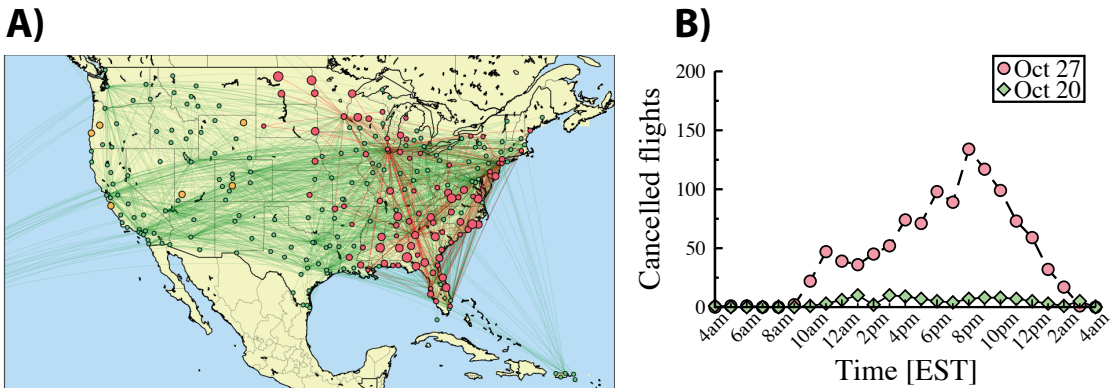


Figure 6.1: A) Map of the US airport network for October 27 showing the daily congested cluster in red. The airport color codes are: green (airport not congested), orange (congested airport not belonging to the largest cluster) and red (congested airport belonging to the largest cluster). B) Number of canceled flights per hour for October 27 and 20.

also showed a low average flight delay equal to 24 minutes (less than half of the value for the 27^{th}). From Figure 6.1 B we can conclude that the rise of network congestion, if compared with a low congested regime, induced an important rate of flight cancellation.

6.2

Modeling external interventions & perturbations in the system

The intraday evolution of the size of the congested cluster per hour is displayed in Figure 6.2. The cluster size shows an initial steep growing phase that starts at early morning hours and continues up to 5 pm followed by a declining phase from 5pm onwards. We first used a simple model scenario with a fitted α value of 0.26 and introducing the initial conditions as given in the data to the schedule of October 27^{th} (green dots Figure 6.2). In this basic scenario, we consider every airport working at a nominal capacity ($\beta = 1$) throughout the day. Although the maximum cluster size is reproduced well, the simulation results show a slower increase in the growing phase that does not match the real evolution of the cluster size. In fact, it seems that this particular day morning hours are crucial to understand the development of the congested regime. In addition, the declining phase decays earlier in the empirical data. This difference regarding the dynamics could be due to the fact that the basic model does not take into account external perturbations to the system. As stated before, the severe weather conditions enhanced the delay spreading by affecting some airports capacity. As mentioned in (Fleurquin et al., 2013) events of this kind can be included in the model by modifying the capacity parameter β. Although a change in β is not able

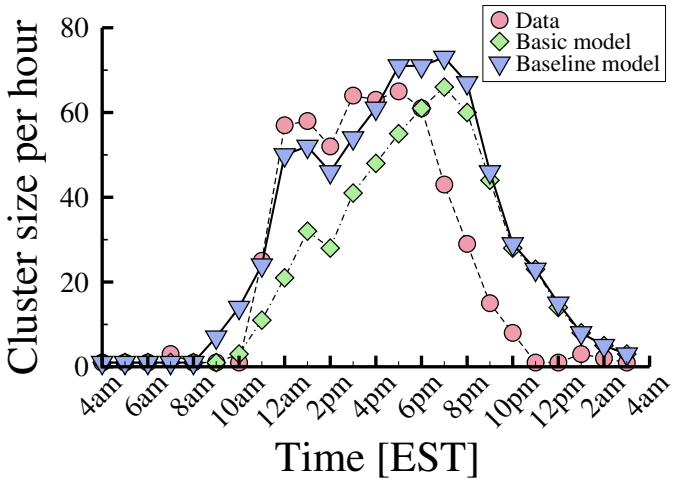


Figure 6.2: Evolution of the largest cluster size per hour. Comparison between reality (data), model without perturbation (basic) and model including perturbation (baseline).

to generate a network-wide spreading of the delays, it could be in fact a source of new "primary" delays that later on will propagate throughout the system. The new delay results from an increase in the length of the waiting queue mechanism. In this sense, we mimic the capacity shortfall of 4 airports (ATL, JFK, LGA and EWR), mentioned in the news report, by setting to 0 the capacity parameter in the morning hours (from 8 am to 10 am local time). This modification introduced to the basic model is shown in Figure 6.2 as the baseline model, while the connectivity factor remains constant. Our assumptions regarding the modeling of weather impacts through the β coefficient are confirmed, remarkably improving the model results in the growing phase (Figure 6.2). Even though the baseline model already shows acceptable results, we analyze even further the results by exploring different scenarios. We call these scenarios variants of the model and will go from 1 to 4 (see Table 6.1 for a summary of the details). We are interested in improving specially the plateau and the timing of the declining phase of the curves representing the cluster evolution along the day. So far we have introduced the weather disturbances into the baseline model by temporarily reducing the capacity of a few major airports, the next step is to simulate the human interventions on the schedule. Naturally, one way of tackling the congestion problem is flight cancellation. According to our modeling approach, this external intervention should affect the airport network connectivity. The passengers and crew of the canceled flight will not be able to connect with following flights in their destination airports. One way of translating this effect into our model is to temporarily modify the α parameter. Figure 6.3 A shows the cluster size evolution for what we refer to as variant 1 of the model and compares the results with real observations and the baseline model. In particular, variant 1 of the model incorporated a connectivity parameter α set to 0 between

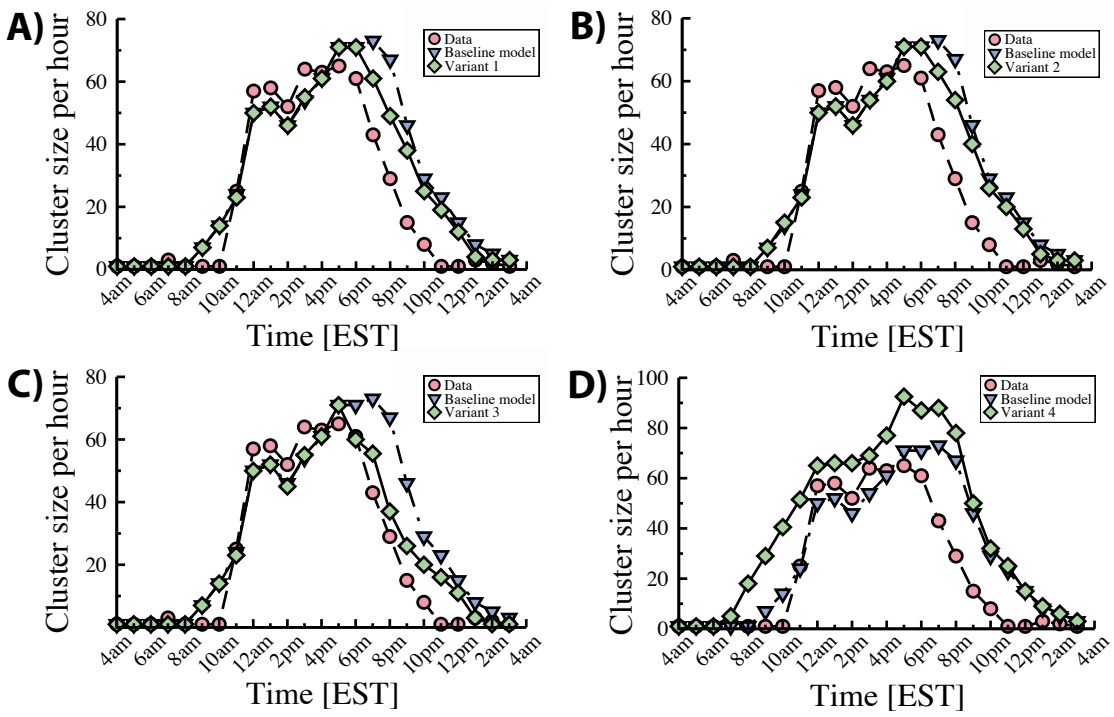


Figure 6.3: Evolution of the largest cluster size per hour. Comparison between model variants 1 to 4.

7 and 9 pm EST ($\Delta\, T^{\alpha}$). This time window ($8pm \pm 1$ hour) is selected because it corresponds to the time with the maximum number of flight cancellations (see Figure 6.1 B). As can be seen in Figure 6.3 A, the simulation results improve for the declining phase, verifying our assumption of modeling the intervention through a decrease of the network connectivity. We also check the sensitivity of the simulation output to a change in the connectivity factor α and the time window $\Delta\, T^{\alpha}$, variant 2 and variant 3 of the model, respectively. In Figure 6.3 B, we consider for variant 2 a decrease of α in one half, instead of setting it to zero, and this slightly increases the congestion on the declining phase if compared to variant 1. The effect of an increase in $\Delta\, T^{\alpha}$ is even more significant as shown in Figure 6.3 C for variant 3 of the model. In this case we fix α to 0 (as in variant 1) increasing the time window by two hour between 6 pm and 10 pm. Therefore, we can observe a refinement of the declining phase matching. It is important to note that after the $\Delta\, T^{\alpha}$ period the cluster size slightly grows and this effect is not seen in the empirical data. This could be due to the fact that queue congestion at the airports does not ease off by this intervention, triggering the propagation of delays when the connectivity is reestablished. The above results show how with slight changes one can gradually improve the capacity of the model to forecast congestion. After exploring the effect of changing the connectivity on the model, we consider now the implementation of a further response element: the so-called Ground Stops. This consists in preventing the departure of flights

Variant of the model	Characteristics (changes implemented)
Basic model	The airport capacity β equals 1 (nominal capacity) for every airport in the network and the connectivity factor α is 0.26. These values remain constant througout the day.
Baseline model (basic model + perturbation)	The airport capacity β equals 1 except for ATL, JFK, LGA and EWR that the value is set to 0 between 8*am* to 10*am* (local time). Connectivity α is 0.26.
Variant 1 (baseline model + intervention in the network)	Same β conditions as the Baseline model. Connectivity α equals 0.26 except for the time period between 7*pm* and 9*pm* (Estearn Time) that drops to 0.
Variant 2 (baseline model + intervention in the network)	Same β conditions as the Baseline model. Connectivity α equals 0.26 except for the time period between 7*pm* and 9*pm* (Estearn Time) that drops to 0.13.
Variant 3 (baseline model + intervention in the network)	Same β conditions as the Baseline model. Connectivity α equals 0.26 except for the time period between 6*pm* and 10*pm* (Estearn Time) that drops to 0.
Variant 4 (baseline model modified)	Connectivity α is 0.26. The airport capacity β equals 1 for every airport in the network. In this case the perturbation is included by issuing Ground Stops to flights whose destination is ATL, JFK, LGA and EWR between 8*am* to 10*am* (local time).

Table 6.1: Modifications of the model accounting for external perturbations and interventions.

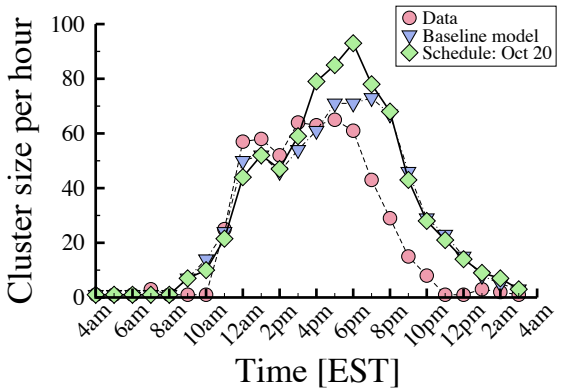


Figure 6.4: Evolution of the largest cluster size per hour. Green dots, using the schedule of October 20 with the initial conditions of October 27.

on origin when the destination airport has problems[4]. In our case, this measure affects the flights with destination in one of the four airports with reduced capacity and with scheduled arrival time between 8 and 10 am Eastern Time (same period of time when β is set to 0 in the baseline mode). The connectivity factor and the airport capacity remain constant throughout the day and equal to the values of the basic model (respectively 0.26 and 1). This scenario corresponds to the variant 4 of the baseline model (Table 6.1). The results of the simulation for the evolution of the largest congested cluster size can be seen in Figure 6.3 D. The congestion starts earlier because the delays surge before in time and the largest clusters extension considerably increases. The effect of an early onset of Ground Stops is shown to be devastating for the delay dynamics. Still, generalized Ground Stops is not likely to happen without early palliative interventions such as flight cancellations or diversions. In our simulations, we have used the schedule of October 27 as observed in the real operations including flight diversions and cancellations. One may thus wonder which is the effect of this particular configuration of the schedule on the final spreading of the delays and how much the changes in the schedule help to reduce congestion. We do not have access to the unperturbed day plan of the airlines but we can still use for the sake of comparison the schedule of October 20. This day the system showed a low level of congestion and so the interventions in the schedule must have been minimal. The variant 4 of the model is thus the baseline scenario, plus the initial conditions of October 27 but implemented in the schedule of October 20. The results of the simulations are depicted in Figure 6.4. It is clear that the schedule of the 27^{th} was not the reason for the unfolding of a large congestion since an important congested cluster of airports still appears. We can thus blame the weather disturbances for most of the congestion of October 27. The differences between the evolution of the unperturbed schedule of variant 4 and the baseline model speak in favor of the intervention measures taken on October 27.

[4]Description of the Ground Stop Operations: http://www.fly.faa.gov/Products/AIS_ORIGINAL/shortmessage.html

In this chapter, we have introduced the weather impacts by varying the airport capacity parameter β to some airports in the network. This change produces a drop in the airport capacity service rate enlarging the airport queue. On the other hand, we implement flight cancellations by affecting the network connectivity parameter α, thus reducing the delay propagation dynamics. Our simulations evidence that weather impacts produce system congestion independently of the day considered, as it is the case when the initial conditions and same input perturbations are introduce to the schedule of October 20. Next we will deepen the understanding of system stability by defining some metrics to measure the system robustness to perturbations and the impact that each element has on the system.

CHAPTER 7

Addresing the dynamical robustness of the air transportation system

Recent works in the area of Complex Systems have addressed the robustness of networks such as power grids, social groups and the Internet. The robustness is evaluated against an external perturbation that can be different in nature depending on the particular network. For instance, the failure of power stations that can trigger a nationwide blackout or a general shutdown of routers and the consequent connectivity loss. In this chapter, we introduce metrics inspired by Complexity Science to explore the robustness of the air transportation system in the US with respect to delay propagation. We use the model to assess the effect of disruptions in the network. These disruptions are introduced as initial conditions and can affect single flights or full airports. The model is then run with and without disruptions and the outcome is compared to quantify the system robustness. Our results indicate that large hubs (in the sense of number of offered destinations) are more vulnerable to flight delays than small or medium sized airports. However, the impact in the whole network of delays initiated in an airport does not depend on whether it is a hub or not. We also detect a set of high impact flights and explore the drivers that generate these long tail extreme events.

7.1 Background

Robustness against external perturbations is an important feature of the networks, bringing together the system structure and dynamics. It can be defined as the system ability to continue primary functions after a perturbation occurs. Perturbations can be modeled in different ways. Reference (Albert et al., 2000) studied how the network is affected to the removal of a fraction of nodes. Under certain conditions this may produce the network fragmentation, therefore severely damaging the communications between its components. In this work, the authors distinguish between random removals (errors) and targeted attacks

to its most connected nodes (hubs). While networks with heterogeneous degree distributions (scale-free networks) are robust against random errors they are likely to fragment into smaller clusters if a critical fraction of its hubs is removed. As noted in Ref. (Cohen et al., 2000) this can be understood as a percolation process, thus as shown in Ref. (Callaway et al., 2000) it is possible to derive exact analytical solutions for node and edge percolation (removal of a fraction of edges). Initial disturbance can trigger a cascade of subsequent failures (Barzel and Barabási, 2013), such is the case in power grids (Albert et al., 2004; Dobson et al., 2007) or air transportation networks (Cardillo et al., 2013; Lordan et al., 2014). This dynamic effect is enhanced by networks coupled together (Buldyrev et al., 2010; Cardillo et al., 2013); where the failure of elements in one network can lead to a branching process affecting elements of other networks in a recursive way.

In this chapter, we tackle the problem of the air transportation network robustness using US performance data. Instead of a structural view, we focus here on the robustness of the system dynamics. In this case, the initial disruption is given by one or several delayed flights (when considered the airport disruption) that later, as the flight operations continue, can spread and multiply producing a cascade of reactionary delays. We therefore consider the initial disruption as a primary delay (Belobaba et al., 2009; Allan et al., 2001) and the subsequent cascade as reactionary delays (Beatty et al., 1999; AhmadBeygi et al., 2008). As shown in (Fleurquin et al., 2013) this ripple effect is boosted by the network connectivity through the aircraft rotation and crew and passenger connections between flights. Based on these findings, in the following section we define metrics able to assess the robustness of the network when a delay impact is produced by an individual delayed flight or a congested airport. Given that the events in the model are fully traceable, we develop a cause-effect analysis allowing us to reconstruct the trees of reactionary delay (Beatty et al., 1999).

Under these assumptions, we use data from the 13th of July 2012. This day showed a high level of congestion, which according to the news was not imputable to meteorological[1], technical or labour causes.

7.2

Assessing network robustness and impact of delays

One way of evaluating the response of the system to a perturbation, is by exploring the delay $D_i(t)$ induced in flight or airport i in response to a primary delay in a flight/airport j, $D_j^0(t_0)$, at time t_0. Therefore, we can measure the

[1]National Oceanic and Atmospheric Administration, web page: http//:www.noaa.gov

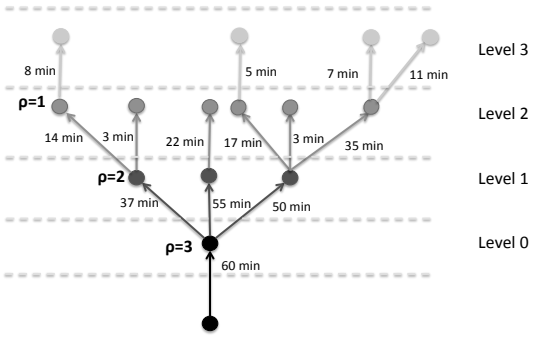


Figure 7.1: Example of a tree of reactionary delay with 4 levels. In this case the delay per flight diminish downstream. ρ is the reproductive number of each flight.

response of an element of the system to an induced perturbation as:

$$s_{ij}(t|t_0) = \frac{dD_i(t)}{dD_j^0(t_0)} \,. \tag{7.1}$$

Following the previous definition, we can construct the response matrix S_{ij} with each entry measuring how responsive an element i is to a perturbation in j ($s_{ij}(t|t_0)$). Regarding delay propagation dynamics, a perturbation could be to delay all departing and arriving flights at an airport within a certain time period and then measuring the delay generated at an airport throughout the day. To better understand the airports response to perturbation, we begin by defining the delay impact of an airport i in the system as:

$$I_i^t = \sum_{j=1}^{N} S_{ij}^T \,. \tag{7.2}$$

Due to the temporal variability of traffic patterns during the day, the perturbation outcome strongly depends on the time of the day that it is generated. Therefore, an airport perturbed at a certain hour t_0 might not have the same consequences in another hour.

In addition, it is possible to revert the argument and measure how robust an airport i is to a perturbation generated in airport j. A way of evaluating the robustness of i (or lack of it, i.e. vulnerability) is by means of S_{ij} as:

$$R_i^t = \frac{1}{\sum_{j=1}^{N} S_{ij}} \,. \tag{7.3}$$

Hence, the robustness of an airport captures its response to perturbations in other airports. A large value of R_i indicates that the airport is very robust.

As a way of measuring the delay impact of a single flight, we make use of the concept of trees of reactionary delay as described in Ref. (Beatty et al., 1999).

The perturbation starts by setting a primary delay of 1 hour to an initial flight. The tree can contain one flight if there is enough slack time in the subsequent flight legs or connections given by the schedule. If this is not the case the delay propagates following a cascade-like effect as it is shown in Figure 7.1. In this example the perturbation branches by delaying the departures of the connecting flights and legs. Because of the complex pattern of connectivity in the network each reactionary tree will have different characteristics. We are interested, in measuring the total impact (I_i) and the average reproductive number ($\bar{\rho}_i$) or branching rate of the tree generated by flight i. These two measures give an idea of the extent of the perturbation in the system. The reproductive number is defined as the number of flights delayed by a delayed flight. For instance, the initial flight with 60 min has $\rho = 3$ because it affects three more flights in the level immediately downstream. The average reproductive number encapsulates the effect of the branching process in a tree. The total impact measures the fraction of minutes generated by the branching process over the initial delay. It could happen that a tree might be large enough in number of levels but with relatively low impact because the were enough slack time in the schedule to absorb the delays. Because of the inherent causality of the tree, we can also capture the influence of 3 important system components: the destination airport, the arrival time and the airline of the perturbed flight. Worth noting is that both impact variants (airport and tree impact) are conceptually equal, one at the level of nodes and the other at the level of edges.

7.3

System response to airport perturbations

We begin by evaluating the system response to a perturbation of one hour at each airport of the network. To do so, the model is run delaying all incoming and departing flights of each airport, one airport at a time, for each hour of

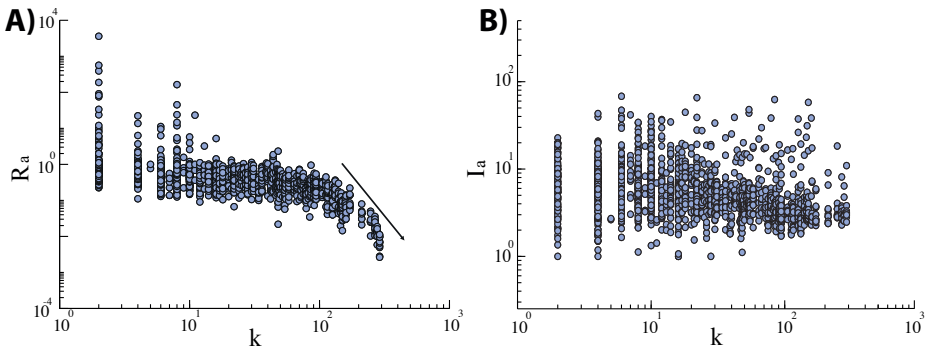


Figure 7.2: Airport impact and robustness as a function of the number of connections. Perturbation is produced delaying all incoming and departing flights from the airport at each our of the day.

the day. Figure 7.2 shows the results concerning airport impact and robustness as a function of the number of connection of each airport (k). Surprisingly, the robustness has a steep decline for the airports with largest degree (Fig 7.2 A). This unexpected result evinces the vulnerability of network hubs. In other words, large airports (in number of connections) are strongly affected by perturbations originated throughout the system. This is not the case for the rest, as it is clear that the relation is almost flat. One might expect a priori that hubs are robust enough to absorb delays due to its excess of capacity, but this result clearly contradicts this vision. A plausible explanation could be because of the reinforcement caused by flights that repeatedly go to a spoke and then return to the hub, magnifying the delay on the hubs. Figure 7.2 B depicts the impact that each airport has on the network, evincing that it does not depend on the node degree. It is important to say though, that while hubs strongly perturb the system generating a large amount of delayed flights, the induced total initial delay at the airport is also high. Therefore, the I_i^t for network hubs is not as high as one might expect. In any case, what this result reflects is that there is no relation whatsoever with respect to the airport size (in number of connections). In principle, this suggest that there are more subtle effects in the dynamics that might be related with other system features. In the next section we explore other possibilities that may affect the system response, accounting for individual flight impact.

7.4

System response to flight perturbations

Following a top-down analysis we keep going further down into the microlevel features of the airport network. If impact is indistinguishable at the airport level may be individual flights produce very different responses. We explore the system response by randomly selecting a flight for each airline, airport and schedule arrival time combinations, thus we generate simulations for 7658 different initial perturbations. In Figure 7.3A we evaluate this possibility by plotting the cumulative impact distribution of the trees generated by 60 minutes delayed flights. As previously explained, each flight might produce a cascading effect developing trees of reactionary delays. The broad distribution of tree impacts signals the heterogeneities present in the system with respect to flights. In this sense, there are few highly impact flights among many low impact ones. But the distribution shows that highly impact flights are likely to occur. Figure 7.3B shows the probability distribution of the average reproductive number $\bar{\rho}$. Similarly, there are quite a few trees with $\bar{\rho} < 1$, the most probable are flights with no impact at all $\rho = 0$. Needless to say that the delay propagation is boosted by trees with average reproductive number larger than one. In addition there exist flights that produce a large cascade with $\bar{\rho} > 2$.

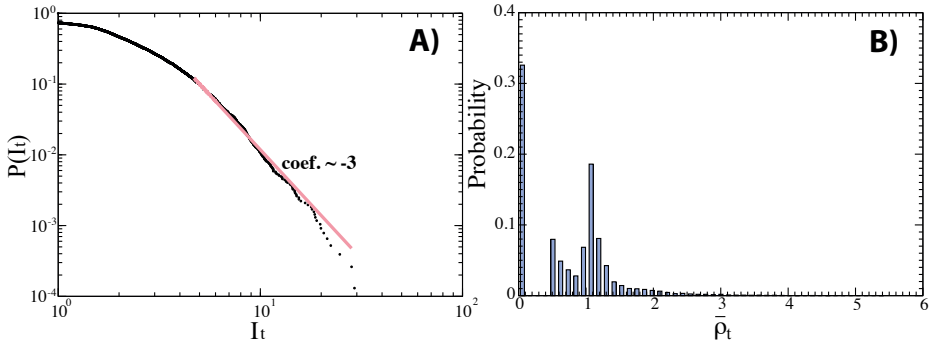


Figure 7.3: Cumulative probability distribution of tree impact and probability distribution of the average reproductive number of the tree.

Spotted how susceptible the system is to individual flights, we remain to understand what are the main characteristics of flight impacts. To do so, we analyze the tree impact with respect to the time of the day the perturbation starts and according to the airline the flight belongs to. Figure 7.4A depicts the tree impact that generates each perturbation in relation to the airline of the initially delayed flight. The first thing to notice is that some airlines (code: 20366, 19790, 20398, 20304, 19393) are likely to produce high impact flights evincing a long tail of extreme events. In addition the average airline impact is slightly different from one another. The main reason for this result should be a combination regarding the airline's pattern of connectivity and the planned slack time within flights. Figure 7.4 B depicts the flight impact regarding the schedule arrival time of the first flight of the tree. Not surprisingly the impact decreases along the day because the temporal cone of events (downstream flights that could be affected by the perturbation) is smaller as the time passes. Nevertheless, hours with highest impact time happens in the morning, not in early hours because, although with larger temporal cone, they also have the lowest average reproductive number (Figure 7.4 D). According to the results shown in Figures 7.4 B and 7.4 D there is a trade-off between the temporal horizon and the reproductive number. While the cone of events decreases through time the connectivity increases, therefore impact in between is the highest. Finally, Figure 7.4 C confirms that $\bar{\rho}$ is closely related to the extent of the impact on the system of each airline with high dense schedules combined with high impact trees. However, there some exceptions (code: 20437, 19790) where the branching rate is relatively high compared to the impact measured. As mentioned before these are large trees but with low impact because there was enough slack time in the schedule to absorb the delays produced.

To further understand the drivers behind the long tail of extreme events, we compare the data for the top 100 highest impact flights, those generating the largest impact trees, with other 100 flights selected at random. Regarding the destination airport of the initially perturbed flights, 67 airports are the initial target for the highest impact trees while also roughly 68 airports are the desti-

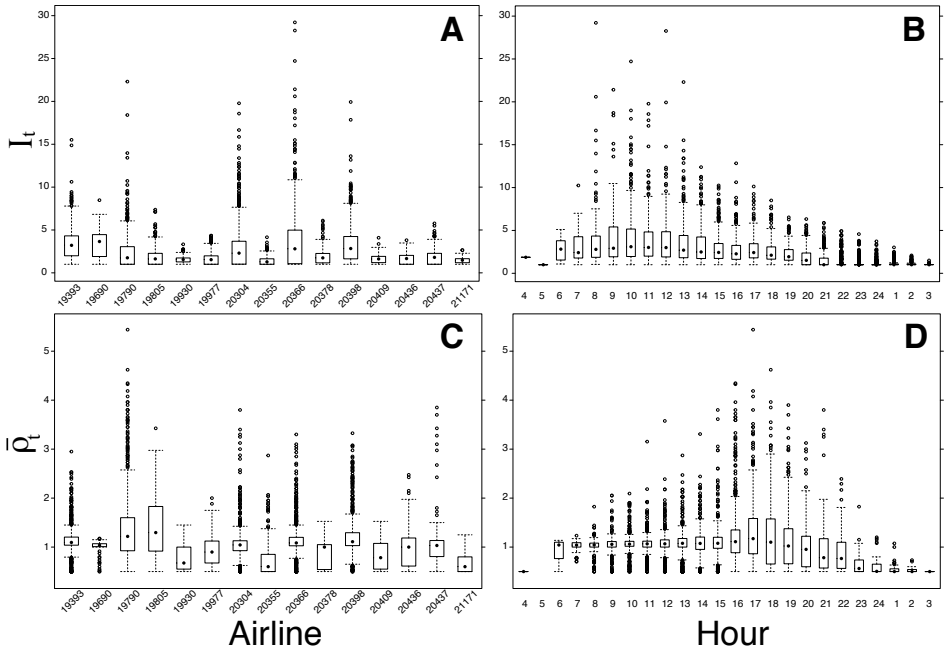


Figure 7.4: Assessing the impact of individual flights (A,B) and the average reproductive number of reactionary trees (C,D) refers to the airline the flight belongs to and the time of the day when the perturbation starts, respectively.

nation of the randomly chosen 100 flights. This represents roughly the 23% of all possible airports for July 13 for the two sets and there is no statistical difference between both. On the other hand, when the airline and scheduled arrival hour of the initial flights are taken into account there is a noticeable difference when compared to the random case. With respect to the hour of the day when the perturbation starts, the top impact tree flights are concentrated in the early morning hours (only 34% of the all daily operational hours). On the contrary, when flights are randomly selected this percentage increases up to 67%. Most striking are the results regarding the airline that the initial flight belongs to. In this case the statistical difference almost triples, with 33% for the top impact trees (5 airlines) while the randomly selected flights affect to 95% of the airlines. These results merely confirm what was observed in Figure 7.4, although we would have expected the destination airport to be another key aspect of high impact flights.

7.5

Dynamical patterns of flight delays

We still lack an explanation of the role of networks hubs in delay-spreading. We have shown that these are more vulnerable to perturbations throughout the system, and we proposed that the key driver behind this observation may be

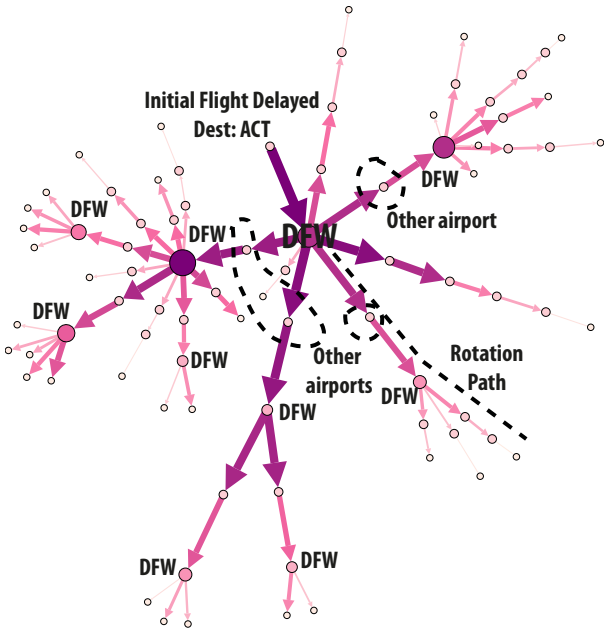


Figure 7.5: 4rd highest impact tree for July 13 with an initial delay of 60 min. Schedule arrival: 12 : 05 pm EST (local time: 11 : 05 am CDT). Airline code: 20398.

the reinforcement pattern caused by flights that repeatedly go to a spoke (other airports that are not hubs in the system) and then return to the hub, magnifying the delay on the hubs and subsequently on the whole system. To explore this issue we reconstructed the trees of reactionary delay for the top 5 highest impact ones. Figures 7.6 and 7.5 show, respectively, the 1st and the 4rd highest impact trees, the rest of the top 5 impact trees are shown in Appendix **??**. Each tree node represents a different flight and there is a link if one flight is connection of the other. The edge width and the color intensity are proportional to the flight delay. Each Figure contains information on which was the destination of the initially delayed flight and the rotation path that was followed by the aircraft of this first delay. It is also shown the destination airport of some flights, when the destination is a network hub, and when it is not a hub it is labeled as "other airports". The Figures shown here were chosen to show trees of different airline, that consequently repeatedly go and return to a different hub, in the examples ATL (Atlanta Hartsfield-Jackson international airport) and DFW (Dallas/Fort Worth interantional airport). The other trees on the Appendix use as hubs ATL and LAX (Los Angeles international airport). In all there is a recurrent pattern: flights continuously go back and forth to the different hubs, and is in the hubs where the spreading mostly takes part by multiplying the branching process. This result is conclusive showing that the role of hubs is critical for boosting the delay and act as a delay multiplier magnifying the number of flights delayed in the system. Clearly the pattern occurs as a result of how most airlines base their operations and develop their connectivity network as "hub and spoke" for

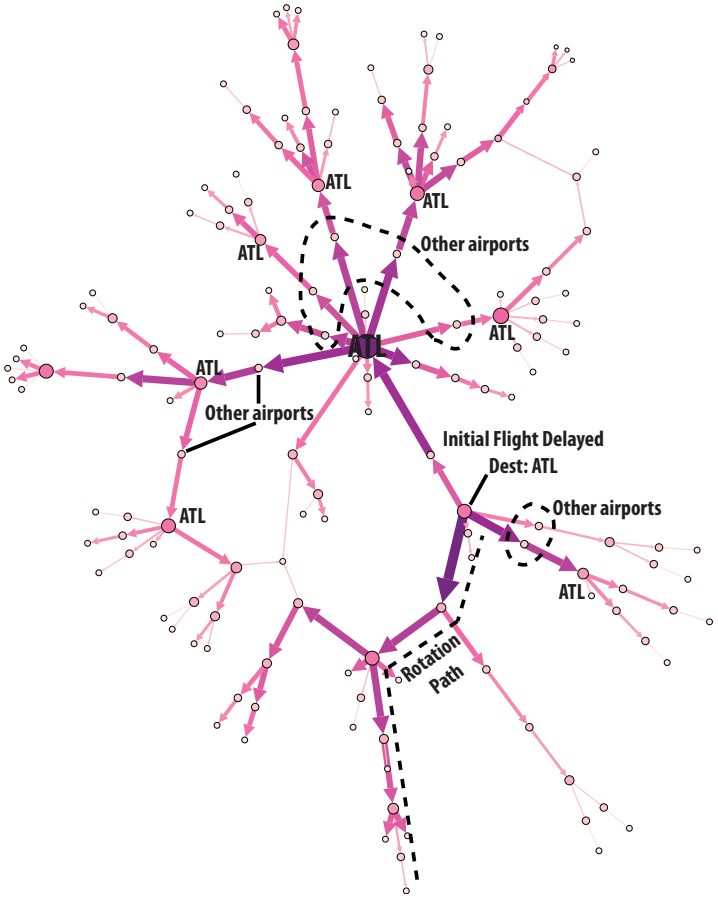


Figure 7.6: Highest impact tree found for July 13 with an initial delay of 60 min. Schedule arrival: 8 : 55 am EST. Airline code: 20366.

increasing the efficiency of connections throughout the system, but this, as has been shown has clear implications on delay spreading.

In summary, we have shown that hubs are more vulnerable to perturbations throughout the system than medium and small sized airports. Among other results, we explore the influence of the airline and time of the day of the initially perturbed flights. Results display a dependence on the airline the flight belongs to, specially with regards to high impact flights. Also perturbations that start in the morning have higher impact than those in the afternoon because of a larger temporal cone of events. However, perturbations that start in the early hours of the day are an exception because of a relatively lower average reproductive number. Thus, we can conclude that the interplay between the airline connectivity pattern and the time of the day are two major causes of the different behaviors encountered in dynamics of delay propagation at the individual flight level.

Part VI

Discussion and Outlook

CHAPTER 8

Discussion and Outlook

Next we summarize the most important findings following the approach used in this Dissertation. We begin by summing up the results related to the WAN and US topology. We then, describe the key aspects of the model developed to explore delay propagation in air transportation and its results related to the internal mechanisms responsible for the delay dynamics and the effects of perturbations in the system. Finally, we discuss the possibility of extending the present approach to other transport systems and asses the viability of using the techniques used here to develop tools for real management problems.

8.1 Air transportation system: a topological approach

In summary, we have analyzed the structural characteristics of the WAN and US airport network paying special attention to the aspects related to delay propagation of secondary delays. We begun by exploring the meso-scale structure of the WAN finding dissimilar partitions that highly depend on the methods used. In this dissertation we applied widely used methods: OSLOM, INFOMAP and modularity optimization techniques (Fast Greedy & Louvain algorithms). To overcome this problem we developed a methodology that extricates the most significant and alike communities between the applied methods. The proposed methodology is based on a systematic comparison of each algorithm result with a null model and examine whether these clusters are consistent across the algorithms used. The results, for instance, could be of used for redefining the IATA regions according to traffic patterns encoded in the WAN topology. Furthermore, the present methodology can be extended not only for the air-transportation system but, also, to systems of a very different nature such as biological and social systems.

We then focused on the US Aiport Network a system for which we have performance data regarding flight delays. Therefore, a daily network has been built by taking into account the direct flights connecting couples of airports and its topological properties such as the distribution of number of flights per airport or

the number of destination per airport. In addition to the topology, we consider also the properties of the airplane rotation in the network and of the real delays observed. The airplane rotation shows a complicated and highly heterogeneous profile with some airplanes covering essentially back and forth routes and others not closing the routes in a simple periodic way. The heterogeneity of the rotation procedures can play a role in the development and propagation of delays.

8.2

Modeling delay dynamics as a cascading failure mechanism

Regarding the real delays, we initially focus on basic properties as, for instance, delay distributions, which show long decays both for arrival and departure delays. The long tails are usually indicative of the complex nature of the mechanisms contributing to the propagation of delays. Similar distributions have been, for example, observed in the size of human epidemics when the infectiveness is close to the propagation threshold (Balcan et al., 2009a). In this case, the system is not necessarily working under critical conditions but the combined action of several factors such as connecting passengers, a predetermined schedule and the geographical distance of the airports can contribute to reach a similar system state at a global level. We study also the properties of the flights with a delay higher than 12 hours, those in the tail of the delay distribution. The evolution of the number of these long-delayed flights as a function of the hours of the day shows a relative concentration early in the morning or late in the afternoon. The destination airport seems to be important to understand which are the elements influencing the surge of the flights with long delay. These results are relevant in order to understand the mechanism behind the propagation of real delays. To do so, we introduce a measure for the level of network-wide extension of the delays by defining when an airport is considered as congested and studying how congested airports form connected clusters in the network. The size of the largest congested cluster displays in the data a high variability from one day to the next. This feature is due to the re-start that the system suffers at the end of each day and points toward the relevance of the daily schedule to define the delay propagation patterns.

In addition, we introduce an agent-based data-driven model able to reproduce the delay evolution observed in the data. It works as a queuing model at the level of airports and reproduces cascading failure dynamics through the connection between flights. In this regard, the model includes three main mechanisms to spread delays: plane rotation, flight connections of either passenger or crews and airport congestion. The last two processes can be modulated at will to understand the role that each one of them plays in delay propagation. Our simulations evidence that passenger and crew connections is the most effective

single mechanism leading to network congestion. The model not only has explanatory value but also predictive power as well. We show how it can be used to study the relevance of primary delay localization for the evolution of congestion in the network.

8.3

System response to perturbations

After signaling the most influential internal mechanisms for the propagation of delays we adapted the model to assess the system ability to deal with an increase in the number of disruptive events. We therefore analyze the effects that external disturbances and interventions produce in the US air traffic network. In particular, our analysis focuses on October 27 2010 because a large meteorological disruption occurred that day. We present the results as a function of the level of network-wide extensions of the delays. By computing the evolution of the largest congested cluster size of the day, we compare the empirical results with the delay dynamics observed in the model and find good agreement when weather impacts and canceled flights are considered as input variables. With a data-driven model we are able to implement the schedule information, weather effects and different intervention measures that took place that day. We introduce weather impacts by varying the airport capacity parameter β of some airports. A variation of β produces a drop in the airport capacity service rate thus enlarging the airport queue. Flight cancellations are implemented by affecting the network connectivity parameter α, thus reducing the delay propagation dynamics. Our simulations evidence that weather impacts could produce system congestion independently of the day considered, as it is the case when the initial conditions and same input perturbations are introduce to the schedule of October 20, an a priori low congested day.

Encouraged by these results, we define a set of measures to assess the susceptibility of the different elements that compose the air transportation system. We explore the response of the system by means of impact and robustness to perturbations at different system levels; namely, at the airport and at the individual flight level. We find that the airport impact on the system has no clear relation to the airport degree, measured as the number of different destinations of the flights leaving the airport. This, as explained in the corresponding chapter, occurs if the impact is normalized by the number of flights leaving the airport. It is thus clear that traffic influences the impact that a perturbation has, but the definition of airport impact was constructed to look for other topological mechanisms other than traffic load. In addition, we show that hubs are more vulnerable to perturbations throughout the system than medium and small sized airports. Among other results, we explore how the airline and time of the day of the initially perturbed flights affects the spread of congestion across the system. Results show a dependency on the airline, specially with regards to high impact flights.

Also, naturally, perturbations that start in the morning have higher impact than those in the afternoon because they have more opportunities to propagate the delays. However, perturbations that start in the early hours of the day are an exception because of a relatively lower average reproductive number. Therefore, we can conclude that there is a trade-off between the airline connectivity and the time of the day, and the interplay between this two are the major causes of the different behaviors encountered in dynamics of delay propagation at the individual flight level. We have also identified an important aspect regarding the cascading mechanism behind delay propagation, namely, the reinforcement pattern of back and forth flights on the network hubs. This dynamical pattern is the reason behind high impact flights and a major cause for delay spreading and reinforcement. Finally, a next step regarding the study of perturbations could be to explore the relationship between the system's resilience and network percolation.

8.4

Outlook for transportation systems

The methodology employed here generates results rich in details that can be used as a predictive tool. Furthermore, our model offers the possibility of evaluating different policy decisions before their real implementation. We show a way of introducing different external inputs that can be used at the strategic planning level to assess possible delay management tools for airports, airlines and the whole network. There are, of course, many possible interventions whose efficiency could be assessed. To give an example, we can quantify the sensitivity of airports to delays, or which ones are most prone to magnify delays, and design palliative measures customized for each airport concentrated, for instance, in increasing the slack time in turnarounds. In addition, with the identification of the reinforcement dynamical patterns, we can then explore how slight modification to these patterns can make the system more efficient. Motivated by these results, researchers have pursued on adapting the model for the European setting. This in principle is not trivial due to the fact that operations at the level of airports are very dissimilar with respect to the US system, and the first results have shown interesting and different behaviors (Campanelli et al., 2014; Ciruelos et al., 2015).

Finally, our results focus on the air transportation system but the concepts and techniques employed can be easily extrapolated to the analysis of the performance of a generic transport system. Flight delays represent failures to meet constraints imposed by a daily schedule. Its propagation in the network is a paradigmatic example of the way in which a distributed transport system moves toward collapse. Its translation to other airport networks is, of course, straightforward, and even though the modeling of other transportation systems may require some particular details, the applicability of the metrics defined

to measure network-wide congestion based on clustering is universal. The framework developed in this work is thus of easy extension to systems with dynamics regulated by predefined schedules. A possibility worth exploring, outside transportation systems is what it has been recently defined as Complex Projects (Remington and Pollack, 2007; Williams, 1999). The increasing complexity of endeavors such as the AIRBUS 380 construction, the Panama Canal Expansion or the National Broadband Network in Australia introduce serious problems to meet deadlines and budget. Recent studies have shown that projects unable to meet time and budget, leading to project failure, are an everyday situation, especially in IT (Yeo, 2002; Al Neimat, 2005; Evans, 2005). Despite the clear difference in nature with the object of study in this Thesis they all are a time-constrained effort to efficiently manage a myriad of connected events, activities and stakeholders that interact between different layers driving the system through a predefined schedule.

Bibliography

AhmadBeygi, S., Cohn, A., Guan, Y., and Belobaba, P. (2008). Analysis of the potential for delay propagation in passenger airline networks. *Journal of air transport management*, 14(5):221–236.

Al Neimat, T. (2005). Why it projects fail. *The project perfect white paper collection*, 8.

Albert, R., Albert, I., and Nakarado, G. L. (2004). Structural vulnerability of the north american power grid. *Physical review E*, 69(2):025103.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

Aldecoa, R. and Marín, I. (2013). Exploring the limits of community detection strategies in complex networks. *Scientific reports*, 3.

Allan, S., Beesley, J., Evans, J., and Gaddy, S. (2001). Analysis of delay causality at newark international airport. In *4th USA/Europe Air Traffic Management R&D Seminar*.

Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., and Vespignani, A. (2009a). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489.

Balcan, D., Hu, H., Goncalves, B., Bajardi, P., Poletto, C., Ramasco, J. J., Paolotti, D., Perra, N., Tizzoni, M., Broeck, W. V., et al. (2009b). Seasonal transmission potential and activity peaks of the new influenza a (h1n1): a monte carlo likelihood analysis based on human mobility. *BMC medicine*, 7(1):45.

Banavar, J. R., Maritan, A., and Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature*, 399(6732):130–132.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439):509–512.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752.

Barthélemy, M. (2011). Spatial networks. *Physics Reports*, 499(1):1–101.

BIBLIOGRAPHY

Barzel, B. and Barabási, A.-L. (2013). Universality in network dynamics. *Nature physics*, 9(10):673–681.

Batty, M. (2007). *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.

Bazzan, A. L. C. and Klugl, F. (2009). *Multi-Agent Systems for Traffic and Transportation Engineering*. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA.

Beatty, R., Hsu, R., Berry, L., and Rome, J. (1999). Preliminary evaluation of flight delay propagation through an airline schedule. *Air Traffic Control Quarterly*, 7(4):259–270.

Belobaba, P., Odoni, A., and Barnhart, C. (2009). *The global airline industry*, volume 23. John Wiley & Sons.

Bianconi, G., Gulbahce, N., and Motter, A. E. (2008). Local structure of directed networks. *Physical review letters*, 100(11):118701.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., and Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298.

Bonnefoy, P. A. and Hansman, R. J. (2007). Scalability and evolutionary dynamics of air transportation networks in the united states. In *Procs. of 7th AIAA Aviation Technology and Operations Conference*.

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. (2008). On modularity clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 20(2):172–188.

Buhl, J., Gautrais, J., Reeves, N., Solé, R., Valverde, S., Kuntz, P., and Theraulaz, G. (2006). Topological patterns in street networks of self-organized urban settlements. *The European Physical Journal B-Condensed Matter and Complex Systems*, 49(4):513–522.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., and Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028.

Burghouwt, G. and de Wit, J. (2005). Temporal configurations of european airline networks. *Journal of Air Transport Management*, 11(3):185–198.

Burghouwt, G., Hakfoort, J., and van Eck, J. R. (2003). The spatial configuration of airline networks in europe. *Journal of Air Transport Management*, 9(5):309–323.

Callaway, D. S., Newman, M. E., Strogatz, S. H., and Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical review letters*, 85(25):5468.

Campanelli, B., Fleurquin, P., Eguıluz, V., Ramasco, J., Arranz, A., Extebarria, I., and Ciruelos, C. (2014). Modeling reactionary delays in the european air transport network. In *The Fourth SESAR Innovation Days (Madrid)*.

Cardillo, A., Scellato, S., Latora, V., and Porta, S. (2006). Structural properties of planar graphs of urban street patterns. *Physical Review E*, 73(6):066107.

Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., del Amo, A. J. G., and Boccaletti, S. (2013). Modeling the multi-layer nature of the european air transport network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, 215(1):23–33.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591.

Chang, K., Howard, K., Oiesen, R., Shisler, L., Tanino, M., and Wambsganss, M. C. (2001). Enhancements to the faa ground-delay program under collaborative decision making. *Interfaces*, 31(1):57–76.

Christakis, N. A. and Fowler, J. H. (2010). *Connected: the amazing power of social networks and how they shape our lives*. HarperPress London.

Churchill, A., Lovell, D., and Ball, M. (2007). Examining the temporal evolution of propagated delays at individual airports: case studies. In *Proceedings of the 7th USA/Europe Air Traffic Management R&D Seminar*.

Ciruelos, C., Arranz, A., Etxebarria, I., Peces, S., Campanelli, B., Fleurquin, P., Eguiluz, V., and Ramasco, J. (2015). Modelling delay propagation trees for scheduled flights. In *Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)*.

Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):066111.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical review letters*, 85(21):4626.

Colizza, V., Barrat, A., Barthélemy, M., and Vespignani, A. (2006a). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7):2015–2020.

Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. (2006b). Detecting rich-club ordering in complex networks. *Nature physics*, 2(2):110–115.

Colizza, V. and Vespignani, A. (2007). Invasion threshold in heterogeneous metapopulation networks. *Physical Review Letters*, 99(14):148701.

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., Sanchez, A., et al. (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, 214(1):325–346.

Cook, A. (2007). *European air traffic management: principles, practice, and research*. Ashgate Publishing, Ltd.

da Rocha, L. E. (2009). Structural evolution of the brazilian airport network. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(04):P04020.

Danon, L., Diaz-Guilera, A., Duch, J., and Arenas, A. (2005). Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(09):P09008.

Dobson, I., Carreras, B. A., Lynch, V. E., and Newman, D. E. (2007). Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103.

Duch, J. and Arenas, A. (2005). Community detection in complex networks using extremal optimization. *Physical review E*, 72(2):027104.

ERDdS, P. and R&WI, A. (1959). On random graphs i. *Publ. Math. Debrecen*, 6:290–297.

Evans, M. (2005). Overdue and over budget, over and over again. *The Economist*, 6.

Ferrara, E. (2012). A large-scale community structure analysis in facebook. *EPJ Data Science*, 1(1):1–30.

Fleurquin, P., Ramasco, J. J., and Eguiluz, V. M. (2013). Systemic delay propagation in the us airport network. *Scientific reports*, 3.

Folkes, V. S., Koletsky, S., and Graham, J. L. (1987). A field study of causal inferences and consumer reaction: the view from the airport. *Journal of consumer research*, pages 534–539.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3):75–174.

Fortunato, S. and Barthélemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41.

Gantz, J. and Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the Future*, 2007:1–16.

Gastner, M. T. and Newman, M. (2006). Optimal design of spatial distribution networks. *Physical Review E*, 74(1):016117.

Gautreau, A., Barrat, A., and Barthélemy, M. (2009). Microdynamics in stationary complex networks. *Proceedings of the National Academy of Sciences*, 106(22):8847–8852.

Girvan, M. and Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

Grabowicz, P. A., Ramasco, J. J., Moro, E., Pujol, J. M., and Eguiluz, V. M. (2012). Social features of online networks: The strength of intermediary ties in online social media. *PloS one*, 7(1):e29358.

Guimera, R. and Amaral, L. A. N. (2004). Modeling the world-wide airport network. *The European Physical Journal B-Condensed Matter and Complex Systems*, 38(2):381–385.

Guimera, R., Mossa, S., Turtschi, A., and Amaral, L. N. (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences*, 102(22):7794–7799.

Guimera, R., Sales-Pardo, M., and Amaral, L. A. N. (2004). Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, 70(2):025101.

Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, 402:C47–C52.

Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of modern physics*, 73(4):1067.

Heppenheimer, T. A. and Heppenheimer, T. (1995). *Turbulent skies: the history of commercial aviation*. J. Wiley & Sons.

Hric, D., Darst, R. K., and Fortunato, S. (2014). Community detection in networks: Structural communities versus ground truth. *Physical Review E*, 90(6):062805.

Hu, Y. and Zhu, D. (2009). Empirical analysis of the worldwide maritime transportation network. *Physica A: Statistical Mechanics and its Applications*, 388(10):2061–2071.

BIBLIOGRAPHY

Huberman, B. A., Pirolli, P. L., Pitkow, J. E., and Lukose, R. M. (1998). Strong regularities in world wide web surfing. *Science*, 280(5360):95–97.

Huffman, D. A. et al. (1952). A method for the construction of minimum redundancy codes. *Proceedings of the IRE*, 40(9):1098–1101.

Hufnagel, L., Brockmann, D., and Geisel, T. (2004). Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15124–15129.

Janić, M. (2005). Modeling the large scale disruptions of an airline network. *Journal of transportation engineering*, 131(4):249–260.

Jeong, H., Mason, S. P., Barabási, A.-L., and Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, 411(6833):41–42.

Jetzki, M. (2009). The propagation of air transport delays in europe. *PhD diss., Department of Airport and Air Transportation Research RWTH, Aachen University*.

Jiao, X., Chang, S., Li, C.-h., Wang, C.-x., et al. (2007). Construction and application of the weighted amino acid network based on energy. *Physical Review E*, 75(5):051903.

Jonsson, P. F., Cavanna, T., Zicha, D., and Bates, P. A. (2006). Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. *BMC bioinformatics*, 7(1):2.

Kaluza, P., Kölzsch, A., Gastner, M. T., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, page rsif20090495.

Kernighan, B. W. and Lin, S. (1970). An efficient heuristic procedure for partitioning graphs. *Bell system technical journal*, 49(2):291–307.

Klein, A., Jehlen, R., and Liang, D. (2007). Weather index with queuing component for national airspace system performance assessment. In *7th FAA/Eurocontrol ATM Seminar, Barcelona, Spain*.

Klein, A., Kavoussi, S., and Lee, R. S. (2009). Weather forecast accuracy: study of impact on airport capacity and estimation of avoidable costs. In *Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)*.

Kurant, M. and Thiran, P. (2006). Extraction and analysis of traffic and topologies of transportation networks. *Physical Review E*, 74(3):036114.

Lacasa, L., Cea, M., and Zanin, M. (2009). Jamming transition in air transportation networks. *Physica A: Statistical Mechanics and its Applications*, 388(18):3948–3954.

Lämmer, S., Gehlsen, B., and Helbing, D. (2006). Scaling laws in the spatial structure of urban road networks. *Physica A: Statistical Mechanics and its Applications*, 363(1):89–95.

Lancichinetti, A., Fortunato, S., and Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical review E*, 78(4):046110.

Lancichinetti, A., Radicchi, F., and Ramasco, J. J. (2010). Statistical significance of communities in networks. *Physical Review E*, 81(4):046110.

Lancichinetti, A., Radicchi, F., Ramasco, J. J., and Fortunato, S. (2011). Finding statistically significant communities in networks. *PloS one*, 6(4):e18961.

Latora, V. and Marchiori, M. (2002). Is the boston subway a small-world network? *Physica A: Statistical Mechanics and its Applications*, 314(1):109–113.

Lenormand, M., Tugores, A., Colet, P., and Ramasco, J. J. (2014). Tweets on the road. *PloS one*, 9(8):e105407.

Li, W. and Cai, X. (2004). Statistical analysis of airport network of china. *Physical Review E*, 69(4):046106.

Lordan, O., Sallan, J. M., and Simo, P. (2014). Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda. *Journal of Transport Geography*, 37:112–120.

Luczkovich, J. J., Borgatti, S. P., Johnson, J. C., and Everett, M. G. (2003). Defining and measuring trophic role similarity in food webs using regular equivalence. *Journal of Theoretical Biology*, 220(3):303–321.

Malone, K. M. (1995). *Dynamic queueing systems: behavior and approximations for individual queues and for networks*. PhD thesis, Massachusetts Institute of Technology.

Mayer, C. and Sinai, T. (2003). Network effects, congestion externalities, and air traffic delays: Or why all delays are not evil. *American Economic Review*, 93(4):1194–1215.

McAfee, A. and Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, (90):60–6.

M'Kendrick, A. (1925). Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130.

Moreno, J. L., Jennings, H. H., et al. (1934). Who shall survive?

BIBLIOGRAPHY

Newman, M. (2010). *Networks: an introduction*. Oxford University Press.

Newman, M. E. (2002). Assortative mixing in networks. *Physical review letters*, 89(20):208701.

Newman, M. E. (2004a). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5200–5205.

Newman, M. E. (2004b). Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6):066133.

Newman, M. E. (2006a). Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, 74(3):036104.

Newman, M. E. (2006b). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582.

Opsahl, T., Colizza, V., Panzarasa, P., and Ramasco, J. J. (2008). Prominence and control: the weighted rich-club effect. *Physical review letters*, 101(16):168702.

Palla, G., Derényi, I., Farkas, I., and Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 435(7043):814–818.

Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Physical review letters*, 87(25):258701.

Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1):51–59.

Pyrgiotis, N., Malone, K. M., and Odoni, A. (2013). Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies*, 27:60–75.

Radicchi, F. (2014). A paradox in community detection. *EPL (Europhysics Letters)*, 106(3):38001.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., and Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2658–2663.

Remington, K. and Pollack, J. (2007). *Tools for complex projects*. Gower Publishing, Ltd.

Rosenberger, J. M., Schaefer, A. J., Goldsman, D., Johnson, E. L., Kleywegt, A. J., and Nemhauser, G. L. (2002). A stochastic model of airline operations. *Transportation science*, 36(4):357–377.

Rosvall, M. and Bergstrom, C. T. (2007). An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences*, 104(18):7327–7331.

Rupp, N. G. (2007). Further investigations into the causes of flight delays. Technical report, Working paper, Department of Economy, East Carolina University. Available online at http://www. ecu. edu/cs-educ/econ/upload/ecu0707. pdf.

Sales-Pardo, M., Guimera, R., Moreira, A. A., and Amaral, L. A. N. (2007). Extracting the hierarchical organization of complex systems. *Proceedings of the National Academy of Sciences*, 104(39):15224–15229.

Schaefer, L. and Millner, D. (2001). Flight delay propagation analysis with the detailed policy assessment tool. In *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, volume 2, pages 1299–1303. IEEE.

Scott, J. and Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. SAGE publications.

Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P., Mukherjee, G., and Manna, S. (2003). Small-world properties of the indian railway network. *Physical Review E*, 67(3):036106.

Shah, A., Pritchett, A., Feigh, K., Kalarev, S., Jadhav, A., Corker, K., Holl, D., and Bea, R. (2005). Analyzing air traffic management systems using agentbased modeling and simulation. In *Proceedings of the 6th USA/Europe Seminar on Air Traffic Management Research and Development*.

Sienkiewicz, J. and Hołyst, J. A. (2005). Statistical analysis of 22 public transport networks in poland. *Physical Review E*, 72(4):046127.

Snyder, D. and Kick, E. L. (1979). Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions. *American Journal of Sociology*, pages 1096–1126.

Sridhar, B., Wang, Y., Klein, A., and Jehlen, R. (2009). Modeling flight delays and cancellations at the national, regional and airport levels in the united states. In *8th USA/Europe ATM R&D Seminar, Napa, California (USA)*.

Strehl, A. and Ghosh, J. (2003). Cluster ensembles—a knowledge reuse framework for combining multiple partitions. *The Journal of Machine Learning Research*, 3:583–617.

Tizzoni, M., Bajardi, P., Poletto, C., Ramasco, J. J., Balcan, D., Gonçalves, B., Perra, N., Colizza, V., and Vespignani, A. (2012). Real-time numerical forecast of global epidemic spreading: case study of 2009 a/h1n1pdm. *BMC medicine*, 10(1):165.

BIBLIOGRAPHY

Ugander, J., Backstrom, L., Marlow, C., and Kleinberg, J. (2012). Structural diversity in social contagion. *Proceedings of the National Academy of Sciences*, 109(16):5962–5966.

Vespignani, A. (2012). Modelling dynamical processes in complex socio-technical systems. *Nature Physics*, 8(1):32–39.

Von Ferber, C., Holovatch, T., Holovatch, Y., and Palchykov, V. (2009). Public transport networks: empirical analysis and modeling. *The European Physical Journal B-Condensed Matter and Complex Systems*, 68(2):261–275.

Wang, P. T., Schaefer, L. A., and Wojcik, L. A. (2003). Flight connections and their impacts on delay propagation. In *Digital Avionics Systems Conference, 2003. DASC'03. The 22nd*, volume 1, pages 5–B. IEEE.

Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge university press.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-worldnetworks. *nature*, 393(6684):440–442.

Williams, T. M. (1999). The need for new paradigms for complex projects. *International journal of project management*, 17(5):269–273.

Woolley-Meza, O., Thiemann, C., Grady, D., Lee, J., Seebens, H., Blasius, B., and Brockmann, D. (2011). Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo-ship movements. *The European Physical Journal B*, 84(4):589–600.

Wu, C.-L. and Caves, R. E. (2000). Aircraft operational costs and turnaround efficiency at airports. *Journal of Air Transport Management*, 6(4):201–208.

Wu, Z., Braunstein, L. A., Havlin, S., and Stanley, H. E. (2006). Transport in weighted networks: partition into superhighways and roads. *Physical review letters*, 96(14):148702.

Wuellner, D. R., Roy, S., and DSouza, R. M. (2010). Resilience and rewiring of the passenger airline networks in the united states. *Physical Review E*, 82(5):056101.

Xu, N., Donohue, G., Laskey, K. B., and Chen, C.-H. (2005). Estimation of delay propagation in the national aviation system using bayesian networks. In *6th USA/Europe Air Traffic Management Research and Development Seminar*. Citeseer.

Yang, J. and Leskovec, J. (2013). Overlapping community detection at scale: a nonnegative matrix factorization approach. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 587–596. ACM.

Yeo, K. T. (2002). Critical failure factors in information system projects. *International Journal of Project Management*, 20(3):241–246.

Zhang, J., Cao, X.-B., Du, W.-B., and Cai, K.-Q. (2010). Evolution of chinese airport network. *Physica A: Statistical Mechanics and its Applications*, 389(18):3922–3931.

Zhou, S. and Mondragón, R. J. (2004). The rich-club phenomenon in the internet topology. *Communications Letters, IEEE*, 8(3):180–182.

Zou, B., Elke, M., and Hansen, M. (2013). Evaluating air carrier fuel efficiency and co2 emissions in the us airline industry. In *Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013)*.