Critical Cooperation Range to Improve Spatial Network Robustness

V. H. P. Louzada,1,* N. A. M. Araújo,1 T. Verma,1 F. Daolio,2 H. J. Herrmann,1,3 and M. Tomassini2

1Computational Physics, IfB, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland
2Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland
3Departamento de Física, Universidade Federal do Ceará, 60451-970 Fortaleza, Ceará, Brazil

(Dated: April 11, 2014)

A robust worldwide air-transportation network (WAN) is one that minimizes the number of stranded passengers under a sequence of airport closures. Here we address how WAN’s robustness can profit from local cooperation between airports. For this, we reroute a series of flights among airports within a certain distance, a cooperation range, and describe the improvement of WAN’s robustness with distance as a continuum transition. We calculate the critical exponents and identify a critical cooperation range below which improvement is negligible. We propose a network model that falls into the same universality class in which the probability to connect two nodes decays algebraically with their distance. Practical implications of this result are also discussed.

INTRODUCTION

The construction of a new terminal at the Schenzen airport, Southeast China, has been used to question the current strategies of infrastructure growth in developing countries [1, 2]. Schenzen is a large city, but its airport is directly connected by an eight kilometer ferry to the Hong Kong international airport, which can handle twice as many passengers. Is it reasonable to invest more than one billion dollars increasing the capacity of such a large infrastructure with another one nearby? Opponents to this investment classify it as a white elephant and as one example of the misbegotten infrastructure growth in developing countries. The ones in favor, argue that not only the costs of investing in Schenzen are lower than in Hong Kong but also global connectivity can profit from cooperation between stakeholders of air-transportation systems in the region. Here we address, from a network science perspective, how such proximity and cooperation between airports is key to the robustness of the worldwide air-transportation network (WAN).

The simplified representation of complex systems as a network of nodes and links has provided important insights into the design of a variety of systems, such as power grids [3, 4], maritime commerce [5], and communication networks such as the Internet [6–8]. We describe the WAN as a complex network where nodes are airports and links are flights connecting them. Nodes are spatially embedded according to the geographical coordinates of the airports. This simplification allows us to focus on the topological aspects of the system and to easily extend our results to other applications.

As an ubiquitous infrastructure system, airports are used daily to transport people and freight throughout the planet [9]. It is paramount that this system works in an extremely reliable and efficient fashion, as any temporary airspace closure, such as the one caused by the eruption of the volcano Eyjafjallajökull in 2010, may cause huge losses worldwide [10]. With that in mind, we characterize the robustness of the WAN as its capacity to maintain global connectivity under a sequence of node removals and describe a strategy based on local cooperation to improve robustness under realistic constraints. We discover a continuum transition when changing the distance for which airports are allowed to reroute flights, a cooperation range. We calculate the critical exponents of this transition and show that the same set of exponents can be obtained for an artificial network if the connection probability decays algebraically with the distance between nodes.

MODEL

Complex Networks have been used to study the WAN and its local subdivisions [11–18]. To focus on its structural aspects, we summarize data provided by Open-
FIG. 2. Robustness increase of the WAN. Dependence of the optimized robustness on the cooperation range (blue) and variance over samples (red) with a maximum at 910 ± 90 km.

Flights in the year 2011 as a single static network with 3237 airports (nodes) and around eighteen thousand connections (links) [19]. A summary of the network characteristics is included in Section I of the Supplementary Material. Links are undirectional, assuming that each flight should return to its origin, and weighted according to the number of alternative flights between two airports. Airports are weighted by the number of passengers transported. An analysis of its static properties, such as presence of communities and core-periphery relationships, reveals that the WAN has a peculiar hierarchical structure wherein the core is small and has many redundant connections and the periphery is vast and weakly connected [20].

As a self-organized system, in which preferential attachment is expected to play a pivotal role, the WAN is quite fragile to targeted attacks, i.e., intentional removal of the most connected nodes causing the collapse of the giant connected component [20]. The aim of our optimization strategy is to create a robust yet economically feasible WAN. Since robustness can be defined in different ways, we consider that a robust WAN should be capable of transporting passengers even in face of a targeted attack to its most important hubs. Previous works have proposed topological features to characterize robustness where only the size of the largest connected component is considered [21–26]. We simulate a sequence of airport closures (node removal) and quantify robustness $r$ as:

$$r = \frac{1}{\Pi(0)} \sum_{n=1}^{N} \Pi \left( \frac{n}{N} \right),$$

where $N$ is the total number of nodes, $n$ is the number of nodes removed from the network, and $\Pi(q)$ is the number of passengers in the largest component after a fraction $q = n/N$ of nodes were removed. Closures are executed from the most to the least connected node.

The location of airports are mostly determined by economical forces, such as to cater to local demands or increase a region’s business attractiveness. In many cases, airports are located within a short distance from each other, sometimes only a few kilometers away as, e.g., airports in the Schenzen-Hong Kong area, or a few hundred kilometers but still easily reachable, such as the airports in the northeast of the United States. We assume that a flight rerouted to an airport within a cooperation range $v$ of the original destination has a similar attractiveness. If need be, a passenger landing at a different airport could easily take another means of transportation, such as the local train network or a shuttle bus, to go to the desired destination. In this work cooperation range is defined as the geographical distance between airports (details in Section Methods), but it is important to note that it could be easily extended to other metrics, such as travel time or operational cost.

In order to increase network robustness, our procedure is based on flight rerouting. Instead of adding or deleting connections, rerouting does not change the transportation capacity of the system: no airport would have to be expanded or mothballed as the number of flights does not change. Moreover, a flight is only rerouted to an airport within the cooperation range of the original destination. As an example, flights could be distributed between Hong Kong and Schenzen or among the airports surrounding the city of London. Lastly, the probability to swap a route is inversely proportional to the number of alternative flights, the weight of an undirected link $e_{ij}$ between airports $i$ and $j$, so that important connections are affected with less priority.

FIG. 3. Critical cooperation range as a function of the nodes’ coverage. The critical cooperation $v^*$ is positively correlated with $\sqrt{A/N}$, where $A$ is the total area in which $N$ nodes are embedded. The three different symbols represent the types of infrastructure networks in which the geo swap is applied.
calculates the distance between nodes $i$ and $j$, the following swap strategy is performed:

1. Select a node $i$ randomly having at least one neighbor;
2. Select a neighbor $j$ of node $i$ with probability inversely proportional to the weight of $e_{ij}$;
3. Select a pair of connected nodes $k$ and $l$ so that $d(i, k) \leq v$ and $d(j, l) \leq v$ with probability inversely proportional to the weight of $e_{kl}$;
4. Remove links $e_{ij}$ and $e_{kl}$ and create links $e_{il}$ and $e_{jk}$.

An illustration of this strategy is provided in Fig. 1. Swaps change the network robustness $r$ but we only perform swaps that increase $r$. From this point on, we call such a swap, a geo swap. A fixed number of geo swaps of the order of the number of links is executed. To compare networks of different sizes and populations, we normalize the robustness as $R = (r - r_{\min})/(r_{\max} - r_{\min})$, in which $r_{\min}$ is the value of $r$ for $v = 50$ km, the minimum value for which a geo swap will be considered in the WAN, and the maximum robustness $r_{\max}$ obtained for $v = 18 \times 10^3$ km, which is approximately half of the planet perimeter. This strategy builds on top of previous work which focused on different acceptance mechanisms [22, 25] or topological characteristics [26], but differs significantly by focusing on geographic limitations and low-weight links.

**RESULTS**

The cooperation range $v$ limits the search area of possible swaps to guarantee that a geographically acceptable change is performed. A too small value of $v$ does not provide sufficient room for robustness improvement. While a too large value of $v$ leads to reroutes of connections to impractically far away airports. In the limit of $v$ being comparable to the perimeter of the planet one recovers the strategy proposed by Schneider et al. [21] where links are swapped randomly without any geographic limitation. By tuning the values of $v$, we observe a critical value of the cooperation range $v^* = 910 \pm 90$ km at which a significant improvement in the WAN is first registered. This range in fact yields the highest variance of robustness increase among all possibilities, as shown in Fig. 2. We divide the WAN into continents by considering only flights in which both end-points are in the same continent (details in Section I of Supplementary Material) and, together with other spatially embedded networks, observe that $v^*$ is positively correlated with $\sqrt{A/N}$ (Spearman’s correlation coefficient $\rho = 0.78$), where $A$ is the total area in which the $N$ nodes are embedded (Fig. 3). Artificially generated random networks also confirm this relationship (Section II of Supplementary Material). Being the combination of local and intercontinental flights, the WAN lies slightly off the trend, but in general we can conclude that the typical radius served by an airport is correlated to the minimum distance at which swaps become effective. Additionally, other topological characteristics of the optimized networks are detailed in Section V of Supplementary Material.

Close to the critical cooperation range, the evolution of $R$ scales with $v - v^*$ for the continental networks. Applying the finite-size scaling,

$$R = N^{-\beta/\nu} F\left((v - v^*)N^{\nu/\beta}\right),$$

(2)

where $\nu$ and $\beta$ are critical exponents and $F[x]$ is a scaling function, we collapse the data for different system sizes. For $v = v^*$, as shown in the inset of Fig. 4, $R$ scales with $N^{-\beta/\nu}$, with $\beta/\nu = 0.08 \pm 0.01$, as expected for a continuous transition. This allows us to calculate the exponents in the main panel of Fig. 4 as $\beta = 0.20 \pm 0.02$ and $1/\nu = 0.40 \pm 0.03$. The data suggest that the construction of airports and the creation of flight connections follow a similar mechanism in all continents, though the limited system size of each continent prevents us from making stronger conclusions. However, the evolution of $R$ for Australia significantly differs from the others. Because a great number of islands in Oceania have many small airports, sometimes being the only feasible connection between remote areas, we assume that airports and flights in this continent were established following a different mechanism from the remaining WAN.

In the WAN, three main properties define its charac-
teristics: nodes’ position, degree distribution, and connection pattern. A simplified model – where nodes are assigned random positions, the degree distribution is a Poisson distribution, and connections are randomly assigned without any spatial/degree bias – displays different critical exponents (Section IV of Supplementary Material). However, if the probability \( P(i,j) \) that nodes \( i \) and \( j \) are connected decays algebraically with the distance between \( i \) and \( j \),

\[
P(i,j) \propto \frac{1}{d(i,j)\alpha},
\]

where \( \alpha \in \mathbb{R} \) is the decay exponent, we obtain with an adequate \( \alpha \) exponents that are numerically consistent with the ones in Fig. 4 for the real data. Based on a simplification of the gravity model, commonly used to describe connections between geographically distributed nodes \([5, 27, 28]\), we call Eq. (3) a distance-decay model as it does not take into account the degree/weight of the nodes to calculate \( P(i,j) \). We observe that changes in the value of \( \alpha \) affect consistently the slope in the data-collapse (Panel b of Fig. 5). We find a value of \( \beta \) similar to the one of the continents, without Australia, for \( \alpha \in [1.8, 2.0] \). For \( \alpha = 2.0 \), the finite-size scaling in Fig. 5c allows us to estimate: \( \beta = 0.23 \pm 0.02 \) and \( \beta/\nu = 0.08 \pm 0.01 \) (Panels a and c of Fig. 5). Further analysis also show that when \( \alpha \approx 2 \) the ratio between the average distance traveled by flights (link length) and the average geographical distance between two airports is similar to that found for the continents (Section III of Supplementary Material). Interestingly, the empirical probability distribution of link lengths in the WAN is a power law of exponent \( \alpha = 2.2 \pm 0.2 \) (Section VI of Supplementary Material). This suggests that correlations as the ones developed in the distance-decay model are consistent with the ones found for the WAN and characteristic for the universality class.

**DISCUSSION**

In order to provide applicable insights, any network modification strategy should take into account realistic constraints naturally imposed by the problem. With that in mind, the geo swap contains a simple set of rules specifically designed to improve the robustness of infrastructures, in particular of the WAN. It is important to note however that a probabilistic approach, in which nodes and links are chosen randomly, is more a guidance than a closed optimization recipe. We expect that future procedures built on top of our strategy, but carefully tailored to the underlying problem, present the same basic properties of the simplified model explored here.

In an infrastructure perspective, the swap of links, which considers that flights are able to land in different airports, has two crucial implications. Firstly, a second transportation system should be used to connect nearby airports, in line with recent works dealing with the coupling of infrastructure networks \([29]\). By taking into account the critical cooperation range, different for each multilevel network, these couplings could be designed or improved for distances close to \( v^* \). Secondly, a local level of cooperation is necessary between airports. Flights in Schenzen and Hong Kong could be rerouted to attend different yet nearby airports worldwide, further increasing the reliability of the local service and the overall WAN robustness.

Our rewiring strategy is also able to show that the continents, with one exception, follow the same universality class regarding possible robustness improvements, mostly because the probability that two airports are connected decays quadratically with their distance. Being a continent with its own geographical idiosyncrasies, Australia does not fit our analysis. It is not clear why this is the case, but studies from other contexts suggest that economic forces are the major drivers of new connec-
tions [30]. Therefore, while Australia might have flights as the only practical way to link its sparse territory, the remaining WAN follows other market constraints. This similarity might be explored in the future to create unified robustness optimization procedures that work in different regions of the planet.

**METHODS**

Nodes are modeled as points distributed across the surface of a sphere with distance calculated accordingly (Haversine formula). Figure 2 results from the application of $10^4$ tentative geo swaps for each value of $v$ in the WAN. Each blue circle is the average over 250 samples while red circles stand for the variance over samples.

In Fig. 3, data for the European Power Grid (red) is retrieved from Ref. [31] and the Rail transportation network was manually assembled using public data. The power grid network has 1254 nodes and 1812 links, and the rail network has 39 nodes and 70 links. For continents, the power grid, and the rail network, a total of $10^4$ tentative geo swaps are executed for several cooperation range values. The value of $v^*$ is selected as the highest variance point over 100 samples, with error bars representing the values where variance is equal to $0.75\sigma^2(v^*)$. The same data for the continents is used to construct Fig. 4, in which symbols are larger than the standard deviation.

Networks presented in Fig. 5 are generated as follows. Node positions are uniformly distributed across an Earth spherical cap of area $A$ and in order to keep $\sqrt{A/N} \approx 195$, the same value of the WAN, the network system size is calculated accordingly. Node degree follows a Poisson distribution of average degree 12. Node weights are chosen according to the equation $W(i) = 102.6k_i^{1.1}$, where $k_i$ is the degree of node $i$, which is a fit of the relationship between node weights and degree of the WAN. Link weights are randomly distributed from $[1, 14]$, in which 14 is the maximum link weight on the WAN. For all panels in Fig. 5, a maximum of $10^4$ tentative geo swaps are executed for several cooperation ranges. Each point represents the average over 100 samples, with symbols being larger than standard deviation in Panel c. The critical cooperation range is defined as $v^* = 240 \pm 10$ km.

**Acknowledgments.** Authors would like to thank the CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil, the European Research Council advanced grant FP7-319968-flowCSS, and the ETH Zurich Risk Center for financial support. We would also like to thank I. Nikolakopoulos and K. J. Schrenk for providing the caffeine for the discussion in which the idea of this paper was conceived.

* Correspondence and requests for materials should be addressed to V. H. P. Louzada (louzada@ethz.ch)


Supplementary Material

Critical Cooperation Range to Improve Spatial Network Robustness

V. H. P. Louzada, 1, ∗ N. A. M. Araújo, 1 T. Verma, 1 F. Daolio, 2 H. J. Herrmann, 1, 3 and M. Tomassini 2

1 Computational Physics, IfB, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland
2 Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland
3 Departamento de Física, Universidade Federal do Ceará, 60451-970 Fortaleza, Ceará, Brazil

(Dated: April 9, 2014)

I - THE WORLD AIR-TRANSPORTATION NETWORK

The network that represents the WAN was carefully described in Ref. [1]. Information about continents, contained in the original data set, was used to further break it down. Continental networks contain flights in which both end-points are in the same continent. Australia includes all islands in the Pacific ocean. For simplification, Russia is entirely part of Europe, and Turkey is entirely in Asia. A summary of the main characteristics of the continents and the WAN is present in the table below.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nodes</th>
<th>Links</th>
<th>Passengers (daily)</th>
<th>Flights (daily)</th>
<th>Average Degree</th>
<th>Average Passengers (daily)</th>
<th>Average Flights (daily)</th>
<th>Average Distance btw Airports (km)</th>
<th>Average Flight Distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>269</td>
<td>642</td>
<td>220,481</td>
<td>3,023</td>
<td>4.77 ± 0.43</td>
<td>819.63 ± 129.59</td>
<td>2.35 ± 0.10</td>
<td>3,759.92 ± 10.52</td>
<td>1,139.23 ± 44.35</td>
</tr>
<tr>
<td>Asia</td>
<td>773</td>
<td>3,911</td>
<td>2,409,160</td>
<td>23,406</td>
<td>10.12 ± 0.64</td>
<td>3,116.64 ± 281.14</td>
<td>2.99 ± 0.07</td>
<td>4,089.57 ± 4.16</td>
<td>1,338.13 ± 19.76</td>
</tr>
<tr>
<td>Australia</td>
<td>288</td>
<td>567</td>
<td>178,447</td>
<td>2,499</td>
<td>3.94 ± 0.39</td>
<td>619.61 ± 132.87</td>
<td>2.20 ± 0.11</td>
<td>3,450.12 ± 10.14</td>
<td>810.18 ± 39.03</td>
</tr>
<tr>
<td>Europe</td>
<td>602</td>
<td>5,188</td>
<td>1,907,980</td>
<td>25,587</td>
<td>17.24 ± 1.07</td>
<td>3,169.40 ± 391.04</td>
<td>2.47 ± 0.07</td>
<td>2,410.24 ± 4.05</td>
<td>1,250.84 ± 12.11</td>
</tr>
<tr>
<td>North America</td>
<td>1,006</td>
<td>4,289</td>
<td>2,344,470</td>
<td>20,593</td>
<td>8.53 ± 0.62</td>
<td>2,330.48 ± 245.49</td>
<td>2.40 ± 0.08</td>
<td>3,386.44 ± 2.66</td>
<td>1,150.44 ± 15.67</td>
</tr>
<tr>
<td>South America</td>
<td>299</td>
<td>762</td>
<td>380,348</td>
<td>3,984</td>
<td>5.10 ± 0.44</td>
<td>1,272.07 ± 154.33</td>
<td>2.61 ± 0.10</td>
<td>2,719.93 ± 6.84</td>
<td>793.98 ± 27.81</td>
</tr>
<tr>
<td>World</td>
<td>3,237</td>
<td>18,125</td>
<td>7,440,880</td>
<td>94,644</td>
<td>11.20 ± 0.42</td>
<td>2,298.70 ± 127.80</td>
<td>2.61 ± 0.04</td>
<td>8,678.58 ± 1.92</td>
<td>1,734.64 ± 14.58</td>
</tr>
</tbody>
</table>

II - CRITICAL COOPERATION RANGE IN RANDOM NETWORKS

![FIG. S1. The critical cooperation $v^*$ is correlated with $\sqrt{A/N}$, where $A$ is the total area in which $N$ nodes are embedded. Data is based on artificially generated random networks, similar to the ones used for the distance-decay model in Fig. 5 of the main text, but with links randomly assigned without any bias. Each point represent the average over 100 randomly generated networks of 500 nodes. A total of $10^4$ tentative geo swaps are executed for several cooperation range values. The value of $v^*$ is selected as the point with highest variance, with error bars representing the values where the variance is equal to $0.75\sigma^2(v^*)$.](image-url)
III - DISTANCE RATIO IN THE MODEL

FIG. S2. Distance-decay model reproduces the same $\beta$ for different ratios between the average distance traveled by the flights (link length) and the average geographical distance between two airports of the WAN. Plot shows the impact on $\beta$ for the distance-decay model with different values of $\alpha$ (blue) in comparison with data for the continents (green). The data used is the same from Fig. 4 (continents) and Fig. 5b (distance-decay model) in the main text.

IV - FINITE SIZE SCALING OF RANDOM NETWORKS

FIG. S3. Data collapse of the robustness evolution for $v - v^* > 0$ after successive applications of the geo swap in random networks. Curves in the main panel represent each continent scaled with $1/\nu = 0.28$ and $\beta/\nu = 0.20$. The inset shows the size dependence of $R$ at $v = v^*$, scaling as $R \sim N^{-\beta/\nu}$, with $\beta/\nu = -0.20 \pm 0.01$, where $N$ is the total number of nodes. Data is based on artificially generated random networks, similar to the ones used for the distance-decay model in Fig. 5 of the main text, but with links randomly assigned without any bias. A total of $10^4$ tentative geo swaps are executed for several cooperation range values. Each point represents the average over 200 samples, with symbols being larger than standard deviation in the main panel. The critical cooperation range is defined as $v^* = 240 \pm 10$. 
FIG. S4. Changes on networks characteristics after successive geo swaps with $v = v^*$. The strategy makes the networks more assortative, more onion-like, but also more random, as clustering coefficient and modularity decrease. Plots show the change of several features of the airport networks after $10^4$ tentative geo swaps. a, Degree assortativity ($a_{deg}$) [2]. b, Neighbors’ degree correlation ($k_{nn}$) [3]. c, Weighted clustering coefficient ($CC^w$) [4], weighted by the number of passengers per airport. d, Clustering coefficient ($CC$). e, Onion-likeness ($Sk$) [5]. f, Modularity ($Q$) [6]. The subscript 0 represents the value of the feature without any optimization. Box plots are used to represent the quantities computed for 100 networks, according to: lower whisker (horizontal trace below and on top of the box) for the lowest observation still within 1.5 IQR of the lower quartile (25% percentile of the distribution), bottom of the box for the lower quartile, white trace for the median, top of the box for the upper quartile, and upper whisker for the highest value still within 1.5 IQR of the upper quartile.
FIG. S5. Flight distance distribution (link length) of the WAN. The empirical distribution of flight length shows a power law decay with exponent $-2.2 \pm 0.2$. Both raw (turquoise) and binned (purple) data are shown in the plot. For simplicity, an exponential cutoff is not considered.

* Correspondence and requests for materials should be addressed to V. H. P. Louzada (louzada@ethz.ch)