Eur. Phys. J. B **38**, 381–385 (2004)
DOI: 10.1140/epjb/e2004-00131-0

THE EUROPEAN
PHYSICAL JOURNAL B

Modeling the world-wide airport network

R. Guimerà[a] and L.A.N. Amaral[b]

Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA

Received 12 November 2003 / Received in final form 14 January 2004

Abstract. Recently, we have presented the first exhaustive analysis of the world-wide airport network. Two important results of that study are that: (i) the world-wide airport network is a small-world network with power-law decaying degree and betweenness centrality distributions; (ii) the most connected cities (largest degree) are typically not the most central cities (largest betweenness centrality). This second finding is particularly significant because of results demonstrating that nodes with high betweenness tend to play a more important role in keeping networks connected than those with high degree. Here, we investigate if current network models can explain this finding and we show that they cannot. Thus, we propose a new model that explains this behavior in terms of the geo-political constraints that affect the growth of the airport network. We further hypothesize that in other infrastructures, affected by similar geo-political constraints, critical locations might not coincide with highly-connected hubs.

PACS. 89.75.Fb Structures and organization in complex systems – 89.75.Da Systems obeying scaling laws – 89.40.Dd Air transporation

1 Introduction

The importance of the world-wide airport network extends beyond the convenience it may provide travelers. Airports and national airline companies are often times associated with the image a country or region wants to project [1–4], and have an enormous economic impact on local, national, and international economies [5]. For these reasons, many measures—including, total number of passengers, total number of flights, or total amount of cargo—quantifying the importance of the world airports are compiled and publicized [6].

As for any critical infrastructure, failures or inefficiencies of the system have large economic costs. For example, it was calculated by the European Organisation for the Safety of Air Navigation that flight delays cost the countries in Europe 150–200 billion Euro in 1999 alone [5].

The airport network is also responsible for the mobility of millions of people everyday —for example, O'Hare international airport, in Chicago, is used by approximately 200,000 people every day— and, indirectly, for the propagation of certain diseases such as influenza and, recently, SARS. The airport network thus plays for certain diseases a role that is analogous to that of the web of human sexual contacts [7] for the propagation of AIDS and other sexually-transmitted infections [8].

In a recent article [9], we have reported the results of the first exhaustive analysis of the structure of the world-wide airport network. We considered 3883 cities and more than 500,000 non-stop flights [10], corresponding to 27,051 distinct city pairs having non-stop connections, and investigated the overall properties, efficiency, and growth mechanisms of the resulting network. A number of results of that study are worth pointing out.

First, the world-wide airport network is a small-world network [11,12] for which the number of direct connections k to a given city (degree) has a cumulative distribution $P(>k)$ that decays as a truncated power-law

$$P(>k) \propto k^{-\alpha}\ f(k/k_{\times})\,, \tag{1}$$

where $\alpha = 1.0 \pm 0.1$ is the power-law exponent, $f(u)$ is a truncation function, and $k_{\times}$ is a crossover value that depends on the size S of the network as $k_{\times} \sim S^{0.4}$.

Second, the number of shortest paths b going through a given city (betweenness centrality [13,14]) has a distribution that also decays as a truncated power-law

$$P(>b) \propto b^{-\nu}\ g(b/b_{\times}) \tag{2}$$

where $\nu = 0.9 \pm 0.1$ is the power-law exponent, $g(u)$ is a truncation function, and $b_{\times}$ is a crossover value that depends on the size of the network as $b_{\times} \sim S^{0.5}$.

Third, unlike for most other networks reported in the literature [15–18], the most connected cities (largest degree) are typically not the most central cities (largest betweenness centrality) —this behavior is observed both at

---

[a] e-mail: `rguimera@northwestern.edu`
[b] e-mail: `amaral@northwestern.edu`

the world-wide level and when considering regional airport networks. In other words, in the airport network there are cities with very small degree but very large betweenness. This result turns out to be very important as it has been shown that nodes with high betweenness tend to play a more important role than those with high degree in keeping networks connected [16] —and therefore might also play a key role in the propagation of diseases.

Here, we address the issue of identifying the mechanism by which central nodes that are not hubs can emerge. Recently, it has been suggested that assortative networks —that is, networks in which nodes belongs to *classes* and in which nodes tend to connect to other nodes of the same *class* [19]— might naturally display this property [17]. First, we show that current models that consider preferential attachment and geographical distance constraints cannot reproduce the observed behavior, even though they generate networks in which airports tend to be connected with other airports that are geographically close.

We solve the large-betweenness/small-degree puzzle by considering a new type of mechanism that takes into account geo-political constraints. This type of constraints are present in the growth of the airport network and also in the growth of a number of other critical infrastructures, such as the power grid. Thus, we surmise that, as for the airport network, for such infrastructures the critical locations might not coincide with the hubs.

2 The roles of preferential attachment and distance on the structure of the airport network

First, we consider a minimum model that takes into account preferential attachment and geographical distance constraints. Preferential attachment is known to generate a scale-free degree distribution, as observed in the real airport network [9], while the distance constraints account for the fact that airports are typically connected to other airports that are close geographically.

The model is as follows. At each time step, one of two events can take place [20]:

- With probability p, we establish a new link between two nodes already in the network.
- With probability $1-p$, we add a new node and connect it to m nodes already in the network.

In both cases, when establishing the connections we take into account the degrees of the nodes and the effect of geographical distance between nodes. Airports, when created, are placed in locations that correspond to actual airport locations and the size of the model network is always the same as the size of the real network[1]. When creating a link between a new node i and an existing node j, j is selected with probability

$$\Pi_j \propto \frac{k_j}{F(d_{ij})}, \tag{3}$$

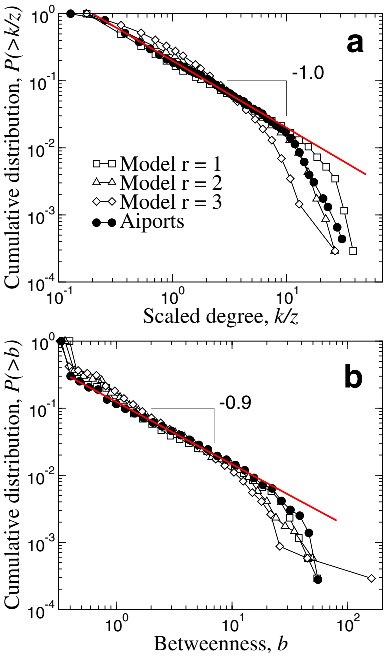

Fig. 1. Case $F(d) = d^r$. Filled symbols correspond to the airport network and open symbols to the model. The simulations are carried out for $p = 0.65$ (so that the exponent of the degree distribution is in agreement with the observed one) and $m = 1$ (so that the average degree is as close as possible to the average degree of the airport network). The number of cities in the network and their geographical coordinates are taken from the data. We present results for $r = 1$, $r = 2$, and $r = 3$. (a) Degree distribution of the real networks and of networks generated with the model. For convenience, the degree k is scaled with the average degree z of the network. (b) Betweenness distribution of the real networks and of networks generated with the model. The betweenness of the nodes is scaled with the average betweenness of the network.

where $F(d_{ij})$ is an increasing function of the geographical distance d_{ij} between the two cities. Similarly, when creating a new link between two existing nodes i and j, they are selected according to

$$\Pi_{ij} \propto \frac{k_i k_j}{F(d_{ij})}. \tag{4}$$

We investigate two different functional forms for the function $F(d)$: (i) a power-law $F_1(d) = d^r$ [21]; (ii) an exponential $F_2(d) = \exp(d/d_\times)$ [22], where $d_\times$ is a characteristic distance.

The preferential attachment mechanism leads to a power-law degree distribution —whose exponent can be tuned by changing the value of p [20]— while $F(d)$ gives rise to a truncation of the power-law decay. When $F(d)$ increases very rapidly, the power-law decay regime can disappear altogether [22].

2.1 Power-law dependence on the distance

The power-law dependence on the distance has been used to model the topology of the Internet [21], where distance constraints are present but appear not to be very strong. We show the results of the model with power-law distance

---

[1] The locations of the airports are chosen in random order. Identical results are obtained if airports are added using some alternative order–for example, adding first the airports with a higher load of passengers.

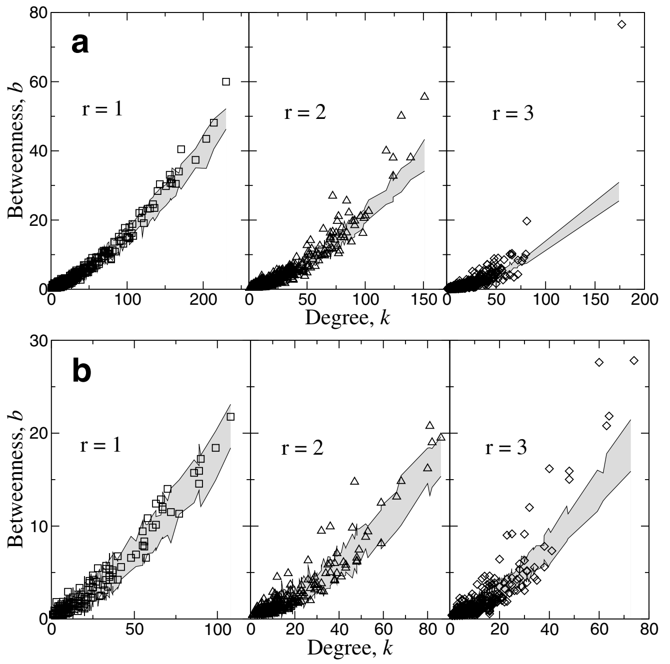

Fig. 2. Case $F(d) = d^r$. The simulations are carried out with the same parameters as in Figure 1. The points correspond to simulations of the model and the shaded regions represent the 95% confidence intervals for random networks with exactly the same degree distributions as for the model networks [23,24]. (a) Betweenness of the nodes as a function of their degree for a model world-wide airport network. (b) Betweenness of the nodes as a function of their degree for a model North American airport network.

constraints in Figure 1. We fix the parameter $p = 0.65$ so that the exponent of the degree distribution is in agreement with the observed one. Similarly, we fix $m = 1$ so that the average degree is as close as possible to the average degree of the world-wide airport network.

Our results show that the model is able to reproduce both the degree distribution $P(> k)$ (Fig. 1a) —whose exponent is imposed at the outset— and the betweenness distribution $P(> b)$ (Fig. 1b) —which is an outcome of the model.

However, the preferential attachment rule and the distance constraints cannot explain the fact that some nodes with very small degree have a very high betweenness (Fig. 2). Although larger r exponents in the distance constraint generate some fluctuations in the betweenness of nodes as a function of their degree, these fluctuations are clearly smaller than the fluctuations observed in the real data [9].

2.2 Exponential dependence on the distance

Next, we consider the case $F(d) = \exp(d/d_\times)$. We show the results of the model for this case in Figure 3. As in the case of $F(d) = d^r$, we use $p = 0.65$ and $m = 1$, so that the degree distribution is as similar as possible to the real one.

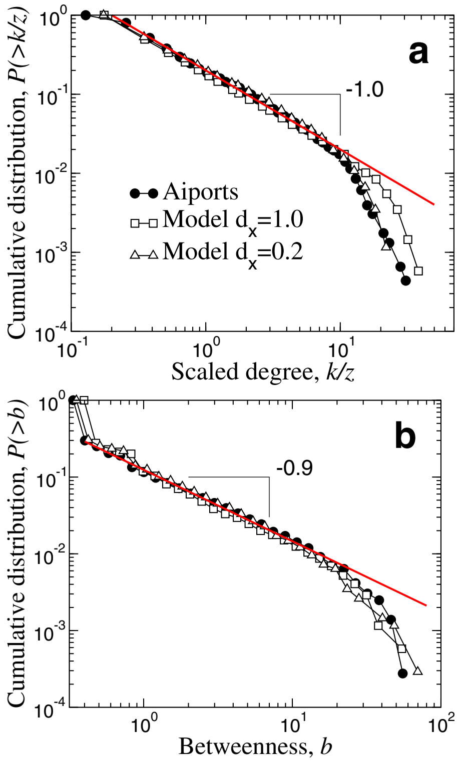

Fig. 3. Case $F(d) = \exp(d/d_\times)$. Filled symbols correspond to the airport network and open symbols to the model. The simulations are carried out with the same parameter values as in Figure 1. Distances are expressed in terms of the radius of the Earth R_T. We plot results for $d_\times = 0.2\,R_T \approx 1,300$ km and $d_\times = R_T \approx 6,500$ km. (a) Degree distribution of the real network and of networks generated with the model. (b) Betweenness distribution of the real network and of networks generated with the model.

The degree and betweenness distributions (Figs. 3a and b, respectively) are again in good agreement with the data. However, the introduction of a characteristic length has some important consequences. Namely, $F(d)$ only affects the structure of the network when considering regions much larger than $d_\times$. When $d_\times \approx 1,300$ km we observe important fluctuations of $b(k)$ for the world-wide model network and the emergence of cities with relatively small degree and high betweenness (Fig. 4a). However, with the same characteristic length such cases are not observed at all for a model of the North American region (Fig. 4b).

3 Role of geo-political constraints on the structure of the airport network

From the analysis of the role of preferential connections and distance constraints, one can conclude that some other mechanism must be at play in the formation and evolution of the airport network. Preferential attachment and distance constraints appear to explain the degree and the betweenness distributions but fail to account for the fact that some central airports have small degree. We hypothesize that there is an additional constraint that arises

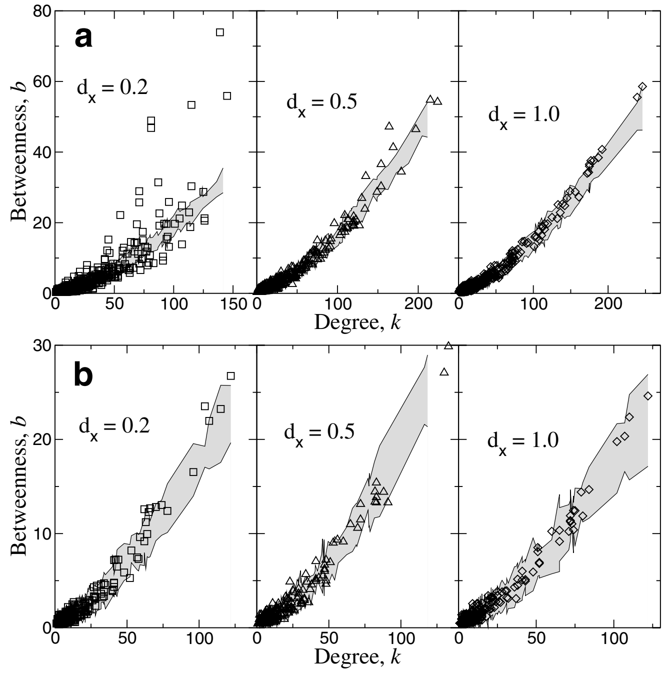

Fig. 4. Case $F(d) = \exp(d/d_\times)$. The simulations are carried out with the same parameters as in Figure 3. The points correspond to simulations of the model and the shaded regions represent the 95% confidence intervals for random networks with exactly the same degree distributions as for the model networks [23,24]. (a) Betweenness of the nodes as a function of their degree for a model world-wide airport network. (b) Betweenness of the nodes as a function of their degree for a model North American airport network.

from geo-political considerations. Namely, only a few airports in each country are connected to airports in other countries. The other airports are only permitted to connect to airports within the same country, even though the geographical distance of some of the flights can be longer than to cities in neighboring countries.

In order to take this effect into consideration, we modify the model in the previous section in the following way: Most cities are only allowed to establish connections with other cities within the same country and only a few are allowed to establish international connections. The precise rules of the model are the following. First, 10% of the nodes are added exactly as described in Section 2. We use $F(d) = d^r$ with $r = 1$ to avoid the mentioned problem of introducing a characteristic distance.

After this initial fraction of nodes is added, the remaining are added again as described in Section 2 but with the additional constraint that only connections between cities in the same country are permitted. Some international connections will still be added whenever a node is already connected to all the cities in its country.

As we show in Figure 5, this model generates central nodes with small degree, as observed in the real airport network. Moreover, this behavior is observed both at the global level and at the regional level.

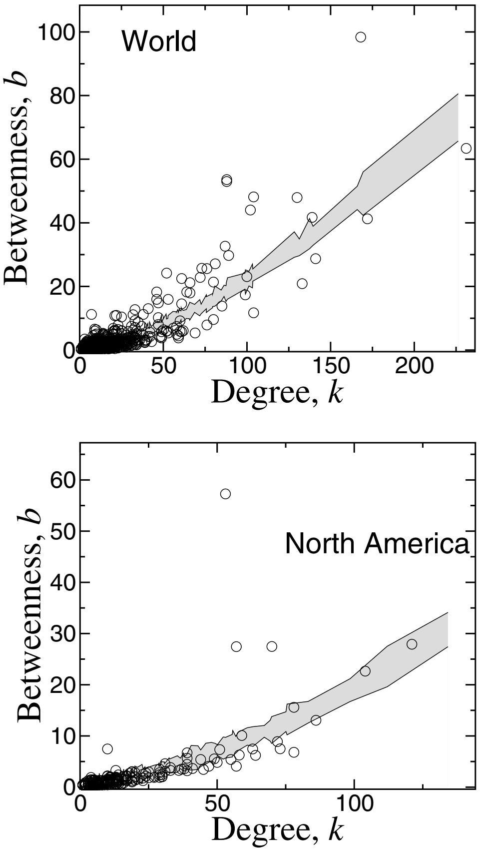

Fig. 5. Case with geo-political constraints. The simulations are carried out with the same parameters as in Figure 1. The initial fraction of nodes added without political constraints is 10%. Central nodes with small degree appear both at the world-wide level and at the regional level.

4 Conclusions

A recent analysis of the world-wide airport network has uncovered the surprising finding that the most connected cities are typically not the most central cities [9].

Here, we have shown that models existing in the literature cannot account for this fact. In particular, we have considered a model with preferential attachment and geographical distance constraints. For this model, large degree nodes tend to have also large betweenness and vice versa. Therefore, we have proposed a new model with geo-political constraints. In our model, only a few cities in each country are allowed to establish connections to cities in other countries. This model explains the existence of large-betweenness/small-degree nodes.

Importantly, we note that such geo-political constraints are also present in the formation and growth of other critical infrastructures such as the power grid. Therefore, we surmise that, as for the airport network, such infrastructures might have critical locations —“Achilles’ heels”— that do not coincide with the hubs.

We thank A. Arenas, A. Barrat, M. Barthelémy, A. Díaz-Guilera, A.A. Moreira, S. Mossa, R. Pastor-Satorras, M. Sales,

D. Stouffer, and A. Vespignani for stimulating discussions and helpful suggestions. We also thank OAG for making their electronic database of airline flights available to us, and Landings.com for providing us with the geographical coordinates of the world airports.

References

1. A. Bisseur, F. Alamdari, Transportation **25**, 331 (1998)
2. L.P. Dana, D. Vignali, Int. Marketing Rev. **16**, 278 (1999)
3. B.J. Turton, C.C. Mutambirwa, Tourism Management **17**, 453 (1996)
4. K. Raguraman, Tourism Management **19**, 533 (1998)
5. Cost of the air transportation delay in Europe, *Tech. rep.*, EUROCONTROL – European Organisation for the Safety of Air Navigation (2000); [`http://www.eurocontrol.int/prc/reports/prr2/index.html`]
6. ACI Annual Worldwide Airports Traffic Reports, *Tech. rep.*, Airport Council International, Geneva (1999). [`http://www.airports.org`]
7. F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, Y. Aberg, Nature **411**, 907 (2001)
8. F. Liljeros, C.R. Edling, L.A.N. Amaral, Microbes Infect. **5**, 189 (2003)
9. R. Guimerà, S. Mossa, A. Turtschi, L.A.N. Amaral, `arxiv/cond-mat 0312535` (2003)
10. OAG MAX Database, OAG, London (2000)
11. D.J. Watts, S.H. Strogatz, Nature **393**, 440 (1998)
12. L.A.N. Amaral, A. Scala, M. Barthelémy, H.E. Stanley, Proc. Nat. Acad. Sci. USA **97**, 11149 (2000)
13. L.C. Freeman, Sociometry **40**, 35 (1977)
14. M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001)
15. A. Vázquez, R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 066130 (2002)
16. P. Holme, B.J. Kim, Phys. Rev. E **65**, 056109 (2002)
17. K.-I. Goh, E. Oh, B. Kahng, D. Kim, Phys. Rev. E **67**, 017101 (2003)
18. M. Barthelémy, submitted `arxiv/cond-mat 0309436` (2003)
19. M. Newman, Phys. Rev. Lett. **89**, 208701 (2002)
20. S. Dorogovtsev, J.F.F. Mendes, Europhys. Lett. **52**, 33 (2000)
21. S.-H. Yook, H. Jeong, A.-L. Barabási, Proc. Nat. Acad. Sci. USA **99**, 13382 (2002)
22. M. Barthelémy, Europhys. Lett. **63**, 915 (2003)
23. M. Molloy, B. Reed, Random Structures and Algorithms **6**, 161 (1995)
24. M. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. E **64**, 026118 (2001)