
THE STRUCTURAL SIMILARITY OF NEIGHBOURHOODS IN URBAN STREET NETWORKS: a case of London

093

Hoon-Tae Park

The Bartlett School of Graduate Studies, UCL

Keywords:

Local analysis
Neighbourhoods
Structural similarity
Scaling

Hoon-Tae Park

*The Bartlett School of Graduate
Studies, 1-19 Torrington Place,
University College London (UCL)
Gower St. London WC1E
hoon-tae.park@ucl.ac.uk*

Abstract

This paper studies the structural similarity of neighbourhoods in urban street networks represented by axial maps. Here r -th neighbourhood of a node is defined as a subgraph induced by nodes encountered within radius r from the reference node. Through the case study of London, we identify three distinct types of scaling between radius and neighbourhood size: the power-law, the exponential-law, and the super-power. Individual nodes are grouped by a form of similarity represented by each of those scaling laws. First, the power-law with scaling dimension 2.73 is found to be the dominant type of scaling. Under the power-law, it is shown that the "urban correlation", redefined in this paper as the homogenisation of mean depth, should be the normal state of affairs in local analysis. Until recently, particularly in network theory where the completion of formalism matters, it is the presence of such homogeneity that has been paid most attention. However, we also observe that the other minor-types of scaling seem to introduce potentially important, life-like local variations underneath the apparent homogeneity. This findings will lead us to some new insights on the genuinely complex and heterogeneous structure of urban street networks we study. In parallel, we also report that the street network of London is neutral in terms of the degree-degree correlation, redefined in this paper as the homogenisation of control. Neutrality, or the absence of the degree-degree correlation, has been so far regarded mainly as the property of random networks. However, random networks would never exhibit the power-law scaling of neighbourhood size. From this we propose the formula of 'being neutral without being random' to define a phenomenon that is specifically 'urban.'

Introduction

Local analysis has been introduced in space syntax as a heuristic way of overcoming "edge effects" (Penn et al. 1998). Here edge effects refer to phenomena in which nodes close to the edge of a graphically mapped area are less central purely because of our selection of the boundary for analysis. Every global analysis is thus affected to some extent by edge effects. The problem will be resolved, ultimately and ideally, by making the analysis independent from the subjective procedure of drawing boundaries.

In a recent paper, Dalton (2005) has observed that local analysis leads to a kind of homogeneity in which nodes are all equally central as they have more or less the same mean depth. This observation, which he dubbed “urban correlation”, is repeatedly confirmed for almost any ‘axial’ maps analysed, but not for randomly generated maps; so specifically “urban” phenomenon it seems to be. It follows, as he shows, that each node is differentiated only by means of the size of the neighbourhood defined by a certain depth from it.

Then, we are entitled to assume that neighbourhood size may be redundant as it is aptly reducible to the degree (or connectivity) of nodes. That is to say, the neighbourhood of a node should be large if it is locally well connected to others. On what basis can we say that neighbourhood size gives us any other additional information than degree? Perhaps, is this reduction to degree what we should expect by eliminating edge effects? Is the knowledge of degree distribution the ultimate result of local analysis?

In this paper, we study the structural similarity of neighbourhoods in urban street networks. Through this, we aim to rediscover, independently of degree, the true differences of neighbourhoods beneath the apparent homogeneity, and derive a criterion that enables us to distinguish what is unique and thus irreducible in local analysis. These subject issues are explored systematically through the axial map of Inner London. We also discuss, in line with the new results from Hillier et al. (2007), some possible implications the results in this paper may have in understanding the structure and functionality of urban street networks.

Local Variables

First, we define the necessary basics on local analysis and fix the notation. Let G denote a simple undirected graph consisting of a set of N nodes and a set of E edges. Depth (or distance) $d = d(v, u)$ between a pair of nodes v and u in G is defined as the discrete length of geodesic path connecting them.

This path metric allows us to introduce a neighbourhood concept. The r th neighbourhood of v , denoted by $\Gamma_r(v)$, is a subgraph induced by nodes u with $d(v, u) \leq r$, and the boundary of the neighbourhood, denoted by $\partial\Gamma_r(v)$, is the set of nodes u having exactly $d(v, u) = r$. The boundary parameter r is called the ‘radius’ of the neighbourhood. Let us denote the orders of the r th neighbourhood of v and its boundary by $N_r(v) = |\Gamma_r(v)|$ and $n_r(v) = |\partial\Gamma_r(v)|$ respectively, hence formulating the cumulative relationship $\sum_{d=0}^r n_d(v) = N_r(v)$. By contrast, if the whole graph is accounted for, we will have $\sum_{d=0}^{e(v)} n_d(v) = N$, with $e(v)$ being the eccentricity of v , i.e. the depth to a node farthest from v . We note that the two definitions provide us with a first important distinction between local and global analyses: the global becomes the local through a trade-off in which N is localised into $N_r(v)$ while $e(v)$ is standardised into r .

The differential relationship between $N_r(v)$ and $n_r(v)$ is given by:

$$n_r(v) = N_r(v) - N_{r-1}(v) \quad (1)$$

with $n_0(v) = N_0(v) = 1$, corresponding to the reference node itself, for all v in G . We further note that,

$n_1(v) = N_1(v) - N_0(v) = N_1(v) - 1$ is equivalent to the degree (or connectivity) of v , i.e., the number of nearest neighbours of v .

The total depth, $D_r(v)$, of v is defined as:

$$D_r(v) = \sum_{d=1}^r d \cdot n_d(v) \quad (2)$$

If we substitute $n_r(v)$ in (1) into (2), the equation for total depth becomes:

$$\begin{aligned} D_r(v) &= \sum_{d=1}^r d \cdot \{N_d(v) - N_{d-1}(v)\} \\ &= N_1(v) - N_0(v) + 2N_2(v) - 2N_1(v) + \dots + rN_r(v) - rN_{r-1}(v) \\ &= rN_r(v) - N_0(v) - N_1(v) - \dots - N_{r-1}(v) \end{aligned}$$

and this simplifies to:

$$D_r(v) = rN_r(v) - \sum_{d=0}^{r-1} N_d(v) \quad (3)$$

Dividing $D_r(v)$ by $N_r(v)$, we have the mean depth, $\bar{d}_r(v)$, relative to v :

$$\bar{d}_r(v) = r - \sum_{d=0}^{r-1} \frac{N_d(v)}{N_r(v)} \quad (4)$$

from which it is obvious that $\bar{d}_r(v)$ must be strictly less than radius applied. Note also that this equation is true of both local and global analyses, since global mean depth is just $\bar{d}_{e(v)}(v)$.

Closeness centrality, $C_r(v)$, of v measures the extent to which the node is close to all other nodes in $\Gamma_r(v)$ and has been defined simply as the reciprocal of mean depth (Freeman 1979), that is:

$$C_r(v) = \frac{1}{\bar{d}_r(v)} \quad (5)$$

In space syntax, closeness centrality is modified to reflect the 'assumption' that mean depth tends to increase with neighbourhood size. The modified measure is called 'integration', $I_r(v)$, of v , which we can write in its simplistic form as:

$$I_r(v) = \frac{\log N_r(v)}{\bar{d}_r(v)} \quad (6)$$

where $\log N_r(v)$ has been introduced as a correction factor (Park 2005). The correction factor thus specifies that 'mean depth increases, in average, following the logarithm of neighbourhood size.' This scaling is popularly known as 'small-world' behaviour (Watts & Strogatz 1998).

Mean depth depends solely on the depth-based partition of the set of nodes and does not reflect the distribution of edges within the neighbourhoods. For this reason, we consider a different kind of local

measure in space syntax, called 'control' (Hillier & Hanson 1984), which takes into account a complete description of connectivity structure around v . The control, $ct(v)$, of v measures the extent to which the nearest neighbours u of v are accessible to each other only by way of v , given their degree distribution. It has been defined as:

$$ct(v) = \sum_{u \in \partial\Gamma_1(v)} \frac{1}{n_1(u)} \quad (7)$$

where $n_1(u)$ is the degree of the nearest neighbours of v .

093-04

Note first that the control is completely independent of the depth metric. But this original definition of control suffers from degree-degree correlation biases (Soffer & Vázquez 2004), and consequently tends to give values that are often trivially related to the degree of v . For instance, if a graph is organised in such a way that high-degree nodes are linked to nodes with low-degree, high-degree nodes will always have high control values. We can fix this problem by relativising the control against its possible maximum, given the degree of v . The maximum $ct_{\max} = n_1(v)$ will be achieved when $n_1(u) = 1$ for all u in $\partial\Gamma_1(v)$. Then, the proper definition of control, without the effects of degree correlation, is:

$$ct'(v) = \frac{ct(v)}{ct_{\max}} = \frac{1}{n_1(v)} \sum_{u \in \partial\Gamma_1(v)} \frac{1}{n_1(u)} \quad (8)$$

Note also that $ct'(v)$ in (8) is simply $1/H_{mn}(v)$, that is, the reciprocal of the 'harmonic mean' of the degree of the nearest neighbours of v . We can thus say that the higher average degree of the nearest neighbours of v , the less control it will exert over communication or movement among its neighbours. The control may be viewed just as a local version of what is usually called "betweenness centrality" at the global scale (Freeman 1979).

The Domain of Local Analysis

The aim of this section is to delimit the domain of local analysis where the urban correlation prevails and to clarify its implications for local analysis. As originally reported by Dalton (2005), the urban correlation refers to a phenomenon in which total depth scales linearly with neighbourhood size, i.e. $D_r(v) \sim N_r(v)$. This in turn implies that mean depth, $\bar{d}_r(v) = D_r(v)/N_r(v)$, remains homogeneous across an entire graph. Dalton has therefore maintained that local mean depth cannot be related to the diversity of urban phenomena observed in reality.

Consequently, with the urban correlation, closeness centrality becomes:

$$C_r(v) = \frac{1}{\bar{d}_r(v)} \sim \text{constant} \quad (9)$$

By contrast, for integration, we have:

$$I_r(v) = \frac{\log N_r(v)}{\bar{d}_r(v)} \sim \log N_r(v) \quad (10)$$

The equations (9) and (10) state that every node is equally central under the urban correlation, wherever it is located; but not equally integrated as it takes advantage of differences in neighbourhood size.

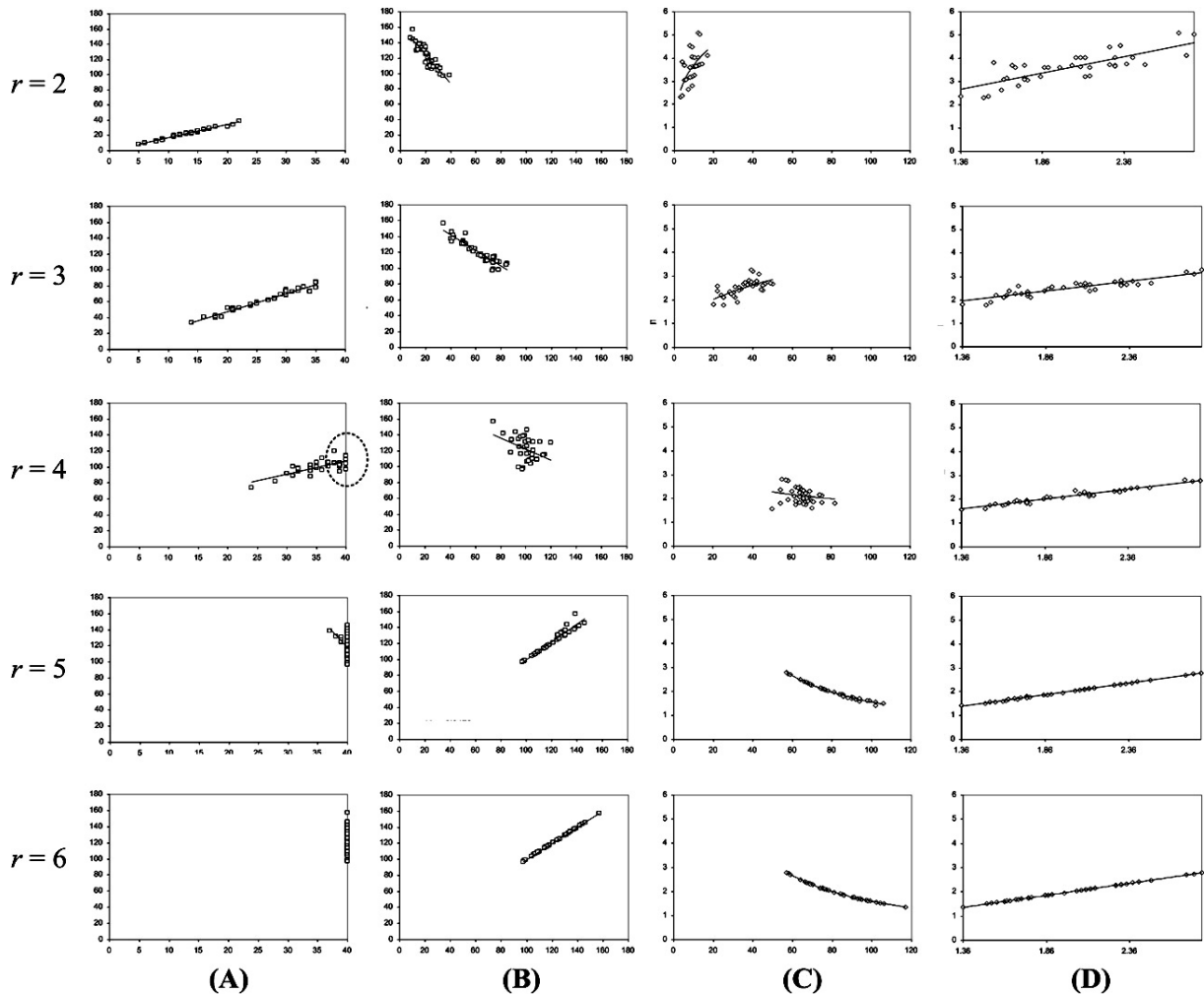
The difference and similarity of closeness centrality and integration can be now clarified. In the local, a node is integrated not because it has small mean depth but because it has large neighbourhood size relative to a given radius, so densely populated. In this case, $\log N_r(v)$ in the formula of integration does not function merely as a correction factor; but actually re-defines the measure, independently of mean depth, as that which must be conceptually distinguished from closeness centrality. In the global, however, every neighbourhood has the same order N , and consequently, integration becomes differentiated not by neighbourhood size but by mean depth alone. In this case, there must not be any conceptual differences between closeness centrality and integration.

To illustrate how integration changes its nature dynamically as radius increases, we take Gassin (Hillier & Hanson 1984) as an example (Fig. 1). Gassin is small enough to allow us to trace the dynamic process over the entire range of radii.

Figure 1:

Correlation matrix for Gassin. Column (A): neighbourhood size vs. total depth. Column (B): local total depth vs. global total depth. Column (C): total depth vs. integration. Column (D): local integration vs. global integration

093-05



(A) The strong 'positive' correlation between neighbourhood size and total depth is observed at the local scale ($r \leq 3$), namely, the urban correlation. The slope of the regression should approximate the common value of local mean depth. Yet, as radius increase globally, the urban correlation breaks down; hence the differentiation of global

093-06

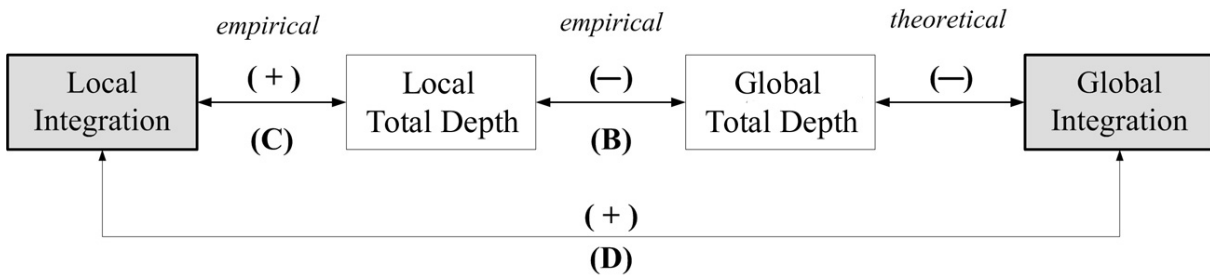
mean depth. The edge effect begins to take place from $r = 4$ (indicated by a dotted circle).

(B) The strong ‘negative’ correlation between local total depth and global total depth is also observed ($r \leq 3$). This suggests that total depth behaves more like degree at the local scale — recall that $D_1(v) = n_1(v)$. Yet there is a turning point, $r = 4$, beyond which this tendency becomes completely reversed. In other words, as radius increases globally, the regression rotates anticlockwise and eventually changes its sense; hence the differentiation of total depth from degree.

(C) There is a ‘positive’ correlation between total depth and integration. But the correlation seems relatively weak even when $r = 2$. It becomes weaker and then disappears completely at the turning point $r = 4$, beyond which begins to increase rapidly. The higher correlation found in the global is trivial as it is what we have predetermined as such through the definition of integration (or closeness centrality). But their difference in the local is worth bearing in mind. It is not the case that high integration signifies low total depth automatically.

(D) There is a consistent ‘positive’ correlation between local and global integration, which is often called ‘synergy’ in space syntax. But it would only conceal the “paradox of integration” unless we see the causal factor underneath the association. That is, such a consistent positive correlation results from double negation, as it were, which include the empirical results in (B) and (C), plus the predetermined or theoretical negative relation between global total depth and global integration (Fig. 2).

Figure 2:
Diagram explaining the apparent positive correlation, or ‘synergy’, between local and global integration



In effect, we can divide the entire range of radii into three subsections: (1) the ‘local’ where the urban correlation prevails; (2) the ‘non-local’ where the measures lose their consistent senses completely (which we must thus try to avoid in any practice of analysis); and (3) the ‘global’ where the senses of measures are completely reversed from those in the local.

The local can be then safely marked by so-called “radius-radius” in space syntax. Radius-radius, denoted by r_c , is a critical radius set to the mean depth of most ‘globally’ integrated node, that is, $r_c = \lfloor \min \bar{d}_{e(v)}(v) \rfloor$ (e.g. $r_c = 2$ in the case of Gassin). As Hillier (1996, p.163) remarks, “The effect of a radius-radius analysis is to maximise the globality of the analysis without inducing ‘edge effect’, [...]” Throughout this paper, we shall have the notion of radius-radius define the domain of local analysis, i.e. $1 \leq r \leq r_c$ for all v in G .”

Within the domain of local analysis, total depth is reduced to neighbourhood size and mean depth loses its discriminating power. However, beyond the domain, without the urban correlation, it is mean depth that becomes all powerful.

Structural Similarity

Scaling Dimension

Now we demonstrate that the urban correlation, or the homogenisation of mean depth in the local, can be observed if there exists any kind of scaling between radius and neighbourhood size. Just to illustrate: if $N_r(v)$ is a function of r , i.e. $N_r(v) = f(r)$, then, it follows from (4) that local mean depth becomes also a function of r only. Applying the same radius to all v in G will, therefore, entail the same local mean depth. The aim of the following discussion is then to specify a model function for $N_r(v)$ and to deduce its consequences in a manner appropriate to test it.

093-07

The two simple scaling laws shall be contrasted:

$$N'_r(v) = r^\alpha \quad (11)$$

and

$$N'_r(v) = e^{\beta(r-1)} \quad (12)$$

where $N'_r(v) = N_r(v) / N_1(v)$ denotes relative neighbourhood size with respect to the degree of v , and $\alpha > 0$ and $\beta > 0$ are parameters that need to be estimated empirically. Note that relative neighbourhood size has been introduced to make scaling functions independent of degree. Obviously, $N'_1(v) = 1$ for all v in G .

Taking the logarithms of the both sides of equations (11) and (12), we obtain:

$$\log N'_r(v) = \alpha \log r \quad (13)$$

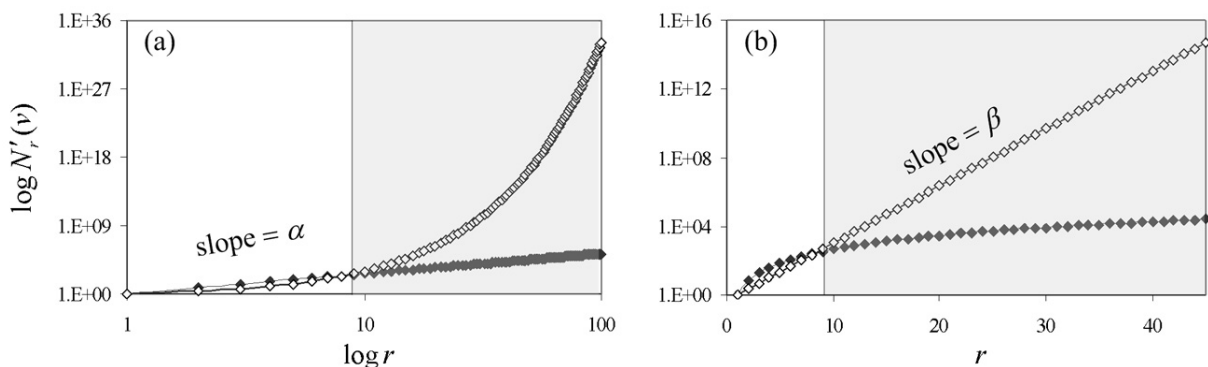
and

$$\log N'_r(v) = \beta(r-1) \quad (14)$$

Both scaling laws state that neighbourhood size must grow with radius; but, for small radii, more rapidly according to the 'power law' (11) than to the 'exponential law' (12). However, for large radii, the increasing rate of change of neighbourhood size is reversed so that it grows much more rapidly according to the exponential law than to the power law. If we thus plot $N'_r(v)$ against r , a straight line will emerge in doubly logarithmic scale for the power law (slope α), while only in semi-logarithmic scale for the exponential law (slope β) (Fig. 3).

Figure 3:

Two different forms of scaling between radius and relative neighbourhood size: the power-law (●) and the exponential-law (○), plotted in (a) double logarithmic plot and (b) semi logarithmic plot. Here we assume that the graph is infinitely large, so does not undergo any edge effects



The power law can be observed in 'scale-free' graphs, while the exponential law in 'random' or 'small-world' graphs (Csányi & Szendroi 2004). Theoretically, the power law is known to reflect that

the short-cut effects forming random or small-world graphs are significantly diminished by some strong 'spatial' constraints such as geographical embedding (ben-Avraham et al. 2003). This means that long-range connections leading to small-world behaviour cannot occur in a purely random fashion. By contrast, the exponential law is typically found for such 'trans-spatial' relationships as the WWW, the scientific collaboration network, the network of corporate elites and the like.

If the power law is observed, the parameter α in (11) shall define the 'scaling dimension' starting from v , i.e. $\alpha = \alpha(v)$. If α is relatively stable across all v , we will also say that G has the scaling dimension $\alpha = \alpha(G)$. Clearly, this definition is the discrete analogue of 'fractal dimension.' Nowotny and Requardt (1998) have remarked that such a concept is intrinsic as it is determined by the geometric or relational organisation of the graph itself rather than by the ambient space in which the graph is embedded. Furthermore, they showed the structural stability of scaling dimension under local perturbations. That is, no insertion of a finite number of additional edges between nodes could change the scaling dimension.

Now by substituting (11) into (4), we can show that $\bar{d}_r(v)$ is expressed under the power law as:

$$\bar{d}_r(v) \approx \frac{\alpha}{\alpha + 1} r + \varphi \quad (15)$$

with $\varphi = (1/2)(1 - \alpha/2r)$ which converges on 0.5 as r increasesⁱⁱⁱ. Alternatively, from $r = N'_r(v)^{1/\alpha}$ in (11), we obtain:

$$\bar{d}_r(v) \approx \frac{\alpha}{\alpha + 1} N'_r(v)^{1/\alpha} + \varphi \quad (16)$$

The equations (15, 16) clearly show that mean depth under the power-law scaling is linearly proportional to r or some power of $N'_r(v)$. On the one hand, if $\alpha = \alpha(G)$, it is obvious that $\bar{d}_r(v)$ becomes independent of v , which will lead us to an observation of the urban correlation. On the other, if $\alpha = \alpha(v)$, $\bar{d}_r(v)$ may vary from one node to one another. But we note that the equation (15) is rather insensitive to the small variation of α , so that the urban correlation can be still preserved to some significant level. For instance, as we will see, $\alpha(v)$ of urban street networks vary normally in the interval of [2, 3]. If this is the case, $\bar{d}_r(v)$ will vary in the remarkably narrow interval of [0.67r, 0.75r], approximately.

Next, the same procedure can be considered for the exponential law (12). If the exponential law is followed, $\bar{d}_r(v)$ becomes:

$$\bar{d}_r(v) = r - \frac{1}{e^{\beta r}} \cdot \frac{1 - e^{\beta r}}{1 - e^{\beta}} \quad (17)$$

Furthermore, since $r = (1/\beta) \log N'_r(v) + 1$ from (14), we get alternatively:

$$\bar{d}_r(v) = \frac{1}{\beta} \log N'_r(v) + \xi \quad (18)$$

where the substitution $\xi = 1 - (1/e^{\beta r})[(1 - e^{\beta r})/(1 - e^{\beta})] < 1$ has been made. We note that ξ converges to $(2 - e^{\beta})/(1 - e^{\beta})$ as r increases^{iv}.

From (17, 18), we can see that mean depth under the exponential law is also linearly proportional to r , as under the power law; but now to the logarithm of $N'_r(v)$. This implies that the urban correlation can be observed in principle for random or small-world graphs, and thus, against Dalton's conjecture, it may not be a specifically "urban" phenomenon after all. However, we point out that, in contrast to the stability of mean depth under the power-law, mean depth under the exponential law, as shown in (17, 18), is very sensitive to the individual variations of $\beta = \beta(v)$. This seems to suggest that the power-law scaling between radius and neighbourhood size is more likely to be the source of the observed urban correlation. To verify this claim is then what remains to be done.

Assortativity by Degree

In complex network theory, it has been recently found that the 'assortativity by degree', a preference for high-degree nodes to associate with other high-degree nodes, captures a fundamental feature of graph structure. In this subsection, we briefly introduce the concept, with a particular emphasis on its possible relationship to our measure of control.

A graph is called 'assortative' (or 'dissortative'), if high-degree nodes tend to be attached with nodes with high (or low) degrees; and 'neutral', if there is no such tendency. It is said that assortativity reflects well-defined "core-periphery" structure, while dissortativity highlights "hierarchical" structure.

Pastor-Satorras, Vazquez & Vespignani (2001) and Newman (2002) have observed that social networks are typically assortative, while technological and biological networks, such as Internet, WWW, protein interactions, neural network, are typically dissortative. In urban contexts, Porta, Crucitti & Latora (2004) have studied 1-square-mile samples of urban street networks to find out that those are mainly dissortative. Xulvi-Brunet and Sokolov (2005) have demonstrated that the characteristic path length of a graph having a fixed degree sequence increases both with assortativity and dissortativity, while achieves minimum when it is neutral. In other words, neutral structure seems to provide a best solution in optimising "natural movement" in a network.

There are several ways of measuring assortativity by degree, one of which is to calculate the degree-degree correlation. This is done by introducing a new quantity $\langle A_{nn}(k) \rangle$, which is the average degree of the nearest neighbours of v whose degree is k . Note the quantity is a function of degree, not of nodes, which can be written formally as conditional average:

$$\langle A_{nn}(k) \rangle = \frac{1}{|D_k|} \sum_{v \in D_k} A_{nn}(v) \quad (19)$$

where D_k is a set of nodes having degree k and $A_{nn}(v)$ is the 'arithmetic' mean of the degree of the nearest neighbours of v . Then, $\langle A_{nn}(k) \rangle$ will increase according to the power law $\langle A_{nn}(k) \rangle \sim k^\phi$ ($\phi > 0$) if graphs are assortative; decrease ($\phi < 0$) if they are

dissortative; and flat ($\phi = 0$) if neutral (Pastor-Satorras et al. 2001; Cantanzaro, Caldarelli & Pietronero 2004).

In this paper we replace the arithmetic mean $A_m(v)$ in (19) by the harmonic mean $H_m(v)$ in (8), to have a new measure for the degree-degree correlation, that is:

$$\langle H_m(k) \rangle = \frac{1}{|D_k|} \sum_{v \in D_k} H_m(v) = \frac{1}{|D_k|} \sum_{v \in D_k} \frac{1}{ct'(v)} \quad (20)$$

The replacement will not affect the sense of the degree-degree correlation, although it does not corroborate the stochastic model assumed in the references for the degree correlation^v. The advantage of using the harmonic mean, if any, is thus practical. In particular, if the degree sequence of G is 'scale-free', i.e. $|D_k| \sim k^{-\gamma}$ (Li, Alderson, Tanaka, Doyle & Willinger 2005), $H_m(v)$ will be much more efficient than $A_m(v)$, since the former has a strong tendency to mitigate the influence of high-degree nodes.

Now, suppose first that $H_m(v)$ is homogeneous across all v in G . Then, we will have $\langle H_m(k) \rangle$ that remains uncorrelated whatsoever with degree. The graph is neutral in the sense that it is neither of core-periphery nor hierarchical structure; but at least a mixture of the two. This will ensure, together with the homogenisation of mean depth, the structural similarity of neighbourhoods. On the other, if $\langle H_m(k) \rangle$ increases (or decreases) with degree, i.e. assortative (or dissortative), we cannot say that, in spite of the homogenisation of mean depth, neighbourhoods are structurally similar to one another.

Analysis: Inner London Case

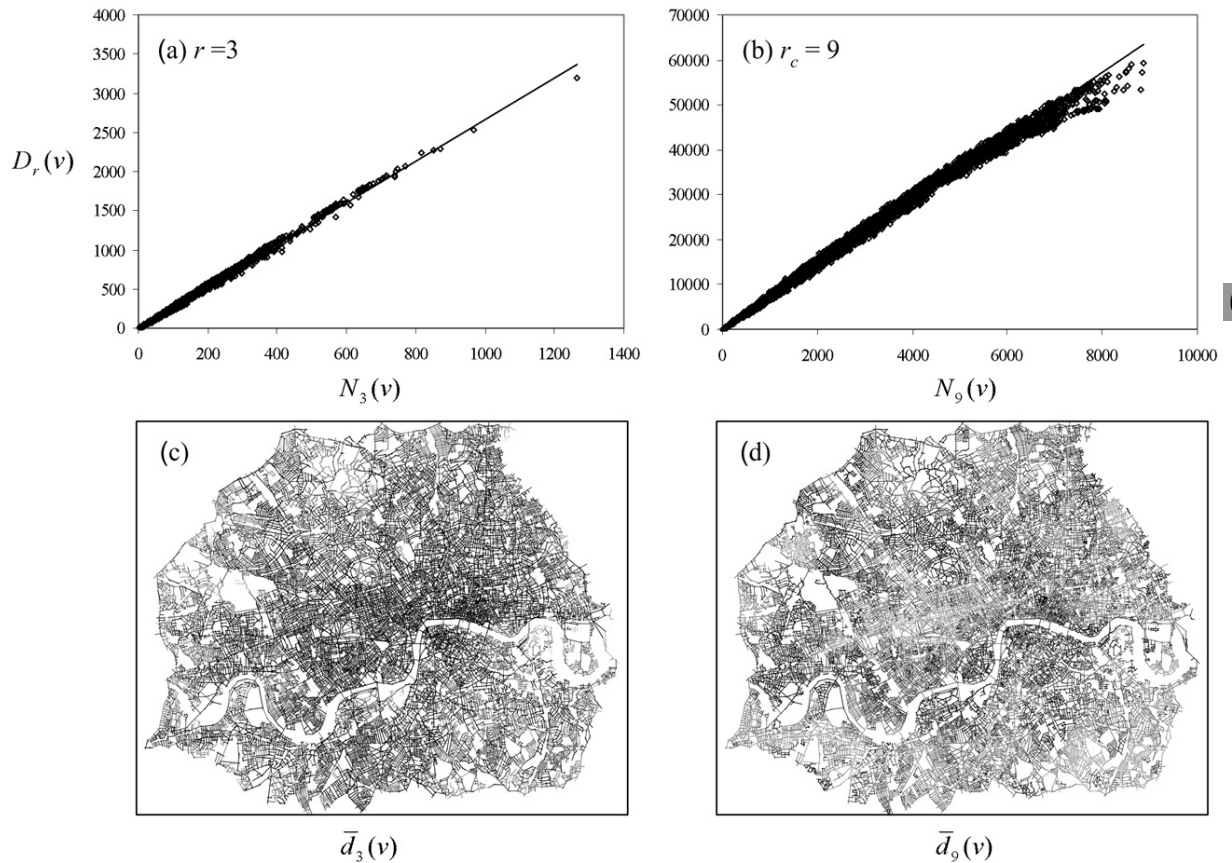
Now we are sufficiently prepared to apply what we have reviewed so far to the real case of Inner London ($N = 17,321$, $E = 36,452$). The radius-radius of Inner London is $r_c = 9$ (i.e. the mean depth of Oxford street). For reference only, the maximum radius for global analysis is $r_{\max} = 45$.

5.1 Scaling Dimension

We start from probing the urban correlation. Fig. 4(a) and (b) show the linear regression of $D_r(v)$ on $N_r(v)$ for (a) $r = 3$ and (b) $r_c = 9$.

The urban correlation clearly exists with (a) $R^2 = 0.997$ and (b) $R^2 = 0.994$. The values of mean depth remain accordingly homogeneous with (a) $\bar{d}_3(v) = 2.50 \pm 0.22$ and (b) $\bar{d}_9(v) = 7.27 \pm 0.33$. Relative variations diminish as radius increases, although we can see that the urban correlation already begins to break down at the larger radius due to the edge effects. Then, Fig. 4(c) and (d) show the spatial distribution of those minute variations in mean depth for (c) $r = 3$ and (d) $r_c = 9$ respectively. In the former, streets with higher local mean depth (i.e. locally less central) are located mainly in the global centre of the bounded area, while, in the latter, form a 'ring' around the global centre. But it is not entirely clear

what these spatial distributions reveal to us as for the structure of the network in question.



093-11

Having observed the urban correlation, we now try to identify its cause. Fig. 5(a) shows $\langle N'_r(v) \rangle$, i.e. the average of $N'_r(v)$ over all v in G , as a function of r , forming a straight line when plotted in doubly logarithmic scale. Based on this plot we can infer that it is the power-law scaling that has induced the urban correlation observed in Fig. 4. The scaling dimension, equivalent to the slope estimate, is $\alpha = \alpha(G) = 2.728$. This also implies that we need at least three-dimensional space in which to embed the neighbourhoods. Notice the intercept is uninteresting as the fitting line must pass through the origin for any G .

Then, from equation (17) and (18), it can be deduced that the average of mean depth over v should increase following $\langle \bar{d}_r(v) \rangle = 0.73 r + 0.5 - 0.68/r$ and $\langle \bar{d}_r(v) \rangle = 0.73 N'_r(v)^{0.37} + 0.5 - 0.68/r$. In Fig. 5(b) and (c), we compare the expected values of mean depth given by those formulae with the observed values over the entire range of radii. They are in good agreement in the local (white region), while diverge widely in the global due to the edge effect. The graph as a whole has therefore a structure that is far from being random or small-world. Importantly, this implies, against common critics, that axial graphs somehow internalise geometric or geographical constraints of the space from which it has been drawn out (Hillier, 1999).

Although average neighbourhood size scales with radius following the power-law, there are important individual variations that cannot be treated as mere exceptions. In other words, the existence of scaling dimension $\alpha(G)$ at the collective level does not guarantee the

Figure 4:

The urban correlation in London. (a)

$$D_3(v) = 2.67 N_3(v)$$

($R^2 = 0.997$). (b)

$$D_9(v) = 7.17 N_9(v)$$

($R^2 = 0.994$). The images show the spatial variations of mean depth, where darker lines indicate higher values

existence of scaling dimension $\alpha(v)$ at the individual level (although the inverse is true). We thus now need to investigate further scaling behaviour starting from each node.

093-12

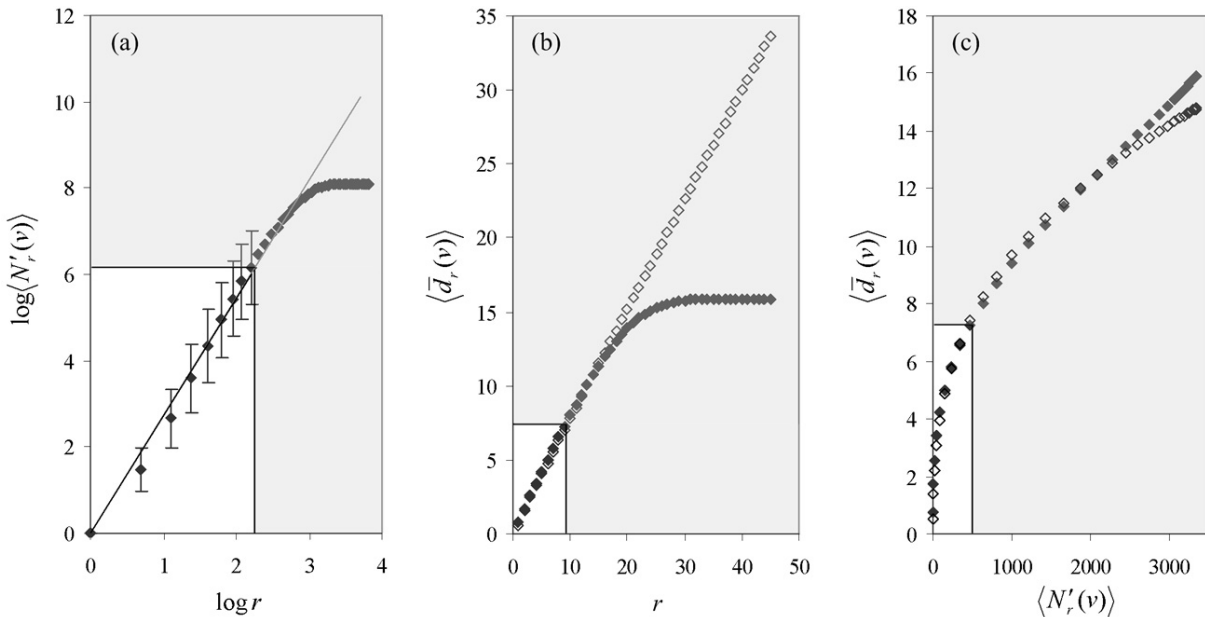


Figure 5:

(a) The power-law scaling between radius and average neighbourhood size. (b) Average mean depth as a function of radius. (c) Average mean depth as a function of neighbourhood size (●: the observed values, ○: the expected values)

Fig. 6(a)-(c) show $N_r(v)$ as a function of r , also plotted in double logarithmic scale, for some typical nodes of Inner London. They represent three distinct types: (a) one which follows the power law (straight line); (b) one which follows the exponential law (downward curvature); (c) one which we may call ‘super-power’ scaling as it lies well beyond what can be predicted by the power-law (upward curvature). We estimate about 62% belongs to the first type, 26% to the second, and 12% to the third^{vi}.

For the first type with the power-law scaling, we find that the scaling dimension $\alpha(v)$ of v is $(2.66, 0.40^2)$ normally distributed, which traps safely the scaling dimension $\alpha(G) = 2.73$ of G as a whole. On the other hand, for the second type with the exponential scaling, we find $\beta(v)$ is $(0.77, 0.13^2)$ normally distributed. We can also see from Fig. 6(d)-(e) that the first-type nodes form a well-connected body of the street network, while the second-type nodes are scattered as rather isolated islands of different size (e.g. Primrose Hill, Barbican etc.)^{vii}. At least, this seems to tell us that a mere collection of ‘small-worlds’ cannot build a city.

The neighbourhood size of the first and the second-type nodes can be then estimated using the formulae (11) and (12) respectively. Fig. 7(a) compares those observed values with the estimated, when $r = 3$ (without the edge effects). As expected, the agreements are good enough for both types of nodes (power: $R^2 = 0.83$; exponential:). This means that neighbourhood size can be predicted, to some significant extent, only by knowing degree under each scaling law. Note also that the neighbourhood size of the second-type nodes is in average considerably smaller than that of the first-type.

The solid points are translated vertically for better visual clarity.

Next, Fig. 7(b) compares the observed values of mean depth with those estimated by the formulae (15) and (17) respectively. They are

relatively well-fitted with the power scaling ($R^2 = 0.76$), while highly fluctuant with the exponential-scaling ($R^2 = 0.43$). Moreover, note that the mean depth of the second-type nodes is in average well below the common value of mean depth (indicated by vertical lines). This means that the exponential scaling is mainly responsible for the dissipation of the urban correlation by inducing minute variations to it.

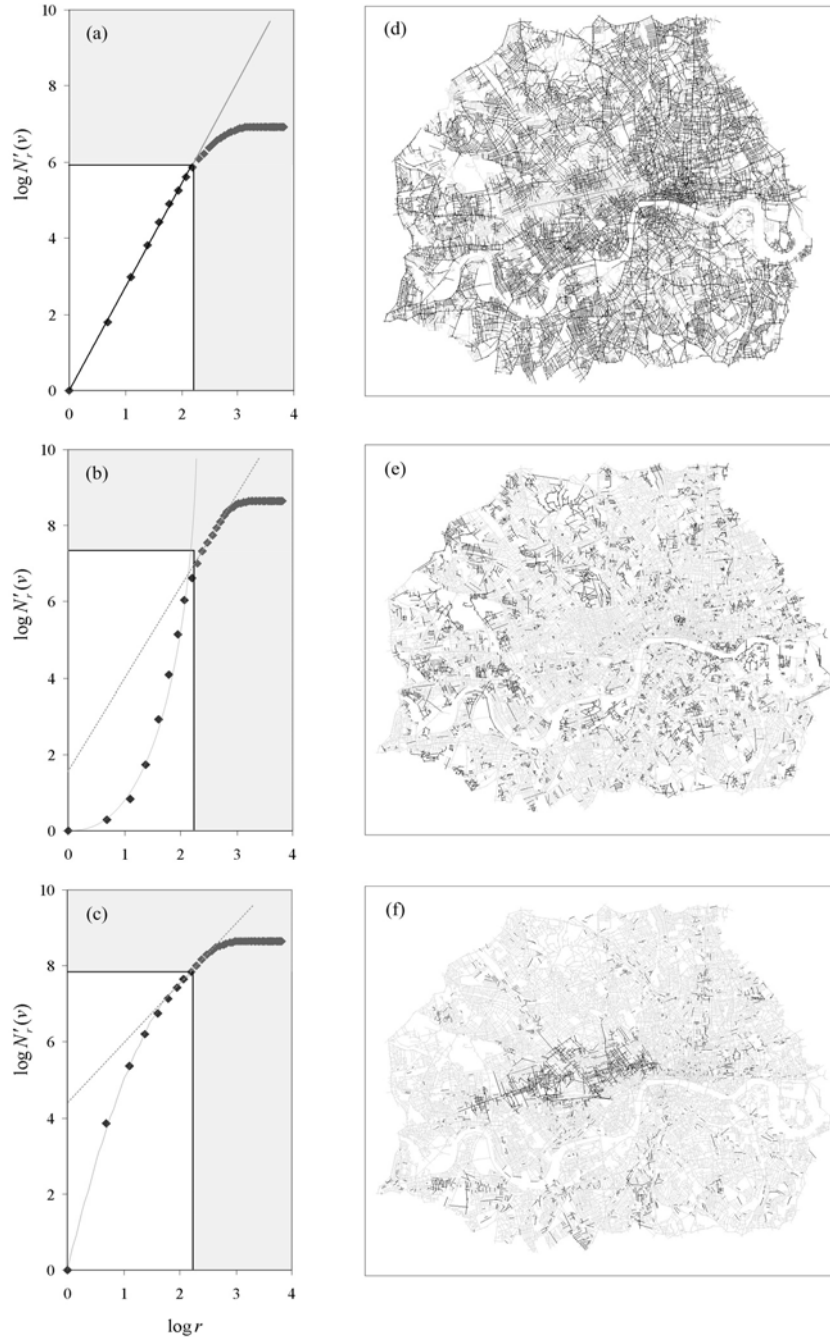
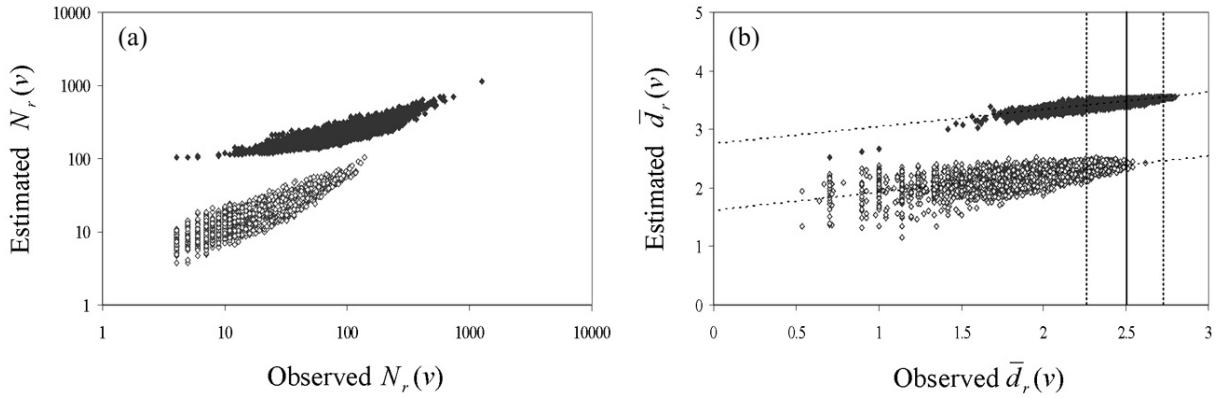


Figure 6:

The spatial distribution of individual nodes following (a) the power-law (62%) (b) the exponential-law (26%), and (c) the super-power scaling (12%)

093-13

For the third type of the super-power scaling, neither the power nor the exponential law can explain the actual data in any meaningful way. But at least we can observe that the nodes of this type tend to have the highest values for both neighbourhood size and mean depth. This happens due to the overwhelming dominance of close nodes over distant nodes. Fig. 6(f) shows the third-type nodes are mainly found in the areas close to the arterial streets with high degree, so that easily accessible to other areas, without being themselves arterial streets (e.g. local distributors directly connected to Oxford Street).



093-14

Figure 7:

The estimated and the observed values of neighbourhood size under each scaling law. (b) The estimated and the observed values of mean depth under each scaling law (the power-law: ●, the exponential-law: ○). Both are for $r = 3$

Assortativity by Degree

Fig. 8(a) shows the degree-degree correlation plotted in double logarithmic scale. It is found that $\phi = -0.039$, meaning that the network in question is slightly dissortative. However, such dissortativity has been caused mainly by nodes with degree 1 (i.e. no through streets) as they tend to be connected to nodes with much higher degree. If we disregard those outliers, it will be safe to conclude that the street network of London is neutral. This is also evidenced by studying separately the relationship between each scaling-type of streets and its degree correlation. As shown in Fig. 8(a), no particular type displays any significant positive or negative degree correlation. We also confirm from Fig. 8(b) that the spatial distribution of relative control, i.e. $1/H_m(v)$, does not seem to reveal any interesting structural features of the network, except that those segregated streets tend naturally to have higher control values. All these results tell us that, in terms of local connections, the street network of London is quite homogeneous.

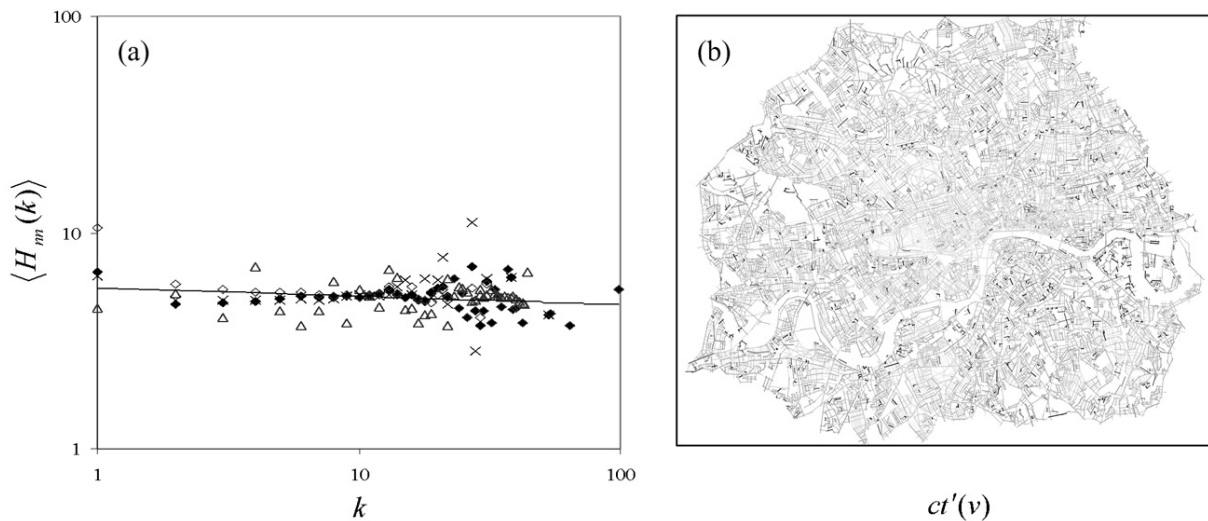


Figure 8:

(a) The degree-degree correlation of London, showing no tendency of the degree-degree correlation (●: all types, ○: the power-law, Δ: the exponential-law, ×: the super-power). (b) The spatial distribution of relative control $ct'(v)$

Discussion

First, we have shown through the case study of London that the urban correlation, redefined in this paper as the homogenisation of mean depth, is the normal state of affairs in local analysis. Accordingly, it is suggested that we may utilise the urban correlation as a method of detecting the presence of the edge effects and thus of delimiting the domain of local analysis. Radius-radius, employed so far as a rule of thumb for the same purpose, is not always sufficiently small to eliminate the edge effects, so that warrants more cautious use. It is not because those effects could not have anything to do with realities

but because they make measures lose their consistent meanings altogether.

However, to put too much emphasis on the urban correlation itself is perhaps misleading, since it tends to conceal the truly local differences underneath the apparent homogeneity. Those differences are represented by three distinct types of scaling laws the sequence of each neighbourhood size follow with increasing radius. In the case of London, it is the power-law that is the dominant type of scaling. But other minor types of scaling cannot be simply neglected as they are found to articulate some interesting features of networks, such as the existence of 'small-worlds' in a city. Taken together, the minor types of scaling are observed to induce numerically minute but typically heterogeneous variations against what seems otherwise quite homogeneous. Without such variations, it would be indeed hard to reject the assumption of the primacy of degree in local analysis.

093-15

We have also reported that the street network of London is neutral in terms of the degree-degree correlation, redefined in this paper as the homogenisation of control. This indicates that the network as a whole possesses neither a unifying core nor overarching hierarchical structure. This observation is valuable as we believe that the absence of assortativity is the least necessary condition for multi-centrality, or what Hillier (2007) envisage as 'periodical area-isation', to emerge in networks. Also, as Xulvi-Brunet and Sokolov (2005)'s previous work suggests, such neutral structure facilitates necessary communication or movement in the network and thus may reflect a long-term optimisation process with which the city has evolved. So far the idea of neutrality has been generally associated with randomness; but random networks do never exhibit the power-law scaling. 'Being neutral without being random' – this is, in effect, a phenomenon that we find specifically 'urban.'

References

- Ben-Avraham, D., Rozenfeld, A.F., Cohen, R., Havlin, S., 2003, "Geographical Embedding of Scale-Free Networks", *Physica A*, Vol. 330, pp. 107.
- Cantanzaro, M, Caldarelli, G., Pietronero, L 2004, "Assortative Model for Social Networks", *Physical Review E*, Vol. 70:037101.
- Csányi, G., Szendroi, B., 2004, "The Fractal/Small-World Dichotomy in Real-World Networks", *Physical Review E*, Vol. 70: 016122.
- Dalton, N., 2005, "Toward General Relativisation in Space Syntax", A. van Nes (Ed.), *Proceedings*, 5th International Space Syntax Symposium, Delft University of Technology.
- Freeman, L.C., 1979, "Centrality in Social Networks: Conceptual Classification", *Social Network*, Vol. 1, pp.215-239.
- Hillier, B., Hanson, J, 1984, *The Social Logic of Space*, Cambridge University Press, Cambridge.
- Hillier, B., 1999, "The Hidden Geometry of Deformed Grids: Or, Why Space Syntax Works, When it Looks as Though it Shouldn't", *Environment and Planning B: Planning and Design*, Vol. 26, pp.169-191.
- Hillier, B., Turner, A., Yang, T., Park, H. T., 2007, "Metric and Topo-Geometric Properties of Urban Street Networks: Some Convergences, Divergence, and New Results", A.S. Kubat, Ö. Ertekin, Y.İ. Güney, E. Eyüboğlu (Eds.), *Proceedings*, 6th International Space Syntax Symposium, Istanbul Technical University, Cenkler, Istanbul, pp.001.01-22.
- Li, L., Alderson, D., Tanaka, R., Doyle, J.C., Willinger, W., 2005, "Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications (Extended Version)", arXiv.org:cond-mat/0501169.
- Newman, M.E.J., 2002, "Assortative Mixing in Networks", *Physical Review Letter*, Vol. 89: 208701.

Nowotny, T., Requardt, M., 1998, "Dimension Theory of Graphs and Networks", *Journal of Physics A*, Vol. 31: 2447.

Park, H.T., 2005, "Before Integration: A Critical Review of Integration Measure in Space Syntax", A. van Nes (Ed.), *Proceedings, 5th International Space Syntax Symposium*, Delft University of Technology, Delft.

Pastor-Satorras, R., Vazquez, A., Vespignani, A., 2001, "Dynamical and Correlation Properties of the Internet", *Physical Review Letters*, Vol. 87:258701.

Penn, A., Hillier, B., Banister, D., Xu, J., 1998, "Configurational Modelling of Urban Movement Networks", *Environment and Planning B: Planning and Design*, Vol. 25, pp.59-84.

Porta, S., Crucitti, P., Latora, V., 2004, "The Network Analysis of Urban Streets: A Dual Approach", arXiv.org:cond-mat/0411241.

Soffer, S.N., Vázquez, A., 2005, "Clustering Coefficient without Degree Correlations Biases", *Physical Review E*, Vol. 71:057101.

Turner, A., 2007, "Depthmap", ver. 7.0116b, Spatial Network Analysis Software, University College London, London.

Watt, D.J., Strogatz, S.H., 1998, "Collective Dynamics of 'Small World' Network", *Nature*, Vol. 393:440.

Xulvi-Brunet, R. Sokolov, I., 2005, "Changing Correlations in Networks: Assortativity and Dissortativity", *ACTA Physica Polonica B*, Vol. 36, no. 5.

Yang, T., Hillier, B., 2007, "The Fuzzy Boundary: The Spatial Definition Of Urban Areas", A.S. Kubat, Ö. Ertekin, Y.I. Güney, E. Eyüboğlu (Eds.), *Proceedings, 6th International Space Syntax Symposium*, Istanbul Technical University, Cenkler, Istanbul, pp.001.01-16.

- i. $N_r(v)$ is given directly by Depthmap, a space syntax software (Turner 2007), under the field name of "Node Count." We can then deduce $n_r(v)$ easily according to the equation (1).
- ii. However, we point out here that the determination of the local in this way is contingent upon that of the global.
- iii. By substituting (11) into (4), we have:

$$\begin{aligned} \bar{d}_r(v) &= r - \sum_{d=0}^{r-1} \frac{N_1(v)d^\alpha}{N_1(v)r^\alpha} \\ &= r - \frac{1}{r^\alpha} \sum_{d=0}^{r-1} d^\alpha \end{aligned}$$

It is mathematically hard to express $\sum_{d=0}^{r-1} d^\alpha$ in the above equation in terms of r and α . But at least we know its lower and upper bounds from replacing it with integrals:

$$\sum_{d=0}^{r-1} d^\alpha \geq \int_0^{r-1} d^\alpha \, dd = \frac{1}{\alpha+1} (r-1)^{\alpha+1}$$

$$\sum_{d=0}^{r-1} d^\alpha \leq \int_0^r d^\alpha \, dd = \frac{1}{\alpha+1} r^{\alpha+1}$$

With these bounds we get:

$$\frac{1}{\alpha+1} r \left(1 - \frac{1}{r}\right)^{\alpha+1} \leq \frac{1}{r^\alpha} \sum_{d=1}^{r-1} d^\alpha \leq \frac{1}{\alpha+1} r$$

Hence, for mean depth:

$$r - \frac{1}{\alpha + 1}r \leq \bar{d}_r(v) \leq r - \frac{1}{\alpha + 1}r \left(1 - \frac{1}{r}\right)^{\alpha + 1}$$

$$\Rightarrow \frac{\alpha}{\alpha + 1}r \leq \bar{d}_r(v) \leq \frac{\alpha}{\alpha + 1}r + \left(1 - \frac{\alpha}{2r}\right)$$

where, for the upper bound, we have expanded out $(1 - 1/r)^{\alpha + 1}$ to include only the first three terms. Consequently, we estimate the values of mean depth by the midpoint of the interval.

iv. Similarly, by substituting (12) to (4), we have:

$$\bar{d}_r(v) = r - \frac{\sum_{d=0}^{r-1} N_1(v)e^{\beta(d-1)}}{\sum_{d=0}^{r-1} N_1(v)e^{\beta(r-1)}}$$

$$= r - \frac{1}{e^{\beta r}} \sum_{d=0}^{r-1} e^{\beta d}$$

$$= r - \frac{1}{e^{\beta r}} \cdot \frac{1 - e^{\beta r}}{1 - e^{\beta}}$$

093-17

- v. In the traditional model of preferential attachment for scale-free graphs it has been assumed that the probability of a new node v to be connected to a previous node u is a function of the degree of u only, independently of the degree of v . The idea of assortativity has arisen by calling this assumption into question, and considers the degrees of v and u at the same time in measuring the probability of an edge $e = uv$. With the harmonic mean, however, it is somewhat awkward to express such connection probabilities.
- vi. For all individual nodes, the r-squared values are calculated from the linear regression of (1) $\log N'_r(v)$ on $\log r$ (the power-law), (2) $\log N'_r(v)$ on r (the exponential-law), and (3) $\log N'_r(v)$ on $\log(\log r)$ (the super-power). We can then assign a type of scaling to each node by choosing the highest value of r-squared.
- vii. In segment maps equipped with a metric distance function, this second-type nodes tend to produce a more distinct pattern of "patchworks." See Yang & Hillier (2007), in which they study the rate of change of neighbourhood size (or node count). Notably, the second-type nodes with the exponential scaling will have the lowest rate of change at small radii, while the third-type with the super-power scaling will have the highest. However, such a difference in the rate of change is reversed at large radii. The similar relationship between the rate of change of neighbourhood size and metric mean depth is also reported in Hillier et al. (2007).

093-18