Analysis of urban complex networks

D.Volchenkov*
Bielefeld Bonn Stochastic Research Center (BiBoS), University of Bielefeld, Postfach 100131, D–33501, Bielefeld, Germany

Received January 31, 2008

We analyze the dual graph representation of urban textures by the methods of complex network theory and spectral graph theory. We present the empirical diagrams of distributions of the nearest and far-away neighbors in the several European compact urban patterns and the spectra of normalized Laplace operator defined on their dual graphs.

Key words: complex networks, urban studies
PACS: 89.75.Fb, 89.75.-k, 89.90.+n

1. Introduction to city networks studies

An inexorable worldwide trend toward urbanization presents an urgent need in developing a quantitative theory of urban organization and sustainable development. Town planning is not just a practice of manipulating the space, but rather an art with a far-reaching influence on society. Many aspects of urban networks substantially differ from other complex networks and call for an alternative type of analysis.

In most investigations, the relations between different components of an urban structure are frequently measured along the streets and routes considered as edges of a planar graph, while the traffic ultimate destination points and street junctions are treated as nodes. The *primary* graph representation of an urban network is based on relations between junctions through their streets. Being embedded into the geographical and economical landscapes, the resulting planar graphs reveal their multiple fingerprints. Among the main factors featuring the structure of graphs there are high costs of maintenance of long range connections and a scarce availability of physical space. In contrast, the *dual* graph representation of an urban network traces the relations between the streets through their junctions; the streets are treated as nodes, while junctions between them are represented by edges. They have been extensively studied within the concept of *space syntax*, an architectural theory developed in the late 1970s that seeks to reveal the mutual effects of complex spatial systems on society and vice versa, [1]. Its basic suggestion is that a spatial configuration (the system of spaces) is a driving force for human activity and cognition within urban environments, [2].

Our paper is devoted to the analysis of morphology and complexity of urban networks in their graph representations. We study several compact urban patterns: two medieval German cities (*Bielefeld* in Westfalia and *Rothenburg ob der Tauber* in Bavaria), the patterns of city canals in *Venice* and in *Amsterdam* playing the role of transport networks, and the modern urban development of *Manhattan*, a borough of New York City, planned in grid. A part of the results reported in the presented paper was published earlier in [23] – we reproduce them here for clarity, while the issues of general connectivity and portraits of the cities as complex networks have never been discussed before. The fact that a connected undirected graph can be embedded into \((N – 1)\) dimensional Euclidean space worked out in section 5 is also a novelty.

*E-mail: VOLCHENK@Physik.Uni-Bielefeld.DE
2. Complex networks

The paper written by Leonard Euler on the Seven Bridges of Königsberg published in 1736 is regarded as the first paper in the history of graph theory, [3]. Ever since, urbanistics has been a permanent source of inspiration for graph theory researchers.

Nowadays, cities contain many networks such as the municipal transport network, power grids and water supply which are entirely complex.

Complex networks appear as a result of the interplay between topology determined by a connected graph $G(V, E)$, dynamics described by a set of linear operators $\mathfrak{A}$, and the embedding space specified by $\mathfrak{M}$, a set of measures defined on the set of vertices $V$ and the weights assigned to edges $E$.

In the context of complex network theory created by physicists [4], the non-trivial topological structure of large networks is investigated by means of various statistical distributions. The structure and properties of complex networks essentially depend on the way the nodes get connected to each other. In the majority of complex networks emerging in society and in technology, each node has a certain feature attracting others. Although the investigations into the statistical properties of graphs could uncover their hierarchical structure, they will be futile if a detailed information on the structure of graphs is of primary interest since many graphs, characterized by the similar statistics of node degrees and the shortest path lengths, can be of dramatically different structures.

In mathematics, the automorphism groups of a graph are studied. They characterize its symmetries, and turn out to be very useful in determining its properties. The sets of graph automorphisms define a permutation group of vertices $S_N$. A transitive permutation group may be represented graphically, and the converse is also true: a graph gives rise to a permutation group. It is important to note that some graphs may constitute a Euclidean topological space with the natural topology on $E(N – 1)$ induced by the metric defined on $G$.

In particular, the Euclidean metric related to dynamics can be defined on some graphs by means of linear operators $\mathfrak{A}$ that remain invariant under permutations of nodes and satisfy some supplementary conservation properties. These operators describe certain dynamical processes defined on graphs such as random walks and diffusions.

Recent developments in random walks theory on graphs [5–7] and in spectral graph theory [8] are based on the use of self-adjoint operators $A \in \mathfrak{A}$ with respect to the measure $\mu = \sum_{j \in V} \text{deg}(j) \delta(j)$. Properties of graphs can be studied in relationship to the spectral properties of self-adjoint operators defined on them.

3. Space syntax theory

Studies of spatial configuration of urban textures are instrumental in predicting human behavior such as pedestrian and vehicle movement through urban environments [13]. The set of theories and techniques for the spatial analysis is called space syntax theory [2]. Space syntax is established on a quite sophisticated speculation that the evolution of the built form can be explained in analogy to the way biological forms unravel [14]. There has been developed as a method for analyzing space in an urban environment capturing its quality as being comprehensible and easily navigable [13]. Although, in its initial form, space syntax was focussed mainly on patterns of pedestrian movement in cities, later the various space syntax measures of urban configuration were found to be correlated with the different aspects of social life, [15]. Space syntax approach has subsequently been extended to a number of other aspects, such as modelling urban traffic, predicting air pollution levels, assessing the occurrence of burglaries in different neighborhoods, and estimating the potential for retail development in streets. Over the past decade, space syntax techniques have been used to make a research in the fields as diverse as archaeology, information technology, urban and human geography, and anthropology.

Dual graph representations have recently been used in the studies devoted to the information city network [16] and in the analysis of the public transport networks [17].

The order of street layouts in an urban environment can be retrieved directly from a city map.
and is immediately available for perception consisting of similar geometrical elements in repetition. In contrast to it, the structure of urban space is experienced through moving about a spatial system [9] and can be perceived as a sequence of navigation instructions. Every sequence of this kind describes a single path, perceived in real time, and contributes to the creation of a mental map reflecting an imaginable geometry of the city. The entire mental map amounts to a compilation of all navigation sequences which could be used while travelling through the city into a single, dual graph that in general is substantially distinct from the actual city map.

In an urban environment, dual graphs formalize the relations between open spaces which are treated as nodes.

Relations between the primary graph representation of an urban network and its dual counterpart have been studied in [10] in detail. It is worth mentioning that the use of a dual representation has never been discussed in the analysis of complex networks so far.

Open spaces are all interconnected, so that one can travel within open spaces from everywhere to everywhere else. One of the basic conceptions of space organization that has been discussed in space syntax is an axial space formed by all axial lines [11], the fewest and longest lines of sight that pass through every open space comprising any system. There is obviously an infinite number of these lines. Axial maps [12], a particular one constructed for the city of Rothenburg ob der Tauber (see figure 1) is shown in figure 2.

In figure 3, we have presented the non-planar dual city graph for the axial map of the city of Rothenburg (figure 2). The transition to a dual graph is a topologically non-trivial transformation of a planar primary graph into a non-planar one which encapsulates the hierarchy and structure of the urban area.

The dual transformation replaces the 1D open segments (streets) with the zero-dimensional...
nodes. Geometrical elements found in repetition and establishing the order in the urban developments are converted into the twins nodes, the pairs of nodes such that any other is adjacent either to them both or to neither of them.

4. Degree statistics and generalized connectivity in compact cities

The degree distribution, the probability that a node selected at random among all nodes of the graph has exactly \( k \) links,

\[
P(k) = \Pr \{ i \in G \mid \deg(i) = k \},
\]

has become an important concept in complex network theory describing the topology of complex networks. It originates from the study of random graph by Erdős and Rényi, [19]. The importance of degree distributions in networks of all kinds was recognized long time ago, [18]. It is used for the purpose of classifying graphs into categories, such as random graphs, for which \( P(k) \) asymptotically follows a Poisson distribution [20]. An alternative way of presenting the degree data is to compute the cumulative degree distribution, [4],

\[
\mathcal{P}(k) = \sum_{k' = k}^{N} P(k'),
\]

which is the probability that the degree is greater than or equal to \( k \). The presentation of degree data by the cumulative degree distribution has an advantage over the simple degree distribution, since it reduces the noise in the distribution tail.

The results on the probability degree statistics for the dual graphs of compact urban patterns analyzed in accordance with the above described street identification principle have been reported in [21,23]. In general, compact city patterns do not provide us with sufficient data to conclude on the universality of degree statistics.

![Figure 4](image.png)

**Figure 4.** The probability degree statistics for the dual graph of Manhattan (the log-log plot). \( NP(k) \) (circles) represent the numbers of streets in Manhattan having precisely \( k \) junctions and the solid line the cumulative distribution (2).

To give an example, in figure 4, we display the log-log plot of the numbers of streets \( NP(k) \) via the number of their junctions \( k \) in Manhattan. These numbers are shown by points, and the solid line is associated with cumulative distribution, \( N\mathcal{P}(k) \).

It is remarkable that the probability degree distributions for the dual graphs correspondent to the compact city patterns are broad and have a clearly expressed maximum and a long right tail which could decay faster than power law. It has been noted [22] that irregular shapes and faster decays in the tails could characterize the scale-dependent distributions. In the compact urban patterns, natural topographical or geographical factors can limit the path lengths in several cases,
such as on an island. In addition, long lines often cross peripheral or non-urban areas, where the number of potential connections is noticeably reduced. The presence of a noticeable maximum in the probability degree distributions indicates that there is the most probable number of junctions an average street has in a given city. The long right tails of distributions correspond to the highly connected nodes of dual graphs, just a few “broadways”, embankments, and belt roads crossing many more streets than an average street in the city, [23].

The statistics of distant neighbors can be presented in the form of cumulative distribution functions $P_d(n)$ quantifying the probabilities that the number of neighbors a node has at a topological distance $d$ is greater than or equal to $n$. Being monotonous functions of $n$, cumulative distributions reduce the noise in the distribution tails. However, the adjacent points on their plots are not statistically independent.

Usual networks constitute the graphs in which most nodes are neighbors of one another, and the number of far-away neighbors grows up fast with the distance apart from a node. The portraits of German medieval cities are the examples (figure 5).

![Figure 5](image1.png)

**Figure 5.** The cumulative distributions of far-away neighbors in the dual graphs of German medieval cities. The cumulative degree distribution are shown in the first row. The cumulative distributions of far-away neighbors are encoded in the second and forthcoming rows. Probability is ranked from zero (black, on the right) to 1 (grey, on the left).

![Figure 6](image2.png)

**Figure 6.** The cumulative distributions of far-away neighbors in the dual graphs of Venetian canal network and of the street grid in Manhattan. Probability is ranked from zero (black, on the right) to 1 (grey, on the left).

It is seen from the red profile in figure 5 that most streets in these organic cities are just by three syntactic steps away from each other. In contrast to them, the cumulative distributions of far-away neighbors in the dual city graphs displayed in figure 6 show that while for the street grid in Manhattan and for the canal network of Venice most nodes can still be reached in at most three syntactic steps, there are also a noticeable fraction of relatively segregated nodes few steps deeper than the others.
The structure of cumulative distributions shown in figure 6 evidences in favor of the corresponding dual graphs having high representation of cliques, and subgraphs that miss just a few edges of being cliques. In complex network theory, this phenomenon is referred to as a small world. Small-world networks are characterized by a high cluster coefficient having connections between almost any two nodes within them. Hubs - nodes in the network with a high number of connections serving as the common connections mediating the short path lengths between other edges – are commonly associated with small-world networks, [24].

The small world phenomenon indicated on the city portraits in figure 6 acquires an elegant explanation in the framework of space syntax theory.

In [13,25], it has been suggested that the spatial structure of organic cities is shaped by the public processes ordered in such a way as to maximize the presence of people in the central areas. In such a context, the compact structure of German burgs (figure 5) uncovers its historical functional pertinence.

The tendency to shorten syntactic distances in the fabric of urban space induced by the public processes is complemented in the “small world” cities (like those portrayed in figure 6) with the residential process which shapes the relations between inhabitants and strangers preserving the original residential culture against unsanctioned invasion of privacy. While most streets and canals characterized by an excellent accessibility promotes commercial activities and intensifies cultural exchanges, certain districts of such cities stipulate the alternative tendency having the residential areas relatively segregated from the rest of the urban fabric.

5. Undirected graphs and linear operators defined on them

The human mind thinks about the relations between things, agents, concepts and graphs [3] as a natural mathematical tool for facilitating the analysis.

A graph $G = (V, E)$ contains a set of vertices $V$ and a collection of edges $E$ that connect pairs of vertices. For each graph $G(V, E)$, there exists a unique, up to permutations of rows and columns, adjacency matrix $A_G$, the $N \times N$ matrix defined by $A_{Gij} = 1$ if there is an edge from $i \in V$ to $j \in V$, $i \sim j$, but $A_{Gij} = 0$ otherwise. Then, the degree of node $i$ is $\deg(i) = \sum_{j \in V} A_{Gij}$. If the graph is undirected, its adjacency matrix is symmetric and therefore has a complete set of real eigenvalues and an orthogonal basis of eigenvectors.

It can be proven that the Laplace operator is the only harmonic linear function,

$$\sum_{j \in V} L(A_G)_{ij} = 0, \quad \forall i \in V,$$

which remains invariant under permutations of vertices of the graph $G$, [26]. An important special case is the normalized Laplace operator [8],

$$L_{ij} = \delta_{ij} - \frac{A_{Gij}}{\sqrt{\deg(i) \deg(j)}},$$

(self-adjoint with respect to the measure

$$\mu = \sum_{j \in V} \deg(j) \delta(j)$$

associated with the random walk on undirected graphs. The matrix (4) is a symmetric, positive semi-definite matrix associated with the transition matrix of ordinary random walks, [8]. Eigenvalues of the normalized Laplace operator (4) satisfy $0 \leq \lambda_i \leq 2$.

The spectra of the normalized Laplace operator defined on the dual graphs of compact cities are presented in figure 7. The similarity of spectra allows us to divide these four compact urban patterns into two categories: i) the medieval cities: Bielefeld and Rothenburg o.d.T.; ii) the canal patterns: Venice and Amsterdam.
Analysis of urban complex networks

Figure 7. The spectra of the normalized Laplace operator defined on the dual graphs representing the route networks of four compact cities.

Random walkers are massless and do not interact with each other. Nevertheless, their flows can have nontrivial thermodynamical properties induced by the complex topology of the graph. The obvious advantage of statistical mechanics is that statistical moments of random walks would acquire a “physical interpretation” in the framework of this thermodynamic formalism [23].

The spectra of the relevant Laplace operators could have gaps indicating the time scales separation, that is, there are only a few “slow” time scales at which the transport network is meta-stable, with many “fast” modes describing the transient processes toward the slow modes.

The methods of spectral graph theory permit to detect and separate “slow” and “fast” time scales giving rise to the component analysis of networks in which the primary eigenvalues play an essential role. The statistical mechanics approach provides us with another efficient tool for the study of network components. Since the spectra of self-adjoint Laplace operators defined on the undirected graphs are non-negative they can be investigated by means of characteristic functions that discriminate contributions from the largest eigenvalues of Laplace operator in favor of those from the minimal ones.

All eigenvalues \( \lambda_k \) of the self-adjoint Laplace operator defined on the undirected graph \( G \) are positive and bounded [8]; they are characterized by the spectral density function,

\[
\rho(\lambda) = \sum_{k=1}^{N} \delta(\lambda - \lambda_k).
\]  

(6)

It is more convenient to study the properties of non-negative spectra by means of the characteristic functions.

The first characteristic function is given by the Laplace transform of the spectral density (6),

\[
Z(\beta) = \int_{0}^{\infty} e^{-\beta \lambda} \rho(\lambda) d\lambda, \quad \beta > 0,
\]

(7)

known as the canonical partition function\(^1\), the trace of the digitalized heat kernel:

\[
Z(\beta) = \sum_{k=1}^{N} e^{-\beta \lambda_k},
\]

(8)

that encodes the statistical properties of a system in thermodynamic equilibrium and converges to a continuous limit as \( N \to \infty \). The inverse temperature parameter \( \beta \) which can be considered.

\(^1\)In statistical mechanics, the partition function is a measure of the number of states accessible to the system at a given temperature \( \beta^{-1} \).
either as an effective time scale in the problem (the number of spaces in the dual graph a walker passes through in one time step) or as the laziness parameter.

Statistical mechanics provides a framework for relating the microscopic properties of a system to its macroscopic or bulk properties. In particular, it can be used to calculate the thermodynamic properties of the system from its spectrum [23].

Alternatively, if we require that the permutation invariant linear function should be stochastic, i.e.

$$\sum_{j \in V} T(A_G)_{ij} = 1, \quad \forall i \in V,$$

then the only possibility takes the form of generalized random walks,

$$T_{ij} = (1 - \beta_i) \delta_{ij} + \beta_i \left( D^{-1} A_G \right)_{ij},$$

in which $$\{\beta_i\}_{i=1}^N$$ is a set of arbitrary parameters and where $$D = \text{diag}(\{\deg(i)\}, \ i \in V), [27].$$

Random walks turn a connected undirected graph into the $$(N - 1)$$-dimensional Euclidean space. In order to estimate the mutual accessibility between any two nodes $$i, j \in G$$ in a finite undirected graph, we can use the statistical moments of random walks. The first statistical moment $$H_{ij}^{(1)}$$, the access time$^2$ (or the first hit time), has been extensively studied in random walks theory [5],[6]. Given the usual random walks defined by the transition operator, the matrix of access times $$H_{ij}^{(1)}$$ is computed for each pair of nodes $$(i, j)$$ by

$$H_{ij}^{(1)} = 2M \sum_{s=2}^{N} \frac{1}{1 - \mu_s} \left( \frac{\varphi_{sj}}{\deg(j)} - \frac{\varphi_{si}\varphi_{sj}}{\sqrt{\deg(i)\deg(j)}} \right),$$

in which $$\mu_1 > \mu_2 \geq \ldots \geq \mu_N \geq -1$$ are the eigenvalues and $$\varphi_s$$ are the corresponding eigenvectors of the symmetric transition matrix (10) for $$\forall \beta_i = 1.$$ The sums of reciprocal first access times, $$K_{ij} = H_{ij}^{(1)} + H_{ji}^{(1)}$$

are called the commute times. They give the expected number of steps a random walk starting at $$i$$ visits $$j$$ and then the node $$i$$ is reached again [5].

With the use of the spectral representation (11) we immediately obtain

$$K(i, j) = 2M \sum_{k=2}^{N} \frac{1}{1 - \mu_k} \left( \frac{\varphi_{ki}}{\sqrt{\deg(i)}} - \frac{\varphi_{kj}}{\sqrt{\deg(j)}} \right)^2,$$

where $$\mu_k$$ are the eigenvalues and $$\varphi_k$$ are the eigenvectors of the symmetric transition matrix of random walks. It is clear from (13) that the commute times $$K(i, j)$$ are nothing else than the distances in the $$(N - 1)$$-dimensional Euclidean space associated to random walks,

$$K(i, j) = \sum_{s=1}^{N-1} (\phi_{si} - \phi_{sj})^2 \equiv ||\vec{i} - \vec{j}||_{RW}^2,$$

in which

$$\phi_{ki} = \frac{\varphi_{k+1,i}}{\pi_i(1 - \mu_{k+1})}, \quad k = 1, \ldots, N - 1,$$

is the coordinates of vector $$\vec{i}, \ i \in G,$$ and $$\pi_i = \deg(i)/2M$$ is the stationary distribution of random walks on undirected graph.

The physical meaning of the Euclidean “norm” $$||\vec{i}||^2$$ of the vector $$\vec{i}$$ is the expected number of steps a random walker reaches the node $$i \in G$$ if started from a node randomly chosen among all the nodes of $$G$$ with respect to the stationary distribution $$\pi.$$

The Euclidean space representation of undirected graphs can be used in order to compare the accessibility properties of certain nodes and subgraphs.

$^2$The first time at which the stochastic process “hits” the node $$j$$ if started from $$i.$
6. Discussion and conclusion

We conclude that in contrast to spectral graph theory proper, in which the properties of a given graph are studied in relationship to the characteristic polynomial, eigenvalues, and eigenvectors of the graph adjacency matrix or the matrix of canonical Laplacian, we have applied the graph theory methods to dual graphs.

Although the dual transformation preserves connectivity, the main topological property of any graph, other structural and spectral properties can be dramatically different form those of the primary pre-images. In particular, the sizes of primary and dual graphs may be different; the number of streets in a city is not usually equal to the number of junctions. While the graph-theoretic distance counts the number of hops between any two nodes in a primary graph, the space syntax distance, the depth between two open spaces, defined on the dual counterpart accounts for the number of elementary navigation actions to be performed in order to travel between two open spaces in the city.

The spectral analysis of dual representations of transport networks provides us with the important information on their navigability and morphology. The dual transformation removes the network graph from the physical space in which it is embedded being represented by the primary graph and immerses it into the cognitive space, resulting in a unique categorization of all urban places with respect to their structural intelligibility.

Let us note that the same spectral graph theory methods can also be applied directly to a city map in order to analyze the primary graph representation of an urban transport network. In such a framework, the Euclidean distance between any two street junctions can be naturally interpreted as a weight of the edge of the primary graph. Random walks defined on such a weighted graph categorize every node of it in accordance with the level of its accessibility for random walkers.

However, there is a well-known criticism towards the interpretation of random walks flows along streets such as the transport flows of vehicles and pedestrians following down the same itineraries: although human behavior seems not predictable, it is also by no means random.

7. Acknowledgements

The author is grateful to the Volkswagen Foundation (Germany) supporting his research in complex networks. We thank Ph. Blanchard for the fruitful discussions and multiple comments on our work.

References


Аналіз урбаністичних складних мереж

Д.Волченков
Дослідницький центр BiBoS, Університет Білефельда, Німеччина

Отримано 31 січня 2008 р.

Ми аналізуємо урбаністичні структури за допомогою дуально-графічного представлення і застосування методів теорії комплексних мереж та спектральної теорії графів. Ми представляємо емпіричні діаграми розподілів ближніх та віддалених сусідів у компактних урбаністичних архівах ряду європейських міст, а також спектри операторів Лапласа, означених на їх дуальних графах.

Ключові слова: складні мережі, урбаністичні дослідження

PACS: 89.75.Fb, 89.75.-k, 89.90.+p