



Delay causality network in air transport systems

Wen-Bo Du^a, Ming-Yuan Zhang^{a,b}, Yu Zhang^c, Xian-Bin Cao^{a,*}, Jun Zhang^{d,*}

^a School of Electronic and Information Engineering, Beihang University, Beijing 100191, PR China

^b Shen Yuan Honors College, Beihang University, Beijing 100191, PR China

^c Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620, USA

^d Beijing Institute of Technology, Beijing 100081, PR China



ARTICLE INFO

Keywords:

Flight delay
Delay propagation
Delay causality
Network analysis

ABSTRACT

To better understand the mechanism of flight delay propagation at the system-level, we built a delay causality network (DCN) based on the Granger causality test. Through topological analysis of DCNs, we found that only about a quarter of airports were involved in delay propagation during the peak travel period and large airports affected by many upstream airports impact fewer downstream airports. Furthermore, temporal analysis of DCNs indicates that the culprits of delay propagation in the air transport system are not a fixed set of airports; instead, they vary daily depending on the operational environment.

1. Introduction

With increasing globalization, the world's civil aviation industry has been growing at a fast pace (Barnhart et al., 2009). The flight delay problems that have resulted from the rapid development of the civil aviation industry have become a worldwide challenge (Czerny, 2010). Flight delays have negative impacts on several aspects, such as passengers, airlines, and air transport systems (Ahmadbeygi et al., 2008). Delayed flights throw travel plans into disarray, often making passengers dissatisfied with the airlines (Britto et al., 2012). Airline companies also suffer, not only paying for the resource waste caused by delays but also having to invest more to improve passenger satisfaction (Zou and Hansen, 2014). Due to flight delays, air transport systems are faced with reduced efficiency and increased security risk, leading to economic loss and environmental pollution (JEC report, 2008). A recent study reported that the total direct cost induced by flight delays was nearly \$28.9 billion in the United States in 2007 (Ball et al., 2010) (see Table 1).

An initial flight delay can be attributed to several reasons, such as air carrier issues, extreme weather, air traffic control, etc. (BTS report, 2017). However, a propagated delay occurs because of connected resources (Kafle and Zou, 2016). The most common resource is aircraft (Zou and Hansen, 2014). Because the same aircraft flies multiple flight legs, the delay of an earlier flight can affect the subsequent flights of the same aircraft (Lan et al., 2006). If passengers are not free from a previous delayed flight, the next flight will be delayed by waiting for it. Flight crews also switch between different aircraft, causing the delay from one flight to propagate across multiple flights (Beatty et al., 1999; Wang et al., 2017). For these reasons, a small initial delay may lead to larger delays later, inducing much worse situations (Li et al., 2014; Meng and Zhou, 2011). Therefore, research on the mechanism of delay propagation is timely yet challenging.

* Corresponding authors.

E-mail addresses: xbcao@buaa.edu.cn (X.-B. Cao), buaazhangjun@vip.sina.com (J. Zhang).

<https://doi.org/10.1016/j.tre.2018.08.014>

Received 7 February 2018; Received in revised form 28 August 2018; Accepted 29 August 2018

1366-5545/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1
The Granger test results.

| Pairwise test | F-statistic | P-value | Hypothesis | GC relationship |
|---------------|--------------|--------------|---------------|-----------------|
| 1–2 | 2.256 | 0.127 | Accept | Not exist |
| 1–3 | 1.764 | 0.200 | Accept | Not exist |
| 1–4 | 1.139 | 0.368 | Accept | Not exist |
| 2–1 | 0.223 | 0.879 | Accept | Not exist |
| 2–3 | 1.069 | 0.394 | Accept | Not exist |
| 2–4 | 0.368 | 0.777 | Accept | Not exist |
| 3–1 | 4.644 | 0.019 | Reject | Exist |
| 3–2 | 0.955 | 0.441 | Accept | Not exist |
| 3–4 | 0.558 | 0.651 | Accept | Not exist |
| 4–1 | 0.840 | 0.494 | Accept | Not exist |
| 4–2 | 0.358 | 0.784 | Accept | Not exist |
| 4–3 | 0.829 | 0.500 | Accept | Not exist |

1.1. Literature review

The traditional Approximate Network Delays (AND) model was originally conceptualized in a prototype form of three-airport networks (Malone, 1995). Pyrgiotis et al. (2013) enriched the AND model and investigated the delay propagation based on 34 US airports. Their results showed that delay propagation tends to mitigate daily airport demand profiles and push more demands into late evening hours. Zhang and Nayak (2010, 2011) used the multivariate simultaneous equation regression (MSER) model to study the impact of a single airport on the others, and vice versa. Their results revealed that major airports have a higher impact on the average delay. Later, Hao et al. (2014) used the MSER model and the Federal Aviation Administration (FAA) system-wide analysis capability (SWAC) simulation model to quantify the impact of the three airports in the New York area on delays throughout the airport network, finding that the delays within the New York area are lower than expected. Fleurquin et al. (2013) developed the maximum connected subgraph of congested airports for assessing the level of delays across the entire system. They also introduced a model that comprehends aircraft rotation, passenger connectivity, and airport congestion as well as crew rotation to simulate the propagation of delays. This model can simulate the congestion of the system accurately. Then, they proposed a new model involving slot reallocation and swapping to simulate the propagation of reactionary delays in Europe (Campanelli et al., 2014). Afterward, Campanelli et al. (2016) used these two models to simulate flight delay propagation and assessed the effect of disruptions in the US and European aviation networks.

Despite the advances in understanding flight delay propagation, few studies have investigated delay propagation by considering the interdependence relationship of delay time-series. Thus, a systematic framework probing the causal relationship among airports continues to be elusive. Recent years have witnessed a growing interest in the inference of causal interactions (Wahl et al., 2017; Stokes and Purdon, 2017) in complex systems. Thanks to theoretical innovation, the application fields of causality tests have grown to include biology (Stokes and Purdon, 2017), ecology (Sugihara et al., 2012), social sciences (Frank et al., 2018), physics (Martin et al., 2016) and economics (Song et al., 2008). All of these studies have a common theme: the details of the temporal mutual influence of units are difficult to understand, and causality tests are used to detect the interaction patterns in dynamical systems by time series analysis. All of the prior results have shown that causality tests yield new insights into large-scale complex systems.

The air transportation system is also a typical large-scale complex system. Due to its complexity, the mechanisms of delay propagation are not fully understood, especially for the interdependencies of different airports. Causal analysis may provide a new perspective on this problem. In this study, we adopted Granger causality (Granger, 1969) as the main method due to its primary advance on the causation problem (Frank et al., 2018). Then we built a delay causality network (DCN) based on the Granger causality test and investigated the topological and temporal properties of the DCNs, offering insights into the features of specific airports.

1.2. Contributions and outline

We apply a theoretical framework of causality test to study the delay propagation of the complex airport system. By considering delay propagation problem from the perspective of delay time-series interdependence, this approach can capture the interaction patterns of delay between different airport pairs. Due to the large number of airports and their complex relationships, the features of delay propagation cannot be understood from information at the individual airport level alone. We construct DCNs to characterize the global structure and dynamics of delay propagation, revealing the direction and range of delay propagation at the network-level. Although air transport system have been abstracted to a directed/undirected, weighted/unweighted network in previous studies (Wang et al., 2016), existing studies have mostly considered static graphs, meaning that the dynamics of the network were could not explicitly be considered explicitly (Ren and Li, 2018). In our research, the edges of DCN are the results of daily temporal interactions, representing the functional connectivity and underlying operational conditions. Theory and application of complex networks are used to further reveal the properties of DCNs. We use the degree, reciprocity, clustering coefficient, community, and other metrics of complex network to describe the situation of delay propagation and find that only about a quarter of airports in China face delay propagation during the peak travel period in China. The results also reveal that large airports affected by many upstream airports actually impact fewer downstream airports. By studying the connected clusters formed by high-degree airports, we find that the

culpits of delay propagation in the air transport system are not a fixed set of airports; instead, they vary daily depending on the operational environment.

These findings not only help us understand the complex aviation system better, but also provide support to air traffic managers on decision-making. According to the findings, the air traffic managers can develop effective countermeasures to prevent the delay propagation in particular links to alleviate network-wide flight delay. Furthermore, through causality test, if one variable acts as the cause for another one, actively intervening on the first would lead to the changes in the second. Thus, policy makers could potentially adopt the proposed method to analyze the interaction of airports in order to identify critical ones. It will help them to make decisions on resource allocation for improving airport capacity.

The remainder of the paper is organized as follows: Section 2 introduces the methodology, including the Granger causality test, construction of the DCN, and network analysis; Section 3 reports the results, and Section 4 is the conclusions and discussions.

2. Methodology

2.1. Delay time series

We utilize delay time series to represent the on-time performance of an airports. We focus on daily time series because daily interactions are the finest temporal resolution in the flight data set. For airport i , we construct its delay time series Y_i by splitting one day into 24 time intervals. The value of each time interval represents the average delay $d_i(t)$. The average delay for airport i is defined as

$$d_i(t) = \frac{u_i(t) + c_i(t)*h}{s_i(t)}, t \in \{0, 1, 2, \dots, 23\} \tag{1}$$

where $u_i(t)$ represents the total delay of departure flights at airport i during $(t, t + 1)$, and $c_i(t)$ and $s_i(t)$ represent the numbers of cancelled flights and scheduled departure flights at airport i during $(t, t + 1)$, respectively. Traditional methods do not consider flight cancellations. However, not considering flight cancellations, under extreme conditions, may generate bias in the airport operations. Cancellations should be taken into account as a delay metric for assessing the performance of air transport systems (Xiong and Hansen, 2009). According to the regulations of the FAA, the Civil Aviation Administration of China (CAAC), and the European Aviation Safety Agency (EASA), h can be the equivalent delay time of a cancellation ($h = 180$). With this method, a daily delay time series can be constructed as a series of data points of $d_i(t)$ indexed sequentially as a function of time.

2.2. Causality test

In our research, causality reveals the impacts between airport pairs and reflects the interaction of airport delays. If the delays observed at one airport can explain the delays appearing at a second one after several hours, there exists a causality relationship. Here, Granger causality (GC) will help understand the existence and direction of the influence between two airports based on the delay time-series. A time series Y_j is said to cause the time series Y_i if it can be shown that the values in Y_j provide statistically significant information about the future values in Y_i .

First, the Granger causality test uses an unrestricted regression equation to obtain the residual sum of squares:

$$y_i^T = \sum_{m=1}^{p_{ij}} a^m y_i^{T-m} + \sum_{m=1}^{p_{ij}} b^m y_j^{T-m} + e^T \tag{2}$$

where y_i^T is the current value of delay time series Y_i , y_i^{T-m} is the past value of time series Y_i , y_j^{T-m} is the past value of time series Y_j , e^T is the error term, and a^m and b^m are coefficients. In addition, p_{ij} stands for the lag, indicating that the current value should be regressed with the values in the past p_{ij} hours. Then, the null hypothesis that j does not cause i is defined as

$$b^1 = b^2 = \dots = b^{p_{ij}} = 0 \tag{3}$$

Second, we apply a restricted regression equation to obtain the residual sum of squares RSS_R :

$$y_i^T = \sum_{m=1}^{p_{ij}} a^m y_i^{T-m} + e^T \tag{4}$$

Finally, F-statistic and p-value are adopted to test the null hypothesis:

$$F = \frac{(RSS_R - RSS_{UR})/p_{ij}}{RSS_{UR}/(w - p_{ij})} \tag{5}$$

where w is the sample size of each time series. When the p-value is less than the chosen significance level ∂ (5% by default), the null hypothesis is rejected. If item Y_j belongs to this regression, Y_j is the cause of Y_i ; hence, the value in i is partly attributed to j .

2.3. Individual test

Through the causality test process above, we can assess the delay time series of each airport pair for each day. However, before

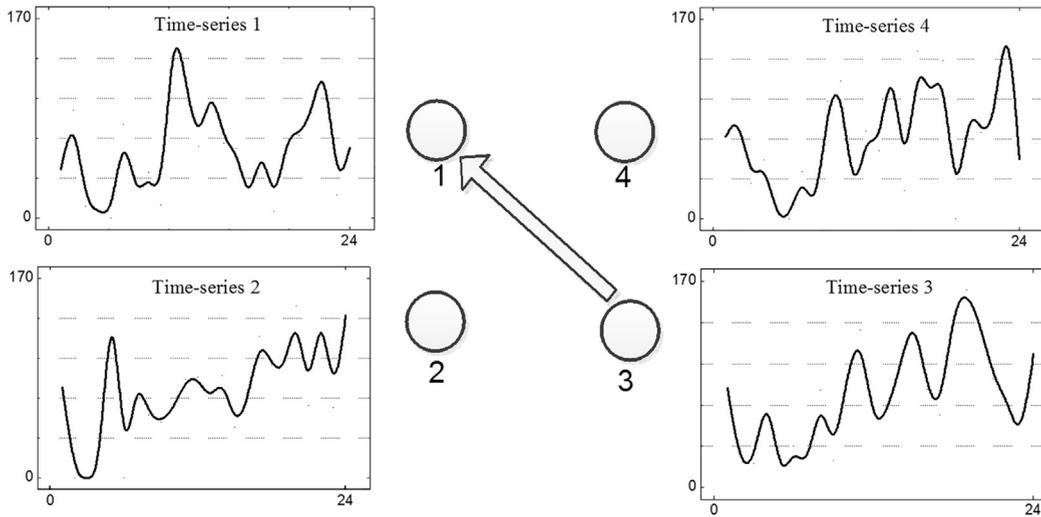


Fig. 1. Individual test.

applying the method to delay propagation, the following criteria need to be met:

- Due to the regression basis of causality detection, the current delays of an airport should not be independent but closely related to its delays that occurred during earlier time periods as well as other airports. This is reasonable because current delays in one airport cannot break off relations with delays in previous periods of itself or other upstream airports.
- Lag p_{ij} equals the average flying time between the airport pair plus turnaround time in airport j . Therefore, airport activities are represented by departure delay time series, interactions of which are captured by the pairwise causality test. We postulate that the impact is still effective until the aircraft of the connecting flight (from airport i to airport j) departs from airport j , i.e., executing its subsequent flight. Statistical analysis shows that the turnaround time of nearly 95% of the flights is less than 120 min. Therefore, we postulate that the lag p_{ij} equals the average flying time between the airport pair plus 120 min.

By satisfying these two criteria, the causality relationship of an airport pair can be determined. Here, we utilize four airports as examples to demonstrate the method for determining. For an airport pair, we construct their departure delay time series and obtain the average flight time of an airport pair. Then, GC is used to conduct the pairwise test. According to Fig. 1, when the p-value is less than 5%, the causality relationship can be accepted, and a directed edge connects these two airports. The causality relationship between airport 1 and airport 3 represents the interactions caused by delay propagation, i.e., delays in airport 1 are affected by airport 3. This relationship is not monopolistic. Delays in airport 1 could be partly attributed to not only airport 3 but also other airports. Moreover, there are many airports in air transport system; thus, determining the relationships are very complicated among different airports. Pairwise analysis cannot handle the complexity of the system-level delay propagation.

2.4. Network analysis

As aforementioned, due to the large number of airports and complex interactions, the features of delay propagation cannot be understood from the information at the individual airport level alone. Complex network theory and its associated metrics and tools present an apposite approach to study the air transport system beyond what is offered by classical techniques (Cook et al., 2015). Thus, a network-level analysis is adopted to capture the global structure of the functional interaction.

We construct a DCN of one day using the above method and analyze it with the tools provided by network science (Cardillo et al., 2013; Wang et al., 2014; Du et al., 2016). Fig. 2(a) presents a sample network to demonstrate the network analysis method. $A_{ij} = (a_{ij})_{N \times N}$ is the adjacency matrix of the sample network. Here, $a_{ij} = 1$ if, and only if, there exists an edge from node i to node j , otherwise $a_{ij} = 0$. In this sample network, the total number of airports is $N = 8$ and the total number of edges is $M = \sum_{i,j=1}^N a_{ij} = 12$. Some topologies with practical significance are introduced here to help analyze a system-level delay propagation.

Degree of an airport reflects the number of airports that have delay propagation links with it. In a directed network, airport i has in-degree ($k_i^{in} = \sum_{j=1}^N a_{ji}$) and out-degree ($k_i^{out} = \sum_{j=1}^N a_{ij}$) links, denoting the number of airports affected by airport i and the number of airports affecting airport i , respectively. The total degree of node i is $k_i = k_i^{in} + k_i^{out}$. For a certain node in Fig. 2(b), airport 1 is affected by airport 3 and affects airports 2 and 3. Thus, k_1^{in} is one, whereas k_1^{out} is two. The average degree of the network in Fig. 2(a) is $\langle k^{in} \rangle = \langle k^{out} \rangle = M/N = 1.5$, meaning that each airport affects 1.5 others on average.

Reciprocity Parameter represents the bidirectional nature of delay propagation links between airport pairs. As shown in Fig. 2(c), reciprocity means that airport i affects airport j , whereas airport j also affects airport i ($a_{ij} = a_{ji} = 1$). The parameter R is used to measure the overall symmetry of a directed network (da Rocha, 2009). It is defined as

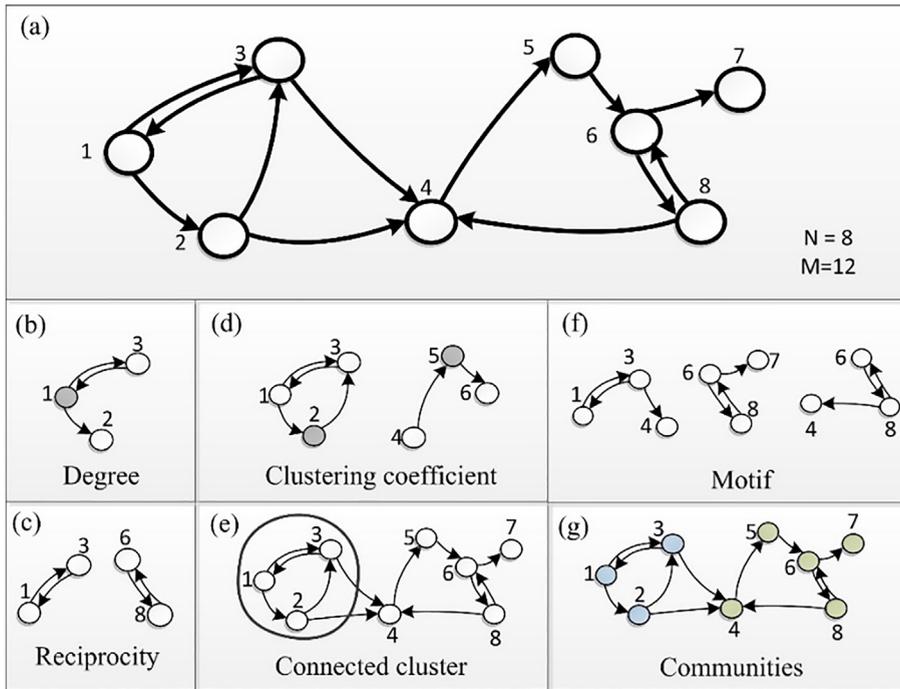


Fig. 2. Features of a sample network.

$$R = \frac{\sum_{i \neq j}^N (a_{ij} - \bar{a})(a_{ji} - \bar{a})}{\sum_{i \neq j}^N (a_{ij} - \bar{a})^2} \tag{6}$$

where

$$\bar{a} = \frac{\sum_{i \neq j}^N a_{ij}}{N(N-1)} \tag{7}$$

The maximum R is 1, implying that the delay propagation between all airport pairs is bidirectional. Larger R values indicate that a network is more symmetric.

Clustering coefficient is used to qualify the inherent cluster tendency of airports (Watts and Strogatz, 1998; Fagiolo, 2007). The clustering coefficient of an airport is the fraction of pairs of its neighbor airports (airports with delay propagation links to an airport) that have a direct delay propagation link (i.e., the number of triangles in the network). In Fig. 2(d), two neighbor airports of Airport 2 interact with each other, forming a clique, whereas the two neighbor airports of Airport 5 have no relationship. Thus, the clustering coefficient of node 2 is larger than node 5. For a network, the overall clustering coefficient is calculated as

$$C^D = \frac{1}{n} \sum_{i=1}^n \frac{\left(\frac{1}{2}\right) \sum_j \sum_h (a_{ij} + a_{ji})(a_{ih} + a_{hi})(a_{jh} + a_{hj})}{[d_i^{dot} (d_i^{dot} - 1) - 2d_i^{leftrightarrow}] } \tag{8}$$

where

$$d_i^{dot} = \sum_{j \neq i} a_{ji} + \sum_{j \neq i} a_{ij} \tag{9}$$

and

$$d_i^{leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji} \tag{10}$$

C^D is 0.177 for the network in Fig. 1(e).

Largest Connected Cluster is introduced to represent the extent of delay propagation (i.e., disaster area). The largest connected cluster (Broder et al., 2000) is a group in which airports are connected by propagation links. To make it represent the disaster area of delay propagation, we set an effective baseline for the members of a cluster. An airport is considered when it affects several other airports (k_i^{out} exceeds a threshold). If we postulate that the threshold is $\langle k_i^{out} \rangle = 1.5$, Fig. 2(e) shows that the largest connected cluster is formed by airports 1, 2, and 3. The size of largest connected cluster is M_d is 3.

To measure the cluster similarity of daily networks, the Jaccard index J is introduced, which is defined as: $J = |A \cap B| / |A \cup B|$,

where both A and B are finite sample sets comprising the airport members. J equals one if the airports sets of clusters are the same, and zero if they are strictly different.

Community is used to evaluate whether the delay propagation among airports can be divided into several sub-regions in which each sub-region of airports have dense delay propagation links internally and sparse links with the rest of the system. (Arenas et al., 2007). Fig. 2(g) shows the two communities of the sample network. Furthermore, Modularity is designed to measure the strength of division of a network into communities (Newman, 2006). Q_d in a directed DCN is defined as

$$Q_d = \frac{1}{M} \sum_{ij} \left(A_{ij} - \frac{k_i^{out} k_j^{in}}{M} \right) \delta(c_i, c_j) \tag{11}$$

The δ function yields one if vertices i and j are in the same community, and is zero otherwise. M is the total number of edges, and Q_d is 0.280 for the network in Fig. 1.

Network motifs are utilized to show the local relationship pattern among any three airports. It is defined as recurrent sub-graphs G' in network G (Shenorr et al., 2002). For example, Fig. 2(f) shows that the three airport groups have the similar relationship pattern. One important tool to evaluate the significance level of motifs is Z-score, which is defined as follows:

$$Z(G') = \frac{F_G(G') - \mu_R(G')}{\sigma_R(G')}, \tag{12}$$

where $F_G(G')$ is the frequency of G' in network G , and $\mu_R(G')$ and $\sigma_R(G')$ represent the mean and standard deviation of the frequency in a set of N random networks.

Network Randomization is used to conduct a random network that is compared with DCN. Starting with a certain number of nodes, i.e., 8 nodes, equal to that in the simple network, and zero edge, at each step, one new edge is uniformly chosen from the set of possible edges until the number of edges is equal to 12. During the process of randomization, self-connections and duplicated edges are prohibited.

3. Case study of china air transport system

3.1. Data description

The dataset analyzed in this paper was provided by the CAAC, comprising all flight information in July 2012 in China. July is the peak travel period in China and flight delays are typically severe. The database contained 219,845 domestic scheduled flights connecting 224 airports. The average delay of all flights during the study period was 42 min; July 1 was the best day, with an average delay of 23.1 min, and July 22 was the worst day, with an average delay of 65.1 min. Details of the data set are shown in Table 2.

3.2. Analysis of China air transport system

3.2.1. Basic properties of DCN

To perform a system-level analysis of delay propagation, we build a DCN using the pairwise GC test based on the flight data described in Table 2. We focus on the daily time series and construct DCN_d on d day. There are 224 airports in China air transport system. For one DCN of each day, 50,176 ($224 * 224$) times GC tests are performed. July 22 was the worst day in terms of flight delays. DCN_{22} acts as a typical network to show the delay propagation properties. After removing airports with no connections in DCN_{22} , we find that DCN_{22} only contains 53 nodes and 242 edges (Fig. 3), which means that only about a quarter of airports have delay propagation links with other airports. According to the airport categories definition of the FAA, airports in DCN_{22} consist of 32 large airports, 16 medium airports, and 5 small airports (FAA, 2017). Although small airports represent the majority of the system, and they rarely involved in the spread of delays. Large and medium airports are easily embroiled in delay propagation.

For this specific DCN_{22} , we can answer the following questions about delay propagation by using network analysis tools.

1. How many airports does each airport affect or affected by? For a certain airport i in DCN_{22} , k_i^{in} is the number of airports, each of which partly results in a delay in airport i , while k_i^{out} is the number of airports whose delay is partly caused by airport i . Here, $\langle k^{out} \rangle = \langle k^{in} \rangle = 4.57$, indicating that each airport affects approximately 5 airports and be affected by about 5 airports.

Table 2
Available information for each flight.

| Information for each flight | |
|-----------------------------|---------------------------|
| Flight number | Scheduled arrival airport |
| Flight date | Scheduled arrival time |
| Mission type | Real departure airport |
| Aircraft type | Real departure time |
| Scheduled departure airport | Real arrival airport |
| Scheduled departure time | Real arrival time |

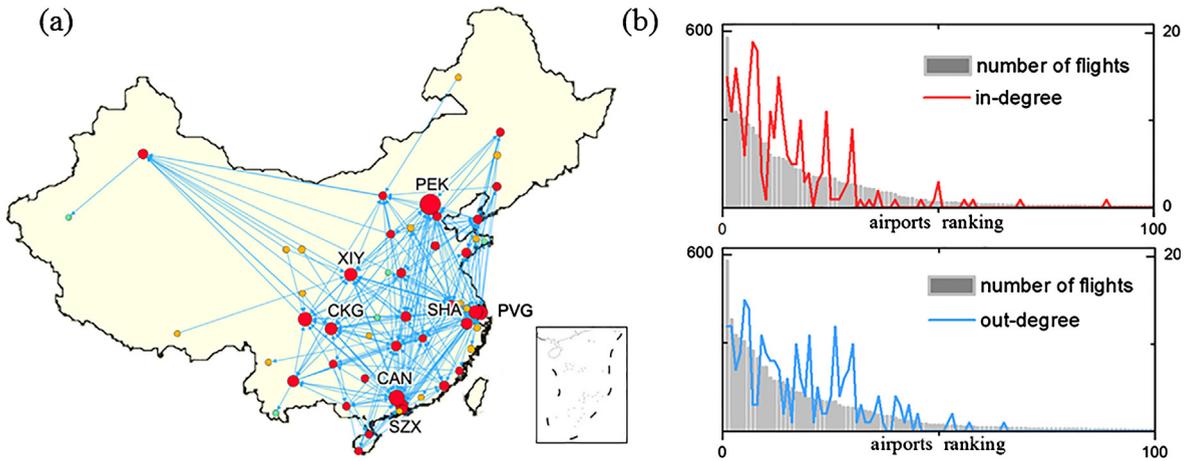


Fig. 3. (a) Features of a DCN. In DCN_{22} , directed edges represent the relationship of delay propagation between airport pairs, and the node size is proportional to its traffic flow. The top 3 airports in the ranking of k_i^{in} are SZX, CKG, and PVG. The top 3 airports in the ranking of k_i^{out} are SHA, XIY, and PEK. The top 3 airports in the ranking of k_i are XIY, PEK, and CAN. (b) k_i^{in} (k_i^{out}) of each airport and its traffic flow on July 22. Airports are ranked by traffic flow.

2. Are the delay propagation links between airport pairs bidirectional? The reciprocity parameter of DCN_{22} is $R = 0.20$. For randomized networks with the same number of nodes and edges — 1000 networks generated by network randomization technology for comparison purposes — the average R' of the randomized networks is only 0.01, which is much less than $R = 0.20$. Thus, the DCN_{22} is more symmetric. One possible reason is that two-way flights between airport pairs result delay propagating in both directions.
3. What is the clustering tendency of airports? The overall clustering coefficient C^D is 0.191, larger than twice that of a randomized network ($C^{D'} = 0.088$), indicating that airports on July 22 had a tendency to cluster.
4. Can the delay propagation among airports be divided into several sub-regions? Community detection algorithm is adopted to analyze the DCN_{22} . Modularity is used to evaluate the strength of division of a network into communities (Newman, 2006), larger value of modularity, more obvious community structure. The modularity value of DCN_{22} is 0.219 and the average modularity value of 1000 randomized networks is 0.213. Thus, no evidence suggests that delay propagation on July 22 could be clearly divided into sub-regions. Moreover, the average modularity value of other 30 DCNs is 0.244, indicating that delay propagation of other days in July is more regional.
5. How serious is the delay propagation? We use the largest connected cluster to represent the seriousness of delay propagation. Members of the connected cluster were selected through a k_i^{out} threshold so that the cluster contains sets of high out-degree airports whose delays affect many other airports. We define k_i^{out} threshold such that it should be larger than average degree of the network ($\langle k^{out} \rangle = 4.57, k^{out} \in Z$) representing the airports that affect more than or equal to five other airports. M_d of DCN_{22} is 26, indicating that 26 are contained in the disaster area of delay propagation.

The interactions among airports are mainly caused by delayed connecting flights. Thus, in theory, more flights at a particular airport should result in a higher probability of interactions. To investigate the relationship between the number of flights and influence times, we consider the k_i^{in} (k_i^{out}) of airport i and its number of flights on July 22 in Fig. 4(b). The results show that airports with more flights tend to have a larger k_i^{in} (k_i^{out}). Due to the randomness of daily operations, the daily flow and degree of airports only has a similar trend. Then, we further consider the correlation over the entire month. The average times is calculated and denoted by $\overline{k_i^{out}}$ ($\overline{k_i^{in}}$). The correlations between $\overline{k_i^{out}}$ ($\overline{k_i^{in}}$) and the average daily traffic flow q_i of each airport is shown in Fig. 4a and b. Both $\overline{k_i^{in}}$ and $\overline{k_i^{out}}$ are positively correlated with q_i , proving that airports with more flights could have a high probability to affect others or be affected. Moreover, Fig. 4a shows a linear relationship on a log-linear graph. Fig. 4b exhibits a linear relationship on a log-log graph. Hence, we further explored the relationship between k_i^{in} and k_i^{out} by defining

$$r = (\overline{k_i^{out}} + 1) : (\overline{k_i^{in}} + 1). \tag{13}$$

Fig. 4c displays the traffic flow of the airports, larger circle higher flow, as well as the r values of various airports, with red color indicting r value less than 1 and blue more than 1. Note that for airport i , $r < 1$ means that outward links of airport i are fewer than its inward links, indicating that airports whose delays affect airport i are more than those whose delays are partly caused by airport i , while $r > 1$ the opposite. From Fig. 4c, we can conclude that airports with larger flows tend to have smaller r , i.e., larger airports are affected by many upstream airports but they impact fewer downstream airports. The conclusion for medium and small airports, however, is exactly the opposite. The same information is presented by the histogram in Fig. 4d, where the airports are ordered according to their flow ranking. Nevertheless, such findings are contrary to expectations. The largest airports are located in developed cities and bear the pressure of heavy traffic. Why are these largest airports able to affect fewer downstream airports and reduce the number of propagation paths of delays? Recently, the CAAC has implemented a no take-off limit regulation for the largest airports.

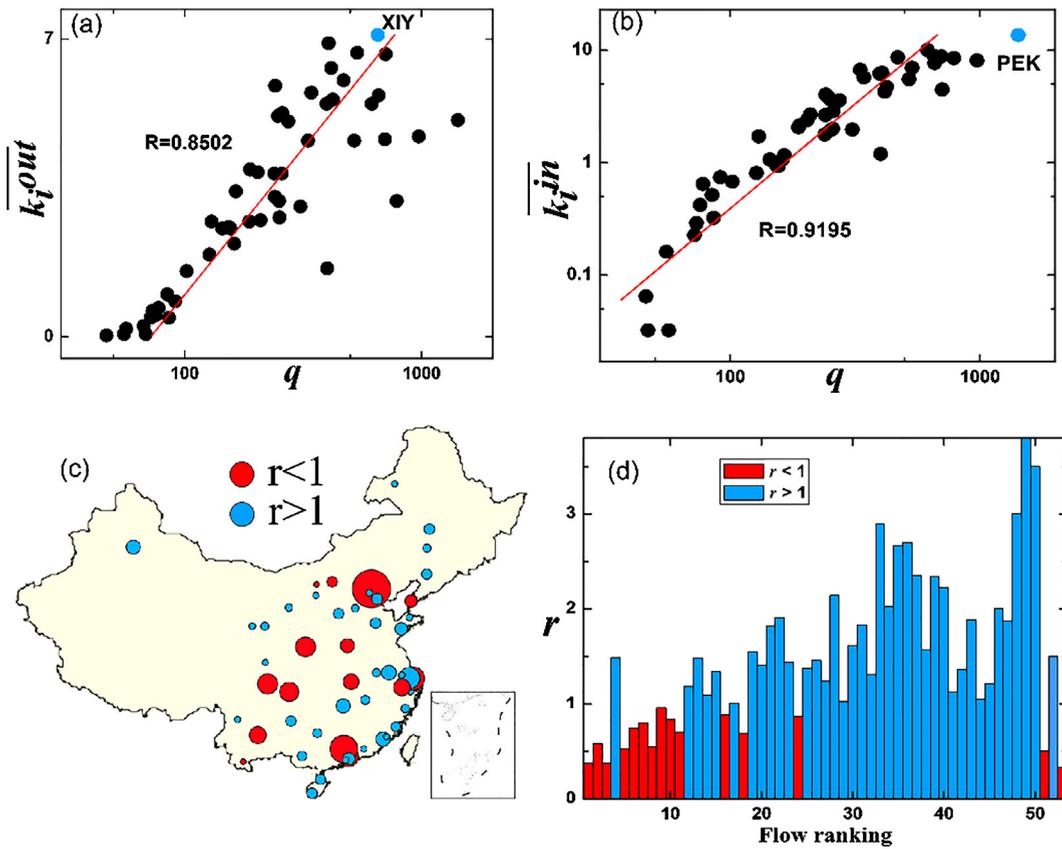


Fig. 4. Relationship between degree and traffic flow. (a) The relationship between \overline{k}_i^{out} and q_i . (b) The relationship between \overline{k}_i^{in} and q_i . (c) Airports with different traffic flows and different values of r . The node size is proportional to the traffic flow. (d) Airports are ranked according to their traffic flows in descending order. The height of histograms represents the value of r .

Flights departing from these airports are not subject to traffic flow management initiatives. This measure has avoided delays in these largest airports from spreading out. Moreover, because the largest airports are already close to saturation, ATC often publishes the traffic flow management initiatives that can postpone upstream flights to later hours, ensuring that the traffic demands of the busiest airports are compatible with their capacities. The result of such traffic flow management initiatives is that upstream are delayed, which brings delays to the largest airports. However, the operations at largest airports are relatively normal and stable. This two-pronged strategy for arrival and departure flights results in significant differences between the largest airports and others.

3.2.2. Temporal properties of DCNs

Another significant problem is the following, what features does the DCN present when the system is seriously delayed? To answer this question, we evaluate the Pearson correlation coefficient, PCC (Pearson, 2006), and P-value (Wasserstein and Lazar, 2017) between daily topologies of the DCN and the average daily delay of all flights D_d during July 2012. Table 3 shows that there exists a rather weak correlation between D_d and topologies except for the size of the largest cluster M_d , which confirms the association between the connected cluster formed by high out-degree airports in the DCN and the level of delay.

We took this point to the next level by analyzing temporal properties in the DCNs. Fig. 5a and b shows the clusters in the DCNs of

Table 3
Correlation analysis.

| Topology | PCC($\rho_{X,Y}$) | P-value |
|--|---------------------|---------|
| Clustering coefficient (C^D) | 0.31 | 0.09 |
| Degree ($\langle k^{in} \rangle$ ($\langle k^{out} \rangle$)) | 0.46 | 0.01 |
| Modularity (Q^D) | -0.29 | 0.11 |
| Reciprocity (R) | 0.19 | 0.31 |
| Size of largest cluster (M_d) | 0.70 | 0.00 |

The first column represents topologies of the DCN, and the second and third columns represent the PCC and P-value, respectively.

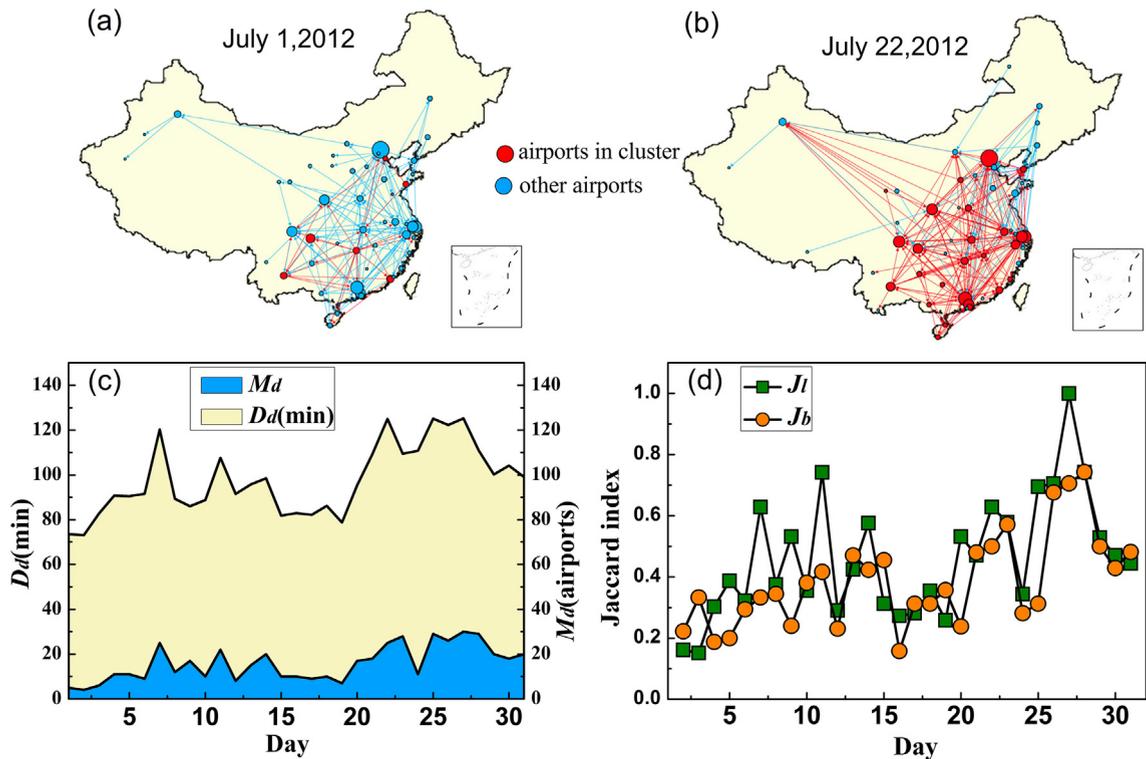


Fig. 5. Connected cluster. (a, b) Connected cluster in the DCNs on July 1 and July 22, 2012, respectively. Nodes in red represent airports belonging to the connected cluster, while blue nodes represent other airports. Directed edges indicate the relationship of delay propagation among airports. (c) Size of largest cluster (M_d) and average delay (D_d) as functions of time. To maintain consistency, we adopt $k_i^{out} \geq 5$, $k_i^{out} \in Z$ as the threshold. (d) Similarity of airports belonging to connected clusters of different days. J_i is the Jaccard index between an airport set of connected cluster on each day and that on July 22 (whose connected cluster has the most airports); J_b is the Jaccard index between the airport set of connected cluster on each day and the prior day. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

July 1 and July 22, respectively. The former is the least congested day, whereas the latter is the most congested, measured by flight delays in China’s aviation system. The largest cluster in DCN changes dramatically. On July 1, only six airports belong to the connected cluster, whereas on July 22, the large cluster covers most airports. Fig. 5c shows M_d and D_d on each day of July, exhibiting the consistency between the two curves. This result implies the delay propagation induces high-level delays, and high-level delays facilitate delay propagation, making the delay more severe. Fig. 5d shows the similarity of the largest clusters on different days with the Jaccard index. Interestingly, the index is relatively low for most days, implying that the airports belonging to the largest cluster change substantially over time. Thus, although the clusters formed by airports affecting several others in DCN are highly correlated with D_d , the culprits of the delay propagation are not always the same set of airports.

3.2.3. Local properties of DCNs

The global features of delay propagation are analyzed by the DCNs’ properties, yet delay propagation cannot be fully understood without analyzing the local relationship pattern. Network motif is a useful tool to fill this gap. It is defined as recurrent relationship patterns among any three airports. $Z(G)$ is used to evaluate the significance level of each motif. As shown in Fig. 6, an acute daily variance occurs for each type of motif over time, which can be attributed to the ever-changing system state. The average traffic flow at each position of the motifs is also shown, and the color of the nodes represents the size of the airport. Furthermore, a notable observation is that the average Z-score of G'_{12} is the highest on most of days, which reveals a situation with three large airports having bidirectional interactions more likely on a given day. In fact, the motif is highly heterogeneous in DCNs, although some local properties of DCNs are revealed. From G'_1 , G'_3 , and G'_9 in Fig. 5, we can see that medium airports appear frequently in the original location of these motifs. For the problem “who is the first piece of delay propagation dominoes?”, this analysis turns our attention to medium airports, although large airports still play the dominant role in delay propagation.

4. Conclusion

In this study, we investigated the mechanism of delay propagation among airports from a new perspective, i.e., building a delay causality network (DCN) based on the relationship of delay time series of each airport and applying network analysis tools to reveal the macroscopic appearance of delay propagation. To demonstrate the method, we built a DCN with the data set of flight delays of

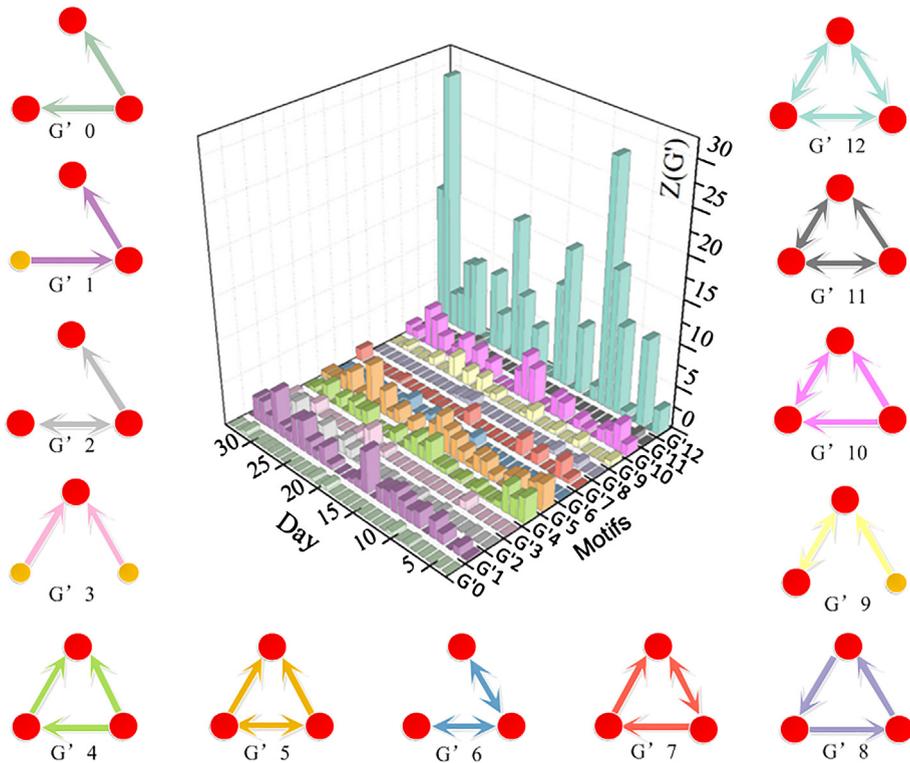


Fig. 6. Motifs of the DCNs. The three-dimensional coordinate system shows $Z(G')$ of each 3 airports' motif for each day during July 2012. Red nodes indicate that the average traffic flow of airports at these positions meets the standard of large airports and yellow nodes correspond to medium airports. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Chinese commercial airports in July 2012. As a typical bad weather day, July 22 was chosen to present the network analysis results of DCNs. We found that on average each airport affected approximately 5 airports and was affected by about 5 airports as well. However, airports of different sizes were not alike. Large airports, i.e. airports with high traffic flow, were affected by more airports than what they impacted in the downstream. Medium and small airports were the opposite. The relationship between in-degree and out-degree of airports proves that some of the largest airports reduced delay propagation paths. Reciprocity parameter showed the bidirectional nature of delay propagation paths between airport pairs of DCN. Community and modularity indicated that delay propagation on July 22 could not be divided into several sub-regions. Nevertheless, we did find clusters in the analysis, which were formed by high-degree airports. We also found that the average daily delay of all flights was highly correlated with the largest connected cluster. Moreover, the airports belonging to the connected cluster varied substantially with time, suggesting that the culprits of delay propagation were not a fixed set of airports.

Understanding the delay propagation mechanism is very important for both air traffic management and aviation planning. For air traffic managers, identifying delay propagation paths could help them weigh the impacts of traffic management initiatives, for example slot controls and schedule optimization, and choose the best combinations to improve the efficiency of aviation system. For aviation planners, by applying the proposed method, they can identify the critical airports in the network in terms of delay propagation. Such information will help them make decisions on network capacity expansion and resources allocation. The counter-intuitive result of large airports impacting fewer downstream airports would attract aviation planners' attention towards medium and small airports regarding delay mitigation, which may have been neglected previously.

The study of using DCN to study flight delay propagation could be extended further. For example, our DCN was an unweighted network, although edge weights could be taken into account, as they reflect the degree of causation. It is also interesting to compare the DCNs of the aviation systems in different countries and investigate the practices of the countries. We may come up with insights for mitigating flight delays from such international comparison.

Acknowledgements

This study was supported by the National Key Research and Development Program of China (Grant No. 2016YFB1200100) and the National Natural Science Foundation of China (Grant Nos. 61425014, 61521091, 91538204, 61671031, and 61722102).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.tre.2018.08.014>.

References

- Ahmadbeygi, S., Cohn, A., Guan, Y., Belobaba, P., 2008. Analysis of the potential for delay propagation in passenger airline networks. *J. Air Transport Manage.* 14 (5), 221–236.
- Arenas, A., Duch, J., Fernandez, A., Gomez, S., 2007. Size reduction of complex networks preserving modularity. *New J. Phys.* 9 (26).
- Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., et al., 2010. Total delay impact study. NEXTOR Final Report, prepared for the U.S. Federal Aviation Administration.
- Barnhart, C., Belobaba, P., Odoni, A., 2009. The global airline industry. *Aviation Saf.* 78 (4), 425–439.
- Beatty, R., Hsu, R., Berry, L., Rome, J., 1999. Preliminary evaluation of flight delay propagation through an airline schedule. *Air Traffic Control Q.* 7 (4).
- Britto, R., Dresner, M., Voltes, A., 2012. The impact of flight delays on passenger demand and societal welfare. *Transp. Res. Part E* 48 (2), 460–469.
- Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., et al., 2000. Graph structure in the web. *Comput. Netw.* 33 (1), 309–320.
- BTS report. Airline on-time performance and causes of flight delays. <https://www.bts.gov/explore-topics-and-geography/topics/airline-time-performance-and-causes-flight-delays>. Accessed July 15, 2017.
- Campanelli, B., Fleurquin, P., Eguiluz, V.M., et al., 2014. Modeling reactionary delays in the European air transport network. In: *Proceedings of the 4th SESAR Innovation Days*.
- Campanelli, B., Fleurquin, P., Arranz, A., Etxebarria, I., Ciruelos, C., Eguiluz, V.M., et al., 2016. Comparing the modeling of delay propagation in the US and European air traffic networks. *J. Air Transport Manage.* 56, 12–18.
- Cardillo, A., Zanin, M., Gómezgardeñes, J., Romance, M., Amo, A.J.G.D., Boccaletti, S., 2013. Modeling the multi-layer nature of the European air transport network: resilience and passengers re-scheduling under random failures. *Eur. Phys. J. Special Topics* 215 (1), 23–33.
- Cook, A., Blom, H., Lillo, F., Mantegna, R., Micciché, S., Rivas, D., et al., 2015. Applying complexity science to air traffic management. *J. Air Transport Manage.* 42, 149–158.
- Czerny, A., 2010. Airport congestion management under uncertainty. *Transp. Res. Part B* 44 (3), 371–380.
- da Rocha, L.E., 2009. Structural evolution of the Brazilian airport network. *J. Stat. Mech: Theory Exp.* 2009, P04020.
- Du, W., Zhou, X., Lordan, O., Wang, Z., Zhao, C., Zhu, Y., 2016. Analysis of the Chinese airline network as multi-layer networks. *Transp. Res. Part E* 89, 108–116.
- FAA. Airport categories. https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/categories/ Accessed November 15, 2017.
- Fagiolo, G., 2007. Clustering in complex directed networks. *Phys. Rev. E* 76 (2), 026107.
- Fleurquin, P., Ramasco, J.J., Eguiluz, V.M., 2013. Systemic delay propagation in the US airport network. *Sci. Rep.* 3 (3), 1159.
- Frank, M., Obradovich, N., Sun, L., Woon, W., LeVeck, B., Rahwan, I., 2018. Detecting reciprocity at a global scale. *Sci. Adv.* 4 (1).
- Granger, C., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37 (3), 424–438.
- Hao, L., Hansen, M., Zhang, Y., Post, J., 2014. New York, New York: Two ways of estimating the delay impact of New York airports. *Transp. Res. Part E* 70 (1), 245–260.
- JEC report. Your flight has been delayed again: Flight delays cost passengers, airlines, and the U.S. economy billions. https://www.jec.senate.gov/public/index.cfm/democrats/2008/5/your-flight-has-been-delayed-again_1539. Accessed July 7, 2017.
- Kafle, N., Zou, B., 2016. Modeling flight delay propagation: A new analytical-econometric approach. *Transp. Res. Part B* 93 (A), 520–542.
- Lan, S., Clarke, J.P., Barnhart, C., 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Sci.* 40 (1), 15–28.
- Li, D., Jiang, Y., Rui, K., Havlin, S., 2014. Spatial correlation analysis of cascading failures: Congestions and blackouts. *Sci. Rep.* 4 (4), 5381.
- Malone, K.M., 1995. *Dynamic Queueing Systems: Behavior and Approximations for Individual Queues and for Networks*. Massachusetts Institute of Technology.
- Martin, R., Christina, G., Rafael, C., Fabio, C., White, A., Alessandro, F., 2016. Experimental test of nonlocal causality. *Sci. Adv.* 2 (8).
- Meng, L., Zhou, X., 2011. Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon solution approach. *Transp. Res. Part B* 45 (7), 1080–1102.
- Nayak, N., Zhang, Y., 2011. Estimation and comparison of impact of single airport delay on national airspace system with multivariate simultaneous models. *Transp. Res. Rec.* 2206 (1), 52–60.
- Newman, M., 2006. Modularity and community structure in networks. *PNAS* 103 (23), 8577–8582.
- Pearson, K., 2006. Note on regression and inheritance in the case of two parents. *Proc. Royal Soc. London* 58, 240–242.
- Pyrgiotis, N., Malone, K.M., Odoni, A., 2013. Modelling delay propagation within an airport network. *Transp. Res. Part C* 27 (2), 60–75.
- Ren, P., Li, L., 2018. Characterizing air traffic networks via large-scale aircraft tracking data: a comparison between China and the US networks. *J. Air Transport Manage.* 67, 181–196.
- Shenorr, S., Milo, R., Mangan, S., Alon, U., 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* 31 (1), 64–68.
- Song, C., Chen, C., Zhu, Z., 2008. Economic growth and energy consumption revisited: Evidence from linear and nonlinear Granger causality. *Energy Econ.* 30 (6), 3063–3076.
- Stokes, P., Purdon, P., 2017. A study of problems encountered in Granger causality analysis from a neuroscience perspective. *PNAS* 114 (34), E7063.
- Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., et al., 2012. Detecting causality in complex ecosystems. *Science* 338 (6106), 496–500.
- Wahl, B., Feudel, U., Hlinka, J., Wächter, M., Peinke, J., Freund, J.A., 2017. Conditional Granger causality of diffusion processes. *Eur. Phys. J. B* 90 (10), 197.
- Wang, Y., Bu, J., Han, K., Sun, R., Hu, M., Zhu, C., 2016. A novel network approach to study communication activities of air traffic controllers. *Transp. Res. Part C* 68, 369–388.
- Wang, Y., Cao, Y., Zhu, C., Wu, F., Hu, M., Barzel, B., 2017. Characterizing departure delays of flights in passenger aviation network of United States. arXiv:1701.05556.
- Wang, J., Mo, H., Wang, F., 2014. Evolution of air transport network of China 1930–2012. *J. Transp. Geogr.* 40 (40), 145–158.
- Wasserstein, R., Lazar, N., 2017. The ASA's statement on p-values: Context, process, and purpose. *Am. Stat.* 70 (2), 129–133.
- Watts, D., Strogatz, S., 1998. Collective dynamics of “small-world” networks. *Nature* 393 (6684), 440.
- Xiong, J., Hansen, M., 2009. Value of flight cancellation and cancellation decision modeling ground delay program postoperation study. *Transp. Res. Rec.* 2106 (2106), 83–89.
- Zhang, Y., Nayak, N., 2010. A macroscopic tool for measuring delay performance in the national airspace system. In: *Presented at the 89th Annual Meeting of the Transportation Research Board*.
- Zou, B., Hansen, M., 2014. Flight delay impact on airfare and flight frequency: A comprehensive assessment. *Transp. Res. Part E* 69 (3), 54–74.