Physica A 388 (2009) 3651–3656

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

An optimal network for passenger traffic

A.K. Nandi, K. Bhattacharya, S.S. Manna *

Satyendra Nath Bose National Centre for Basic Sciences, Block-JD, Sector-III, Salt Lake, Kolkata-700098, India

ARTICLE INFO

Article history:
Received 24 February 2009
Received in revised form 6 May 2009
Available online 19 May 2009

PACS:
89.75.Hc
89.75.Fb
05.60.-k
45.10.Db

Keywords:
Networks
Structures
Transport processes
Optimization processes

ABSTRACT

The optimal solution of an inter-city passenger transport network has been studied using Zipf's law for the city populations and the Gravity law describing the fluxes of inter-city passenger traffic. Assuming a fixed value for the cost of transport per person per kilometer we observe that while the total traffic cost decreases, the total wiring cost increases with the density of links. As a result the total cost to maintain the traffic distribution is optimal at a certain link density which vanishes on increasing the network size. At a finite link density the network is scale-free. Using this model the air-route network of India has been generated and an one-to-one comparison of the nodal degree values with the real network has been made.

The identification of certain crucial controlling parameters that ensure the characteristic structure of a random network, be it a network that has been created by a natural process or a network that has evolved due to the social requirements, has been a focal point of research interest for quite some time [1–4]. For example, different algorithms have been proposed to generate the well-known scale-free structures of highly heterogeneous networks which successfully reproduce the statistical features of important networks like the Internet [5], World Wide Web [6] and airport networks [7].

On the other hand, not much attention has been paid to reproducing the structural features specific to a particular network and to making a one-to-one comparison of the real and the model networks. Intuitively it is evident that such a modeling would need information specific to such a network. In this paper we argue that for a network of passenger traffic it is possible to construct an optimized model network of this kind using only two ingredients, namely the node-wise population distribution as well as a guiding rule for the passenger traffic flows.

A transport network should be efficient as well as cost effective. Efficiency is ensured when the communication between an arbitrary pair of nodes takes only a finite and short duration even when the network is very large. This implies that the network must be characterized by 'small-world' features. In addition, the network should be robust with respect to random failures. If a link is down, the transport process should not be grossly affected. This implies that the network must not have a tree structure which is most economic but has extreme sensitivity to failures. In practice the network should be such that when the flow is not possible along a certain path, there must exist alternate paths, even of longer lengths, to maintain the flow. Indeed real-world transport networks are never like tree graphs. Actually they have multiple loops of many different length scales and therefore they are hardly affected by random link or node failures. A prominent example of this is the Internet, and its robustness to random failures in its structure is quite well known [8]. Secondly, the laying cost of the network is another controlling factor. If every node is connected to all other nodes it would be excellent, but that would involve large establishment and maintenance cost. Planners and administrators of railway networks, city bus transport systems, or even postal networks establish and upgrade their networks keeping mainly these two aspects in mind.

* Corresponding author. Tel.: +91 33 2335 5708; fax: +91 33 2335 3477.
E-mail address: manna@bose.res.in (S.S. Manna).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2009.05.017

Recently, optimal networks embedded in Euclidean space have attracted much attention. Given a spatial distribution of human population the locations of the different facilities so that the mean distance is a minimum was discussed in Ref. [9]. Signatures of topology and patterns are explored in Ref. [10]. A minimal spanning tree structure of the optimal network was proposed in Ref. [11].

Here we study a model network for the passenger traffic among different cities. We ask if, given the populations and locations of all cities in a country, can one predict the structure of the network that is optimized with respect to the connection robustness and wiring cost? Our study is based on the framework of Zipf's law [12,13] of city population distribution and the Gravity law [14] of social and economic sciences describing the strength of the passenger traffic between a pair of cities. Finally, we apply this scheme to the Indian air traffic network, which gives good correspondence with the real network.

Zipf's law for the frequency of occurrence of English words has also been applied to the rank-size distribution of city populations [12]. In a country the maximally populated city is assigned rank 1, the second largest is put in rank 2, etc. It is known that the population size $p(r)$ varies inversely with rank r. This also implies that the population distribution is a power law [13].

The passenger traffic among N different cities and towns in a country is given by the well-established Gravity model [14]. In its introductory form the magnitude of the passenger flow from city i to city j is jointly proportional to their individual populations p_i and p_j and at the same time is penalized by an inversely proportional factor which is the square of their distance of separation ℓ_{ij} as $F_{ij} \propto p_i p_j/\ell_{ij}^2$. This equation has been generalized to the following asymmetric parametric form [15]:

$$F_{ij} = p_i^{\alpha} \left(\frac{p_j^{\beta}}{\ell_{ij}^{\theta}} \bigg/ \sum_{k \neq i} \frac{p_k^{\beta}}{\ell_{ik}^{\theta}} \right) \tag{1}$$

where α, β and θ are suitable parameters and k runs over all $N - 1$ nodes except i.

While applying these two laws we assume that not only the total population of the country is conserved but also the individual city populations remain constant. More specifically, we assume that in a certain unit of time F_{ij} tourists travel from city i to city j but they eventually return to their own city i within the same time interval. Of course there are a few who migrate from one city to the other and start living there, but their number must be very small compared to the tourist traffic, and we ignore this component of migratory flow. Therefore neither the city populations nor the inter-city traffic flow changes with time. It seems that our model should also be quite appropriate for a postal distribution network.

In a simple model we take a unit square box on the x–y plane to represent the country and N points distributed at random positions within the box as the locations of different cities. Though the periodic boundary condition has no physical meaning in this context we use it along both the transverse directions on the box to make the data more well behaved. Given the set of coordinates of N points $\{x_i, y_i\}$, $i = 1, N$, all inter-city distances ℓ_{ij} are determined. Cities are then assigned populations p_i, $(i = 1, N)$ (in real numbers) by drawing them from a power law distribution $\text{Prob}(p) \sim p^{-\mu}$ with $\mu = 1$ as per Zipf's law. Using $p_{min} = 0.001$ and $p_{max} = 1$, the city populations are generated using the relation $p = p_{min}\,(p_{max}/p_{min})^{r_1}$, where r_1 is a uniformly distributed random fraction and finally normalized such that $\Sigma_i p_i = 1$.

Knowing the values for the city populations p_i and the mutual inter-city distances, the magnitudes of passenger fluxes F_{ij} and F_{ji}s are calculated using Eq. (1) and for a certain set of values of α, β and θ. By definition this flow pattern is inherently directed. However, we consider only an undirected traffic flux between i and j by considering the net flow $\tilde{F}_{ij} = F_{ij} + F_{ji}$. Let at some arbitrary intermediate stage all N cities be linked by a singly connected network. Since nodes are randomly distributed on a continuous plane, there exists one and only one shortest path between a pair of nodes. We assume that the entire flow $\tilde{F}_{ij}$ passes through the shortest path on the network connecting the nodes i and j and therefore each link on this path is assigned $\tilde{F}_{ij}$. When this assignment process has been completed for all distinct $N(N-1)/2$ node pairs, the net flow through a link measures the net traffic w through that link. The quantity w is like a weighted betweenness centrality, and the net traffic along a link connecting a pair of nodes i' and j' is $w_{i'j'} = \Sigma \tilde{F}_{ij}$, where the summation is taken over the subset of $N(N-1)/2$ node pairs whose shortest paths pass through the link $\{i'j'\}$. On a graph having loops the shortest paths are found using the well-known Dijkstra algorithm [16].

The cost function for this traffic distribution has two competing factors. Given a network there is a cost to maintain the traffic along every link. We assume a fixed value for the cost to transport a unit population through unit distance along every link of the network (e.g., per person per kilometer). Therefore if w_{ij} is the net flow between the two end nodes i and j of a link of length ℓ_{ij} then the total cost involved to maintain the entire traffic flow is $\mathcal{C}_{tra} = \Sigma_{i \neq j} w_{ij} \ell_{ij} a_{ij}$. The second factor is the establishment cost to construct the connections, which is $\mathcal{C}_{net} = \Sigma_{i \neq j} \ell_{ij} a_{ij}$. Here, the a_{ij}s are the elements of the adjacency matrix and $a_{ij} = 1$ if there exists a link between i and j; otherwise it is 0. Therefore the total cost function to maintain the whole traffic distribution of the network is the sum of these two factors:

$$\mathcal{C} = \mathcal{C}_{net} + \lambda \mathcal{C}_{tra} = \Sigma_{i \neq j} (1 + \lambda w_{ij}) \ell_{ij} a_{ij} \tag{2}$$

where λ is a type of conversion factor that makes an equivalence between the two types of cost.

The Minimal Spanning Tree (MST) graph covering all N nodes using the Euclidean distances ℓ_{ij} as the link weights has the minimal value of the networking cost $\mathcal{C}_{net}$. Using Kruskal's algorithm [17] to generate the MST, the whole set of $N(N-1)/2$