

Supporting Information: Revealing the structure of the world airline network

Trivik Verma,^{1,*} Nuno A. M. Araújo,^{2,†} and H. J. Herrmann^{1,3,‡}

¹ETH Zürich, Computational Physics for Engineering Materials,
Institute for Building Materials, Wolfgang-Pauli-Strasse 27, HIT, CH-8093 Zürich, Switzerland

²Departamento de Física, Faculdade de Ciências, Universidade de Lisboa,

P-1749-016 Lisboa, Portugal, and Centro de Física Teórica e Computacional,
Universidade de Lisboa, Avenida Professor Gama Pinto 2, P-1649-003 Lisboa, Portugal

³Departamento de Física, Universidade Federal do Ceará,
Campus do Pici, 60455-760 Fortaleza, Ceará, (Brazil)

S1. FLIGHT MODEL

Figure S1 visually depicts how the world fragments into different parts as the peripheral connections with the least number of flights are removed. Frame 2 of this figure illustrates that 45% of airports fall out of the largest connected component upon removing connections with merely one flight ($\approx 60\%$ connections). In frame three of the figure, it is clear how only the global hubs are connected when connections carrying up to *eight* flights ($\approx 96\%$ connections) have stopped plying. These flights are connecting small fishing industries in remote islands or making chartered connections to sectors outside of economic and political hubs. These also include connections to scarcely populated areas in Siberia, Alaska, Northern Canada, Papua New Guinea, the Sahara Desert, etc.

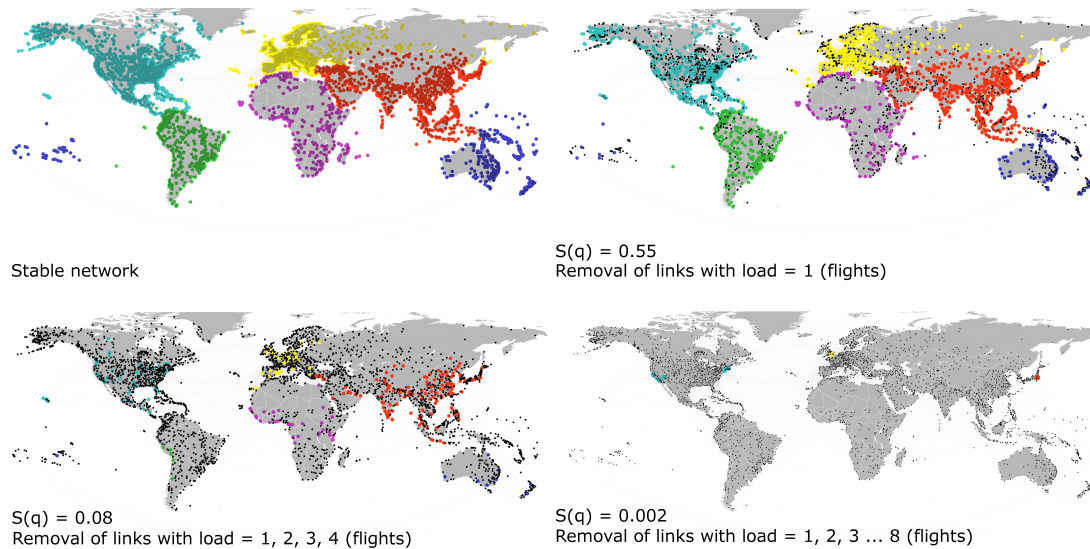


FIG. S1. Subsequent stages of WAN under a *rare* removal of links as ranked by number of *flights*. The colors represent different continents but also only the largest connected component of the network. Black nodes are airports that have been disconnected from the largest connected component but still may be functioning (not completely isolated) within their own small clusters. All maps are produced using *Gephi* [1].

S2. RATIONALE BEHIND DEGREE OF CONNECTIVITY

Figure S2(a) shows node b has zero clustering. Link $b - a$ is extremely important as it is the only connection between b and a . Whereas, in fig. S2(b), a can be reached from b in many ways and link $b - a$ is redundant in a network sense.

* trivik@ethz.ch

† nuno@ethz.ch

‡ hjherrmann@ethz.ch

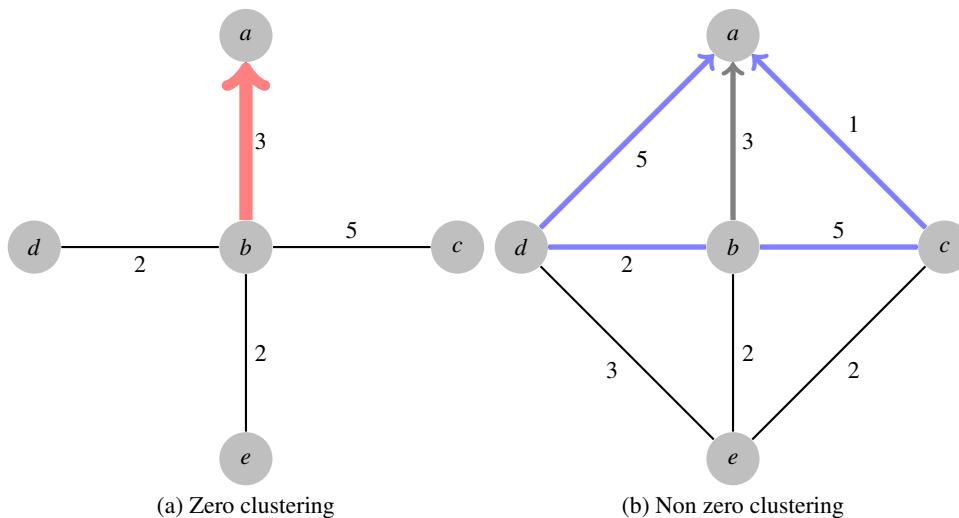


FIG. S2. Illustration of the rationale behind the *degree of connectivity*. Each link is characterized by the number of flights on that connection. In (a), the clustering coefficient of b is zero and hence $b - a$ is a very important link as there is no other way to reach a via b . However, in (b), the clustering coefficient of b is greater than zero, i.e., nodes in the neighborhood of b are also connected to each other. This implies that the higher *degree of connectivity* gives a lower "importance" to the link and also clarifies our rationale behind unconventional removal scenarios.

S3. t-CORE DECOMPOSITION

Figure S3 shows the core of the World Airline Network. It comprises 73 airports that are extremely well-connected and provide a strong foundation for air travel.

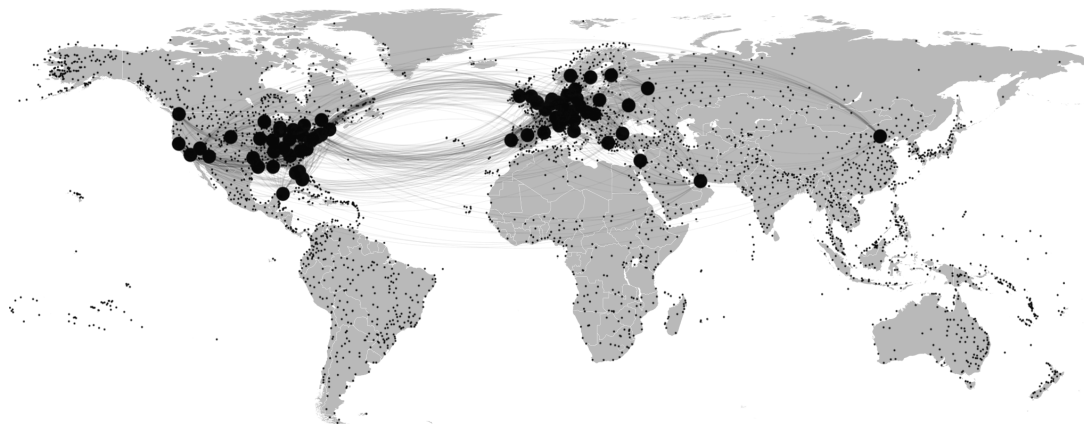


FIG. S3. Seventy three airports that belong to the core of the network. All these airports fall in the t -core with $t = 387$. All maps are produced using *Gephi* [1].

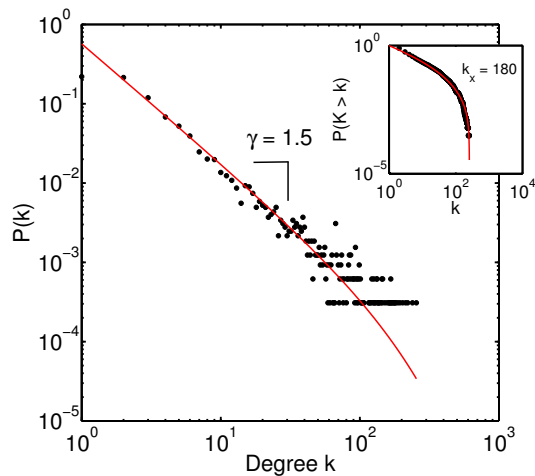


FIG. S4. Degree distribution of WAN plotted in a log-log scale. The degree is a measure of the connections of an airport and follows a scale-free behavior with an exponential cut-off, $P(k) \sim k^{-\gamma} \exp(-k/k_x)$ with an exponent $\gamma = 1.5 \pm 0.1$ and truncates at $k_x = 180 \pm 5$. The **inset** shows the cumulative complementary distribution for the network in a log-log scale with the same exponents as above. There is an exponential decay for the number of airports with large degree.

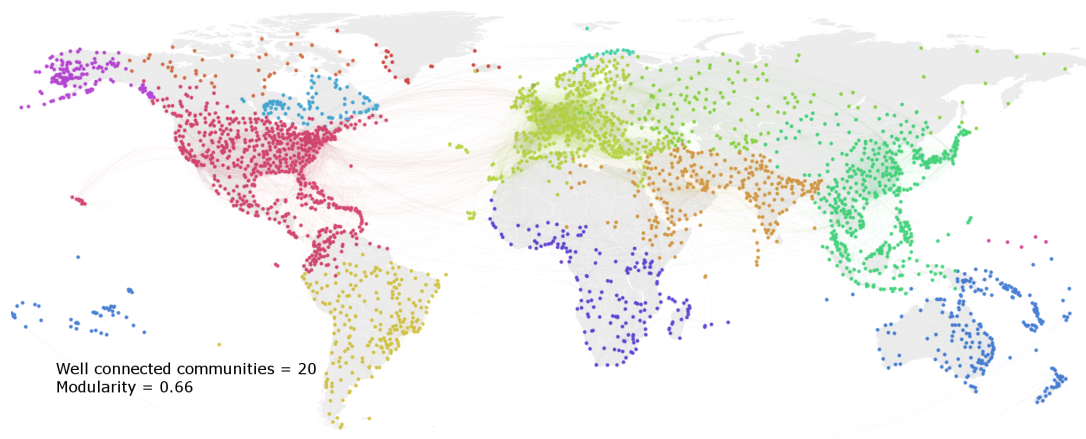


FIG. S5. Closely held communities existing in WAN based on the number of alternate flights between each pair of airports. Nodes are locations of the airports and colors represent different communities. A high value of modularity ($M \rightarrow 1$) indicates presence of well-formed communities. All maps are produced using *Gephi* [1].

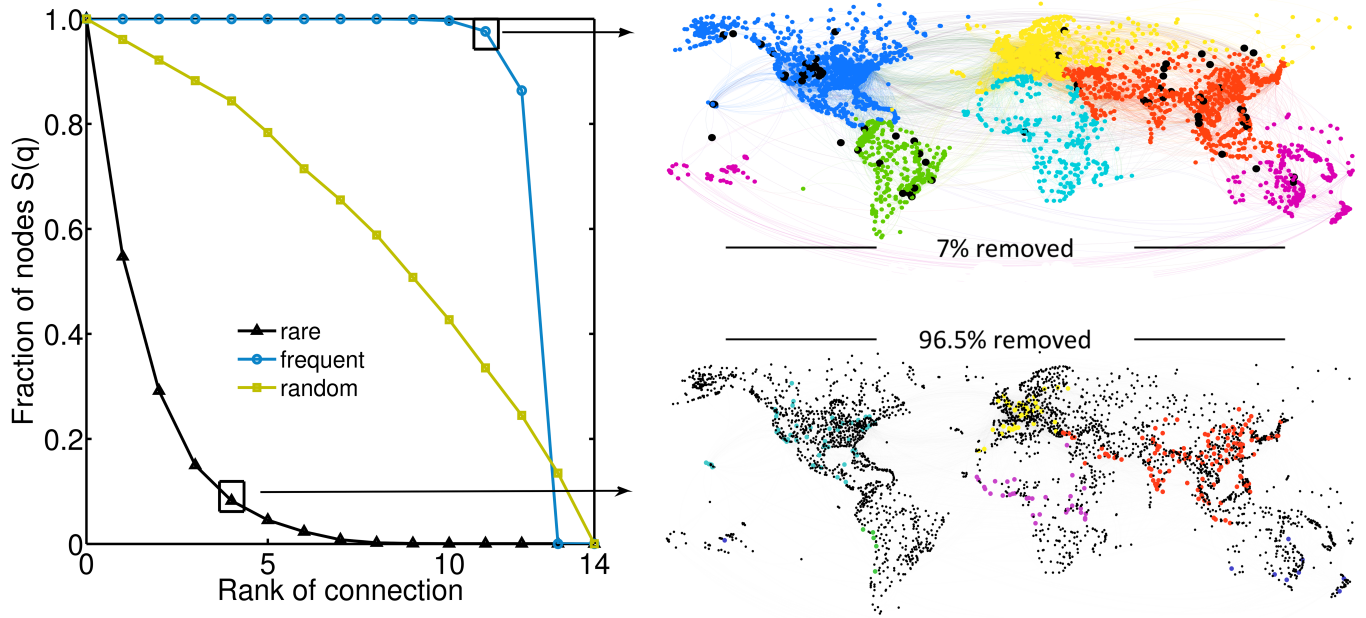


FIG. S6. Drop in the size of the largest connected cluster of WAN against removal of links. Connections are ranked according to their weight given by the number of alternative flights. In *frequent* removal, each subsequent step corresponds to removal of all connections with the highest weight. In *rare* removal, each subsequent step corresponds to removal of all connections with the lowest weight. The random removal strategy is an average over 500 statistically independent simulations and each step removes all connections with a random probability. After removing connections that offer at least four alternatives (7.5% connections), a large part of the network is still connected, shown in the top-right corner. The bottom right map shows that after removing connections that offer at most four alternatives (96.5% connections), the world disintegrates completely, revealing the vulnerable nature of the periphery of the network. The black nodes are not part of the largest connected cluster. The remaining colors represent different continents and show the nodes that are part of the largest connected cluster. All maps are produced using Gephi [1].

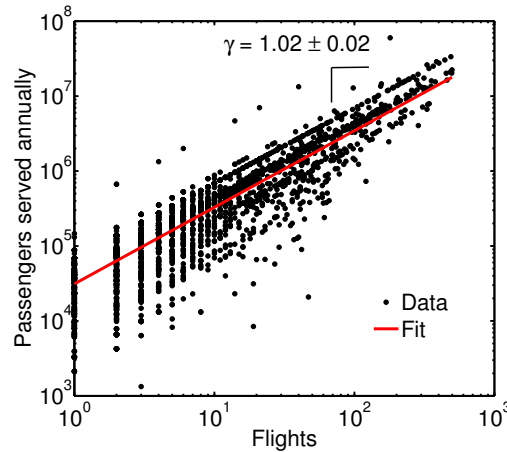


FIG. S7. Scatter plot of the number of passengers served at each airport annually versus the number of available alternative connections at each airport. The plot is in a log-log scale. Several airports with the same number of connections has varied passenger count. We describe this behavior by adding the red line to the curve with the following scale-free exponent, $\gamma = 1.02 \pm 0.02$.

-
- [1] Bastian, M., Heymann, S., and Jacomy, M. Gephi: An open source software for exploring and manipulating networks. *ICWSM* (2009).