Scale-free Segregation in Transport Networks

Ph. Blanchard, D. Volchenkov
Bielefeld-Bonn Stochastic Research Center (BiBoS),
Bielefeld University, Postfach 100131, 33501, Bielefeld, Germany
Email: volchenk@physik.uni-bielefeld.de

February 2, 2008

Abstract

Every route of a transport network approaching equilibrium can be represented by a vector of Euclidean space which length quantifies its segregation from the rest of the graph. We have empirically observed that the distribution of lengths over the edge connectivity in many transport networks exhibits scaling invariance phenomenon. We give an example of the canal network of Venice to demonstrate our result. The method is applicable to any transport network.

PACS: 89.65.Lm, 89.75.Fb, 05.40.Fb, 02.10.Ox

Transport networks are used to model the flow of commodity, information, viruses, opinions, or traffic. They typically represent the networks of roads, streets, pipes, aqueducts, power lines, or nearly any structure which permits either vehicular movement or flow of some commodity, products, goods or service. The major aim of the analysis is to determine the structure and properties of transport networks that are important for the emergence of complex flow patterns of vehicles (or pedestrians) through the network such as the Braess paradox [1]. This counter-intuitive phenomenon occurs when adding more resources to a transportation network (say, a new road or a bridge) worsens the quality of traffic by creating longer delays for the drivers, rather than alleviate it. The Braess paradox has been observed in the street traffic of New York City and Stuttgart, [2].

In the present Letter, we show that while approaching equilibrium, a transport network can be embedded into Euclidean space $\mathbb{R}^{N-1}$, $N$ being a number of vertices. Then, every edge of the network is represented by a vector which length quantifies its segregation from the rest of the graph. We have empirically observed that the distribution of lengths over the edge connectivity in urban transport networks exhibits scaling invariance phenomenon. The relation between the connectivity of city spaces and their centrality known as intelligibility is a key determinant of human behaviors in urban environments, [3].

In most of researches devoted to the improving of transport networks, a primary graph representation of urban networks is used in which streets and routes are considered as edges of a planar graph, while the traffic end points and street junctions are treated as nodes. The usual city map based on Euclidean geometry can be considered as an example of primary city graphs.

However, another graph representation can be useful if we are interested in describing the transport network at equilibrium. Given a connected undirected graph $G(V,E)$, in which $V$ is the set of nodes and $E$ is the set of edges, we introduce the traffic function $f : E \to (0, \infty]$ through every edge $e \in E$. It then follows from the Perron-Frobenius theorem [4] that
the linear equation
\[ f(e) = \sum_{e' \sim e} f(e') \exp(-h \ell(e')) , \] (1)
where the sum is taken over all edges \( e' \in E \) which have a common node with \( e \), has a unique positive solution \( f(e) > 0 \), for every edge \( e \in E \), for a fixed positive constant \( h > 0 \) and a chosen set of positive metric length distances \( \ell(e) > 0 \). This solution is naturally identified with the traffic equilibrium state of the transport network defined on \( G \), in which the permeability of edges depends upon their lengths. The parameter \( h \) is called the volume entropy of the graph \( G \), while the volume of \( G \) is defined to be the exponential growth of the balls in a universal covering tree of \( G \) with the lifted metric, \[ \text{Vol}(G) = \frac{1}{2} \sum_{e \in E} \ell(e) . \]

The degree of a node \( i \in V \) is the number of its neighbors in \( G \), \( \deg_G(i) = k_i \). It has been shown that among all undirected connected graphs of normalized volume, \( \text{Vol}(G) = 1 \), which are not cycles and for which \( k_i \neq 1 \) for all nodes, the minimal value of the volume entropy, \( \min(h) = \frac{1}{2} \sum_{i \in V} k_i \log(k_i - 1) \) is attained for the length distances
\[ \ell(e) = \frac{\log((k_i - 1)(k_j - 1))}{2 \min(h)} , \] (2)
where \( k_i \) and \( k_j \) are the degrees of the nodes linked by \( e \in E \). It is then obvious that substituting (2) and \( \min(h) \) into (1) the operator \( \exp(-h \ell(e')) \) is given by a symmetric Markov transition operator,
\[ f(e) = \sum_{e' \sim e} f(e') \frac{f(e')}{\sqrt{(k_i-1)(k_j-1)}} \] (3)
where \( i \) and \( j \) are the nodes linked by \( e' \in E \), and the sum in (3) is taken over all edges \( e' \in E \) which share a node with \( e \). The symmetric operator (3) rather describes time reversible random walks over edges rather than over nodes. In other words, we are invited to consider random walks described by the symmetric operator defined on the dual graph \( G^* \).

The Markov process (3) represents the conservation of the traffic volume through the transport network, while other solutions of (1) (with \( h > \min(h) \)) are related to the possible termination of travels along edges. If we denote the number of neighbor edges the edge \( e \in E \) has in the dual graph \( G^* \) as \( q_e = \deg_{G^*}(e) \), then the simple substitution shows that \( w(e) = \sqrt{q_e} \) defines an eigenvector of the symmetric Markov transition operator defined over the edges \( E \) with eigenvalue 1. This eigenvector is positive and being properly normalized determines the relative traffic volume through \( e \in E \) at equilibrium. Eq.(3) relates the equilibrium transport flows on the graph \( G \) to the stationary distribution of random walks defined on its dual counterpart \( G^* \) and emphasizes that the degrees of nodes are a key determinant of the transport networks properties.

The notion of traffic equilibrium had been introduced by J.G. Wardrop in [9] and then generalized in [10] to a fundamental concept of network equilibrium. Wardrop’s traffic equilibrium is strongly tied to the human apprehension of space since it is required that all travellers have enough knowledge of the transport network they use. Dual city graphs are extensively investigated within the concept of space syntax, a theory developed in the late 1970s, that seeks to reveal the effect of spatial configurations on the human perception of places and behavior in urban environments, [11, 3]. Spatial perception that shapes peoples understanding of how a place is organized determines eventually the pattern of local movement, which is quantified by the space syntax measure being nothing else, but an element of a transition probability matrix of a Markov chain [12], with surprising accuracy [13]. Random walks embed connected undirected graphs into the Euclidean space \( \mathbb{R}^{N-1} \). This embedding can be used in order to compare nodes with respect to the quality of paths they provide for random walkers and to construct the optimal coarse-graining representations.

While analyzing a graph, whether it is primary or dual, we assign the absolute scores to all nodes based on their properties with respect to a transport process defined on that. Indeed, the nodes of \( G(V,E) \) can be weighted with respect to any measure \( m = \sum_{i \in V} m_i 1_i \), specified by a set of positive numbers \( m_i > 0 \). The space \( L^2(m) \) of square-assumable functions with respect to the measure \( m \) is the Hilbert space \( \mathcal{H}(V) \). Among all measures which
can be defined on \( V \), the set of normalized measures (or densities), \( 1 = \sum_{i \in V} \pi_i \mathbf{1}_i \), are of essential interest since they express the conservation of a quantity, and therefore may be relevant to a physical process.

The fundamental physical process defined on a graph is generated by the subset of its linear automorphisms which share the property of probability conservation; it can be naturally interpreted as random walks. The linear automorphisms of the graph are specified by the symmetric group \( S_N \) including all admissible permutations \( p \in S_N \) taking \( i \in V \) to \( p(i) \in V \) and preserving all of its structure.

Markov’s operators on Hilbert space form the natural language of transport networks theory. Being defined on a connected aperiodic graph, the transition matrix of random walks \( T_{ij} \) is a real positive stochastic matrix, and therefore, in accordance to the Perron-Frobenius theorem [4], its maximal eigenvalue is \( 1 \), and it is simple. The left eigenvector \( \pi^T = \pi \) associated with the eigenvalue \( 1 \) is interpreted as a unique stationary state \( \pi \) (the stationary distribution of random walks). The orthonormal ordered set of real eigenvectors \( \{ \psi_i \}_{i=1}^N \) belonging to the eigenvalues \( 1 > \mu_2 \geq \ldots \mu_N \geq -1 \) describe the *global* connectedness of the graph. For example, the eigenvector corresponding to the second eigenvalue \( \mu_2 \) is used to define the spectral bisection of graphs; it is called the Fiedler vector if related to the Laplacian matrix of a graph [14].

Markov’s symmetric transition operator \( \hat{T} \) defines a projection any density \( \sigma \in \mathcal{H}(V) \) on the eigenvector \( \psi_1 \) of the stationary distribution \( \pi \),

\[
\sigma \hat{T} = \psi_1 + \sigma_\perp \hat{T}, \quad \sigma_\perp = \sigma - \psi_1, \tag{6}
\]

in which \( \sigma_\perp \) is the vector belonging to the orthogonal complement of \( \psi_1 \). In space syntax, we are interested in a comparison between the densities with respect to random walks defined on the graph \( G \). Since all components \( \psi_1,i > 0 \), it is convenient to rescale the density \( \sigma \) by dividing its components by the components of \( \psi_1 \),

\[
\tilde{\sigma}_i = \frac{\sigma_i}{\psi_1,i}. \tag{7}
\]

Thus, it is clear that any two rescaled densities \( \tilde{\sigma}, \tilde{\rho} \in \mathcal{H}(V) \) differ with respect to random walks only by their dynamical components, \( (\tilde{\sigma} - \tilde{\rho}) \hat{T}^t = (\tilde{\sigma}_\perp - \tilde{\rho}_\perp) \hat{T}^t \), for all \( t > 0 \). Therefore, we can define the distance \( ||\ldots||_\mathcal{F} \) between any two densities established by random walks by

\[
||\sigma - \rho||_\mathcal{F}^2 = \sum_{t \geq 0} \langle \tilde{\sigma}_\perp - \tilde{\rho}_\perp | \hat{T}^t | \tilde{\sigma}_\perp - \tilde{\rho}_\perp \rangle, \tag{8}
\]

or, using the spectral representation of \( \hat{T} \),

\[
||\sigma - \rho||_\mathcal{F}^2 = \sum_{s=2}^N \frac{\langle \tilde{\sigma}_\perp - \tilde{\rho}_\perp | \psi_s \rangle \langle \psi_s | \tilde{\sigma}_\perp - \tilde{\rho}_\perp \rangle}{1 - \mu_s}, \tag{9}
\]

where we have used Dirac’s bra-ket notations especially convenient for working with inner products and rank-one operators in Hilbert space.

If we introduce a new inner product for densities \( \sigma, \rho \in \mathcal{H}(V) \) by

\[
(\sigma, \rho)_T = \sum_{t \geq 0} \sum_{s=2}^N \langle \tilde{\sigma}_\perp | \psi_s \rangle \langle \psi_s | \tilde{\rho}_\perp \rangle \frac{1}{1 - \mu_s}, \tag{10}
\]
then \( (9) \) is nothing else but \( \| \sigma - \rho \|^2_T = \| \sigma \|^2_T + \| \rho \|^2_T - 2 (\sigma, \rho)_T \), where

\[
\| \sigma \|^2_T = \sum_{s=2}^{N} \frac{\langle \bar{\sigma}_s | \psi_s \rangle \langle \psi_s | \bar{\sigma}_s \rangle}{1 - \mu_s},
\]

is the square of the norm of \( \sigma \in \mathcal{H}(V) \) with respect to random walks defined on the graph \( G \).

We finish the description of the Euclidean space structure of \( G \) induced by random walks by mentioning that given two densities \( \sigma, \rho \in \mathcal{H}(V) \), the angle between them can be introduced in the standard way,

\[
\cos \angle (\rho, \sigma) = \frac{(\sigma, \rho)_T}{\| \sigma \|_T \| \rho \|_T},
\]

The cosine of an angle calculated in accordance to \( (12) \) has the structure of Pearson’s coefficient of linear correlations. The notion of angle between any two nodes of the graph arises naturally as soon as we become interested in the strength and direction of a linear relationship between the flows of random walks moving through them. If the cosine of an angle \( (12) \) is 1 (zero angles), there is an increasing linear relationship between the flows of random walks through both nodes. Otherwise, if it is close to -1 (\( \pi \) angle), there is a decreasing linear relationship. The correlation is 0 (\( \pi/2 \) angle) if the variables are linearly independent. It is important to mention that as usual the correlation between nodes does not necessarily imply a direct causal relationship (an immediate connection) between them.

In order to illustrate the approach, we have studied five different patterns of compact urban transport networks. Two of them are situated on islands: the street network in Manhattan and the network of Venetian canals. We have also considered two medieval German cities developed within the fortresses: Rothenburg ob der Tauber in Bavaria and the downtown of Bielefeld in Eastern Westphalia. To supplement the study of urban canal networks, we have also examined it in the city of Amsterdam. In all analyzed transport networks, we observe that the segregation of a node measured by \( (11) \) scales negatively with its connectivity; the slopes of the regression lines slightly exceed 2.

In the present Letter, we describe the Euclidean space structure of space syntax for the spatial network of 96 Venetian canals which serve the function of roads in the ancient city that stretches across 122 small islands. While identifying a canal over the plurality of water routes on the city map of Venice, the canal-named approach has been used, in which two different arcs of the city canal network were assigned to the same identification number provided they have the same name.

In accordance to \( (11) \), the density \( 1_i \), which equals 1 at \( i \in V \) and zero otherwise, acquires the norm \( \| 1_i \|_T \) associated to random walks. Its square,

\[
\| 1_i \|_T^2 = \frac{1}{\pi_i} \sum_{s=2}^{N} \frac{\psi_{s,i}^2}{1 - \mu_s},
\]

expresses the access time to a target node in random walk theory \([14]\) quantifying the expected number of steps required for a random walker to reach the node \( i \in V \) starting from an arbitrary node chosen randomly among all other nodes with respect to the stationary distribution \( \pi \).

The notion of spatial segregation acquires a statistical interpretation with respect to random walks by means of \( (13) \). In urban spatial networks encoded by their dual graphs, the access times \( (13) \) strongly vary from one open space to another and could be very large for statistically segregated spaces. It is remarkable that the norm a canal of Venice acquires with respect to random walks scales with its connectivity (see Fig. 1).

The Euclidean distance between \( 1_i \) and \( 1_j \) induced by random walks,

\[
\| 1_i - 1_j \|^2_T = \sum_{s=2}^{N} \frac{1}{1 - \mu_s} \left( \frac{\psi_{s,i}}{\sqrt{\pi_i}} - \frac{\psi_{s,j}}{\sqrt{\pi_j}} \right)^2,
\]

is the commute time in theory of random walks being equal to the expected number of steps required for a random walker starting at \( i \in V \) to visit \( j \in V \) and then to return back to \( i \). \([14]\).

Indeed, the structure of vector space \( \mathbb{R}^{N-1} \) induced by random walks cannot be represented visually, however if we choose a node of the graph as a point of reference, we can draw the 2-dimensional projection
The scatter plot of the connectivity vs. the norm a node in the dual graph representation of 96 Venetian canals acquires with respect to random walks. Three data points characterized by the shortest access times represent the main water routes of Venice: the Lagoon of Venice, the Giudecca canal, and the Grand canal. Four data points of the worst accessibility are for the canal sub-network of Venetian Ghetto. The slope of the regression line equals 2.07.

The 2-dimensional projection of the Euclidean space of Venetian canals set up by random walks drawn for the Grand Canal of Venice (the point (0, 0)) is shown in Fig. 2. Nodes of the dual graph representation of the canal network in Venice are shown by disks with radiuses taken equal to the degrees of the nodes. All distances between the chosen origin and other nodes of the graph have been calculated in accordance to (14) and (12) has been used in order to compute angles between nodes. Canals positively correlated with the Grand Canal of Venice are set under negative angles (below the horizontal), and under positive angles (above the horizontal) if otherwise.

The radiuses of disks display the equilibrium configuration of flows along the Venetian canals when the traffic volume is conserved. It is evident from Fig. 2 that disks of smaller radiuses demonstrate a clear tendency to be located far away from the origin being characterized by the excessively long commute times with the reference point (the Grand canal of Venice), while the large disks which stand in Fig. 2 for the main water routes are settled in the closest proximity to the origin that intends an immediate access to them.

Probably, the most important conclusion of space syntax theory is that the adequate level of the positive relationship between the connectivity of city spaces and their integration property (vs. segregation) called intelligibility encourages peoples way-finding abilities [3]. Intelligibility of Venetian canal network reveals itself quantitatively in the scaling of the norms of nodes with connectivity shown in Fig. 1 and qualitatively in the tendency of smaller disks to be located on the outskirts of the Venetian space syntax displayed in Fig. 2.

The support from the Volkswagen Foundation (Germany) in the framework of the project "Network formation rules, random set graphs and generalized epidemic processes" is gratefully acknowledged.
References

[1] D. Braess, "Über ein Paradoxon aus der Verkehrsplanung". Unternehmensforschung 12, 258-268 (1968).

[2] G. Kolata, "What if they closed the 42nd Street and nobody noticed?", The New York Times, Dec. 25 (1990).

[3] B. Hillier, Space is the Machine: A Configurational Theory of Architecture. Cambridge University Press. ISBN 0-521-64528-X (1999).

[4] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, 1990 (chapter 8).

[5] A. Manning, Ann. of Math. 110, 567-573 (1979).

[6] M. Bourdon, L’Enseign. Math., 41 (2), 63-102 (1995) (in French).

[7] T. Roblin, Ann. Inst. Fourier (Grenoble) 52, 145-151 (2002) (in French).

[8] S. Lim, Minimal Volume Entropy on Graphs. Preprint arXiv:math.GR/050621, (2005).

[9] J.G. Wardrop, Proc. of the Institution of Civil Engineers 1 (2), pp. 325-362 (1952).

[10] M.J. Beckmann, C.B. McGuire, C.B. Winsten, Studies in the Economics of Transportation. Yale University Press, New Haven, Connecticut (1956).

[11] B. Hillier, J. Hanson, The Social Logic of Space. Cambridge University Press. ISBN 0-521-36784-0 (1984).

[12] M. Batty, A New Theory of Space Syntax, UCL Centre For Advanced Spatial Analysis Publications, CASA Working Paper 75 (2004).

[13] A. Penn, Space Syntax and Spatial Cognition. Or, why the axial line? In: Peponis, J. and Wineman, J. and Bafna, S., (eds). Proc. of the Space Syntax 3rd International Symposium, Georgia Institute of Technology, Atlanta, May 7-11 2001.

[14] L. Lovász, Bolyai Society Mathematical Studies 2: Combinatorics, Paul Erdös is Eighty, Keszthely (Hungary), p. 1-46 (1993).

[15] F. Chung, Lecture notes on spectral graph theory, AMS Publications Providence (1997).