

LETTER

## Optimal traffic networks

Marc Barthélemy<sup>1,2</sup> and Alessandro Flammini<sup>1</sup>

<sup>1</sup> School of Informatics and Biocomplexity Center, Indiana University, Eigenmann Hall, 1900 East Tenth Street, Bloomington IN 47406, USA

<sup>2</sup> CEA-Centre d'Etudes de Bruyères-le-Châtel, Département de Physique Théorique et Appliquée BP12, 91680 Bruyères-Le-Châtel, France  
E-mail: [marc.barthelemy@cea.fr](mailto:marc.barthelemy@cea.fr) and [aflammin@indiana.edu](mailto:aflammin@indiana.edu)

Received 29 March 2006

Accepted 10 July 2006

Published 24 July 2006

Online at [stacks.iop.org/JSTAT/2006/L07002](http://stacks.iop.org/JSTAT/2006/L07002)

[doi:10.1088/1742-5468/2006/07/L07002](https://doi.org/10.1088/1742-5468/2006/07/L07002)

**Abstract.** Inspired by studies on the airports' network and the physical Internet, we propose a general model of weighted networks via an optimization principle. The topology of the optimal network turns out to be a spanning tree that minimizes a combination of topological and metric quantities. It is characterized by strongly heterogeneous traffic, non-trivial correlations between distance and traffic and a broadly distributed centrality. A clear spatial hierarchical organization, with local hubs distributing traffic in smaller regions, emerges as a result of the optimization. Varying the parameters of the cost function, different classes of trees are recovered, including in particular the minimum spanning tree and the shortest path tree. These results suggest that a variational approach represents an alternative and possibly very meaningful path to the study of the structure of complex weighted networks.

**Keywords:** random graphs, networks, communication, supply and information networks

Transportation and communication infrastructures such as the airports' network and the physical Internet are characterized by broad distributions of traffic [1, 2], betweenness centrality [3] and in some case also of degree [4]. Strong, nonlinear traffic–distance and traffic–connectivity correlations have also been reported [1, 2]. Modelization attempts with ingredients such as random weights, dynamical rules or weight–topology coupled dynamics (see e.g. [2] and references therein) have mainly focused on growth processes, and variational approaches have instead mainly been used only in practical problems by road traffic engineers [5]. Both the problem of optimal traffic [6] on a network and of optimal networks [7] have a long tradition in mathematics and physics. It is well known, for example, that the laws that describe the flow of currents in a resistor network [8] can be derived by minimizing the energy dissipated by the network [9]. On the other hand, optimal networks have shown to be relevant in the study of the mammalian circulatory system [10], food webs [11], general transportation networks [12], metabolic rates [13], river networks [14], and gas pipelines or train tracks [15]. All these studies share the fact that the nodes of the network are embedded in a  $d$ -dimensional euclidean space, which implies that the degree is almost always limited and the connections are restricted to ‘neighbours’ only. A second broad class of optimal networks where spatial constraints are absent has also been investigated recently. It has been shown, for example, that optimization of both the average shortest path and the total length can lead to small-world networks [16] and, more generally, degree correlations [17] or scale-free features [18] can emerge from an optimization process. Cancho and Sole [19] showed that the minimization of the average shortest path and the link density leads to a variety of networks, including exponential-like graphs and scale-free networks. Guimera *et al* [20] studied networks with minimal search cost and found two classes of networks: star-like and homogeneous networks. Finally, Colizza *et al* [21] studied networks with the shortest route and the smallest congestion and showed that this interplay could lead to a variety of networks when the number of links per node is changed.

In this letter, mostly inspired by studies on airports' networks [1, 4], we investigate the important case for which nodes are embedded in a two-dimensional plane but links are not constrained (as for air routes) to connect ‘neighbours’. We propose a cost function that depends both on the length and the traffic carried by the links and show that the resulting optimal network is hierarchically organized in space and displays a complex traffic structure. We consider a set of  $N$  points (‘airports’) randomly distributed in a square of unitary area and we would like to build a network (air routes) that connects all points. The cost or weight  $w_{ij}$  associated with ‘travelling’ along a link  $(i, j)$  is a function of both the length  $d_{ij}$  of the link and of the traffic  $t_{ij}$  that it carries. In the air-network analogy, the quantity  $t_{ij}$  represents the number of passengers on the link  $(i, j)$  and is symmetrical,  $t_{ij} = t_{ji}$ . To travel from a generic node  $i_0$  to another generic node  $i_p$  along a specific path  $\{i_0, i_1, i_2, \dots, i_{p-1}, i_p\}$ , the cost to be paid is the sum of the weights  $w_{i_k, i_{k+1}}$  associated with the links that compose the path, and when more than one path is available we assume that the most economical one is chosen:

$$\mathcal{C}_{i_0, i_p} = \min_{p \in \mathcal{P}(i_0, i_p)} \sum_{e \in p} w_e \quad (1)$$

where the minimization is over all paths  $p$  belonging in the set of paths  $\mathcal{P}(i_0, i_p)$  going from  $i_0$  to  $i_p$  ( $w_e$  is the weight of the edge  $e$ ). If there is no path between a pair of nodes,

the corresponding cost will be taken as being equal to infinity. This choice ensures that the optimal network will be connected. The global quantity  $\mathcal{E}_0$  that we wish to minimize is then the average cost to pay to travel from a generic node to another:

$$\mathcal{E}_0(\{t_{ij}\}) = \frac{2}{N(N-1)} \sum_{i<j} c_{i,j}. \quad (2)$$

Our purpose is therefore to find the traffic  $\{t_{ij}^*\}$  carried by the links and which minimizes (2), with the only constraints that all  $t_{ij} \geq 0$  and that the total traffic  $T = \sum_{i<j} t_{ij}$  is fixed. In this letter, we choose as the weight of a link  $e$  the ratio of its length to traffic:  $w_e = d_e/t_e$  (with this choice, the value of  $T$  fixes the scale of traffic and does not affect the topology of the optimal network, since a rescaling of the total energy by a constant factor will not affect the minimization<sup>3</sup>). Although this choice is not the most general, it naturally verifies the expectation that the weight increases with  $d_{ij}$  and decreases with  $t_{ij}$ . This last condition can be easily understood in the case of transportation networks and means that it is more economic to travel on links with large traffic, reducing the effective distance—or the marginal cost—of the connection. We search for the minimum-realizing traffic using a zero-temperature metropolis algorithm. The elementary move consists of transferring a random fraction of the traffic carried by a link to another one (the total traffic being fixed). We choose at random two links  $(i, j)$  and  $(i', j')$  and transfer weights between them according to

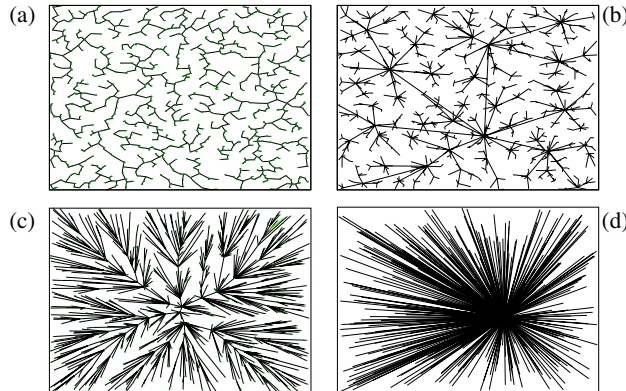
$$t_{ij} \rightarrow t_{ij} - \alpha t_{ij} \quad (3)$$

$$t_{i'j'} \rightarrow t_{i'j'} + \alpha t_{ij} \quad (4)$$

where  $\alpha$  is a uniform random number between zero and one. The sign of  $\alpha$  is positive with probability  $p$  and negative with probability  $1 - p$  (if one of the links has a zero weight, the transfer can only be made in one direction; in other cases  $p = 0.5$  ensures a quick convergence). If, after the transfer, the weight of a link is zero, the corresponding link is deleted. The minimum-cost path between two points is recalculated at any step using Dijkstra's algorithm [7]. We compute the energy difference,  $\Delta = \mathcal{E}'_0 - \mathcal{E}_0$ , and only if it is negative is the transfer accepted. We test a number of order  $\mathcal{O}(N^2)$  of such transfers which converges to an optimal network that minimizes the energy  $\mathcal{E}_0$ . The initial topology is a complete graph with random weights on the links. As we show below, the optimal solution is characterized by a non-trivial topology and spatial organization which results as the compromise of two opposing forces: the need for short routes and traffic concentration on as few paths as possible. The interplay between topology and traffic naturally induces the observed correlations between degree, distance and traffic itself.

Numerical simulations show that the optimal network is a tree. A simple example supporting this finding is obtained by considering an isocetes triangle  $ABC$  with  $d(A, C) = d(B, C) = d$  and  $d(A, B) = d'$ ; optimization leads to the values  $t_{AC} = t_{BC} \approx T/2$  and  $t_{AB} \approx 0$  when  $d \gg d'$ . The minimum energy is thus (at leading order in  $d$ )  $\mathcal{E}_0 \approx d/t_{AC} + d/t_{BC} \approx 4d/T$ . When we remove the link BC and thus kill the loop, the traffic on AC becomes approximately twice as high, i.e.  $t_{AC} \approx T$  (and  $t_{BC} \approx t_{AB} \approx 0$ ), but the minimum energy at leading order is  $\mathcal{E}'_0 \approx 2d/t_{AC} \approx 2d/T$ , which is lower than  $\mathcal{E}_0$ .

<sup>3</sup> We chose  $t_{ij}$  to be integer for faster simulations, but the results are the same with real  $t_{ij}$ .



**Figure 1.** Different spanning trees obtained for different values of  $(\mu, \nu)$  in equation (6) obtained for the same set of  $N = 1000$  nodes. (a) Minimum spanning tree obtained for  $(\mu, \nu) = (0, 1)$ . In this case the total distance is minimized. (b) Optimal traffic tree obtained for  $(\mu, \nu) = (1/2, 1/2)$ . In this case, we have an interplay between centralization and minimum distance resulting in local hubs. (c) Minimum euclidean distance tree obtained for  $(\mu, \nu) = (1, 1)$ . In this case, centrality dominates over distance and a ‘star’ structure emerges with a few dominant hubs. (d) Optimal betweenness centrality tree obtained for  $(\mu, \nu) = (1, 0)$ . In this case, we obtain the shortest path tree which has one star hub (for the sake of clarity, we omitted some links in this last figure).

This example shows that optimization reduces the number of links joining nodes in the same regions and increases the traffic on the remaining links. Loops between nodes in the same neighbourhood then become redundant.

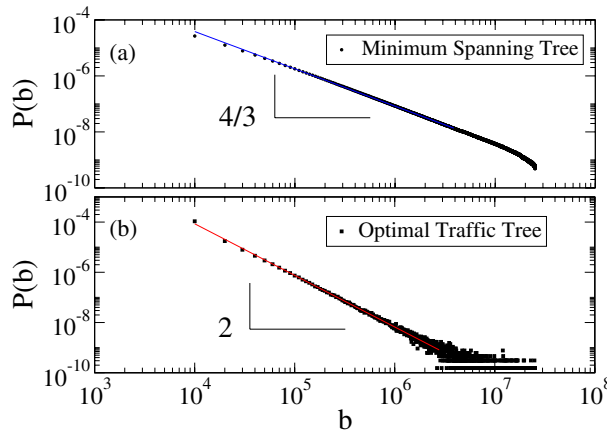
The optimal network being a tree enormously simplifies the computation of the energy. Since only a single path exists between any two nodes in a tree, the energy (2) can be rewritten as  $\mathcal{E}_0 = \sum_{e \in \mathcal{T}} b_e d_e / t_e$ , where  $b_e$  is the edge-betweenness [22] and counts the number of times that  $e$  belongs to the shortest path between two nodes. The optimal traffic (with the same constraints as above) is given by  $t_e = T \sqrt{b_e d_e} / \sum_e \sqrt{b_e d_e}$  and the topology of the ‘optimal traffic tree’ (OTT) can then be obtained by minimizing

$$\mathcal{E} = \sum_{e \in \mathcal{T}} \sqrt{b_e d_e}. \quad (5)$$

The minimal configuration can now be searched by rewiring links. Replacing link  $(i, j)$  by  $(i, j')$  modifies only the centralities along the path between  $j$  and  $j'$ , which implies that our calculation has a complexity of order  $\mathcal{O}(N)$  and allows computations over very large networks (the same algorithm was used in the context of river networks [23]). We expect to obtain something very different from the classical (euclidean) minimum spanning tree (MST) [7], since  $\mathcal{E}$  involves a combination of metric (distance) and topological (betweenness) quantities. The expression equation (5) suggests an interesting generalization given by the optimization of

$$\mathcal{E}_{\mu\nu} = \sum_{e \in \mathcal{T}} b_e^\mu d_e^\nu \quad (6)$$

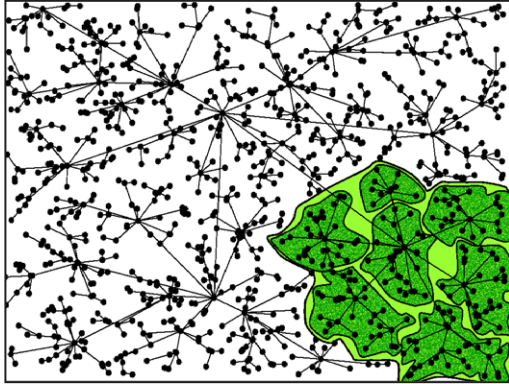
where  $\mu$  and  $\nu$  control the relative importance of distance against topology as measured by centrality. Figure 1 shows examples of spanning trees obtained for different values of



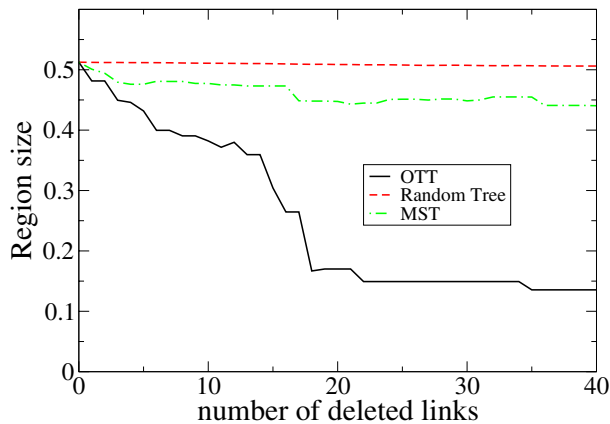
**Figure 2.** Betweenness centrality distribution for the MST and for the OTT. The lines are power-law fits and give the theoretical result  $\delta = 4/3$  for the MST and the value  $\delta \simeq 2.0$  for the OTT ( $N = 10^4$ , 100 configurations).

$(\mu, \nu)$ . For  $(\mu, \nu) = (0, 1)$ , one obtains the euclidean minimum spanning tree (figure 1(a)) which can also be obtained by minimizing the total weight  $\sum_e w_e$  and gives traffic  $t_e = T\sqrt{d_e}/\sum\sqrt{d_e}$ . For  $(\mu, \nu) = (1/2, 1/2)$ , we obtain the OTT (figure 1(b)) which displays an interesting interplay between distance and shortest path minimization (see below). For  $(\mu, \nu) = (1, 1)$ , the energy is proportional to the average shortest weighted path (with weights equal to euclidean distance (figure 1(c))). When  $(\mu, \nu) = (1, 0)$ , the energy (6) is proportional to the average betweenness centrality and therefore to the average shortest path  $\sum_e b_e \propto \ell$ . The tree (1,0) shown in figure 1(d) is thus the shortest path tree (SPT) with an arbitrary ‘star-like’ hub (a small non-zero value of  $\nu$  would select the closest node to the gravity centre as the star). The minimization of equation (6) thus provides a natural interpolation between the MST and the SPT, a problem which was addressed in previous studies [24]. The degree distribution for all cases considered above is not broad (with the possible exception  $(\mu, \nu) = (1, 1)$ —a complete inspection of the plane  $(\mu, \nu)$  is left for future studies), possibly as a consequence of spatial constraints. In particular, the degree distribution for the OTT is well fitted by an exponential function.

It has been shown that trees can be classified in ‘universality classes’ [25, 26] according to the size distribution of the two parts in which a tree can be divided by removing a link (or the sub-basins areas distribution in the language of river network). We define  $A_i$  and  $A_j$  as the sizes of the two parts in which a generic tree is divided by removing the link  $(i, j)$ . The betweenness  $b_{ij}$  of link  $(i, j)$  can be written as  $b_{ij} = \frac{1}{2}[A_i(N - A_i) + A_j(N - A_j)]$ , and the distributions of  $A$ s and  $b$ s can easily be derived one from the other. It is therefore not surprising that the same exponent  $\delta$  characterizes both  $P(A) \sim A^{-\delta}$  and  $P(b)$ . While we obtain the value  $\delta = 4/3$  for the MST [25], for the OTT we obtain (figure 2) an exponent  $\delta \simeq 2$ , which is a value also obtained for trees grown with a preferential attachment mechanism [27] (see also [28] for a supporting argument). Interestingly, most real-world networks are also described by this value,  $\delta \simeq 2$  [3]. The OTT thus tends to have a more uniform centrality with respect to the MST [29], with important consequences on the vulnerability of the network, since there is no clearly designated ‘Achille’s heel’ for the OTT.



**Figure 3.** Hierarchical organization emerging for the optimal traffic tree  $(\mu, \nu) = (1/2, 1/2)$  ( $N = 1000$  nodes). Longer links lead to regional hubs, which in turn connect to smaller hubs distributing traffic in smaller regions.



**Figure 4.** Average euclidean size of the largest cluster remaining after deleting links ranked according to their length (in decreasing order). This plot is obtained for one typical configuration of size  $N = 1000$  and a square of area set to one. The decrease observed for the OTT is consistent with a hierarchical spatial organization, as is visually evident from figure 3.

We now investigate the traffic properties of the OTT and we find that the traffic scales as  $t_{ij} \sim d_{ij}^\tau$  with  $\tau \approx 1.5$  showing that large traffic is carried over large distance and is then dispatched on smaller hubs that distribute it on still smaller regions. Despite the limited range of degrees, we also observe for the strength [1]  $s_i = \sum_j t_{ij}$  a superlinear behaviour with the degree. This result demonstrates the existence of degree–traffic correlations, as observed for the airport network for example [1] which can emerge from a global optimization process. The spatial properties of the OTT are also remarkable, and it displays (figure 3) a hierarchical spatial organization where long links connect regional hubs, which in turn are connected to sub-regional hubs, etc. This hierarchical structure can be probed by measuring the average euclidean distance between nodes belonging to the largest cluster obtained by deleting recursively the longest link. For the OTT (figure 4), we observe a decrease of the region size, demonstrating that longer links connect smaller

regions, a feature absent in non-hierarchical networks such as the MST, the SPT or the random tree (figure 4).

In summary, we showed that the emergence of complex structures in traffic organization could be explained by an optimization principle. In particular, strong correlations between distance and traffic arise naturally as a consequence of optimizing the average weighted shortest path. In the optimal network, long-range links carry large traffic and connect regional hubs dispatching traffic on a smaller scale, ensuring efficient global distribution. These results suggest that the organization of the traffic on complex networks and more generally the architecture of weighted networks could in part result from an evolutionary process. The optimal networks obtained here are trees, but some transportation networks contain loops which reflect the fact that other ingredients are needed in order to describe them. Our results however suggest that some transportation networks could possibly be seen as the superposition of many trees. It would be interesting to check if this is the case for the world-wide airport network which results from the superposition of individual airline company networks which are probably close to trees. Finally, this study led us to propose a generalization of the usual minimum spanning tree by introducing the centrality and allows one to interpolate from the MST to the shortest path tree. This generalization however deserves further study and raises interesting questions such as the crossover from different tree universality classes.

We thank Vittoria Colizza and Alessandro Vespignani for interesting discussions and suggestions.

## References

- [1] Barrat A, Barthélemy M, Pastor-Satorras R and Vespignani A, 2004 *Proc. Nat. Acad. Sci.* **101** 3747
- [2] Barrat A, Barthélemy M and Vespignani A, 2005 *J. Stat. Mech.* **P05003**
- [3] Goh K-I, Oh E, Jeong H, Kahng B and Kim D, 2002 *Proc. Nat. Acad. Sci.* **99** 12583
- [4] Guimera R, Mossa S, Turtschi A and Amaral L A N, 2005 *Proc. Nat. Acad. Sci.* **102** 7794
- [5] Wardrop J G, *Some theoretical aspects of road traffic research*, 1952 *Proc. Inst. Civ. Eng.* II **1** 325
- [6] Ahuja R K, Magnanti T L and Orlin J B, 1993 *Network Flows* (Englewood Cliffs, NJ: Prentice-Hall)
- [7] Jungnickel D, *Graphs, networks, and algorithms*, 1999 *Algorithm and Computation in Mathematics* vol 5 (Berlin: Springer)
- [8] Kirchoff G, 1847 *Ann. Phys. Chem.* **72** 497
- [9] Doyle P G and Snell J L, 1984 *Random Walks and Electric Networks* (USA: The Mathematical Association of America) p 83
- [10] McMahon T A and Bonner J T, 1983 *On Size and Life* (New York: Scientific American Library)
- [11] Garlaschelli D, Caldarelli G and Pietronero L, 2003 *Nature* **423** 165
- [12] Banavar J R, Maritan A and Rinaldo A, 1999 *Nature* **399** 130
- [13] West G B, Brown J H and Enquist B J, 1997 *Science* **276** 122
- [14] Maritan A, Colaiori F, Flammini A, Cieplak M and Banavar J R, 1996 *Science* **272** 984
- [15] Gastner M T and Newman M E J, 2004 *Preprint cond-mat/0409702*  
Gastner M T and Newman M E J, 2004 *Preprint cond-mat/0407680*
- [16] Mathias N and Gopal V, 2001 *Phys. Rev. E* **63** 021117
- [17] Berg J and Lassig M, 2002 *Phys. Rev. Lett.* **89** 228701
- [18] Valverde S, Ferrer i Cancho R and Sole R V, 2002 *Europhys. Lett.* **60** 512
- [19] Ferrer i Cancho R and Sole R V, *Optimization in complex networks*, 2003 *Statistical Mechanics of Complex Networks (Lecture Notes in Physics* vol 625) (Berlin: Springer) p 114
- [20] Guimera R, Diaz-Guilera A, Vega-Redondo F, Cabrales A and Arenas A, 2002 *Phys. Rev. Lett.* **89** 258701
- [21] Colizza V, Banavar J, Maritan A and Rinaldo A, 2004 *Phys. Rev. Lett.* **92** 198701
- [22] Freeman L C, 1977 *Sociometry* **40** 35
- [23] Colaiori F, Flammini A, Maritan A and Banavar J R, 1997 *Phys. Rev. E* **55** 1298
- [24] Khuller S, Raghavachari B and Young N, 1995 *Algorithmica* **14** 305

- [25] Takayasu H, Takayasu M, Provata A and Huber G, 1991 *J. Stat. Phys.* **65** 725
- [26] Manna S S, Dhar D and Majumdar S N, 1992 *Phys. Rev. A* **46** R4471
- [27] Albert R and Barabasi A-L, 2002 *Rev. Mod. Phys.* **74** 47
- [28] De Los Rios P, 2001 *Europhys. Lett.* **56** 898
- [29] Wu Z, Braunstein L A, Havlin S and Stanley H E, 2005 *Preprint* [cond-mat/0511525](https://arxiv.org/abs/cond-mat/0511525)